HANDBOOK OF DIFFERENTIAL EQUATIONS

ORDINARY DIFFERENTIAL EQUATIONS

VOLUME I

Contents

Preface v
List of Contributors vii

1. A survey of recent results for initial and boundary value problems singular in the dependent variable 1
   R.P. Agarwal and D. O'Regan
2. The lower and upper solutions method for boundary value problems 69
   C. De Coster and P. Habets
3. Half-linear differential equations 161
   O. Došlý
4. Radial solutions of quasilinear elliptic differential equations 359
   J. Jacobsen and K. Schmitt
5. Integrability of polynomial differential systems 437
   J. Llibre
6. Global results for the forced pendulum equation 533
   J. Mawhin
7. Ważewski method and Conley index 591
   R. Srzednicki

Author Index 685
Subject Index 693

CHAPTER 1

A Survey of Recent Results for Initial and Boundary Value Problems Singular in the Dependent Variable

Ravi P. Agarwal

Department of Mathematics, Florida Institute of Technology, Melbourne, FL 32901, USA
E-mail: agarwal@fit.edu

Donal O'Regan

Department of Mathematics, National University of Ireland, Galway, Ireland
E-mail: donal.oregan@nuigalway.ie

Contents

1. Introduction . . . 3
2. Singular boundary value problems . . . 8
   2.1. Positone problems . . . 8
   2.2. Singular problems with sign changing nonlinearities . . . 22
3. Singular initial value problems . . . 51
References . . . 67

Abstract

In this survey paper we present old and new existence results for singular initial and boundary value problems. Our nonlinearity may be singular in its dependent variable and is allowed to change sign.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 1
Edited by A. Cañada, P. Drábek and A. Fonda

© 2004 Elsevier B.V. All rights reserved

1. Introduction

The study of singular boundary value problems (singular in the dependent variable) is relatively new. Indeed it was only in the middle 1970s that researchers realized that large numbers of applications [7,11,12] in the study of nonlinear phenomena gave rise to singular boundary value problems (singular in the dependent variable). However, in our opinion, it was the 1979 paper of Taliaferro [20] that generated the interest of many researchers in singular problems in the 1980s and 1990s. In [20] Taliaferro showed that the singular boundary value problem

$$\begin{cases} y'' + q(t)y^{-\alpha} = 0, & 0 < t < 1, \\ y(0) = 0 = y(1), \end{cases} \tag{1.1}$$

has a $C[0,1] \cap C^1(0,1)$ solution; here $\alpha > 0$, $q \in C(0,1)$ with $q > 0$ on $(0,1)$ and $\int_0^1 t(1-t)q(t)\,\mathrm{d}t < \infty$. Problems of the form (1.1) arise frequently in the study of nonlinear phenomena, for example in non-Newtonian fluid theory, such as the transport of coal slurries down conveyor belts [12], and boundary layer theory [11]. It is worth remarking here that we could consider Sturm–Liouville boundary data in (1.1); however since the arguments are essentially the same (in fact easier) we will restrict our discussion to Dirichlet boundary data.

In the 1980s and 1990s many papers were devoted to singular boundary value problems of the form

$$\begin{cases} y'' + q(t)f(t,y) = 0, & 0 < t < 1, \\ y(0) = 0 = y(1), \end{cases} \tag{1.2}$$

and singular initial value problems of the form

$$\begin{cases} y' = q(t)f(t,y), & 0 < t < T(< \infty), \\ y(0) = 0. \end{cases} \tag{1.3}$$

Almost all singular problems in the literature [8–10,14–18,21] up to 1994 discussed positone problems, i.e., problems where $f : [0,1] \times (0,\infty) \to (0,\infty)$. In Section 2.1 we present the most general results available in the literature for the positone singular problem (1.2). In 1999 the question of multiplicity for positone singular problems was discussed for the first time by Agarwal and O'Regan [2]. The second half of Section 2.1 discusses multiplicity. In 1994 [16] the singular boundary value problem (1.2) was discussed when the nonlinearity f could change sign. Model examples are

$$f(t,y) = t^{-1}\mathrm{e}^{\frac{1}{y}} - (1-t)^{-1} \quad \text{and} \quad f(t,y) = \frac{g(t)}{y^\sigma} - h(t), \quad \sigma > 0$$

which correspond to Emden–Fowler equations; here $g(t) > 0$ for $t \in (0,1)$ and $h(t)$ may change sign. Section 2.2 is devoted to (1.2) when the nonlinearity f may change sign. The results here are based on arguments and ideas of Agarwal, O'Regan et al. [1–6], and Habets and Zanolin [16]. Section 3 presents existence results for the singular initial value problem (1.3) where the nonlinearity f may change sign.

The existence results in this paper are based on fixed point theorems. In particular we use frequently Schauder's fixed point theorem and a Leray–Schauder alternative. We begin of course with the Schauder theorem.

THEOREM 1.1. *Let C be a convex subset of a Banach space and $F : C \to C$ a compact, continuous map. Then F has a fixed point in C.*

In applications to construct a set C so that F takes C back into C is very difficult and sometimes impossible. As a result it makes sense to discuss maps F that map a subset of C into C. One result in this direction is the so-called nonlinear alternative of Leray–Schauder.

THEOREM 1.2. *Let E be a Banach space, C a convex subset of E, U an open subset of C and $0 \in U$. Suppose $F : \overline{U} \to C$ (here $\overline{U}$ denotes the closure of U in C) is a continuous, compact map. Then either*

(A1) *F has a fixed point in $\overline{U}$; or*

(A2) *there exists $u \in \partial U$ (the boundary of U in C) and $\lambda \in (0, 1)$ with $u = \lambda F(u)$.*

PROOF. Suppose (A2) does not occur and F has no fixed points in ∂U (otherwise we are finished). Let

$$A = \{x \in \overline{U}\colon\ x = tF(x) \text{ for some } t \in [0, 1]\}.$$

Now $A \neq \emptyset$ since $0 \in A$ and A is closed since F is continuous. Also notice $A \cap \partial U = \emptyset$. Thus there exists a continuous function $\mu : \overline{U} \to [0, 1]$ with $\mu(A) = 1$ and $\mu(\partial U) = 0$. Let

$$N(x) = \begin{cases} \mu(x)F(x), & x \in \overline{U}, \\ 0, & x \in C \backslash \overline{U}. \end{cases}$$

Clearly $N : C \to C$ is a continuous, compact map. Theorem 1.1 guarantees the existence of an $x \in C$ with $x = Nx$. Notice $x \in U$ since $0 \in U$. As a result $x = \mu(x)F(x)$, so $x \in A$. Thus $\mu(x) = 1$ and so $x = F(x)$. □

To conclude the introduction we present existence principles for nonsingular initial and boundary value problems which will be needed in Sections 2 and 3. First we use Schauder's fixed point theorem and a nonlinear alternative of Leray–Schauder type to obtain a general existence principle for the Dirichlet boundary value problem

$$\begin{cases} y'' + f(t, y) = 0, & 0 < t < 1, \\ y(0) = a,\ y(1) = b. \end{cases} \tag{1.4}$$

Throughout this paper $AC[0, 1]$ denotes the space of absolutely continuous functions on $[0, 1]$, $AC_{\mathrm{loc}}(0, 1)$ the space of functions absolutely continuous on each compact subinterval of $(0, 1)$ and $L^1_{\mathrm{loc}}(0, 1)$ the space of functions which are L^1 integrable on each compact subinterval of $(0, 1)$.

THEOREM 1.3. *Suppose the following two conditions are satisfied*:

$$\text{the map } y \mapsto f(t,y) \text{ is continuous for a.e. } t \in [0,1] \tag{1.5}$$

and

$$\text{the map } t \mapsto f(t,y) \text{ is measurable for all } y \in \mathbf{R}. \tag{1.6}$$

(I) *Assume*

$$\begin{cases} \text{for each } r > 0 \text{ there exists } h_r \in L^1_{\mathrm{loc}}(0,1) \text{ with} \\ \int_0^1 t(1-t)h_r(t)\,\mathrm{d}t < \infty \text{ such that } |y| \leqslant r \text{ implies} \\ \bigl|f(t,y)\bigr| \leqslant h_r(t) \text{ for a.e. } t \in (0,1) \end{cases} \tag{1.7}$$

holds. In addition suppose there is a constant $M > |a| + |b|$, *independent of* λ, *with* $|y|_0 = \sup_{t\in[0,1]} |y(t)| \neq M$ *for any solution* $y \in AC[0,1]$ (*with* $y' \in AC_{\mathrm{loc}}(0,1)$) *to*

$$\begin{cases} y'' + \lambda f(t,y) = 0, \quad 0 < t < 1, \\ y(0) = a, \; y(1) = b, \end{cases} \tag*{$(1.8)_\lambda$}$$

for each $\lambda \in (0,1)$. *Then* (1.4) *has a solution* y *with* $|y|_0 \leqslant M$.

(II) *Assume*

$$\begin{cases} \text{there exists } h \in L^1_{\mathrm{loc}}(0,1) \text{ with } \int_0^1 t(1-t)h_r(t)\,\mathrm{d}t < \infty \\ \text{such that } \bigl|f(t,y)\bigr| \leqslant h(t) \text{ for a.e. } t \in (0,1) \text{ and } y \in \mathbf{R} \end{cases} \tag{1.9}$$

holds. Then (1.4) *has a solution.*

PROOF. (I) We begin by showing that solving $(1.8)_\lambda$ is equivalent to finding a solution $y \in C[0,1]$ to

$$\begin{aligned} y(t) = {} & a(1-t) + bt + \lambda(1-t)\int_0^t sf\bigl(s,y(s)\bigr)\,\mathrm{d}s \\ & + \lambda t \int_t^1 (1-s) f\bigl(s,y(s)\bigr)\,\mathrm{d}s. \end{aligned} \tag*{$(1.10)_\lambda$}$$

To see this notice if $y \in C[0,1]$ satisfies $(1.10)_\lambda$ then it is easy to see (since (1.7) holds) that $y' \in L^1[0,1]$. Thus $y \in AC[0,1]$, $y' \in AC_{\mathrm{loc}}(0,1)$ and note

$$y'(t) = -a + b - \lambda \int_0^t sf\bigl(s,y(s)\bigr)\,\mathrm{d}s + \lambda \int_t^1 (1-s) f\bigl(s,y(s)\bigr)\,\mathrm{d}s.$$

Next integrate $y'(t)$ from 0 to x ($x \in (0,1)$) and interchange the order of integration to get

$$y(x) - y(0) = \int_0^x y'(t)\,\mathrm{d}t$$

$$
\begin{aligned}
&= -ax + bx - \lambda \int_0^x \int_0^t s f\big(s, y(s)\big)\, \mathrm{d}s\, \mathrm{d}t \\
&\quad + \lambda \int_0^x \int_t^1 (1-s) f\big(s, y(s)\big)\, \mathrm{d}s\, \mathrm{d}t \\
&= -ax + bx + \lambda(1-x) \int_0^x s f\big(s, y(s)\big)\, \mathrm{d}s \\
&\quad + \lambda x \int_x^1 (1-s) f\big(s, y(s)\big)\, \mathrm{d}s \\
&= -a + y(x),
\end{aligned}
$$

so $y(0) = a$. Similarly integrate $y'(t)$ from x ($x \in (0,1)$) to 1 and interchange the order of integration to get $y(1) = b$. Thus if $y \in C[0,1]$ satisfies $(1.10)_\lambda$ then y is a solution of $(1.8)_\lambda$.

Define the operator $N : C[0,1] \to C[0,1]$ by

$$
\begin{aligned}
Ny(t) &= a(1-t) + bt + (1-t) \int_0^t s f\big(s, y(s)\big)\, \mathrm{d}s \\
&\quad + t \int_t^1 (1-s) f\big(s, y(s)\big)\, \mathrm{d}s.
\end{aligned} \tag{1.11}
$$

Then $(1.10)_\lambda$ is equivalent to the fixed point problem

$$
y = (1-\lambda)p + \lambda N y, \quad \text{where } p = a(1-t) + b. \tag*{$(1.12)_\lambda$}
$$

It is easy to see that $N : C[0,1] \to C[0,1]$ is continuous and completely continuous. Set

$$
U = \big\{u \in C[0,1]:\ |u|_0 < M\big\}, \qquad K = E = C[0,1].
$$

Now the nonlinear alternative of Leray–Schauder type guarantees that N has a fixed point, i.e., $(1.10)_1$ has a solution.

(II) Solving (1.4) is equivalent to the fixed point problem $y = Ny$ where N is as in (1.11). It is easy to see that $N : C[0,1] \to C[0,1]$ is continuous and compact (since (1.9) holds). The result follows from Schauder's fixed point theorem. □

Finally we obtain a general existence principle for the initial value problem

$$
\begin{cases} y' = f(t, y), & 0 < t < T(< \infty), \\ y(0) = a. \end{cases} \tag{1.13}
$$

THEOREM 1.4. *Suppose the following two conditions are satisfied*:

$$
\textit{the map } y \mapsto f(t, y) \textit{ is continuous for a.e. } t \in [0, T] \tag{1.14}
$$

and

$$\text{the map } t \mapsto f(t, y) \text{ is measurable for all } y \in \mathbf{R}. \tag{1.15}$$

(I) *Assume*

$$\begin{cases} \text{for each } r > 0 \text{ there exists } h_r \in L^1[0, T] \text{ such that} \\ |y| \leqslant r \text{ implies } \big|f(t, y)\big| \leqslant h_r(t) \text{ for a.e. } t \in (0, T) \end{cases} \tag{1.16}$$

holds. In addition suppose there is a constant $M > |a|$, *independent of* λ, *with* $|y|_0 = \sup_{t\in[0,T]} |y(t)| \neq M$ *for any solution* $y \in AC[0, T]$ *to*

$$\begin{cases} y' = \lambda f(t, y), \quad 0 < t < T(< \infty), \\ y(0) = a, \end{cases} \tag*{(1.17)$_\lambda$}$$

for each $\lambda \in (0, 1)$. *Then* (1.13) *has a solution* y *with* $|y|_0 \leqslant M$.

(II) *Assume*

$$\begin{cases} \text{there exists } h \in L^1[0, T] \text{ such that } \big|f(t, y)\big| \leqslant h(t) \\ \text{for a.e. } t \in (0, T) \text{ and } y \in \mathbf{R} \end{cases} \tag{1.18}$$

holds. Then (1.13) *has a solution.*

PROOF. (I) Solving $(1.17)_\lambda$ is equivalent to finding a solution $y \in C[0, T]$ to

$$y(t) = a + \lambda \int_0^t f\big(s, y(s)\big)\, \mathrm{d}s. \tag*{(1.19)$_\lambda$}$$

Define an operator $N : C[0, T] \to C[0, T]$ by

$$Ny(t) = a + \int_0^t f\big(s, y(s)\big)\, \mathrm{d}s. \tag{1.20}$$

Then $(1.19)_\lambda$ is equivalent to the fixed point problem

$$y = (1 - \lambda)a + \lambda Ny. \tag*{(1.21)$_\lambda$}$$

It is easy to see that $N : C[0, T] \to C[0, T]$ is continuous and completely continuous. Set

$$U = \big\{u \in C[0, T]\colon\ |u|_0 < M\big\}, \qquad K = E = C[0, T].$$

Now the nonlinear alternative of Leray–Schauder type guarantees that N has a fixed point, i.e., $(1.19)_1$ has a solution.

(II) Solving (1.13) is equivalent to the fixed point problem $y = Ny$ where N is as in (1.20). It is easy to see that $N : C[0, T] \to C[0, T]$ is continuous and compact (since (1.18) holds). The result follows from Schauder's fixed point theorem. □

2. Singular boundary value problems

In Section 2.1 we discuss positone boundary value problems. Almost all singular papers in the 1980s and 1990s were devoted to such problems. In Theorem 2.1 we present probably the most general existence result available in the literature for positone problems. In the late 1990s the question of multiplicity for singular positone problems was raised, and we discuss this question in the second half of Section 2.1. Section 2.2 is devoted to singular problems where the nonlinearity may change sign.

2.1. *Positone problems*

In this section we discuss the Dirichlet boundary value problem

$$\begin{cases} y'' + q(t)f(t,y) = 0, & 0 < t < 1, \\ y(0) = 0 = y(1). \end{cases} \tag{2.1}$$

Here the nonlinearity f may be singular at $y = 0$ and q may be singular at $t = 0$ and/or $t = 1$. We begin by showing that (2.1) has a $C[0,1] \cap C^2(0,1)$ solution. To do so we first establish, via Theorem 1.3, the existence of a $C[0,1] \cap C^2(0,1)$ solution, for each $m = 1, 2, \dots$, to the "modified" problem

$$\begin{cases} y'' + q(t)f(t,y) = 0, & 0 < t < 1, \\ y(0) = \frac{1}{m} = y(1). \end{cases} \tag*{$(2.2)^m$}$$

To show that (2.1) has a solution we let $m \to \infty$; the key idea in this step is the Arzela–Ascoli theorem.

THEOREM 2.1. *Suppose the following conditions are satisfied*:

$$q \in C(0,1), \quad q > 0 \quad \textit{on } (0,1) \quad \textit{and} \quad \int_0^1 t(1-t)q(t)\,\mathrm{d}t < \infty, \tag{2.3}$$

$$f : [0,1] \times (0,\infty) \to (0,\infty) \textit{ is continuous.} \tag{2.4}$$

$$\begin{cases} 0 \leqslant f(t,y) \leqslant g(y) + h(y) \textit{ on } [0,1] \times (0,\infty) \textit{ with} \\ g > 0 \textit{ continuous and nonincreasing on } (0,\infty), \\ h \geqslant 0 \textit{ continuous on } [0,\infty), \textit{ and } \frac{h}{g} \\ \textit{nondecreasing on } (0,\infty), \end{cases} \tag{2.5}$$

$$\begin{cases} \textit{for each constant } H > 0 \textit{ there exists a function } \psi_H \\ \textit{continuous on } [0,1] \textit{ and positive on } (0,1) \textit{ such that} \\ f(t,u) \geqslant \psi_H(t) \textit{ on } (0,1) \times (0,H] \end{cases} \tag{2.6}$$

and

$$\exists r > 0 \quad \textit{with } \frac{1}{\{1 + \frac{h(r)}{g(r)}\}} \int_0^r \frac{\mathrm{d}u}{g(u)} > b_0 \tag{2.7}$$

hold; *here*

$$b_0 = \max\left\{2\int_0^{\frac{1}{2}} t(1-t)q(t)\,\mathrm{d}t,\, 2\int_{\frac{1}{2}}^1 t(1-t)q(t)\,\mathrm{d}t\right\}. \tag{2.8}$$

Then (2.1) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y > 0$ *on* $(0,1)$ *and* $|y|_0 < r$.

PROOF. Choose $\varepsilon > 0$, $\varepsilon < r$, with

$$\frac{1}{\{1+\frac{h(r)}{g(r)}\}}\int_\varepsilon^r \frac{\mathrm{d}u}{g(u)} > b_0. \tag{2.9}$$

Let $n_0 \in \{1,2,\ldots\}$ be chosen so that $\frac{1}{n_0} < \varepsilon$ and let $N_0 = \{n_0, n_0+1,\ldots\}$. To show $(2.2)^m$, $m \in N_0$, has a solution we examine

$$\begin{cases} y'' + q(t)F(t,y) = 0, & 0 < t < 1,\\ y(0) = y(1) = \frac{1}{m}, & m \in N_0, \end{cases} \tag*{$(2.10)^m$}$$

where

$$F(t,u) = \begin{cases} f(t,u), & u \geqslant \frac{1}{m},\\ f\big(t,\frac{1}{m}\big), & u \leqslant \frac{1}{m}. \end{cases}$$

To show $(2.10)^m$ has a solution for each $m \in N_0$ we will apply Theorem 1.3. Consider the family of problems

$$\begin{cases} y'' + \lambda q(t)F(t,y) = 0, & 0 < t < 1,\\ y(0) = y(1) = \frac{1}{m}, & m \in N_0, \end{cases} \tag*{$(2.11)^m_\lambda$}$$

where $0 < \lambda < 1$. Let y be a solution of $(2.11)^m_\lambda$. Then $y'' \leqslant 0$ on $(0,1)$ and $y \geqslant \frac{1}{m}$ on $[0,1]$. Also there exists $t_m \in (0,1)$ with $y' \geqslant 0$ on $(0,t_m)$ and $y' \leqslant 0$ on $(t_m,1)$. For $x \in (0,1)$ we have

$$-y''(x) \leqslant g\big(y(x)\big)\left\{1 + \frac{h(y(x))}{g(y(x))}\right\}q(x). \tag{2.12}$$

Integrate from $t\,(t \leqslant t_m)$ to t_m to obtain

$$y'(t) \leqslant g\big(y(t)\big)\left\{1 + \frac{h(y(t_m))}{g(y(t_m))}\right\}\int_t^{t_m} q(x)\,\mathrm{d}x$$

and then integrate from 0 to t_m to obtain

$$\int_{\frac{1}{m}}^{y(t_m)} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(y(t_m))}{g(y(t_m))}\right\}\int_0^{t_m} xq(x)\,\mathrm{d}x.$$

Consequently

$$\int_{\varepsilon}^{y(t_m)} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(y(t_m))}{g(y(t_m))}\right\} \int_0^{t_m} xq(x)\,\mathrm{d}x$$

and so

$$\int_{\varepsilon}^{y(t_m)} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(y(t_m))}{g(y(t_m))}\right\} \frac{1}{1-t_m} \int_0^{t_m} x(1-x)q(x)\,\mathrm{d}x. \tag{2.13}$$

Similarly if we integrate (2.12) from t_m to t $(t \geqslant t_m)$ and then from t_m to 1 we obtain

$$\int_{\varepsilon}^{y(t_m)} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(y(t_m))}{g(y(t_m))}\right\} \frac{1}{t_m} \int_{t_m}^{1} x(1-x)q(x)\,\mathrm{d}x. \tag{2.14}$$

Now (2.13) and (2.14) imply

$$\int_{\varepsilon}^{y(t_m)} \frac{\mathrm{d}u}{g(u)} \leqslant b_0\left\{1 + \frac{h(y(t_m))}{g(y(t_m))}\right\}.$$

This together with (2.9) implies $|y|_0 \neq r$. Then Theorem 1.3 implies that $(2.10)^m$ has a solution y_m with $|y_m|_0 \leqslant r$. In fact (as above),

$$\frac{1}{m} \leqslant y_m(t) < r \quad \text{for } t \in [0,1].$$

Next we obtain a sharper lower bound on y_m, namely we will show that there exists a constant $k > 0$, independent of m, with

$$y_m(t) \geqslant kt(1-t) \quad \text{for } t \in [0,1]. \tag{2.15}$$

To see this notice (2.6) guarantees the existence of a function $\psi_r(t)$ continuous on $[0,1]$ and positive on $(0,1)$ with $f(t,u) \geqslant \psi_r(t)$ for $(t,u) \in (0,1) \times (0,r]$. Now, using the Green's function representation for the solution of $(2.10)^m$, we have

$$\begin{aligned} y_m(t) = {}& \frac{1}{m} + t\int_t^1 (1-x)q(x)f\bigl(x, y_m(x)\bigr)\,\mathrm{d}x \\ &+ (1-t)\int_0^t xq(x)f\bigl(x, y_m(x)\bigr)\,\mathrm{d}x \end{aligned}$$

and so

$$\begin{aligned} y_m(t) &\geqslant t\int_t^1 (1-x)q(x)\psi_r(x)\,\mathrm{d}x + (1-t)\int_0^t xq(x)\psi_r(x)\,\mathrm{d}x \\ &\equiv \Phi_r(t). \end{aligned} \tag{2.16}$$

Now it is easy to check (as in Theorem 1.3) that

$$\Phi_r'(t) = \int_t^1 (1-x)q(x)\psi_r(x)\,\mathrm{d}x - \int_0^t xq(x)\psi_r(x)\,\mathrm{d}x \quad \text{for } t \in (0,1)$$

with $\Phi_r(0) = \Phi_r(1) = 0$. If $k_0 \equiv \int_0^1 (1-x)q(x)\psi_r(x)\,\mathrm{d}x$ exists then $\Phi_r'(0) = k_0$; otherwise $\Phi_r'(0) = \infty$. In either case there exists a constant k_1, independent of m, with $\Phi_r'(0) \geqslant k_1$. Thus there is an $\varepsilon > 0$ with $\Phi_r(t) \geqslant \frac{1}{2}k_1 t \geqslant \frac{1}{2}k_1 t(1-t)$ for $t \in [0,\varepsilon]$. Similarly there is a constant k_2, independent of m, with $-\Phi_r'(1) \geqslant k_2$. Thus there is a $\delta > 0$ with $\Phi_r(t) \geqslant \frac{1}{2}k_2(1-t) \geqslant \frac{1}{2}k_2 t(1-t)$ for $t \in [1-\delta, 1]$. Finally since $\frac{\Phi_r(t)}{t(1-t)}$ is bounded on $[\varepsilon, 1-\delta]$ there is a constant k, independent of m, with $\Phi_r(t) \geqslant kt(1-t)$ on $[0,1]$, i.e., (2.15) is true.

Next we will show

$$\{y_m\}_{m\in N_0} \quad \text{is a bounded, equicontinuous family on } [0,1]. \tag{2.17}$$

Returning to (2.12) (with y replaced by y_m) we have

$$-y_m''(x) \leqslant g\big(y_m(x)\big)\left\{1 + \frac{h(r)}{g(r)}\right\}q(x) \quad \text{for } x \in (0,1). \tag{2.18}$$

Now since $y_m'' \leqslant 0$ on $(0,1)$ and $y_m \geqslant \frac{1}{m}$ on $[0,1]$ there exists $t_m \in (0,1)$ with $y_m' \geqslant 0$ on $(0,t_m)$ and $y_m' \leqslant 0$ on $(t_m,1)$. Integrate (2.18) from t $(t < t_m)$ to t_m to obtain

$$\frac{y_m'(t)}{g(y_m(t))} \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\}\int_t^{t_m} q(x)\,\mathrm{d}x. \tag{2.19}$$

On the other hand integrate (2.18) from t_m to t $(t > t_m)$ to obtain

$$\frac{-y_m'(t)}{g(y_m(t))} \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\}\int_{t_m}^{t} q(x)\,\mathrm{d}x. \tag{2.20}$$

We now claim that there exists a_0 and a_1 with $a_0 > 0$, $a_1 < 1$, $a_0 < a_1$ with

$$a_0 < \inf\{t_m\colon\ m \in N_0\} \leqslant \sup\{t_m\colon\ m \in N_0\} < a_1. \tag{2.21}$$

REMARK 2.1. Here t_m (as before) is the unique point in $(0,1)$ with $y_m'(t_m) = 0$.

We now show $\inf\{t_m\colon\ m \in N_0\} > 0$. If this is not true then there is a subsequence S of N_0 with $t_m \to 0$ as $m \to \infty$ in S. Now integrate (2.19) from 0 to t_m to obtain

$$\int_0^{y_m(t_m)} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\}\int_0^{t_m} xq(x)\,\mathrm{d}x + \int_0^{\frac{1}{m}} \frac{\mathrm{d}u}{g(u)} \tag{2.22}$$

for $m \in S$. Since $t_m \to 0$ as $m \to \infty$ in S, we have from (2.22) that $y_m(t_m) \to 0$ as $m \to \infty$ in S. However since the maximum of y_m on $[0,1]$ occurs at t_m we have $y_m \to 0$ in $C[0,1]$

as $m \to \infty$ in S. This contradicts (2.15). Consequently $\inf\{t_m\colon m \in N_0\} > 0$. A similar argument shows $\sup\{t_m\colon m \in N_0\} < 1$. Let a_0 and a_1 be chosen as in (2.21). Now (2.19), (2.20) and (2.21) imply

$$\frac{|y'_m(t)|}{g(y_m(t))} \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\} v(t) \quad \text{for } t \in (0,1), \tag{2.23}$$

where

$$v(t) = \int_{\min\{t,a_0\}}^{\max\{t,a_1\}} q(x)\,\mathrm{d}x.$$

It is easy to see that $v \in L^1[0,1]$. Let $I : [0,\infty) \to [0,\infty)$ be defined by

$$I(z) = \int_0^z \frac{\mathrm{d}u}{g(u)}.$$

Note I is an increasing map from $[0,\infty)$ onto $[0,\infty)$ (notice $I(\infty) = \infty$ since $g > 0$ is nonincreasing on $(0,\infty)$) with I continuous on $[0, A]$ for any $A > 0$. Notice

$$\bigl\{I(y_m)\bigr\}_{m\in N_0} \quad \text{is a bounded, equicontinuous family on } [0,1]. \tag{2.24}$$

The equicontinuity follows from (here $t, s \in [0,1]$)

$$\bigl|I\bigl(y_m(t)\bigr) - I\bigl(y_m(s)\bigr)\bigr| = \left|\int_s^t \frac{y'_m(x)}{g(y_m(x))}\,\mathrm{d}x\right| \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\} \left|\int_s^t v(x)\,\mathrm{d}x\right|.$$

This inequality, the uniform continuity of I^{-1} on $[0, I(r)]$, and

$$\bigl|y_m(t) - y_m(s)\bigr| = \bigl|I^{-1}\bigl(I\bigl(y_m(t)\bigr)\bigr) - I^{-1}\bigl(I\bigl(y_m(s)\bigr)\bigr)\bigr|$$

now establishes (2.17).

The Arzela–Ascoli theorem guarantees the existence of a subsequence N of N_0 and a function $y \in C[0,1]$ with y_m converging uniformly on $[0,1]$ to y as $m \to \infty$ through N. Also $y(0) = y(1) = 0$, $|y|_0 \leqslant r$ and $y(t) \geqslant kt(1-t)$ for $t \in [0,1]$. In particular $y > 0$ on $(0,1)$. Fix $t \in (0,1)$ (without loss of generality assume $t \neq \frac{1}{2}$). Now y_m, $m \in N$, satisfies the integral equation

$$y_m(x) = y_m\left(\frac{1}{2}\right) + y'_m\left(\frac{1}{2}\right)\left(x - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{x} (s-x) q(s) f\bigl(s, y_m(s)\bigr)\,\mathrm{d}s$$

for $x \in (0,1)$. Notice (take $x = \frac{2}{3}$) that $\{y'_m(\frac{1}{2})\}$, $m \in N$, is a bounded sequence since $ks(1-s) \leqslant y_m(s) \leqslant r$ for $s \in [0,1]$. Thus $\{y'_m(\frac{1}{2})\}_{m\in N}$ has a convergent subsequence; for convenience let $\{y_m'(\frac{1}{2})\}_{m\in N}$ denote this subsequence also and let $r_0 \in \mathbf{R}$ be its limit. Now for the above fixed t,

$$y_m(t) = y_m\left(\frac{1}{2}\right) + y_m'\left(\frac{1}{2}\right)\left(t - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{t} (s-t)q(s)f\bigl(s, y_m(s)\bigr)\,\mathrm{d}s,$$

and let $m \to \infty$ through N (we note here that f is uniformly continuous on compact subsets of $[\min(\frac{1}{2}, t), \max(\frac{1}{2}, t)] \times (0, r])$ to obtain

$$y(t) = y\left(\frac{1}{2}\right) + r_0\left(t - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{t} (s-t)q(s)f\bigl(s, y(s)\bigr)\,\mathrm{d}s.$$

We can do this argument for each $t \in (0,1)$ and so $y''(t) + q(t)f(t, y(t)) = 0$ for $0 < t < 1$. Finally it is easy to see that $|y|_0 < r$ (note if $|y|_0 = r$ then following essentially the argument from (2.12)–(2.14) will yield a contradiction). $\square$

Next we establish the existence of two nonnegative solutions to the singular second order Dirichlet problem

$$\begin{cases} y''(t) + q(t)\bigl[g\bigl(y(t)\bigr) + h\bigl(y(t)\bigr)\bigr] = 0, & 0 < t < 1, \\ y(0) = y(1) = 0; \end{cases} \tag{2.25}$$

here our nonlinear term $g + h$ may be singular at $y = 0$. Next we state the fixed point result we will use to establish multiplicity (see [13] for a proof).

THEOREM 2.2. *Let $E = (E, \|\cdot\|)$ be a Banach space and let $K \subset E$ be a cone in E. Also r, R are constants with $0 < r < R$. Suppose $A : \overline{\Omega_R} \cap K \to K$ (here $\Omega_R = \{x \in E\colon \|x\| < R\}$) is a continuous, compact map and assume the following conditions hold*:

$$x \neq \lambda A(x) \quad \text{for } \lambda \in [0, 1) \text{ and } x \in \partial_E \Omega_r \cap K \tag{2.26}$$

and

$$\begin{cases} \text{there exists a } v \in K\backslash\{0\} \text{ with } x \neq A(x) + \delta v \\ \text{for any } \delta > 0 \text{ and } x \in \partial_E \Omega_R \cap K. \end{cases} \tag{2.27}$$

Then A has a fixed point in $K \cap \{x \in E\colon r \leqslant \|x\| \leqslant R\}$.

REMARK 2.2. In Theorem 2.2 if (2.26) and (2.27) are replaced by

$$x \neq \lambda A(x) \quad \text{for } \lambda \in [0, 1) \text{ and } x \in \partial_E \Omega_R \cap K \tag{2.26*}$$

and

$$\begin{cases} \text{there exists a } v \in K\backslash\{0\} \text{ with } x \neq A(x) + \delta v \\ \text{for any } \delta > 0 \text{ and } x \in \partial_E \Omega_r \cap K \end{cases} \tag{2.27*}$$

then A has also a fixed point in $K \cap \{x \in E\colon r \leqslant \|x\| \leqslant R\}$.

THEOREM 2.3. *Let $E = (E, \|\cdot\|)$ be a Banach space, $K \subset E$ a cone and let $\|\cdot\|$ be increasing with respect to K. Also r, R are constants with $0 < r < R$. Suppose $A : \overline{\Omega_R} \cap K \to K$ (here $\Omega_R = \{x \in E\colon \|x\| < R\}$) is a continuous, compact map and assume the following conditions hold*:

$$x \neq \lambda A(x) \quad \text{for } \lambda \in [0,1) \text{ and } x \in \partial_E \Omega_r \cap K \tag{2.28}$$

and

$$\|Ax\| > \|x\| \quad \text{for } x \in \partial_E \Omega_R \cap K. \tag{2.29}$$

Then A has a fixed point in $K \cap \{x \in E\colon r \leqslant \|x\| \leqslant R\}$.

PROOF. Notice (2.29) guarantees that (2.27) is true. This is a standard argument and for completeness we supply it here. Suppose there exists $v \in K\backslash\{0\}$ with $x = A(x) + \delta v$ for some $\delta > 0$ and $x \in \partial_E \Omega_R \cap K$. Then since $\|\cdot\|$ is increasing with respect to K we have since $\delta v \in K$,

$$\|x\| = \|Ax + \delta v\| \geqslant \|Ax\| > \|x\|,$$

a contradiction. The result now follows from Theorem 2.2. □

REMARK 2.3. In Theorem 2.3 if (2.28) and (2.29) are replaced by

$$x \neq \lambda A(x) \quad \text{for } \lambda \in [0,1) \text{ and } x \in \partial_E \Omega_R \cap K \tag*{(2.28)*}$$

and

$$\|Ax\| > \|x\| \quad \text{for } x \in \partial_E \Omega_r \cap K. \tag*{(2.29)*}$$

then A has a fixed point in $K \cap \{x \in E\colon r \leqslant \|x\| \leqslant R\}$.

Now $E = (C[0,1], |\cdot|_0)$ (here $|u|_0 = \sup_{t\in[0,1]} |u(t)|$, $u \in C[0,1]$) will be our Banach space and

$$K = \bigl\{y \in C[0,1]\colon\ y(t) \geqslant 0, t \in [0,1] \text{ and } y(t) \text{ concave on } [0,1]\bigr\}. \tag{2.30}$$

Let $\theta : [0,1] \times [0,1] \to [0,\infty)$ be defined by

$$\theta(t,s) = \begin{cases} \frac{t}{s} & \text{if } 0 \leqslant t \leqslant s, \\ \frac{1-t}{1-s} & \text{if } s \leqslant t \leqslant 1. \end{cases}$$

The following result is easy to prove and is well known.

THEOREM 2.4. *Let* $y \in K$ *(as in* (2.30)*). Then there exists* $t_0 \in [0,1]$ *with* $y(t_0) = |y|_0$ *and*

$$y(t) \geqslant \theta(t,t_0)|y|_0 \geqslant t(1-t)|y|_0 \quad \textit{for } t \in [0,1].$$

PROOF. The existence of t_0 is immediate. Now if $0 \leqslant t \leqslant t_0$ then since $y(t)$ is concave on $[0,1]$ we have

$$y(t) = y\left(\left(1-\frac{t}{t_0}\right)0 + \frac{t}{t_0}t_0\right) \geqslant \left(1-\frac{t}{t_0}\right)y(0) + \frac{t}{t_0}y(t_0).$$

That is

$$y(t) \geqslant \frac{t}{t_0}y(t_0) = \theta(t,t_0)|y|_0 \geqslant t(1-t)|y|_0.$$

A similar argument establishes the result if $t_0 \leqslant t \leqslant 1$. □

From Theorem 2.1 we have immediately the following existence result for (2.25).

THEOREM 2.5. *Suppose the following conditions are satisfied*:

$$q \in C(0,1), \quad q > 0 \quad \textit{on } (0,1) \quad \textit{and} \quad \int_0^1 t(1-t)q(t)\,\mathrm{d}t < \infty \tag{2.31}$$

$$g > 0 \quad \textit{is continuous and nonincreasing on } (0,\infty) \tag{2.32}$$

$$h \geqslant 0 \quad \textit{continuous on } [0,\infty) \textit{ with } \frac{h}{g} \textit{ nondecreasing on } (0,\infty) \tag{2.33}$$

and

$$\exists r > 0 \quad \textit{with } \frac{1}{\{1+\frac{h(r)}{g(r)}\}}\int_0^r \frac{\mathrm{d}u}{g(u)} > b_0; \tag{2.34}$$

here

$$b_0 = \max\left\{2\int_0^{\frac{1}{2}} t(1-t)q(t)\,\mathrm{d}t,\, 2\int_{\frac{1}{2}}^1 t(1-t)q(t)\,\mathrm{d}t\right\}. \tag{2.35}$$

Then (2.25) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y > 0$ *on* $(0,1)$ *and* $|y|_0 < r$.

PROOF. The result follows from Theorem 2.1 with $f(t,u) = g(u) + h(u)$. Notice (2.6) is clearly satisfied with $\psi_H(t) = g(H)$. □

THEOREM 2.6. *Assume* (2.31)–(2.34) *hold. Choose* $a \in (0,\frac{1}{2})$ *and fix it and suppose there exists* $R > r$ *with*

$$\frac{Rg(a(1-a)R)}{g(R)g(a(1-a)R) + g(R)h(a(1-a)R)} \leqslant \int_a^{1-a} G(\sigma,s)q(s)\,\mathrm{d}s; \tag{2.36}$$

here $0 \leqslant \sigma \leqslant 1$ *is such that*

$$\int_a^{1-a} G(\sigma, s)q(s)\,\mathrm{d}s = \sup_{t\in[0,1]} \int_a^{1-a} G(t, s)q(s)\,\mathrm{d}s \tag{2.37}$$

and

$$G(t, s) = \begin{cases} (1-t)s, & 0 \leqslant s \leqslant t, \\ (1-s)t, & t \leqslant s \leqslant 1. \end{cases}$$

Then (2.25) *has a solution* $y \in C[0, 1] \cap C^2(0, 1)$ *with* $y > 0$ *on* $(0, 1)$ *and* $r < |y|_0 \leqslant R$.

PROOF. To show the existence of the solution described in the statement of Theorem 2.6 we will apply Theorem 2.3. First however choose $\varepsilon > 0$ and $\varepsilon < r$ with

$$\frac{1}{\{1 + \frac{h(r)}{g(r)}\}} \int_\varepsilon^r \frac{\mathrm{d}u}{g(u)} > b_0. \tag{2.38}$$

Let $m_0 \in \{1, 2, \dots\}$ be chosen so that $\frac{1}{m_0} < \varepsilon$ and $\frac{1}{m_0} < a(1-a)R$ and let $N_0 = \{m_0, m_0 + 1, \dots\}$. We first show that

$$\begin{cases} y''(t) + q(t)\bigl[g\bigl(y(t)\bigr) + h\bigl(y(t)\bigr)\bigr] = 0, & 0 < t < 1, \\ y(0) = y(1) = \frac{1}{m} \end{cases} \tag*{(2.39)^m}$$

has a solution y_m for each $m \in N_0$ with $y_m > \frac{1}{m}$ on $(0, 1)$ and $r \leqslant |y_m|_0 \leqslant R$. To show $(2.39)^m$ has such a solution for each $m \in N_0$, we will look at

$$\begin{cases} y''(t) + q(t)\bigl[g^\star\bigl(y(t)\bigr) + h\bigl(y(t)\bigr)\bigr] = 0, & 0 < t < 1, \\ y(0) = y(1) = \frac{1}{m} \end{cases} \tag*{(2.40)^m}$$

with

$$g^\star(u) = \begin{cases} g(u), & u \geqslant \frac{1}{m}, \\ g\bigl(\frac{1}{m}\bigr), & 0 \leqslant u \leqslant \frac{1}{m}. \end{cases}$$

REMARK 2.4. Notice $g^\star(u) \leqslant g(u)$ for $u > 0$.

Fix $m \in N_0$. Let $E = (C[0, 1], |\cdot|_0)$ and

$$K = \bigl\{u \in C[0, 1]\colon\ u(t) \geqslant 0, t \in [0, 1] \text{ and } u(t) \text{ concave on } [0, 1]\bigr\}. \tag{2.41}$$

Clearly K is a cone of E. Let $A : K \to C[0, 1]$ be defined by

$$Ay(t) = \frac{1}{m} + \int_0^1 G(t, s)q(s)\bigl[g^\star\bigl(y(s)\bigr) + h\bigl(y(s)\bigr)\bigr]\,\mathrm{d}s. \tag{2.42}$$

A standard argument implies $A : K \to C[0,1]$ is continuous and completely continuous. Next we show $A : K \to K$. If $u \in K$ then clearly $Au(t) \geqslant 0$ for $t \in [0,1]$. Also notice that

$$\begin{cases} (Au)''(t) \leqslant 0 & \text{on } (0,1), \\ Au(0) = Au(1) = \frac{1}{m} \end{cases}$$

so $Au(t)$ is concave on $[0,1]$. Consequently $Au \in K$ so $A : K \to K$. Let

$$\Omega_1 = \{u \in C[0,1] \colon |u|_0 < r\} \quad \text{and} \quad \Omega_2 = \{u \in C[0,1] \colon |u|_0 < R\}.$$

We first show

$$y \neq \lambda A y \quad \text{for } \lambda \in [0,1] \text{ and } y \in K \cap \partial\Omega_1. \tag{2.43}$$

Suppose this is false, i.e., suppose there exists $y \in K \cap \partial\Omega_1$ and $\lambda \in [0,1)$ with $y = \lambda A y$. We can assume $\lambda \neq 0$. Now since $y = \lambda A y$ we have

$$\begin{cases} y''(t) + \lambda q(t)\big[g^\star\big(y(t)\big) + h\big(y(t)\big)\big] = 0, & 0 < t < 1, \\ y(0) = y(1) = \frac{1}{m}. \end{cases} \tag{2.44}$$

Since $y'' \leqslant 0$ on $(0,1)$ and $y \geqslant \frac{1}{m}$ on $[0,1]$ there exists $t_0 \in (0,1)$ with $y' \geqslant 0$ on $(0,t_0)$, $y' \leqslant 0$ on $(t_0,1)$ and $y(t_0) = |y|_0 = r$ (note $y \in K \cap \partial\Omega_1$). Also notice

$$g^\star\big(y(t)\big) + h\big(y(t)\big) \leqslant g\big(y(t)\big) + h\big(y(t)\big) \quad \text{for } t \in (0,1)$$

since g is nonincreasing on $(0,\infty)$. For $x \in (0,1)$ we have

$$-y''(x) \leqslant g\big(y(x)\big)\left\{1 + \frac{h(y(x))}{g(y(x))}\right\} q(x). \tag{2.45}$$

Integrate from t ($t \leqslant t_0$) to t_0 to obtain

$$y'(t) \leqslant g\big(y(t)\big)\left\{1 + \frac{h(r)}{g(r)}\right\} \int_t^{t_0} q(x)\,\mathrm{d}x$$

and then integrate from 0 to t_0 to obtain

$$\int_{\frac{1}{m}}^{r} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\} \int_0^{t_0} x q(x)\,\mathrm{d}x.$$

Consequently

$$\int_{\varepsilon}^{r} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\} \int_0^{t_0} x q(x)\,\mathrm{d}x$$

and so

$$\int_{\varepsilon}^{r} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\} \frac{1}{1-t_0} \int_0^{t_0} x(1-x)q(x)\,\mathrm{d}x. \tag{2.46}$$

Similarly if we integrate (2.45) from t_0 to t ($t \geqslant t_0$) and then from t_0 to 1 we obtain

$$\int_{\varepsilon}^{r} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(r)}{g(r)}\right\} \frac{1}{t_0} \int_{t_0}^{1} x(1-x)q(x)\,\mathrm{d}x. \tag{2.47}$$

Now (2.46) and (2.47) imply

$$\int_{\varepsilon}^{r} \frac{\mathrm{d}u}{g(u)} \leqslant b_0 \left\{1 + \frac{h(r)}{g(r)}\right\}, \tag{2.48}$$

where b_0 is as defined in (2.35). This contradicts (2.38) and consequently (2.43) is true.

Next we show

$$|Ay|_0 > |y|_0 \quad \text{for } y \in K \cap \partial\Omega_2. \tag{2.49}$$

To see this let $y \in K \cap \partial\Omega_2$ so $|y|_0 = R$. Also since $y(t)$ is concave on [0, 1] (since $y \in K$) we have from Theorem 2.4 that $y(t) \geqslant t(1-t)|y|_0 \geqslant t(1-t)R$ for $t \in [0,1]$. Also for $s \in [a, 1-a]$ we have

$$g^\star\big(y(s)\big) + h\big(y(s)\big) = g\big(y(s)\big) + h\big(y(s)\big)$$

since $y(s) \geqslant a(1-a)R > \frac{1}{m_0}$ for $s \in [a, 1-a]$. Note in particular that

$$y(s) \in \big[a(1-a)R, R\big] \quad \text{for } s \in [a, 1-a]. \tag{2.50}$$

With σ as defined in (2.37) we have using (2.50) and (2.36),

$$\begin{aligned}
Ay(\sigma) &= \frac{1}{m} + \int_0^1 G(\sigma,s)q(s)\big[g^\star\big(y(s)\big) + h\big(y(s)\big)\big]\,\mathrm{d}s \\
&> \int_a^{1-a} G(\sigma,s)q(s)\big[g^\star\big(y(s)\big) + h\big(y(s)\big)\big]\,\mathrm{d}s \\
&= \int_a^{1-a} G(\sigma,s)q(s)g\big(y(s)\big)\left\{1 + \frac{h(y(s))}{g(y(s))}\right\}\mathrm{d}s \\
&\geqslant g(R)\left\{1 + \frac{h(a(1-a)R)}{g(a(1-a)R)}\right\} \int_a^{1-a} G(\sigma,s)q(s)\,\mathrm{d}s \\
&\geqslant R = |y|_0,
\end{aligned}$$

and so $|Ay|_0 > |y|_0$. Hence (2.49) is true.

Now Theorem 2.3 implies A has a fixed point $y_m \in K \cap (\overline{\Omega_2} \backslash \Omega_1)$, i.e., $r \leqslant |y_m|_0 \leqslant R$. In fact $|y_m|_0 > r$ (note if $|y_m|_0 = r$ then following essentially the same argument from (2.45)–(2.48) will yield a contradiction). Consequently $(2.40)^m$ (and also $(2.39)^m$) has a solution $y_m \in C[0,1] \cap C^2(0,1)$, $y_m \in K$, with

$$\frac{1}{m} \leqslant y_m(t) \quad \text{for } t \in [0,1],\ r < |y_m|_0 \leqslant R \tag{2.51}$$

and (from Theorem 2.4, note $y_m \in K$)

$$y_m(t) \geqslant t(1-t)r \quad \text{for } t \in [0,1]. \tag{2.52}$$

Next we will show

$$\{y_m\}_{m \in N_0} \quad \text{is a bounded, equicontinuous family on } [0,1]. \tag{2.53}$$

Returning to (2.45) (with y replaced by y_m) we have

$$-y_m''(x) \leqslant g\big(y_m(x)\big)\left\{1 + \frac{h(R)}{g(R)}\right\} q(x) \quad \text{for } x \in (0,1). \tag{2.54}$$

Now since $y_m'' \leqslant 0$ on $(0,1)$ and $y_m \geqslant \frac{1}{m}$ on $[0,1]$ there exists $t_m \in (0,1)$ with $y_m' \geqslant 0$ on $(0, t_m)$ and $y_m' \leqslant 0$ on $(t_m, 1)$. Integrate (2.54) from t ($t < t_m$) to t_m to obtain

$$\frac{y_m'(t)}{g(y_m(t))} \leqslant \left\{1 + \frac{h(R)}{g(R)}\right\} \int_t^{t_m} q(x)\,\mathrm{d}x. \tag{2.55}$$

On the other hand integrate (2.54) from t_m to t ($t > t_m$) to obtain

$$\frac{-y_m'(t)}{g(y_m(t))} \leqslant \left\{1 + \frac{h(R)}{g(R)}\right\} \int_{t_m}^{t} q(x)\,\mathrm{d}x. \tag{2.56}$$

We now claim that there exists a_0 and a_1 with $a_0 > 0$, $a_1 < 1$, $a_0 < a_1$ with

$$a_0 < \inf\{t_m\colon\ m \in N_0\} \leqslant \sup\{t_m\colon\ m \in N_0\} < a_1. \tag{2.57}$$

REMARK 2.5. Here t_m (as before) is the unique point in $(0,1)$ with $y_m'(t_m) = 0$.

We now show $\inf\{t_m\colon\ m \in N_0\} > 0$. If this is not true then there is a subsequence S of N_0 with $t_m \to 0$ as $m \to \infty$ in S. Now integrate (2.55) from 0 to t_m to obtain

$$\int_0^{y_m(t_m)} \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(R)}{g(R)}\right\} \int_0^{t_m} x q(x)\,\mathrm{d}x + \int_0^{\frac{1}{m}} \frac{\mathrm{d}u}{g(u)} \tag{2.58}$$

for $m \in S$. Since $t_m \to 0$ as $m \to \infty$ in S, we have from (2.58) that $y_m(t_m) \to 0$ as $m \to \infty$ in S. However since the maximum of y_m on $[0,1]$ occurs at t_m we have $y_m \to 0$ in $C[0,1]$

as $m \to \infty$ in S. This contradicts (2.52). Consequently $\inf\{t_m\colon\ m \in N_0\} > 0$. A similar argument shows $\sup\{t_m\colon\ m \in N_0\} < 1$. Let a_0 and a_1 be chosen as in (2.57). Now (2.55), (2.56) and (2.57) imply

$$\frac{|y_m'(t)|}{g(y_m(t))} \leqslant \left\{1 + \frac{h(R)}{g(R)}\right\} v(t) \quad \text{for } t \in (0, 1), \tag{2.59}$$

where

$$v(t) = \int_{\min\{t,a_0\}}^{\max\{t,a_1\}} q(x)\,\mathrm{d}x.$$

It is easy to see that $v \in L^1[0, 1]$. Let $I : [0, \infty) \to [0, \infty)$ be defined by

$$I(z) = \int_0^z \frac{\mathrm{d}u}{g(u)}.$$

Note I is an increasing map from $[0, \infty)$ onto $[0, \infty)$ (notice $I(\infty) = \infty$ since $g > 0$ is nonincreasing on $(0, \infty)$) with I continuous on $[0, A]$ for any $A > 0$. Notice

$$\bigl\{I(y_m)\bigr\}_{m \in N_0} \quad \text{is a bounded, equicontinuous family on } [0, 1]. \tag{2.60}$$

The equicontinuity follows from (here $t, s \in [0, 1]$)

$$\bigl|I\bigl(y_m(t)\bigr) - I\bigl(y_m(s)\bigr)\bigr| = \left|\int_s^t \frac{y_m'(x)}{g(y_m(x))}\,\mathrm{d}x\right| \leqslant \left\{1 + \frac{h(R)}{g(R)}\right\} \left|\int_s^t v(x)\,\mathrm{d}x\right|.$$

This inequality, the uniform continuity of I^{-1} on $[0, I(R)]$, and

$$\bigl|y_m(t) - y_m(s)\bigr| = \bigl|I^{-1}\bigl(I\bigl(y_m(t)\bigr)\bigr) - I^{-1}\bigl(I\bigl(y_m(s)\bigr)\bigr)\bigr|$$

now establishes (2.53).

The Arzela–Ascoli theorem guarantees the existence of a subsequence N of N_0 and a function $y \in C[0, 1]$ with y_m converging uniformly on $[0, 1]$ to y as $m \to \infty$ through N. Also $y(0) = y(1) = 0$, $r \leqslant |y|_0 \leqslant R$ and $y(t) \geqslant t(1 - t)r$ for $t \in [0, 1]$. In particular $y > 0$ on $(0, 1)$. Fix $t \in (0, 1)$ (without loss of generality assume $t \neq \frac{1}{2}$). Now y_m, $m \in N$, satisfies the integral equation

$$\begin{aligned} y_m(x) = {} & y_m\left(\frac{1}{2}\right) + y_m'\left(\frac{1}{2}\right)\left(x - \frac{1}{2}\right) \\ & + \int_{\frac{1}{2}}^{x} (s - x) q(s)\bigl[g\bigl(y_m(s)\bigr) + h\bigl(y_m(s)\bigr)\bigr]\,\mathrm{d}s \end{aligned}$$

for $x \in (0, 1)$. Notice (take $x = \frac{2}{3}$) that $\{y_m'(\frac{1}{2})\}$, $m \in N$, is a bounded sequence since $rs(1 - s) \leqslant y_m(s) \leqslant R$ for $s \in [0, 1]$. Thus $\{y_m'(\frac{1}{2})\}_{m \in N}$ has a convergent subsequence; for convenience let $\{y_m'(\frac{1}{2})\}_{m\in N}$ denote this subsequence also and let $r_0 \in \mathbf{R}$ be its limit. Now for the above fixed t,

$$\begin{aligned}y_m(t) &= y_m\left(\frac{1}{2}\right) + y_m'\left(\frac{1}{2}\right)\left(t - \frac{1}{2}\right)\\ &\quad + \int_{\frac{1}{2}}^{t} (s-t)q(s)\big[g\big(y_m(s)\big) + h\big(y_m(s)\big)\big]\,\mathrm{d}s,\end{aligned}$$

and let $m \to \infty$ through N (we note here that $g + h$ is uniformly continuous on compact subsets of $[\min(\frac{1}{2}, t), \max(\frac{1}{2}, t)] \times (0, R]$) to obtain

$$y(t) = y\left(\frac{1}{2}\right) + r_0\left(t - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{t} (s-t)q(s)\big[g\big(y(s)\big) + h\big(y(s)\big)\big]\,\mathrm{d}s.$$

We can do this argument for each $t \in (0, 1)$ and so $y''(t) + q(t)[g(y(t)) + h(y(t))] = 0$ for $0 < t < 1$. Finally it is easy to see that $|y|_0 > r$ (note if $|y|_0 = r$ then following essentially the argument from (2.45)–(2.48) will yield a contradiction). □

REMARK 2.6. If in (2.36) we have $R < r$ then (2.25) has a solution $y \in C[0, 1] \cap C^2(0, 1)$ with $y > 0$ on $(0, 1)$ and $R \leqslant |y|_0 < r$. The argument is similar to that in Theorem 2.6 except here we use Remark 2.3.

REMARK 2.7. It is also possible to use the ideas in Theorem 2.6 to discuss other boundary conditions; for example $y'(0) = y(1) = 0$.

REMARK 2.8. If we use Krasnoselski's fixed point theorem in a cone we need more that (2.31)–(2.34), (2.36) to establish the existence of a solution $y \in C[0, 1] \cap C^2(0, 1)$ with $y > 0$ on $(0, 1)$ and $r < |y|_0 \leqslant R$. This is because (2.43) is less restrictive than $|Ay|_0 \leqslant |y|_0$ for $y \in K \cap \partial\Omega_1$.

THEOREM 2.7. *Assume* (2.31)–(2.34) *and* (2.36) *hold. Then* (2.25) *has two solutions* $y_1, y_2 \in C[0, 1] \cap C^2(0, 1)$ *with* $y_1 > 0$, $y_2 > 0$ *on* $(0, 1)$ *and* $|y_1|_0 < r < |y_2|_0 \leqslant R$.

PROOF. The existence of y_1 follows from Theorem 2.5 and the existence of y_2 follows from Theorem 2.6. □

EXAMPLE 2.1. The singular boundary value problem

$$\begin{cases} y'' + \frac{1}{\alpha+1}\big(y^{-\alpha} + y^{\beta} + 1\big) = 0 & \text{on } (0, 1),\\ y(0) = y(1) = 0, & \alpha > 0,\ \beta > 1 \end{cases} \tag{2.61}$$

has two solutions $y_1, y_2 \in C[0, 1] \cap C^2(0, 1)$ with $y_1 > 0$, $y_2 > 0$ on $(0, 1)$ and $|y_1|_0 < 1 < |y_2|_0$.

To see this we will apply Theorem 2.7 with $q = \frac{1}{\alpha+1}$, $g(u) = u^{-\alpha}$ and $h(u) = u^{\beta} + 1$. Clearly (2.31)–(2.33) hold. Also note

$$b_0 = \max\left\{ \frac{2}{\alpha+1} \int_0^{\frac{1}{2}} t(1-t)\,\mathrm{d}t,\ \frac{2}{\alpha+1} \int_{\frac{1}{2}}^{1} t(1-t)\,\mathrm{d}t \right\} = \frac{1}{6(\alpha+1)}.$$

Consequently (2.34) holds (with $r = 1$) since

$$\begin{aligned} \frac{1}{\{1 + \frac{h(r)}{g(r)}\}} \int_0^r \frac{\mathrm{d}u}{g(u)} &= \frac{1}{(1 + r^{\alpha+\beta} + r^{\alpha})} \left(\frac{r^{\alpha+1}}{\alpha+1} \right) \\ &= \frac{1}{3(\alpha+1)} > b_0 = \frac{1}{6(\alpha+1)}. \end{aligned}$$

Finally note (since $\beta > 1$), take $a = \frac{1}{4}$, that

$$\lim_{R\to\infty} \frac{Rg(\frac{3R}{16})}{g(R)g(\frac{3R}{16}) + g(R)h(\frac{3R}{16})} = \lim_{R\to\infty} \frac{R^{\alpha+1}(\frac{3}{16})^{-\alpha}}{(\frac{3}{16})^{-\alpha} + (\frac{3}{16})^{\beta} R^{\alpha+\beta} + R^{\alpha}} = 0$$

so there exists $R > 1$ with (2.36) holding. The result now follows from Theorem 2.7.

2.2. *Singular problems with sign changing nonlinearities*

In this section we discuss the Dirichlet singular boundary value problem

$$\begin{cases} y'' + q(t)f(t, y) = 0, & 0 < t < 1, \\ y(0) = y(1) = 0, \end{cases} \tag{2.62}$$

where our nonlinearity f may change sign. We first present a variation of the classical theory of upper and lower solutions in this section so that (2.62) can be discussed in its natural setting. We assume the following conditions hold:

$$\begin{cases} \text{there exists } \beta \in C[0,1] \cap C^2(0,1), \\ \beta(0) \geqslant 0,\ q(t)f\big(t, \beta(t)\big) + \beta''(t) \leqslant 0 \\ \text{for } t \in (0,1), \text{ and } \beta(1) \geqslant 0, \end{cases} \tag{2.63}$$

$$\begin{cases} \text{there exists } \alpha \in C[0,1] \cap C^2(0,1),\ \alpha(t) \leqslant \beta(t) \\ \text{on } [0,1],\ \alpha(0) \leqslant 0,\ q(t)f\big(t, \alpha(t)\big) + \alpha''(t) \geqslant 0 \\ \text{for } t \in (0,1), \text{ and } \alpha(1) \leqslant 0 \end{cases} \tag{2.64}$$

and

$$\begin{cases} q \in C(0,1) \text{ with } q > 0 \text{ on } (0,1) \text{ and} \\ \int_0^1 t(1-t)q(t)\,\mathrm{d}t < \infty. \end{cases} \tag{2.65}$$

Let

$$f^\star(t,y)=\begin{cases} f\big(t,\beta(t)\big)+r\big(\beta(t)-y\big), & y\geqslant \beta(t),\\ f(t,y), & \alpha(t)<y<\beta(t),\\ f\big(t,\alpha(t)\big)+r\big(\alpha(t)-y\big), & y\leqslant \alpha(t)\end{cases}$$

and $r:\mathbf{R}\to[-1,1]$ is the radial retraction defined by

$$r(x)=\begin{cases} x, & |x|\leqslant 1\,,\\ \frac{x}{|x|}, & |x|>1.\end{cases}$$

Finally we assume

$$f^\star:[0,1]\times\mathbf{R}\to\mathbf{R} \text{ is continuous.} \tag{2.66}$$

THEOREM 2.8. *Suppose* (2.63)–(2.66) *hold. Then* (2.62) *has a solution* y (*here* $y\in C[0,1]\cap C^2(0,1)$) *with* $\alpha(t)\leqslant y(t)\leqslant\beta(t)$ *for* $t\in[0,1]$.

PROOF. To show (2.62) has a solution we consider the problem

$$\begin{cases} y''+q(t)f^\star(t,y)=0, & 0<t<1,\\ y(0)=y(1)=0.\end{cases} \tag{2.67}$$

Theorem 1.4 guarantees that (2.67) has a solution $y\in C[0,1]\cap C^2(0,1)$. The result of our theorem will follow once we show

$$\alpha(t)\leqslant y(t)\leqslant\beta(t)\quad \text{for } t\in[0,1]. \tag{2.68}$$

We now show

$$y(t)\leqslant\beta(t)\quad \text{for } t\in[0,1]. \tag{2.69}$$

Suppose (2.69) is not true. Then $y-\beta$ has a positive absolute maximum at $t_1\in(0,1)$. As a result $(y-\beta)'(t_1)=0$ and $(y-\beta)''(t_1)\leqslant 0$. However since $y(t_1)>\beta(t_1)$ we have

$$\begin{aligned}(y-\beta)''(t_1)&=-q(t_1)\big[f\big(t_1,\beta(t_1)\big)+r\big(\beta(t_1)-y(t_1)\big)\big]-\beta''(t_1)\\ &\geqslant -q(t_1)r\big(\beta(t_1)-y(t_1)\big)>0,\end{aligned}$$

a contradiction. Thus (2.69) is true. Similarly we can show

$$\alpha(t)\leqslant y(t)\quad \text{for } t\in[0,1]. \tag{2.70}$$

Our result follows. □

In general if we construct an upper solution β and a lower solution α to (2.62), usually it is difficult to check (2.66). As a result it is of interest to provide an alternative approach and to provide conditions that are easy to verify in applications.

THEOREM 2.9. *Let $n_0 \in \{1, 2, \ldots\}$ be fixed and suppose the following conditions are satisfied*:

$$f : [0, 1] \times (0, \infty) \to \mathbf{R} \quad \textit{is continuous} \tag{2.71}$$

$$\begin{cases} \textit{let } n \in \{n_0, n_0 + 1, \ldots\} \textit{ and associated with each } n \textit{ we} \\ \textit{have a constant } \rho_n \textit{ such that } \{\rho_n\} \textit{ is a nonincreasing} \\ \textit{sequence with } \lim_{n\to\infty} \rho_n = 0 \textit{ and such that for} \\ \frac{1}{2^{n+1}} \leqslant t \leqslant 1 \textit{ we have } q(t)f(t, \rho_n) \geqslant 0, \end{cases} \tag{2.72}$$

$$q \in C(0, 1), \quad q > 0 \quad \textit{on } (0, 1) \quad \textit{and} \int_0^1 x(1 - x)q(x)\,\mathrm{d}x < \infty, \tag{2.73}$$

$$\begin{cases} \textit{there exists a function } \alpha \in C[0, 1] \cap C^2(0, 1) \\ \textit{with } \alpha(0) = \alpha(1) = 0, \ \alpha > 0 \textit{ on } (0, 1) \textit{ such} \\ \textit{that } q(t)f\big(t, \alpha(t)\big) + \alpha''(t) \geqslant 0 \textit{ for } t \in (0, 1) \end{cases} \tag{2.74}$$

and

$$\begin{cases} \textit{there exists a function } \beta \in C[0, 1] \cap C^2(0, 1) \\ \textit{with } \beta(t) \geqslant \alpha(t) \textit{ and } \beta(t) \geqslant \rho_{n_0} \textit{ for } t \in [0, 1] \\ \textit{with } q(t)f\big(t, \beta(t)\big) + \beta''(t) \leqslant 0 \textit{ for } t \in (0, 1) \textit{ and} \\ q(t)f\big(\frac{1}{2^{n_0+1}}, \beta(t)\big) + \beta''(t) \leqslant 0 \textit{ for } t \in \big(0, \frac{1}{2^{n_0+1}}\big). \end{cases} \tag{2.75}$$

Then (2.62) *has a solution $y \in C[0, 1] \cap C^2(0, 1)$ with $y(t) \geqslant \alpha(t)$ for $t \in [0, 1]$.*

PROOF. For $n = n_0, n_0 + 1, \ldots$ let

$$e_n = \left[\frac{1}{2^{n+1}}, 1\right] \quad \text{and} \quad \theta_n(t) = \max\left\{\frac{1}{2^{n+1}}, t\right\}, \quad 0 \leqslant t \leqslant 1,$$

and

$$f_n(t, x) = \max\big\{f\big(\theta_n(t), x\big), f(t, x)\big\}.$$

Next we define inductively

$$g_{n_0}(t, x) = f_{n_0}(t, x)$$

and

$$g_n(t, x) = \min\big\{f_{n_0}(t, x), \ldots, f_n(t, x)\big\}, \quad n = n_0 + 1, n_0 + 2, \ldots.$$

Notice

$$f(t,x) \leqslant \cdots \leqslant g_{n+1}(t,x) \leqslant g_n(t,x) \leqslant \cdots \leqslant g_{n_0}(t,x)$$

for $(t,x) \in (0,1) \times (0,\infty)$ and

$$g_n(t,x) = f(t,x) \quad \text{for } (t,x) \in e_n \times (0,\infty).$$

Without loss of generality assume $\rho_{n_0} \leqslant \min_{t\in[\frac{1}{3},\frac{2}{3}]} \alpha(t)$. Fix $n \in \{n_0, n_0+1, \ldots\}$. Let $t_n \in [0,\frac{1}{3}]$ and $s_n \in [\frac{2}{3},1]$ be such that

$$\alpha(t_n) = \alpha(s_n) = \rho_n \quad \text{and} \quad \alpha(t) \leqslant \rho_n \quad \text{for } t \in [0,t_n] \cup [s_n,1].$$

Define

$$\alpha_n(t) = \begin{cases} \rho_n & \text{if } t \in [0,t_n] \cup [s_n,1], \\ \alpha(t) & \text{if } t \in (t_n,s_n). \end{cases}$$

We begin with the boundary value problem

$$\begin{cases} y'' + q(t) g^\star_{n_0}(t,y) = 0, \quad 0 < t < 1, \\ y(0) = y(1) = \rho_{n_0}; \end{cases} \tag{2.76}$$

here

$$g^\star_{n_0}(t,y) = \begin{cases} g_{n_0}\big(t,\alpha_{n_0}(t)\big) + r\big(\alpha_{n_0}(t) - y\big), & y \leqslant \alpha_{n_0}(t), \\ g_{n_0}(t,y), & \alpha_{n_0}(t) \leqslant y \leqslant \beta(t), \\ g_{n_0}\big(t,\beta(t)\big) + r\big(\beta(t) - y\big), & y \geqslant \beta(t), \end{cases}$$

with $r : \mathbf{R} \to [-1,1]$ the radial retraction defined by

$$r(u) = \begin{cases} u, & |u| \leqslant 1, \\ \frac{u}{|u|}, & |u| > 1. \end{cases}$$

From Schauder's fixed point theorem we know that (2.76) has a solution $y_{n_0} \in C[0,1] \cap C^2(0,1)$. We first show

$$y_{n_0}(t) \geqslant \alpha_{n_0}(t), \quad t \in [0,1]. \tag{2.77}$$

Suppose (2.77) is not true. Then $y_{n_0} - \alpha_{n_0}$ has a negative absolute minimum at $\tau \in (0,1)$. Now since $y_{n_0}(0) - \alpha_{n_0}(0) = 0 = y_{n_0}(1) - \alpha_{n_0}(1)$ there exists $\tau_0, \tau_1 \in [0,1]$ with $\tau \in (\tau_1, \tau_2)$ and

$$y_{n_0}(\tau_0) - \alpha_{n_0}(\tau_0) = y_{n_0}(\tau_1) - \alpha_{n_0}(\tau_1) = 0$$

and

$$y_{n_0}(t)-\alpha_{n_0}(t)<0,\quad t\in(\tau_0,\tau_1).$$

We now claim

$$(y_{n_0}-\alpha_{n_0})''(t)<0\quad\text{for a.e. } t\in(\tau_0,\tau_1). \tag{2.78}$$

If (2.78) is true then

$$y_{n_0}(t)-\alpha_{n_0}(t)=-\int_{\tau_0}^{\tau_1}G(t,s)\big[y''_{n_0}(s)-\alpha''_{n_0}(s)\big]\,\mathrm{d}s\quad\text{for } t\in(\tau_0,\tau_1)$$

with

$$G(t,s)=\begin{cases}\frac{(s-\tau_0)(\tau_1-t)}{\tau_1-\tau_0}, & \tau_0\leqslant s\leqslant t,\\ \frac{(t-\tau_0)(\tau_1-s)}{\tau_1-\tau_0}, & t\leqslant s\leqslant\tau_1\end{cases}$$

so we have

$$y_{n_0}(t)-\alpha_{n_0}(t)>0\quad\text{for } t\in(\tau_0,\tau_1),$$

a contradiction. As a result if we show that (2.78) is true then (2.77) will follow. To see (2.78) we will show

$$(y_{n_0}-\alpha_{n_0})''(t)<0\quad\text{for } t\in(\tau_0,\tau_1)\text{ provided } t\neq t_{n_0}\text{ or } t\neq s_{n_0}.$$

Fix $t\in(\tau_0,\tau_1)$ and assume $t\neq t_{n_0}$ or $t\neq s_{n_0}$. Then

$$\begin{aligned}(y_{n_0}-\alpha_{n_0})''(t)&=-\big[q(t)\big\{g_{n_0}\big(t,\alpha_{n_0}(t)\big)+r\big(\alpha_{n_0}(t)-y_{n_0}(t)\big)\big\}+\alpha''_{n_0}(t)\big]\\&=\begin{cases}-\big[q(t)\big\{g_{n_0}\big(t,\alpha(t)\big)+r\big(\alpha(t)-y_{n_0}(t)\big)\big\}+\alpha''(t)\big]\\ \quad\text{if } t\in(t_{n_0},s_{n_0}),\\ -\big[q(t)\big\{g_{n_0}\big(t,\rho_{n_0}\big)+r\big(\rho_{n_0}-y_{n_0}(t)\big)\big\}\big]\\ \quad\text{if } t\in(0,t_{n_0})\cup(s_{n_0},1).\end{cases}\end{aligned}$$

Case (i). $t\geqslant\frac{1}{2^{n_0+1}}$.
Then since $g_{n_0}(t,x)=f(t,x)$ for $x\in(0,\infty)$ we have

$$\begin{aligned}(y_{n_0}-\alpha_{n_0})''(t)&=\begin{cases}-\big[q(t)\big\{f\big(t,\alpha(t)\big)+r\big(\alpha(t)-y_{n_0}(t)\big)\big\}+\alpha''(t)\big]\\ \quad\text{if } t\in(t_{n_0},s_{n_0}),\\ -\big[q(t)\big\{f(t,\rho_{n_0})+r\big(\rho_{n_0}-y_{n_0}(t)\big)\big\}\big]\\ \quad\text{if } t\in(0,t_{n_0})\cup(s_{n_0},1)\end{cases}\\&<0,\end{aligned}$$

from (2.72) and (2.74).

Case (ii). $t \in (0, \frac{1}{2^{n_0+1}})$.

Then since

$$g_{n_0}(t,x) = \max\left\{ f\left(\frac{1}{2^{n_0+1}}, x\right), f(t,x) \right\}$$

we have

$$g_{n_0}(t,x) \geqslant f(t,x) \quad \text{and} \quad g_{n_0}(t,x) \geqslant f\left(\frac{1}{2^{n_0+1}}, x\right)$$

for $x \in (0,\infty)$. Thus we have

$$(y_{n_0} - \alpha_{n_0})''(t) \leqslant \begin{cases} -\big[q(t)\big\{f\big(t,\alpha(t)\big) + r\big(\alpha(t) - y_{n_0}(t)\big)\big\} + \alpha''(t)\big] \\ \quad \text{if } t \in (t_{n_0}, s_{n_0}), \\ -\big[q(t)\big\{f\big(\frac{1}{2^{n_0+1}}, \rho_{n_0}\big) + r\big(\rho_{n_0} - y_{n_0}(t)\big)\big\}\big] \\ \quad \text{if } t \in (0, t_{n_0}) \cup (s_{n_0}, 1) \end{cases}$$
$$< 0$$

from (2.72) and (2.74).

Consequently (2.78) (and so (2.77)) holds and now since $\alpha(t) \leqslant \alpha_{n_0}(t)$ for $t \in [0,1]$ we have

$$\alpha(t) \leqslant \alpha_{n_0}(t) \leqslant y_{n_0}(t) \quad \text{for } t \in [0,1]. \tag{2.79}$$

Next we show

$$y_{n_0}(t) \leqslant \beta(t) \quad \text{for } t \in [0,1]. \tag{2.80}$$

If (2.80) is not true then $y_{n_0} - \beta$ would have a positive absolute maximum at say $\tau_0 \in (0,1)$, in which case $(y_{n_0} - \beta)'(\tau_0) = 0$ and $(y_{n_0} - \beta)''(\tau_0) \leqslant 0$. There are two cases to consider, namely $\tau_0 \in [\frac{1}{2^{n_0+1}}, 1)$ and $\tau_0 \in (0, \frac{1}{2^{n_0+1}})$.

Case (i). $\tau_0 \in [\frac{1}{2^{n_0+1}}, 1)$.

Then $y_{n_0}(\tau_0) > \beta(\tau_0)$ together with $g_{n_0}(\tau_0, x) = f(\tau_0, x)$ for $x \in (0,\infty)$ gives

$$\begin{aligned}(y_{n_0} - \beta)''(\tau_0) &= -q(\tau_0)\big[g_{n_0}\big(\tau_0, \beta(\tau_0)\big) + r\big(\beta(\tau_0) - y_{n_0}(\tau_0)\big)\big] - \beta''(\tau_0) \\ &= -q(\tau_0)\big[f\big(\tau_0, \beta(\tau_0)\big) + r\big(\beta(\tau_0) - y_{n_0}(\tau_0)\big)\big] - \beta''(\tau_0) \\ &> 0\end{aligned}$$

from (2.75), a contradiction.

Case (ii). $\tau_0 \in (0, \frac{1}{2^{n_0+1}})$.

Then $y_{n_0}(\tau_0) > \beta(\tau_0)$ together with

$$g_{n_0}(\tau_0, x) = \max\left\{ f\left(\frac{1}{2^{n_0+1}}, x\right), f(\tau_0, x) \right\}$$

for $x \in (0, \infty)$ gives

$$\begin{aligned}(y_{n_0} - \beta)''(\tau_0) &= -q(\tau_0)\Big[\max\Big\{f\Big(\frac{1}{2^{n_0+1}}, \beta(\tau_0)\Big), f\big(\tau_0, \beta(\tau_0)\big)\Big\} \\ &\qquad + r\big(\beta(\tau_0) - y_{n_0}(\tau_0)\big)\Big] - \beta''(\tau_0) \\ &> 0\end{aligned}$$

from (2.75), a contradiction.

Thus (2.80) holds, so we have

$$\alpha(t) \leqslant \alpha_{n_0}(t) \leqslant y_{n_0}(t) \leqslant \beta(t) \quad \text{for } t \in [0, 1].$$

Next we consider the boundary value problem

$$\begin{cases} y'' + q(t) g^\star_{n_0+1}(t, y) = 0, & 0 < t < 1, \\ y(0) = y(1) = \rho_{n_0+1}; \end{cases} \tag{2.81}$$

here

$$g^\star_{n_0+1}(t, y) = \begin{cases} g_{n_0+1}\big(t, \alpha_{n_0+1}(t)\big) + r\big(\alpha_{n_0+1}(t) - y\big), & y \leqslant \alpha_{n_0+1}(t)\ , \\ g_{n_0+1}(t, y), & \alpha_{n_0+1}(t) \leqslant y \leqslant y_{n_0}(t), \\ g_{n_0+1}\big(t, y_{n_0}(t)\big) + r\big(y_{n_0}(t) - y\big), & y \geqslant y_{n_0}(t). \end{cases}$$

Now Schauder's fixed point theorem guarantees that (2.81) has a solution $y_{n_0+1} \in C[0, 1] \cap C^2(0, 1)$. We first show

$$y_{n_0+1}(t) \geqslant \alpha_{n_0+1}(t), \quad t \in [0, 1]. \tag{2.82}$$

Suppose (2.82) is not true. Then there exists $\tau_0, \tau_1 \in [0, 1]$ with

$$y_{n_0+1}(\tau_0) - \alpha_{n_0+1}(\tau_0) = y_{n_0+1}(\tau_1) - \alpha_{n_0+1}(\tau_1) = 0$$

and

$$y_{n_0+1}(t) - \alpha_{n_0+1}(t) < 0, \quad t \in (\tau_0, \tau_1).$$

If we show

$$(y_{n_0+1} - \alpha_{n_0+1})''(t) < 0 \quad \text{for a.e. } t \in (\tau_0, \tau_1), \tag{2.83}$$

then as before (2.82) is true. Fix $t \in (\tau_0, \tau_1)$ and assume $t \neq t_{n_0+1}$ or $t \neq s_{n_0+1}$. Then

$$(y_{n_0+1} - \alpha_{n_0+1})''(t) = \begin{cases} -\big[q(t)\big\{g_{n_0+1}\big(t, \alpha(t)\big) + r\big(\alpha(t) - y_{n_0+1}(t)\big)\big\} + \alpha''(t)\big] \\ \quad \text{if } t \in (t_{n_0+1}, s_{n_0+1}), \\ -\big[q(t)\big\{g_{n_0+1}(t, \rho_{n_0+1}) + r\big(\rho_{n_0+1} - y_{n_0+1}(t)\big)\big\}\big] \\ \quad \text{if } t \in (0, t_{n_0+1}) \cup (s_{n_0+1}, 1). \end{cases}$$

Case (i). $t \geqslant \frac{1}{2^{n_0+2}}$.

Then since $g_{n_0+1}(t,x) = f(t,x)$ for $x \in (0,\infty)$ we have

$$(y_{n_0+1} - \alpha_{n_0+1})''(t) = \begin{cases} -\big[q(t)\big\{f\big(t,\alpha(t)\big) + r\big(\alpha(t) - y_{n_0+1}(t)\big)\big\} + \alpha''(t)\big] \\ \quad \text{if } t \in (t_{n_0+1}, s_{n_0+1}), \\ -\big[q(t)\big\{f(t,\rho_{n_0+1}) + r\big(\rho_{n_0+1} - y_{n_0+1}(t)\big)\big\}\big] \\ \quad \text{if } t \in (0, t_{n_0+1}) \cup (s_{n_0+1}, 1) \end{cases}$$
$$< 0,$$

from (2.72) and (2.74).

Case (ii). $t \in (0, \frac{1}{2^{n_0+2}})$.

Then since $g_{n_0+1}(t,x)$ equals

$$\min\left\{\max\left\{f\left(\frac{1}{2^{n_0+1}}, x\right), f(t,x)\right\}, \max\left\{f\left(\frac{1}{2^{n_0+2}}, x\right), f(t,x)\right\}\right\}$$

we have

$$g_{n_0+1}(t,x) \geqslant f(t,x)$$

and

$$g_{n_0+1}(t,x) \geqslant \min\left\{f\left(\frac{1}{2^{n_0+1}}, x\right), f\left(\frac{1}{2^{n_0+2}}, x\right)\right\}$$

for $x \in (0,\infty)$. Thus we have

$$\begin{aligned} &(y_{n_0+1} - \alpha_{n_0+1})''(t) \\ &\quad \leqslant \begin{cases} -\big[q(t)\big\{f\big(t,\alpha(t)\big) + r\big(\alpha(t) - y_{n_0+1}(t)\big)\big\} + \alpha''(t)\big] \\ \quad \text{if } t \in (t_{n_0+1}, s_{n_0+1}), \\ -\big[q(t)\big\{\min\big\{f\big(\frac{1}{2^{n_0+1}}, \rho_{n_0+1}\big), f\big(\frac{1}{2^{n_0+2}}, \rho_{n_0+1}\big)\big\} \\ \qquad\qquad + r\big(\rho_{n_0+1} - y_{n_0+1}(t)\big)\big\}\big], \\ \quad \text{if } t \in (0, t_{n_0+1}) \cup (s_{n_0+1}, 1) \end{cases} \\ &\quad < 0, \end{aligned}$$

from (2.72) and (2.74) (note $f(\frac{1}{2^{n_0+1}}, \rho_{n_0+1}) \geqslant 0$ since $f(t,\rho_{n_0+1}) \geqslant 0$ for $t \in [\frac{1}{2^{n_0+2}}, 1]$ and $\frac{1}{2^{n_0+1}} \in (\frac{1}{2^{n_0+2}}, 1)$).

Consequently (2.82) is true so

$$\alpha(t) \leqslant \alpha_{n_0+1}(t) \leqslant y_{n_0+1}(t) \quad \text{for } t \in [0,1]. \tag{2.84}$$

Next we show

$$y_{n_0+1}(t) \leqslant y_{n_0}(t) \quad \text{for } t \in [0,1]. \tag{2.85}$$

If (2.85) is not true then $y_{n_0+1} - y_{n_0}$ would have a positive absolute maximum at say $\tau_0 \in (0,1)$, in which case

$$(y_{n_0+1} - y_{n_0})'(\tau_0) = 0 \quad \text{and} \quad (y_{n_0+1} - y_{n_0})''(\tau_0) \leqslant 0.$$

Then $y_{n_0+1}(\tau_0) > y_{n_0}(\tau_0)$ together with $g_{n_0}(\tau_0, x) \geqslant g_{n_0+1}(\tau_0, x)$ for $x \in (0,\infty)$ gives

$$\begin{aligned}
&(y_{n_0+1} - y_{n_0})''(\tau_0)\\
&\quad = -q(\tau_0)\big[g_{n_0+1}\big(\tau_0, y_{n_0}(\tau_0)\big) + r\big(y_{n_0}(\tau_0) - y_{n_0+1}(\tau_0)\big)\big] - y''_{n_0}(\tau_0)\\
&\quad \geqslant -q(\tau_0)\big[g_{n_0}\big(\tau_0, y_{n_0}(\tau_0)\big) + r\big(y_{n_0}(\tau_0) - y_{n_0+1}(\tau_0)\big)\big] - y''_{n_0}(\tau_0)\\
&\quad = -q(\tau_0)\big[r\big(y_{n_0}(\tau_0) - y_{n_0+1}(\tau_0)\big)\big]\\
&\quad > 0,
\end{aligned}$$

a contradiction.

Now proceed inductively to construct $y_{n_0+2}, y_{n_0+3}, \ldots$ as follows. Suppose we have y_k for some $k \in \{n_0+1, n_0+2, \ldots\}$ with $\alpha_k(t) \leqslant y_k(t) \leqslant y_{k-1}(t)$ for $t \in [0,1]$. Then consider the boundary value problem

$$\begin{cases} y'' + q(t) g^{\star}_{k+1}(t,y) = 0, & 0 < t < 1,\\ y(0) = y(1) = \rho_{k+1}; \end{cases} \tag{2.86}$$

here

$$g^{\star}_{k+1}(t,y) = \begin{cases} g_{k+1}\big(t, \alpha_{k+1}(t)\big) + r\big(\alpha_{k+1}(t) - y\big), & y \leqslant \alpha_{k+1}(t),\\ g_{k+1}(t,y), & \alpha_{k+1}(t) \leqslant y \leqslant y_k(t),\\ g_{k+1}\big(t, y_k(t)\big) + r\big(y_k(t) - y\big), & y \geqslant y_k(t). \end{cases}$$

Now Schauder's fixed point theorem guarantees that (2.86) has a solution $y_{k+1} \in C[0,1] \cap C^2(0,1)$, and essentially the same reasoning as above yields

$$\alpha(t) \leqslant \alpha_{k+1}(t) \leqslant y_{k+1}(t) \leqslant y_k(t) \quad \text{for } t \in [0,1]. \tag{2.87}$$

Thus for each $n \in \{n_0+1, \ldots\}$ we have

$$\alpha(t) \leqslant y_n(t) \leqslant y_{n-1}(t) \leqslant \cdots \leqslant y_{n_0}(t) \leqslant \beta(t) \quad \text{for } t \in [0,1]. \tag{2.88}$$

Lets look at the interval $[\frac{1}{2^{n_0+1}}, 1 - \frac{1}{2^{n_0+1}}]$. Let

$$R_{n_0} = \sup\left\{ \big|q(x) f(x,y)\big| \colon x \in \left[\frac{1}{2^{n_0+1}}, 1 - \frac{1}{2^{n_0+1}}\right] \text{ and } \alpha(x) \leqslant y \leqslant y_{n_0}(x) \right\}.$$

The mean value theorem implies that there exists $\tau \in (\frac{1}{2^{n_0+1}}, 1 - \frac{1}{2^{n_0+1}})$ with $|y_n'(\tau)| \leqslant 2 \sup_{[0,1]} y_{n_0}(t)$. As a result

$$\begin{cases} \{y_n\}_{n=n_0+1}^{\infty} \text{ is a bounded, equicontinuous} \\ \text{family on } \left[\frac{1}{2^{n_0+1}}, 1 - \frac{1}{2^{n_0+1}}\right]. \end{cases} \tag{2.89}$$

The Arzela–Ascoli theorem guarantees the existence of a subsequence N_{n_0} of integers and a function $z_{n_0} \in C[\frac{1}{2^{n_0+1}}, 1 - \frac{1}{2^{n_0+1}}]$ with y_n converging uniformly to z_{n_0} on $[\frac{1}{2^{n_0+1}}, 1 - \frac{1}{2^{n_0+1}}]$ as $n \to \infty$ through N_{n_0}. Similarly

$$\begin{cases} \{y_n\}_{n=n_0+1}^{\infty} \text{ is a bounded, equicontinuous} \\ \text{family on } \left[\frac{1}{2^{n_0+2}}, 1 - \frac{1}{2^{n_0+2}}\right], \end{cases}$$

so there is a subsequence N_{n_0+1} of N_{n_0} and a function

$$z_{n_0+1} \in C\left[\frac{1}{2^{n_0+2}}, 1 - \frac{1}{2^{n_0+2}}\right]$$

with y_n converging uniformly to z_{n_0+1} on $[\frac{1}{2^{n_0+2}}, 1 - \frac{1}{2^{n_0+2}}]$ as $n \to \infty$ through N_{n_0+1}. Note $z_{n_0+1} = z_{n_0}$ on $[\frac{1}{2^{n_0+1}}, 1 - \frac{1}{2^{n_0+1}}]$ since $N_{n_0+1} \subseteq N_{n_0}$. Proceed inductively to obtain subsequences of integers

$$N_{n_0} \supseteq N_{n_0+1} \supseteq \cdots \supseteq N_k \supseteq \cdots$$

and functions

$$z_k \in C\left[\frac{1}{2^{k+1}}, 1 - \frac{1}{2^{k+1}}\right]$$

with

$$y_n \text{ converging uniformly to } z_k \text{ on } \left[\frac{1}{2^{k+1}}, 1 - \frac{1}{2^{k+1}}\right]$$

as $n \to \infty$ through N_k, and

$$z_k = z_{k-1} \quad \text{on} \left[\frac{1}{2^k}, 1 - \frac{1}{2^k}\right].$$

Define a function $y : [0, 1] \to [0, \infty)$ by $y(x) = z_k(x)$ on $[\frac{1}{2^{k+1}}, 1 - \frac{1}{2^{k+1}}]$ and $y(0) = y(1) = 0$. Notice y is well defined and $\alpha(t) \leqslant y(t) \leqslant y_{n_0}(t) (\leqslant \beta(t))$ for $t \in (0, 1)$. Next fix $t \in (0, 1)$ (without loss of generality assume $t \neq \frac{1}{2}$) and let $m \in \{n_0, n_0 + 1, \ldots\}$ be such that $\frac{1}{2^{m+1}} < t < 1 - \frac{1}{2^{m+1}}$. Let $N_m^\star = \{n \in N_m \colon n \geqslant m\}$. Now $y_n, n \in N_m^\star$, satisfies the integral equation

$$\begin{aligned}
y_n(x) &= y_n\left(\frac{1}{2}\right) + y_n'\left(\frac{1}{2}\right)\left(x - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{x} (s - x)q(s)g_n^\star\bigl(s, y_n(s)\bigr)\,\mathrm{d}s \\
&= y_n\left(\frac{1}{2}\right) + y_n'\left(\frac{1}{2}\right)\left(x - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{x} (s - x)q(s)f\bigl(s, y_n(s)\bigr)\,\mathrm{d}s
\end{aligned}$$

for $x \in [\frac{1}{2^{m+1}}, 1 - \frac{1}{2^{m+1}}]$. Notice (take $x = \frac{2}{3}$ say) that $\{y_n'(\frac{1}{2})\}$, $n \in N_m^\star$, is a bounded sequence since $\alpha(s) \leqslant y_n(s) \leqslant y_{n_0}(s)(\leqslant \beta(s))$ for $s \in [0, 1]$. Thus $\{y_n'(\frac{1}{2})\}_{n \in N_m^\star}$ has a convergent subsequence; for convenience we will let $\{y_n'(\frac{1}{2})\}_{n \in N_m^\star}$ denote this subsequence also and let $r \in \mathbf{R}$ be its limit. Now for the above fixed t,

$$y_n(t) = y_n\left(\frac{1}{2}\right) + y_n'\left(\frac{1}{2}\right)\left(t - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{t} (s - t)q(s)f\bigl(s, y_n(s)\bigr)\,\mathrm{d}s$$

and let $n \to \infty$ through $N_m^\star$ to obtain

$$z_m(t) = z_m\left(\frac{1}{2}\right) + r\left(t - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{t} (s - t)q(s)f\bigl(s, z_m(s)\bigr)\,\mathrm{d}s,$$

i.e.,

$$y(t) = y\left(\frac{1}{2}\right) + r\left(t - \frac{1}{2}\right) + \int_{\frac{1}{2}}^{t} (s - t)q(s)f\bigl(s, y(s)\bigr)\,\mathrm{d}s.$$

We can do this argument for each $t \in (0, 1)$ and so $y''(t) + q(t)f(t, y(t)) = 0$ for $t \in (0, 1)$. It remains to show y is continuous at 0 and 1.

Let $\varepsilon > 0$ be given. Now since $\lim_{n\to\infty} y_n(0) = 0$ there exists $n_1 \in \{n_0, n_0 + 1, \ldots\}$ with $y_{n_1}(0) < \frac{\varepsilon}{2}$. Since $y_{n_1} \in C[0, 1]$ there exists $\delta_{n_1} > 0$ with

$$y_{n_1}(t) < \frac{\varepsilon}{2} \quad \text{for } t \in [0, \delta_{n_1}].$$

Now for $n \geqslant n_1$ we have, since $\{y_n(t)\}$ is nonincreasing for each $t \in [0, 1]$,

$$\alpha(t) \leqslant y_n(t) \leqslant y_{n_1}(t) < \frac{\varepsilon}{2} \quad \text{for } t \in [0, \delta_{n_1}].$$

Consequently

$$\alpha(t) \leqslant y(t) \leqslant \frac{\varepsilon}{2} < \varepsilon \quad \text{for } t \in (0, \delta_{n_1}]$$

and so y is continuous at 0. Similarly y is continuous at 1. As a result $y \in C[0, 1]$. ☐

Suppose (2.71)–(2.74) hold and in addition assume the following conditions are satisfied:

$$q(t)f(t,y)+\alpha''(t)>0 \quad \text{for } (t,y)\in(0,1)\times\bigl\{y\in(0,\infty)\colon\ y<\alpha(t)\bigr\} \tag{2.90}$$

and

$$\begin{cases} \text{there exists a function } \beta\in C[0,1]\cap C^2(0,1) \\ \text{with } \beta(t)\geqslant\rho_{n_0} \text{ for } t\in[0,1] \text{ and with} \\ q(t)f\bigl(t,\beta(t)\bigr)+\beta''(t)\leqslant 0 \text{ for } t\in(0,1) \text{ and} \\ q(t)f\bigl(\frac{1}{2^{n_0+1}},\beta(t)\bigr)+\beta''(t)\leqslant 0 \text{ for } t\in\bigl(0,\frac{1}{2^{n_0+1}}\bigr). \end{cases} \tag{2.91}$$

Then the result in Theorem 2.9 is again true. This follows immediately from Theorem 2.9 once we show (2.75) holds, i.e., once we show $\beta(t)\geqslant\alpha(t)$ for $t\in[0,1]$. Suppose it is false. Then $\alpha-\beta$ would have a positive absolute maximum at say $\tau_0\in(0,1)$, so $(\alpha-\beta)'(\tau_0)=0$ and $(\alpha-\beta)''(\tau_0)\leqslant 0$. Now $\alpha(\tau_0)>\beta(\tau_0)$ and (2.90) implies

$$q(\tau_0)f\bigl(\tau_0,\beta(\tau_0)\bigr)+\alpha''(\tau_0)>0.$$

This together with (2.91) yields

$$(\alpha-\beta)''(\tau_0)=\alpha''(\tau_0)-\beta''(\tau_0)\geqslant\alpha''(\tau_0)+q(\tau_0)f\bigl(\tau_0,\beta(\tau_0)\bigr)>0,$$

a contradiction. Thus we have

COROLLARY 2.10. *Let* $n_0\in\{1,2,\ldots\}$ *be fixed and suppose* (2.71)–(2.74), (2.90) *and* (2.91) *hold. Then* (2.62) *has a solution* $y\in C[0,1]\cap C^2(0,1)$ *with* $y(t)\geqslant\alpha(t)$ *for* $t\in[0,1]$.

REMARK 2.9. If in (2.72) we replace $\frac{1}{2^{n+1}}\leqslant t\leqslant 1$ with $0\leqslant t\leqslant 1-\frac{1}{2^{n+1}}$ then one would replace (2.75) with

$$\begin{cases} \text{there exists a function } \beta\in C[0,1]\cap C^2(0,1) \\ \text{with } \beta(t)\geqslant\alpha(t) \text{ and } \beta(t)\geqslant\rho_{n_0} \text{ for } t\in[0,1] \\ \text{with } q(t)f\bigl(t,\beta(t)\bigr)+\beta''(t)\leqslant 0 \text{ for } t\in(0,1) \text{ and} \\ q(t)f\bigl(1-\frac{1}{2^{n_0+1}},\beta(t)\bigr)+\beta''(t)\leqslant 0 \text{ for } t\in\bigl(1-\frac{1}{2^{n_0+1}},1\bigr). \end{cases} \tag{2.92}$$

If in (2.72) we replace $\frac{1}{2^{n+1}}\leqslant t\leqslant 1$ with $\frac{1}{2^{n+1}}\leqslant t\leqslant 1-\frac{1}{2^{n+1}}$ then essentially the same reasoning as in Theorem 2.9 establishes the following results.

THEOREM 2.11. *Let* $n_0\in\{1,2,\ldots\}$ *be fixed and suppose* (2.71), (2.73), (2.74) *and the following hold*:

$$\begin{cases} \textit{let } n \in \{n_0, n_0+1, \dots\} \textit{ and associated with each } n \textit{ we} \\ \textit{have a constant } \rho_n \textit{ such that } \{\rho_n\} \textit{ is a nonincreasing} \\ \textit{sequence with } \lim_{n\to\infty} \rho_n = 0 \textit{ and such that for} \\ \frac{1}{2^{n+1}} \leqslant t \leqslant 1 - \frac{1}{2^{n+1}} \textit{ we have } q(t) f(t, \rho_n) \geqslant 0 \end{cases} \tag{2.93}$$

and

$$\begin{cases} \textit{there exists a function } \beta \in C[0,1] \cap C^2(0,1) \\ \textit{with } \beta(t) \geqslant \alpha(t) \textit{ and } \beta(t) \geqslant \rho_{n_0} \textit{ for } t \in [0,1] \\ \textit{with } q(t) f\big(t, \beta(t)\big) + \beta''(t) \leqslant 0 \textit{ for } t \in (0,1) \textit{ and} \\ q(t) f\big(\frac{1}{2^{n_0+1}}, \beta(t)\big) + \beta''(t) \leqslant 0 \textit{ for } t \in \big(0, \frac{1}{2^{n_0+1}}\big) \textit{ and} \\ q(t) f\big(1 - \frac{1}{2^{n_0+1}}, \beta(t)\big) + \beta''(t) \leqslant 0 \textit{ for } t \in \big(1 - \frac{1}{2^{n_0+1}}, 1\big). \end{cases} \tag{2.94}$$

Then (2.62) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0,1]$.

PROOF. In this case let

$$e_n = \left[\frac{1}{2^{n+1}}, 1 - \frac{1}{2^{n+1}}\right] \quad \text{and} \quad \theta_n(t) = \max\left\{\frac{1}{2^{n+1}}, \min\left\{t, 1 - \frac{1}{2^{n+1}}\right\}\right\}.$$

□

COROLLARY 2.12. *Let* $n_0 \in \{1, 2, \dots\}$ *be fixed and suppose* (2.71), (2.73), (2.74), (2.90), (2.93) *and the following hold*:

$$\begin{cases} \textit{there exists a function } \beta \in C[0,1] \cap C^2(0,1) \\ \textit{with } \beta(t) \geqslant \rho_{n_0} \textit{ for } t \in [0,1] \textit{ and with} \\ q(t) f\big(t, \beta(t)\big) + \beta''(t) \leqslant 0 \textit{ for } t \in (0,1) \textit{ and} \\ q(t) f\big(\frac{1}{2^{n_0+1}}, \beta(t)\big) + \beta''(t) \leqslant 0 \textit{ for } t \in \big(0, \frac{1}{2^{n_0+1}}\big) \textit{ and} \\ q(t) f\big(1 - \frac{1}{2^{n_0+1}}, \beta(t)\big) + \beta''(t) \leqslant 0 \textit{ for } t \in \big(1 - \frac{1}{2^{n_0+1}}, 1\big). \end{cases} \tag{2.95}$$

Then (2.62) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0,1]$.

Next we discuss how to construct the lower solution α in (2.74) and (2.90). Suppose the following condition is satisfied:

$$\begin{cases} \text{let } n \in \{n_0, n_0+1, \dots\} \text{ and associated with each } n \text{ we} \\ \text{have a constant } \rho_n \text{ such that } \{\rho_n\} \text{ is a decreasing} \\ \text{sequence with } \lim_{n\to\infty} \rho_n = 0 \text{ and there exists a} \\ \text{constant } k_0 > 0 \text{ such that for } \frac{1}{2^{n+1}} \leqslant t \leqslant 1 \\ \text{and } 0 < y \leqslant \rho_n \text{ we have } q(t) f(t, y) \geqslant k_0. \end{cases} \tag{2.96}$$

We will show if (2.96) holds then (2.74) (and of course (2.72)) and (2.90) are satisfied (we also note that $\frac{1}{2^{n+1}} \leqslant t \leqslant 1$ in (2.96) could be replaced by $0 \leqslant t \leqslant 1 - \frac{1}{2^{n+1}}$

(respectively $\frac{1}{2^{n+1}} \leqslant t \leqslant 1 - \frac{1}{2^{n+1}}$) and (2.74), (2.90) hold with $\frac{1}{2^{n+1}} \leqslant t \leqslant 1$ replaced by $0 \leqslant t \leqslant 1 - \frac{1}{2^{n+1}}$ (respectively $\frac{1}{2^{n+1}} \leqslant t \leqslant 1 - \frac{1}{2^{n+1}}$)).

To show (2.74) and (2.90) recall the following well-known lemma.

LEMMA 2.13. *Let e_n be as described in Theorem* 2.11 (*or Theorem* 2.9) *and let* $0 < \varepsilon_n < 1$ *with* $\varepsilon_n \downarrow 0$. *Then there exists* $\lambda \in C^2[0,1]$ *with* $\sup_{[0,1]} |\lambda''(t)| > 0$ *and* $\lambda(0) = \lambda(1) = 0$ *with*

$$0 < \lambda(t) \leqslant \varepsilon_n \quad \textit{for } t \in e_n \backslash e_{n-1}, n \geqslant 1.$$

PROOF. Let $r : [0,1] \to [0,\infty)$ be such that $r(0) = r(1) = 0$ and $r(t) = \varepsilon_n$ for all $t \in e_n \backslash e_{n-1}$, $n \geqslant 1$. Moreover, let

$$u(t) = \int_0^t r(s)\,\mathrm{d}s, \quad v(t) = \int_0^t u(s)\,\mathrm{d}s \quad \text{and} \quad w(t) = \int_0^t v(s)\,\mathrm{d}s.$$

It is obvious that u, v, and $w : [0, \frac{1}{2}] \to [0,\infty)$ are continuous and strictly increasing, with $w(\frac{1}{4}) < \varepsilon_1$.

Choose a natural number $k \geqslant 2$ with

$$\frac{(4k+1)v(\frac{1}{4}) + 4v'(\frac{1}{4})}{16k(k+1)} \leqslant \varepsilon_1 - w\left(\frac{1}{4}\right).$$

Let

$$\begin{aligned} c_0 &= \frac{4^{2k}[(2k-1)v(\frac{1}{4}) + 4v'(\frac{1}{4})]}{k+1}, \\ c_1 &= -\frac{4^{2(k-1)}[(2k-1)v(\frac{1}{4}) + 4v'(\frac{1}{4})]}{k}, \\ c_2 &= w\left(\frac{1}{4}\right) + \frac{(4k+1)v(\frac{1}{4}) + 4v'(\frac{1}{4})}{16k(k+1)}, \end{aligned}$$

and

$$p(t) = c_0\left(t - \frac{1}{2}\right)^{2(k+1)} + c_1\left(t - \frac{1}{2}\right)^{2k} + c_2.$$

Define $\lambda : [0,1] \to [0,\infty)$ as follows:

$$\lambda(t) = \begin{cases} w(t), & 0 \leqslant t \leqslant \frac{1}{4}, \\ p(t), & \frac{1}{4} \leqslant t \leqslant \frac{3}{4}, \\ w(1-t), & \frac{3}{4} \leqslant t \leqslant 1. \end{cases}$$

Then λ satisfies the conditions of the lemma. □

Let $\varepsilon_n = \rho_n$ (and $n \geqslant n_0$) and let λ be as in Lemma 2.13. From (2.96) there exists $k_0 > 0$ with

$$q(t)f(t,y) \geqslant k_0 \quad \text{for } (t,y) \in (0,1) \times \{y \in (0,\infty)\colon\ y \leqslant \lambda(t)\} \tag{2.97}$$

since if $t \in e_n \backslash e_{n-1}$ $(n \geqslant n_0)$ then $y \leqslant \lambda(t)$ implies $y \leqslant \rho_n$. Let

$$M = \sup_{[0,1]} |\lambda''(t)|, \quad m = \min\left\{1, \frac{k_0}{M+1}\right\} \quad \text{and} \quad \alpha(t) = m\lambda(t), \quad t \in [0,1].$$

In particular since $\alpha(t) \leqslant \lambda(t)$ we have from (2.97) that

$$q(t)f\bigl(t,\alpha(t)\bigr) + \alpha''(t) \geqslant k_0 + \alpha''(t) \geqslant k_0 - \frac{k_0|\lambda''(t)|}{M+1} > 0$$

for $t \in (0,1)$, and also

$$q(t)f(t,y) + \alpha''(t) \geqslant k_0 + \alpha''(t) > 0$$

for $(t,y) \in (0,1) \times \{y \in (0,\infty)\colon\ y \leqslant \alpha(t)\}$. Thus (2.74) and (2.90) hold.

THEOREM 2.14. *Let* $n_0 \in \{1,2,\dots\}$ *be fixed and suppose* (2.71), (2.73), (2.91) *and* (2.96) *hold. Then* (2.62) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y(t) > 0$ *for* $t \in (0,1)$.

If in (2.96) we replace $\frac{1}{2^{n+1}} \leqslant t \leqslant 1$ with $\frac{1}{2^{n+1}} \leqslant t \leqslant 1 - \frac{1}{2^{n+1}}$ then (2.74) and (2.90) also hold. We combine this with Corollary 2.12 to obtain our next result.

THEOREM 2.15. *Let* $n_0 \in \{1,2,\dots\}$ *be fixed and suppose* (2.71), (2.73), (2.95) *and* (2.96) (*with* $\frac{1}{2^{n+1}} \leqslant t \leqslant 1$ *replaced by* $\frac{1}{2^{n+1}} \leqslant t \leqslant 1 - \frac{1}{2^{n+1}}$) *hold. Then* (2.62) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y(t) > 0$ *for* $t \in (0,1)$.

Looking at Theorem 2.14 we see that the main difficulty when discussing examples is the construction of the β in (2.91). Our next result replaces (2.91) with a superlinear type condition. We first prove the result in its full generality.

THEOREM 2.16. *Let* $n_0 \in \{1,2,\dots\}$ *be fixed and suppose* (2.71)–(2.74) *hold. Also assume the following two conditions are satisfied*:

$$\begin{cases} |f(t,y)| \leqslant g(y) + h(y) \textit{ on } [0,1] \times (0,\infty) \textit{ with} \\ g > 0 \textit{ continuous and nonincreasing on } (0,\infty), \\ h \geqslant 0 \textit{ continuous on } [0,\infty), \textit{ and } \frac{h}{g} \\ \textit{nondecreasing on } (0,\infty) \end{cases} \tag{2.98}$$

and

$$\begin{cases} \textit{for any } R > 0,\ \frac{1}{g} \textit{ is differentiable on } (0, R] \textit{ with} \\ g' < 0 \textit{ a.e. on } (0, R] \textit{ and } \frac{g'}{g^2} \in L^1[0, R]. \end{cases} \tag{2.99}$$

In addition suppose there exists $M > \sup_{t\in[0,1]} \alpha(t)$ *with*

$$\frac{1}{\{1 + \frac{h(M)}{g(M)}\}} \int_0^M \frac{\mathrm{d}u}{g(u)} > b_0 \tag{2.100}$$

holding; *here*

$$b_0 = \max\left\{2\int_0^{\frac{1}{2}} t(1-t)q(t)\,\mathrm{d}t,\, 2\int_{\frac{1}{2}}^1 t(1-t)q(t)\,\mathrm{d}t\right\}.$$

Then (2.62) *has a solution* $y \in C[0, 1] \cap C^2(0, 1)$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0, 1]$.

PROOF. Fix $n \in \{n_0, n_0 + 1, \ldots\}$. Choose ε, $0 < \varepsilon < M$ with

$$\frac{1}{\{1 + \frac{h(M)}{g(M)}\}} \int_\varepsilon^M \frac{\mathrm{d}u}{g(u)} > b_0. \tag{2.101}$$

Let $m_0 \in \{1, 2, \ldots\}$ be chosen so that $\rho_{m_0} < \varepsilon$ and without loss of generality assume $m_0 \leqslant n_0$. Let e_n, θ_n, f_n, g_n and α_n be as in Theorem 2.9. We consider the boundary value problem (2.76) with in this case $g^\star_{n_0}$ given by

$$g^\star_{n_0}(t, y) = \begin{cases} g_{n_0}\big(t, \alpha_{n_0}(t)\big) + r\big(\alpha_{n_0}(t) - y\big), & y \leqslant \alpha_{n_0}(t)\ , \\ g_{n_0}(t, y), & \alpha_{n_0}(t) \leqslant y \leqslant M, \\ g_{n_0}(t, M) + r(M - y), & y \geqslant M. \end{cases}$$

Essentially the same reasoning as in Theorem 2.9 implies that (2.76) has a solution $y_{n_0} \in C[0, 1] \cap C^2(0, 1)$ with $y_{n_0}(t) \geqslant \alpha_{n_0}(t) \geqslant \alpha(t)$ for $t \in [0, 1]$. Next we show

$$y_{n_0}(t) \leqslant M \quad \text{for } t \in [0, 1]. \tag{2.102}$$

Suppose (2.102) is false. Now since $y_{n_0}(0) = y_{n_0}(1) = \rho_{n_0}$ there exists either

(i) $t_1, t_2 \in (0, 1)$, $t_2 < t_1$ with $\alpha_{n_0}(t) \leqslant y_{n_0}(t) \leqslant M$ for $t \in [0, t_2)$, $y_{n_0}(t_2) = M$ and $y_{n_0}(t) > M$ on (t_2, t_1) with $y'_{n_0}(t_1) = 0$; or

(ii) $t_3, t_4 \in (0, 1)$, $t_4 < t_3$ with $\alpha_{n_0}(t) \leqslant y_{n_0}(t) \leqslant M$ for $t \in (t_3, 1]$, $y_{n_0}(t_3) = M$ and $y_{n_0}(t) > M$ on (t_4, t_3) with $y'_{n_0}(t_4) = 0$.

We can assume without loss of generality that either $t_1 \leqslant \frac{1}{2}$ or $t_4 \geqslant \frac{1}{2}$. Suppose $t_1 \leqslant \frac{1}{2}$. Notice for $t \in (t_2, t_1)$ that we have

$$-y''_{n_0}(t) = q(t)g^\star_{n_0}\big(t, y_{n_0}(t)\big) \leqslant q(t)\big[g(M) + h(M)\big]; \tag{2.103}$$

note for $t \in (t_2, t_1)$ that we have from (2.98),

$$g^\star_{n_0}\bigl(t, y_{n_0}(t)\bigr) = g_{n_0}(t, M) + r\bigl(M - y_{n_0}(t)\bigr) \leqslant \max\left\{f\left(\frac{1}{2^{n_0+1}}, M\right), f(t, M)\right\}.$$

Integrate (2.103) from t_2 to t_1 to obtain

$$y'_{n_0}(t_2) \leqslant \bigl[g(M) + h(M)\bigr]\int_{t_2}^{t_1} q(s)\,\mathrm{d}s$$

and this together with $y_{n_0}(t_2) = M$ yields

$$\frac{y'_{n_0}(t_2)}{g(y_{n_0}(t_2))} \leqslant \left\{1 + \frac{h(M)}{g(M)}\right\}\int_{t_2}^{t_1} q(s)\,\mathrm{d}s. \tag{2.104}$$

Also for $t \in (0, t_2)$ we have

$$\begin{aligned}-y''_{n_0}(t) &= q(t)\max\left\{f\left(\frac{1}{2^{n_0+1}}, y_{n_0}(t)\right), f\bigl(t, y_{n_0}(t)\bigr)\right\}\\ &\leqslant q(t)\bigl[g\bigl(y_{n_0}(t)\bigr) + h\bigl(y_{n_0}(t)\bigr)\bigr],\end{aligned}$$

and so

$$\frac{-y''_{n_0}(t)}{g(y_{n_0}(t))} \leqslant q(t)\left\{1 + \frac{h(y_{n_0}(t))}{g(y_{n_0}(t))}\right\} \leqslant q(t)\left\{1 + \frac{h(M)}{g(M)}\right\}$$

for $t \in (0, t_2)$. Integrate from t $(t \in (0, t_2))$ to t_2 to obtain

$$\begin{aligned}&\frac{-y'_{n_0}(t_2)}{g(y_{n_0}(t_2))} + \frac{y'_{n_0}(t)}{g(y_{n_0}(t))} + \int_t^{t_2}\left\{\frac{-g'(y_{n_0}(x))}{g^2(y_{n_0}(x))}\right\}\bigl[y'_{n_0}(x)\bigr]^2\,\mathrm{d}x\\ &\quad\leqslant \left\{1 + \frac{h(M)}{g(M)}\right\}\int_t^{t_2} q(x)\,\mathrm{d}x,\end{aligned}$$

and this together with (2.104) (and (2.99)) yields

$$\frac{y'_{n_0}(t)}{g(y_{n_0}(t))} \leqslant \left\{1 + \frac{h(M)}{g(M)}\right\}\int_t^{t_1} q(x)\,\mathrm{d}x \quad \text{for } t \in (0, t_2).$$

Integrate from 0 to t_2 to obtain

$$\int_\varepsilon^M \frac{\mathrm{d}u}{g(u)} \leqslant \int_{\rho_{n_0}}^M \frac{\mathrm{d}u}{g(u)} \leqslant \left\{1 + \frac{h(M)}{g(M)}\right\}\frac{1}{1 - t_1}\int_0^{t_1} x(1 - x)q(x)\,\mathrm{d}x.$$

That is

$$\int_{\varepsilon}^{M}\frac{\mathrm{d}u}{g(u)}\leqslant\left\{1+\frac{h(M)}{g(M)}\right\}2\int_{0}^{\frac{1}{2}}x(1-x)q(x)\,\mathrm{d}x\leqslant b_0\left\{1+\frac{h(M)}{g(M)}\right\}.$$

This contradicts (2.101) so (2.102) holds (a similar argument yields a contradiction if $t_4\geqslant\frac{1}{2}$). Thus we have

$$\alpha(t)\leqslant\alpha_{n_0}(t)\leqslant y_{n_0}(t)\leqslant M\quad\text{for } t\in[0,1].$$

Essentially the same reasoning as in Theorem 2.9 (from (2.82) onwards) completes the proof. □

Similarly we have the following result.

THEOREM 2.17. *Let* $n_0\in\{1,2,\dots\}$ *be fixed and suppose* (2.71), (2.73), (2.74), (2.93), (2.98) *and* (2.99) *hold. In addition assume there exists*

$$M>\sup_{t\in[0,1]}\alpha(t)$$

with (2.100) *holding. Then* (2.62) *has a solution* $y\in C[0,1]\cap C^2(0,1)$ *with* $y(t)\geqslant\alpha(t)$ *for* $t\in[0,1]$.

COROLLARY 2.18. *Let* $n_0\in\{1,2,\dots\}$ *be fixed and suppose* (2.71)–(2.74), (2.90), (2.98) *and* (2.99) *hold. In addition assume there is a constant* $M>0$ *with*

$$\frac{1}{\{1+\frac{h(M)}{g(M)}\}}\int_0^M\frac{\mathrm{d}u}{g(u)}>b_0 \tag{2.105}$$

holding; *here*

$$b_0=\max\left\{2\int_0^{\frac{1}{2}}t(1-t)q(t)\,\mathrm{d}t,\,2\int_{\frac{1}{2}}^{1}t(1-t)q(t)\,\mathrm{d}t\right\}.$$

Then (2.62) *has a solution* $y\in C[0,1]\cap C^2(0,1)$ *with* $y(t)\geqslant\alpha(t)$ *for* $t\in[0,1]$.

PROOF. The result follows immediately from Theorem 2.16 once we show $\alpha(t)\leqslant M$ for $t\in[0,1]$. Suppose this is false. Now since $\alpha(0)=\alpha(1)=0$ there exists either

(i) $t_1,t_2\in(0,1)$, $t_2<t_1$ with $0\leqslant\alpha(t)\leqslant M$ for $t\in[0,t_2)$, $\alpha(t_2)=M$ and $\alpha(t)>M$ on (t_2,t_1) with $\alpha'(t_1)=0$; or

(ii) $t_3,t_4\in(0,1)$, $t_4<t_3$ with $0\leqslant\alpha(t)\leqslant M$ for $t\in(t_3,1]$, $\alpha(t_3)=M$ and $\alpha(t)>M$ on (t_4,t_3) with $\alpha'(t_4)=0$.

We can assume without loss of generality that either $t_1 \leqslant \frac{1}{2}$ or $t_4 \geqslant \frac{1}{2}$. Suppose $t_1 \leqslant \frac{1}{2}$. Notice for $t \in (t_2, t_1)$ that we have from (2.90) and (2.98) that

$$-\alpha''(t) \leqslant q(t)\bigl[g(M) + h(M)\bigr]$$

so integration from t_2 to t_1 yields

$$\frac{\alpha'(t_2)}{g(\alpha(t_2))} \leqslant \left\{1 + \frac{h(M)}{g(M)}\right\} \int_{t_2}^{t_1} q(s)\,\mathrm{d}s. \tag{2.106}$$

Also for $t \in (0, t_2)$ we have from (2.90) and (2.98) that

$$-\alpha''(t) \leqslant q(t)g\bigl(\alpha(t)\bigr)\left\{1 + \frac{h(\alpha(t))}{g(\alpha(t))}\right\} \leqslant q(t)g\bigl(\alpha(t)\bigr)\left\{1 + \frac{h(M)}{g(M)}\right\}.$$

Integrate from t ($t \in (0, t_2)$) to t_2 and use (2.106) to obtain

$$\frac{\alpha'(t)}{g(\alpha(t))} \leqslant \left\{1 + \frac{h(M)}{g(M)}\right\} \int_{t}^{t_1} q(s)\,\mathrm{d}s \quad \text{for } t \in (0, t_2).$$

Finally integrate from 0 to t_2 to obtain

$$\int_0^M \frac{\mathrm{d}u}{g(u)} \leqslant b_0\left\{1 + \frac{h(M)}{g(M)}\right\},$$

a contradiction. □

COROLLARY 2.19. *Let* $n_0 \in \{1, 2, \ldots\}$ *be fixed and suppose* (2.71), (2.73), (2.74), (2.90), (2.93), (2.98) *and* (2.99) *hold. In addition assume there is a constant* $M > 0$ *with* (2.105) *holding. Then* (2.62) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0,1]$.

Combining Corollary 2.18 with the comments before Theorem 2.14 yields the following theorem.

THEOREM 2.20. *Let* $n_0 \in \{1, 2, \ldots\}$ *be fixed and suppose* (2.71), (2.73), (2.96), (2.98) *and* (2.99) *hold. In addition assume there exists* $M > 0$ *with* (2.105) *holding. Then* (2.62) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y(t) > 0$ *for* $t \in (0,1)$.

Similarly combining Corollary 2.19 with the comments before Theorem 2.14 yields the following theorem.

THEOREM 2.21. *Let* $n_0 \in \{1, 2, \ldots\}$ *be fixed and suppose* (2.71), (2.73), (2.96) (*with* $\frac{1}{2^{n+1}} \leqslant t \leqslant 1$ *replaced by* $\frac{1}{2^{n+1}} \leqslant t \leqslant 1 - \frac{1}{2^{n+1}}$), (2.98) *and* (2.99) *hold. In addition assume there exists* $M > 0$ *with* (2.105) *holding. Then* (2.62) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y(t) > 0$ *for* $t \in (0,1)$.

Next we present some examples which illustrate how easily the theory is applied in practice.

EXAMPLE 2.2. Consider the boundary value problem

$$\begin{cases} y'' + \left(At^{\gamma} y^{-\theta} - \mu^2\right) = 0, & 0 < t < 1, \\ y(0) = y(1) = 0 \end{cases} \tag{2.107}$$

with $A > 0, \theta > 0, \gamma > -2$ and $\mu \in \mathbf{R}$. Then (2.107) has a solution $y \in C[0,1] \cap C^2(0,1)$ with $y(t) > 0$ for $t \in (0,1)$.

To see this we will apply Theorem 2.20. We will consider two cases, namely $\gamma \geqslant 0$ and $-2 < \gamma < 0$.

Case (i). $\gamma \geqslant 0$.

We will apply Theorem 2.20 with

$$q(t) = 1, \quad g(y) = Ay^{-\theta} \quad \text{and} \quad h(y) = \mu^2.$$

Clearly (2.71), (2.73), (2.98) and (2.99) hold. Let

$$n_0 = 1, \quad \rho_n = \left(\frac{A}{2^{(n+1)\gamma}(\mu^2 + 1)}\right)^{1/\theta} \quad \text{and} \quad k_0 = 1.$$

Notice for $n \in \{1, 2, \ldots\}$, $\frac{1}{2^{n+1}} \leqslant t \leqslant 1$ and $0 < y \leqslant \rho_n$ that we have

$$q(t) f(t, y) \geqslant \frac{A}{2^{(n+1)\gamma} \rho_n^{\theta}} - \mu^2 = \left(\mu^2 + 1\right) - \mu^2 = 1,$$

so (2.96) is satisfied. Finally notice for $c > 0$ that

$$\frac{1}{\{1 + \frac{h(c)}{g(c)}\}} \int_0^c \frac{\mathrm{d}u}{g(u)} = \frac{1}{\theta + 1} \frac{c^{\theta+1}}{A + \mu^2 c^{\theta}},$$

so

$$\lim_{c \to \infty} \frac{1}{\{1 + \frac{h(c)}{g(c)}\}} \int_0^c \frac{\mathrm{d}u}{g(u)} = \infty.$$

Thus there exists $M > 0$ with (2.105) holding. Existence of a solution is now guaranteed from Theorem 2.20.

Case (ii). $-2 < \gamma < 0$.

We will apply Theorem 2.20 with

$$q(t) = t^{\gamma}, \quad g(y) = Ay^{-\theta} \quad \text{and} \quad h(y) = \mu^2.$$

Clearly (2.71), (2.73), (2.98) and (2.99) hold. Also as in Case (i) there exists $M > 0$ with (2.105) holding. Let

$$n_0 = 1, \quad \rho_n = \left(\frac{A}{n(\mu^2+1)}\right)^{1/\theta} \quad \text{and} \quad k_0 = 1.$$

Notice for $n \in \{1, 2, \ldots\}$, $\frac{1}{2^{n+1}} \leqslant t \leqslant 1$ and $0 < y \leqslant \rho_n$ that we have since $\gamma < 0$,

$$q(t)f(t,y) \geqslant \frac{At^\gamma}{\rho_n^\theta} - \mu^2 \geqslant \frac{A}{\rho_n^\theta} - \mu^2 = n\big(\mu^2+1\big) - \mu^2 \geqslant \big(\mu^2+1\big) - \mu^2 = 1.$$

Thus (2.96) is satisfied. Existence of a solution is now guaranteed from Theorem 2.20.

EXAMPLE 2.3. Consider the boundary value problem

$$\begin{cases} y'' + \left(\frac{t}{y^2} + \frac{1}{32}y^2 - \mu^2\right) = 0, & 0 < t < 1, \\ y(0) = y(1) = 0, \end{cases} \tag{2.108}$$

where $\mu^2 \geqslant 1$. Then (2.108) has a solution $y \in C[0,1] \cap C^2(0,1)$ with $y(t) > 0$ for $t \in (0,1)$.

To see that (2.108) has the desired solution we will apply Theorem 2.14 with

$$q \equiv 1, \quad \rho_n = \left(\frac{1}{2^{n+1}(\mu^2+a)}\right)^{\frac{1}{2}} \quad \text{and} \quad k_0 = a;$$

here $a > 0$ is chosen so that $a \leqslant \frac{1}{8}$. Also we choose $n_0 \in \{1, 2, \ldots\}$ with $\rho_{n_0} \leqslant 1$. Clearly (2.71) and (2.73) hold. Notice for $n \in \{1, 2, \ldots\}$, $\frac{1}{2^{n+1}} \leqslant t \leqslant 1$ and $0 < y \leqslant \rho_n$ that we have

$$q(t)f(t,y) \geqslant \frac{t}{y^2} - \mu^2 \geqslant \frac{1}{2^{n+1}\rho_n^2} - \mu^2 = \big(\mu^2+a\big) - \mu^2 = a,$$

so (2.96) is satisfied. It remains to check (2.91) with

$$\beta(t) = \sqrt{t} + \rho_{n_0}.$$

Now $\beta''(t) = -\frac{1}{4}t^{-\frac{3}{2}}$ and so for $t \in (0,1)$ we have

$$\begin{aligned} \beta''(t) + q(t)f\big(t,\beta(t)\big) &\leqslant -\frac{1}{4}\frac{1}{t^{\frac{3}{2}}} + \left(\frac{t}{t} + \frac{(\sqrt{t}+\rho_{n_0})^2}{32} - \mu^2\right) \\ &\leqslant -\frac{1}{4} + \left(1 + \frac{1}{8} - \mu^2\right) \leqslant 0. \end{aligned}$$

Also for $t \in (0, \frac{1}{2^{n_0+1}})$ we have

$$\begin{aligned}\beta''(t) + q(t)f\big(t, \beta(t)\big) &\leqslant -\frac{1}{4}\frac{1}{t^{\frac{3}{2}}} + \left(\frac{1}{2^{n_0+1}\rho_{n_0}^2} + \frac{(\sqrt{t}+\rho_{n_0})^2}{32} - \mu^2\right) \\ &\leqslant -\frac{1}{4} + \left(\big(\mu^2 + a\big) + \frac{1}{8} - \mu^2\right) \\ &= a - \frac{1}{8} \leqslant 0.\end{aligned}$$

As a result (2.91) holds so existence is now guaranteed from Theorem 2.14.

In the literature nonresonant results [19] have been presented for nonsingular Dirichlet problems, i.e., for (1.2) when qf is a Carathéodory function. Next, by combining some of the ideas in [19] with those above, we present a nonresonant theory for the singular problem (2.62). It is worth remarking here that we could consider Sturm Liouville boundary data in (2.62); however since the arguments are essentially the same (in fact easier) we will restrict our discussion to Dirichlet boundary data.

The results here rely on the following well-known Rayleigh–Ritz inequality.

THEOREM 2.22. *Suppose* $q \in C(0,1) \cap L^1[0,1]$ *with* $q > 0$ *on* $(0,1)$. *Let* λ_1 *be the first eigenvalue of*

$$\begin{cases} y'' + \lambda q y = 0, & 0 < t < 1, \\ y(0) = 0 = y(1). \end{cases} \tag{2.109}$$

Then

$$\lambda_1 \int_0^1 q(t)\big|v(t)\big|^2\,\mathrm{d}t \leqslant \int_0^1 \big|v'(t)\big|^2\,\mathrm{d}t$$

for all functions $v \in AC[0,1]$ *with* $v' \in L^2[0,1]$ *and* $v(0) = v(1) = 0$.

For notational purposes in our next theorem, for appropriate functions u we let

$$\|u\|_2 = \left(\int_0^1 \big|u(t)\big|^2\,\mathrm{d}t\right)^{\frac{1}{2}}, \quad |u|_\infty = \sup_{[0,1]}\big|u(t)\big| \quad \text{and} \quad \|u\|_1 = \int_0^1 \big|u(t)\big|\,\mathrm{d}t.$$

We begin with our main result (in fact a more general result will be presented at the end of this section; see Theorem 2.26)).

THEOREM 2.23. *Let* $n_0 \in \{1, 2, \ldots\}$ *be fixed and suppose* (2.72) *and* (2.74) *hold. In addition assume the following conditions are satisfied*:

$$f:(0,1)\times(0,\infty)\to\mathbf{R}\quad \textit{is continuous} \tag{2.110}$$

$$q\in C(0,1)\cap L^1[0,1]\quad \textit{with } q>0 \textit{ on } (0,1) \tag{2.111}$$

and

$$\begin{cases} \textit{for any } \varepsilon>0, \exists a_0\geqslant 0 \textit{ with } a_0<\lambda_1, b_0\geqslant 0, 0\leqslant\gamma<1,\\ h_\varepsilon\in L^1_q[0,1] \textit{ with } h_\varepsilon\geqslant 0 \textit{ a.e. on } (0,1) \textit{ with}\\ \left|f(t,u)\right|\leqslant a_0u+b_0u^\gamma+h_\varepsilon(t) \textit{ for } t\in(0,1) \textit{ and } u\geqslant\varepsilon; \end{cases} \tag{2.112}$$

here λ_1 *is the first eigenvalue of* (2.109). *Then* (2.62) *has a solution* $y\in C[0,1]\cap C^2(0,1)$ *with* $y(t)\geqslant\alpha(t)$ *for* $t\in[0,1]$.

PROOF. For $n=n_0,n_0+1,\dots$ let e_n, θ_n, f_n and g_n be as in Theorem 2.9. Without loss of generality assume $\rho_{n_0}\leqslant\min_{t\in[\frac{1}{3},\frac{2}{3}]}\alpha(t)$. Fix $n\in\{n_0,n_0+1,\dots\}$. Let $t_n\in[0,\frac{1}{3}]$ and $s_n\in[\frac{2}{3},1]$ be such that

$$\alpha(t_n)=\alpha(s_n)=\rho_n\quad\text{and}\quad\alpha(t)\leqslant\rho_n\quad\text{for } t\in[0,t_n]\cup[s_n,1].$$

Define

$$\alpha_n(t)=\begin{cases}\rho_n & \text{if } t\in[0,t_n]\cup[s_n,1],\\ \alpha(t) & \text{if } t\in(t_n,s_n).\end{cases}$$

We begin with the boundary value problem

$$\begin{cases} y''+q(t)g^\star_{n_0}(t,y)=0, & 0<t<1,\\ y(0)=y(1)=\rho_{n_0}; \end{cases} \tag{2.113}$$

here

$$g^\star_{n_0}(t,y)=\begin{cases} g_{n_0}\big(t,\alpha_{n_0}(t)\big)+r\big(\alpha_{n_0}(t)-y\big), & y\leqslant\alpha_{n_0}(t),\\ g_{n_0}(t,y), & \alpha_{n_0}(t)\leqslant y\leqslant M,\\ g_{n_0}(t,M)+r(M-y), & y\geqslant M,\end{cases}$$

with $r:\mathbf{R}\to[-1,1]$ the radial retraction defined by

$$r(u)=\begin{cases} u, & |u|\leqslant 1,\\ \frac{u}{|u|}, & |u|>1,\end{cases}$$

and $M(\geqslant\sup_{[0,1]}\alpha_{n_0}(t))$ is a predetermined constant (see (2.117)). From Schauder's fixed point theorem we know that (2.113) has a solution $y_{n_0}\in C^1[0,1]\cap C^2(0,1)$ (notice from (2.112) that for any constants $r_1>0$, $r_2>r_1$, $\exists h_{r_1,r_2}\in L^1_q[0,1]$ with $|f(t,u)|\leqslant h_{r_1,r_2}(t)$ for $t\in(0,1)$ and $r_1\leqslant u\leqslant r_2$). Exactly the same analysis as in Theorem 2.9 guarantees that

$$y_{n_0}(t)\geqslant\alpha_{n_0}(t),\quad t\in[0,1], \tag{2.114}$$

and so

$$\alpha(t) \leqslant \alpha_{n_0}(t) \leqslant y_{n_0}(t) \quad \text{for } t \in [0,1]. \tag{2.115}$$

Next we show

$$y_{n_0}(t) \leqslant M \quad \text{for } t \in [0,1]. \tag{2.116}$$

Now (2.112) (with $\varepsilon = \min_{[0,1]} \alpha_{n_0}(t)$) and the definition of $g^\star_{n_0}$ (of course with (2.114)) implies that there exists $h_\varepsilon \in L^1_q[0,1]$ (with $h_\varepsilon \geqslant 0$ a.e. on $(0,1)$) with (note $r : \mathbf{R} \to [-1,1]$)

$$\left|g^\star_{n_0}\big(t, y_{n_0}(t)\big)\right| \leqslant a_0 y_{n_0}(t) + b_0\big[y_{n_0}(t)\big]^\gamma + h(t) + 1 \quad \text{for } t \in (0,1);$$

here $h(t) = \max\{h_\varepsilon(t), h_\varepsilon(\theta_{n_0}(t))\}$ (to see this fix $t \in (0,1)$ and check the cases $y_{n_0}(t) \geqslant M$ and $\alpha_{n_0}(t) \leqslant y_{n_0}(t) \leqslant M$ separately).

Now let $v = y_{n_0} - \rho_{n_0}$. Then v satisfies

$$\begin{cases} v'' + q(t) g^\star_{n_0}(t, v + \rho_{n_0}) = 0, & 0 < t < 1, \\ v(0) = v(1) = 0. \end{cases}$$

In addition since $-vv'' = qvg^\star_{n_0}(t, v + \rho_{n_0})$ we have

$$\begin{aligned} \big(\|v'\|_2\big)^2 &= \int_0^1 q(t)v(t)g^\star_{n_0}\big(t, v(t) + \rho_{n_0}\big)\,\mathrm{d}t \\ &\leqslant \int_0^1 q(t)\big|v(t)\big|\big|g^\star_{n_0}\big(t, v(t) + \rho_{n_0}\big)\big|\,\mathrm{d}t \\ &\leqslant \int_0^1 q(t)\big|v(t)\big|\big[a_0\big(v(t) + \rho_{n_0}\big) + b_0\big(v(t) + \rho_{n_0}\big)^\gamma + h(t) + 1\big]\,\mathrm{d}t \\ &\leqslant a_0 \int_0^1 q|v|^2\,\mathrm{d}t + a_0\rho_{n_0} \int_0^1 q|v|\,\mathrm{d}t + 2^{\gamma-1}b_0 \int_0^1 q|v|^{\gamma+1}\,\mathrm{d}t \\ &\quad + 2^{\gamma-1}b_0\rho^\gamma_{n_0} \int_0^1 q|v|\,\mathrm{d}t + \int_0^1 qh|v|\,\mathrm{d}t + \int_0^1 q|v|\,\mathrm{d}t. \end{aligned}$$

This together with Theorem 2.22 (and Hölder's inequality) yields

$$\begin{aligned} \big(\|v'\|_2\big)^2 &\leqslant \frac{a_0}{\lambda_1}\big(\|v'\|_2\big)^2 + \frac{a_0\rho_{n_0}}{\sqrt{\lambda_1}}\big(\|q\|_1\big)^{\frac{1}{2}}\|v'\|_2 \\ &\quad + \frac{2^{\gamma-1}b_0}{\lambda_1^{\frac{\gamma+1}{2}}}\big(\|q\|_1\big)^{\frac{1-\gamma}{2}}\big(\|v'\|_2\big)^{\gamma+1} + \frac{2^{\gamma-1}b_0\rho^\gamma_{n_0}}{\sqrt{\lambda_1}}\big(\|q\|_1\big)^{\frac{1}{2}}\|v'\|_2 \\ &\quad + \|qh\|_1|v|_\infty + \frac{1}{\sqrt{\lambda_1}}\big(\|q\|_1\big)^{\frac{1}{2}}\|v'\|_2. \end{aligned}$$

Now since $v(0) = v(1) = 0$ it is easy to check that $|v|_\infty \leqslant \frac{1}{\sqrt{2}}\|v'\|_2$ and so

$$\begin{aligned}\left(1 - \frac{a_0}{\lambda_1}\right)\left(\|v'\|_2\right)^2 \leqslant \frac{a_0\rho_{n_0}}{\sqrt{\lambda_1}}\left(\|q\|_1\right)^{\frac{1}{2}}\|v'\|_2 + \frac{2^{\gamma-1}b_0}{\lambda_1^{\frac{\gamma+1}{2}}}\left(\|q\|_1\right)^{\frac{1-\gamma}{2}}\left(\|v'\|_2\right)^{\gamma+1}\\ + \frac{2^{\gamma-1}b_0\rho_{n_0}^{\gamma}}{\sqrt{\lambda_1}}\left(\|q\|_1\right)^{\frac{1}{2}}\|v'\|_2 + \frac{\|qh\|_1}{\sqrt{2}}\|v'\|_2\\ + \frac{1}{\sqrt{\lambda_1}}\left(\|q\|_1\right)^{\frac{1}{2}}\|v'\|_2.\end{aligned}$$

As a result, since $0 \leqslant \gamma < 1$, there exists a constant K_0 (chosen greater than or equal to $\sqrt{2}\sup_{[0,1]}\alpha_{n_0}(t)$) with $\|v'\|_2 \leqslant K_0$. This together with $|v|_\infty \leqslant \frac{1}{\sqrt{2}}\|v'\|_2$ yields $|v|_\infty \leqslant \frac{1}{\sqrt{2}}K_0$, and as a result

$$|y_{n_0}|_\infty \leqslant \frac{1}{\sqrt{2}}K_0 + \rho_{n_0} \equiv M. \tag{2.117}$$

Consequently (2.116) holds and so we have

$$\alpha(t) \leqslant \alpha_{n_0}(t) \leqslant y_{n_0}(t) \leqslant M \quad \text{for } t \in [0,1].$$

Next we consider the boundary value problem

$$\begin{cases} y'' + q(t)g^\star_{n_0+1}(t,y) = 0, & 0 < t < 1,\\ y(0) = y(1) = \rho_{n_0+1}; \end{cases} \tag{2.118}$$

here

$$g^\star_{n_0+1}(t,y) = \begin{cases} g_{n_0+1}\big(t,\alpha_{n_0+1}(t)\big) + r\big(\alpha_{n_0+1}(t) - y\big), & y \leqslant \alpha_{n_0+1}(t),\\ g_{n_0+1}(t,y), & \alpha_{n_0+1}(t) \leqslant y \leqslant y_{n_0}(t),\\ g_{n_0+1}\big(t,y_{n_0}(t)\big) + r\big(y_{n_0}(t) - y\big), & y \geqslant y_{n_0}(t). \end{cases}$$

Schauder's fixed point theorem guarantees that (2.118) has a solution $y_{n_0+1} \in C^1[0,1] \cap C^2(0,1)$. Exactly the same analysis as in Theorem 2.9 guarantees

$$\alpha(t) \leqslant \alpha_{n_0+1}(t) \leqslant y_{n_0+1}(t) \leqslant y_{n_0}(t) \quad \text{for } t \in [0,1]. \tag{2.119}$$

Now proceed inductively to construct $y_{n_0+2}, y_{n_0+3}, \ldots$ as follows. Suppose we have y_k for some $k \in \{n_0+1, n_0+2, \ldots\}$ with $\alpha_k(t) \leqslant y_k(t) \leqslant y_{k-1}(t)$ for $t \in [0,1]$. Then consider the boundary value problem

$$\begin{cases} y'' + q(t)g^\star_{k+1}(t,y) = 0, & 0 < t < 1,\\ y(0) = y(1) = \rho_{k+1}; \end{cases} \tag{2.120}$$

here

$$g^{\star}_{k+1}(t,y)=\begin{cases} g_{k+1}\big(t,\alpha_{k+1}(t)\big)+r\big(\alpha_{k+1}(t)-y\big), & y\leqslant \alpha_{k+1}(t),\\ g_{k+1}(t,y), & \alpha_{k+1}(t)\leqslant y\leqslant y_k(t),\\ g_{k+1}\big(t,y_k(t)\big)+r\big(y_k(t)-y\big), & y\geqslant y_k(t). \end{cases}$$

Now Schauder's fixed point theorem guarantees that (2.120) has a solution $y_{k+1}\in C^1[0,1]\cap C^2(0,1)$, and essentially the same reasoning as in Theorem 2.9 yields

$$\alpha(t)\leqslant \alpha_{k+1}(t)\leqslant y_{k+1}(t)\leqslant y_k(t)\quad \text{for } t\in[0,1].$$

Thus for each $n\in\{n_0+1,\dots\}$ we have

$$\alpha(t)\leqslant y_n(t)\leqslant y_{n-1}(t)\leqslant\cdots\leqslant y_{n_0}(t)\leqslant M\quad \text{for } t\in[0,1].$$

Essentially the same reasoning as in Theorem 2.9 (from (2.89) onwards) completes the proof. □

REMARK 2.10. In (2.72) it is possible to replace $\frac{1}{2^{n+1}}\leqslant t\leqslant 1$ with $0\leqslant t\leqslant 1-\frac{1}{2^{n+1}}$ or $\frac{1}{2^{n+1}}\leqslant t\leqslant 1-\frac{1}{2^{n+1}}$ and the result in Theorem 2.23 is again true; the minor adjustments are left to the reader.

Next we discuss the lower solution α in (2.74). Suppose (2.96) holds. Then the argument before Theorem 2.14 guarantees that there exists a $\alpha\in C[0,1]\cap C^2(0,1)$, $\alpha(0)=\alpha(1)=0$, $\alpha(t)\leqslant\rho_{n_0}$ for $t\in[0,1]$ with (2.74) holding. Combine with Theorem 2.23 to obtain our next result.

THEOREM 2.24. *Let* $n_0\in\{1,2,\dots\}$ *be fixed and suppose* (2.96), (2.110), (2.111) *and* (2.112) *hold. Then* (2.62) *has a solution* $y\in C[0,1]\cap C^2(0,1)$ *with* $y(t)>0$ *for* $t\in(0,1)$.

If in (2.96) we replace $\frac{1}{2^{n+1}}\leqslant t\leqslant 1$ with $\frac{1}{2^{n+1}}\leqslant t\leqslant 1-\frac{1}{2^{n+1}}$ then once again (2.74) holds.

THEOREM 2.25. *Let* $n_0\in\{1,2,\dots\}$ *be fixed and suppose* (2.110), (2.111), (2.112) *and* (2.96) (*with* $\frac{1}{2^{n+1}}\leqslant t\leqslant 1$ *replaced by* $\frac{1}{2^{n+1}}\leqslant t\leqslant 1-\frac{1}{2^{n+1}}$) *hold. Then* (2.62) *has a solution* $y\in C[0,1]\cap C^2(0,1)$ *with* $y(t)>0$ *for* $t\in(0,1)$.

EXAMPLE 2.4. Consider the boundary value problem

$$\begin{cases} y''+\big(At^{\kappa}y^{-\theta}+a_0y+b_0y^{\gamma}-\mu^2\big)=0, & 0<t<1,\\ y(0)=y(1)=0 \end{cases}\tag{2.121}$$

with $A>0$, $\kappa>-1$, $\theta>0$, $0\leqslant\gamma<1$, $0\leqslant a_0<\pi^2$, $b_0\geqslant 0$ and $\mu\in\mathbf{R}$. Then (2.121) has a solution $y\in C[0,1]\cap C^2(0,1)$ with $y(t)>0$ for $t\in(0,1)$.

To see this we will apply Theorem 2.24 with

$$q=1 \quad \text{and} \quad f(t,y)=At^{\kappa}y^{-\theta}+a_0y+b_0y^{\gamma}-\mu^2.$$

Clearly (2.110) and (2.111) hold. In addition (2.112) is immediate with $h_\varepsilon(t)=At^{\kappa}\varepsilon^{-\theta}+\mu^2$; note $\lambda_1=\pi^2$. We will consider two cases, namely $\kappa\geqslant 0$ and $-1<\kappa<0$.

Case (i). $\kappa\geqslant 0$.

Let

$$n_0=1,\quad \rho_n=\left(\frac{A}{2^{(n+1)\kappa}(\mu^2+1)}\right)^{1/\theta} \quad \text{and} \quad k_0=1.$$

Notice for $n\in\{1,2,\ldots\}$, $\frac{1}{2^{n+1}}\leqslant t\leqslant 1$ and $0<y\leqslant\rho_n$ that we have

$$q(t)f(t,y)\geqslant \frac{A}{2^{(n+1)\kappa}\rho_n^{\theta}}-\mu^2=\left(\mu^2+1\right)-\mu^2=1,$$

so (2.96) holds.

Case (ii). $-1<\kappa<0$.

Let

$$n_0=1,\quad \rho_n=\left(\frac{A}{n(\mu^2+1)}\right)^{1/\theta} \quad \text{and} \quad k_0=1.$$

Notice for $n\in\{1,2,\ldots\}$, $\frac{1}{2^{n+1}}\leqslant t\leqslant 1$ and $0<y\leqslant\rho_n$ that we have since $\kappa<0$,

$$q(t)f(t,y)\geqslant \frac{At^{\kappa}}{\rho_n^{\theta}}-\mu^2\geqslant \frac{A}{\rho_n^{\theta}}-\mu^2=n\left(\mu^2+1\right)-\mu^2\geqslant\left(\mu^2+1\right)-\mu^2=1,$$

so (2.96) holds.

Existence of a solution is now guaranteed from Theorem 2.24.

If one uses the ideas in [19, Chapter 11] it is possible to improve the result in Theorem 2.23.

THEOREM 2.26. *Let* $n_0\in\{1,2,\ldots\}$ *be fixed and suppose* (2.72)*,* (2.74)*,* (2.110) *and* (2.111) *hold. In addition assume the following conditions are satisfied*:

$$\begin{cases} \textit{for any } \varepsilon>0, \exists a_0\in C[0,1] \textit{ with } 0\leqslant a_0(t)\leqslant\lambda_1 \\ \textit{on } [0,1] \textit{ and } a_0(t)<\lambda_1 \textit{ on a subset of } [0,1] \\ \textit{of positive measure with } a_0\left(\frac{1}{2^{n_0+1}}\right)<\lambda_1, b_0\geqslant 0, \\ 1\leqslant\gamma<2, h_\varepsilon\in L^1_q[0,1] \textit{ with } h_\varepsilon\geqslant 0 \textit{ a.e. on } (0,1) \\ \textit{with } uf(t,u)\leqslant a_0(t)u^2+b_0u^{\gamma}+uh_\varepsilon(t) \\ \textit{for } t\in(0,1) \textit{ and } u\geqslant\varepsilon \end{cases} \tag{2.122}$$

and

$$\begin{cases} \textit{for any } \varepsilon > 0, \exists c_0 \geqslant 0, 0 \leqslant \tau < 2, \eta_\varepsilon \in L^1_q[0,1] \textit{ with} \\ \eta_\varepsilon \geqslant 0 \textit{ a.e. on } (0,1) \textit{ with } \left|f(t,u)\right| \leqslant c_0 u^\tau + \eta_\varepsilon(t) \\ \textit{for } t \in (0,1) \textit{ and } u \geqslant \varepsilon; \end{cases} \tag{2.123}$$

here λ_1 *is the first eigenvalue of* (2.109). *Then* (2.62) *has a solution* $y \in C[0,1] \cap C^2(0,1)$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0,1]$.

PROOF. Proceed as in Theorem 2.23 and obtain a solution y_{n_0} of (2.113) with

$$\alpha(t) \leqslant \alpha_{n_0}(t) \leqslant y_{n_0}(t) \quad \text{for } t \in [0,1].$$

Next we show

$$y_{n_0}(t) \leqslant M \quad \text{for } t \in [0,1], \tag{2.124}$$

where M ($\geqslant \sup_{[0,1]} \alpha_{n_0}(t)$) is a predetermined constant (see below). Notice (2.123) (with $\varepsilon = \sup_{[0,1]} \alpha_{n_0}(t)$) implies that there exists $\eta_\varepsilon \in L^1_q[0,1]$ (with $\eta_\varepsilon \geqslant 0$ a.e. on $[0,1]$) with

$$\left|g^\star_{n_0}\big(t, y_{n_0}(t)\big)\right| \leqslant c_0\big[y_{n_0}(t)\big]^\tau + \eta(t) + 1 \quad \text{for } t \in (0,1);$$

here $\eta(t) = \max\{\eta_\varepsilon(t), \eta_\varepsilon(\theta_{n_0}(t))\}$. Also notice (2.123) implies that there exists $h_\varepsilon \in L^1_q[0,1]$ (with $h_\varepsilon \geqslant 0$ a.e. on $[0,1]$) with

$$y_{n_0}(t) g^\star_{n_0}\big(t, y_{n_0}(t)\big) \leqslant d_0(t)\big[y_{n_0}(t)\big]^2 + b_0\big[y_{n_0}(t)\big]^\gamma + \big[h(t)+1\big] y_{n_0}(t)$$

for $t \in (0,1)$ where

$$d_0(t) = \max\big\{a_0(t), a_0\big(\theta_{n_0}(t)\big)\big\} \text{ and } h(t) = \max\big\{h_\varepsilon(t), h_\varepsilon\big(\theta_{n_0}(t)\big)\big\};$$

note for fixed $t \in (\frac{1}{2^{n_0+1}}, 1)$ with $y_{n_0}(t) \geqslant M$ that

$$\begin{aligned} y_{n_0}(t) g^\star_{n_0}\big(t, y_{n_0}(t)\big) &= y_{n_0}(t)\big[f(t,M) + r\big(M - y_{n_0}(t)\big)\big] \\ &= \frac{y_{n_0}(t)}{M}\big[M f(t,M) + M r\big(M - y_{n_0}(t)\big)\big] \\ &\leqslant \frac{y_{n_0}(t)}{M}\big[a_0(t) M^2 + b_0 M^\gamma + h_\varepsilon(t) M + M\big] \\ &= y_{n_0}(t)\big[a_0(t) M + b_0 M^{\gamma-1} + h_\varepsilon(t) + 1\big] \\ &\leqslant a_0(t)\big[y_{n_0}(t)\big]^2 + b_0\big[y_{n_0}(t)\big]^\gamma + \big[h_\varepsilon(t)+1\big] y_{n_0}(t) \end{aligned}$$

(the other cases are treated similarly). Let $v = y_{n_0} - \rho_{n_0}$ so

$$-vv'' = q(v + \rho_{n_0})g^{\star}_{n_0}(t, v + \rho_{n_0}) - \rho_{n_0} q g^{\star}_{n_0}(t, v + \rho_{n_0}).$$

As a result we have

$$\begin{aligned}\int_0^1 \bigl[v'(t)\bigr]^2\,\mathrm{d}t \leqslant{}& \int_0^1 q(t)d_0(t)\bigl[v(t) + \rho_{n_0}\bigr]^2\,\mathrm{d}t + b_0 \int_0^1 q(t)\bigl[v(t) + \rho_{n_0}\bigr]^{\gamma}\,\mathrm{d}t\\ &+ \int_0^1 q(t)\bigl[h(t) + 1\bigr]\bigl[v(t) + \rho_{n_0}\bigr]\,\mathrm{d}t\\ &+ c_0\rho_{n_0} \int_0^1 q(t)\bigl[v(t) + \rho_{n_0}\bigr]^{\tau}\,\mathrm{d}t + \rho_{n_0} \int_0^1 q(t)\bigl[\eta(t) + 1\bigr]\,\mathrm{d}t,\end{aligned}$$

so

$$\begin{aligned}&\int_0^1 \bigl[\bigl(v'(t)\bigr)^2 - q(t)d_0(t)v^2(t)\bigr]\,\mathrm{d}t\\ &\quad\leqslant \int_0^1 q(t)d_0(t)\bigl[2v(t)\rho_{n_0} + \rho^2_{n_0}\bigr]\,\mathrm{d}t + b_0 \int_0^1 q(t)\bigl[v(t) + \rho_{n_0}\bigr]^{\gamma}\,\mathrm{d}t\\ &\quad\quad+ \int_0^1 q(t)\bigl[h(t) + 1\bigr]\bigl[v(t) + \rho_{n_0}\bigr]\,\mathrm{d}t + c_0\rho_{n_0} \int_0^1 q(t)\bigl[v(t) + \rho_{n_0}\bigr]^{\tau}\,\mathrm{d}t\\ &\quad\quad+ \rho_{n_0} \int_0^1 q(t)\bigl[\eta(t) + 1\bigr]\,\mathrm{d}t.\end{aligned}$$

Notice $d_0(t) < \lambda_1$ on a subset of $[0, 1]$ of positive measure. The argument in [19, Chapter 11] guarantees that there exists a $\delta > 0$ with

$$\int_0^1 \bigl[\bigl(v'(t)\bigr)^2 - q(t)d_0(t)v^2(t)\bigr]\,\mathrm{d}t \geqslant \delta\biggl(\int_0^1 q(t)v^2(t)\,\mathrm{d}t + \int_0^1 \bigl[v'(t)\bigr]^2\,\mathrm{d}t\biggr).$$

Consequently

$$\begin{aligned}&\delta\biggl(\int_0^1 q(t)v^2(t)\,\mathrm{d}t + \int_0^1 \bigl[v'(t)\bigr]^2\,\mathrm{d}t\biggr)\\ &\quad\leqslant \int_0^1 q(t)d_0(t)\bigl[2v(t)\rho_{n_0} + \rho^2_{n_0}\bigr]\,\mathrm{d}t + b_0 \int_0^1 q(t)\bigl[v(t) + \rho_{n_0}\bigr]^{\gamma}\,\mathrm{d}t\\ &\quad\quad+ \int_0^1 q(t)\bigl[h(t) + 1\bigr]\bigl[v(t) + \rho_{n_0}\bigr]\,\mathrm{d}t + c_0\rho_{n_0} \int_0^1 q(t)\bigl[v(t) + \rho_{n_0}\bigr]^{\tau}\,\mathrm{d}t\\ &\quad\quad+ \rho_{n_0} \int_0^1 q(t)\bigl[\eta(t) + 1\bigr]\,\mathrm{d}t,\end{aligned}$$

and this together with Hölder's inequality and Theorem 2.22 (note $1 \leqslant \gamma < 2$ and $0 \leqslant \tau < 2$) guarantees that there exists a constant $K_0 \geqslant \sqrt{2}\sup_{[0,1]} \alpha_{n_0}(t)$ with $\|v'\|_2 \leqslant K_0$. Now if $M = \frac{1}{\sqrt{2}}K_0 + \rho_{n_0}$ then (2.124) holds. Essentially the same reasoning as in Theorem 2.23 (from (2.118) onwards) completes the proof. □

REMARK 2.11. We can replace (2.72) and (2.74) in Theorem 2.26 with (2.96).

3. Singular initial value problems

In this section we discuss the singular initial value problem

$$\begin{cases} y' = q(t)f(t,y), & 0 < t < T(<\infty), \\ y(0) = 0, \end{cases} \tag{3.1}$$

where our nonlinearity f may change sign. We first present a variation of the classical theory of upper and lower solutions. We will assume the following conditions hold:

$$\begin{cases} \text{there exists } \beta \in C[0,T] \cap C^1(0,T] \text{ with} \\ \beta \in AC[0,T], \beta(0) \geqslant 0, \text{ and} \\ q(t)f\big(t,\beta(t)\big) \leqslant \beta'(t) \text{ for } t \in (0,T), \end{cases} \tag{3.2}$$

$$\begin{cases} \text{there exists } \alpha \in C[0,T] \cap C^1(0,T] \text{ with} \\ \alpha \in AC[0,T], \alpha(t) \leqslant \beta(t) \text{ on } [0,T], \alpha(0) \leqslant 0 \\ \text{and } q(t)f\big(t,\alpha(t)\big) \geqslant \alpha'(t) \text{ for } t \in (0,T) \end{cases} \tag{3.3}$$

and

$$q \in C(0,T] \cap L^1[0,T] \quad \text{with } q > 0 \text{ on } (0,T]. \tag{3.4}$$

Let

$$f^\star(t,y) = \begin{cases} f\big(t,\beta(t)\big) + r\big(\beta(t) - y\big), & y \geqslant \beta(t), \\ f(t,y), & \alpha(t) < y < \beta(t), \\ f\big(t,\alpha(t)\big) + r\big(\alpha(t) - y\big), & y \leqslant \alpha(t), \end{cases}$$

where $r : \mathbf{R} \to [-1,1]$ is the radial retraction. Finally we assume

$$f^\star : [0,T] \times \mathbf{R} \to \mathbf{R} \quad \text{is continuous.} \tag{3.5}$$

THEOREM 3.1. *Suppose* (3.2)–(3.5) *hold. Then* (3.1) *has a solution* y (*here* $y \in C[0,T] \cap C^1(0,T]$ *with* $y \in AC[0,T]$) *with* $\alpha(t) \leqslant y(t) \leqslant \beta(t)$ *for* $t \in [0,T]$.

PROOF. To show (3.1) has a solution we consider the problem

$$\begin{cases} y' = q(t)f^\star(t,y), & 0 < t < T, \\ y(0) = 0. \end{cases} \tag{3.6}$$

Now Theorem 1.4 guarantees that (3.6) has a solution $y \in C[0,T] \cap C^1(0,T]$ with $y \in AC[0,T]$. The result of our theorem will follow once we show

$$\alpha(t) \leqslant y(t) \leqslant \beta(t) \quad \text{for } t \in [0,T]. \tag{3.7}$$

We now show

$$y(t) \leqslant \beta(t) \quad \text{for } t \in [0,T]. \tag{3.8}$$

Suppose (3.8) is not true. Then since $y(0) \leqslant \beta(0)$ there exists $t_1 < t_2 \in [0,T]$ with

$$y(t_1) = \beta(t_1), \quad y(t_2) > \beta(t_2) \quad \text{and} \quad y(t) > \beta(t) \quad \text{on } (t_1, t_2).$$

Thus

$$\begin{aligned} y(t_2) - y(t_1) &= \int_{t_1}^{t_2} q(s)\big[f\big(s, \beta(s)\big) + r\big(\beta(s) - y(s)\big)\big]\,\mathrm{d}s \\ &\leqslant \int_{t_1}^{t_2} \beta'(s)\,\mathrm{d}s + \int_{t_1}^{t_2} q(s) r\big(\beta(s) - y(s)\big)\,\mathrm{d}s \\ &< \beta(t_2) - \beta(t_1), \end{aligned}$$

i.e., $y(t_2) < \beta(t_2)$, a contradiction. Thus (3.8) is true. A similar argument shows

$$\alpha(t) \leqslant y(t) \quad \text{for } t \in [0,T]. \tag{3.9}$$

Our result follows. ☐

Again because of the difficulties encountered with checking (3.5) it is of interest to provide an alternative approach and to present conditions that are easy to verify in applications.

Our main result can be stated immediately.

THEOREM 3.2. *Let* $n_0 \in \{1, 2, \ldots\}$ *be fixed and suppose the following conditions are satisfied*:

$$f : [0,T] \times (0,\infty) \to \mathbf{R} \quad \textit{is continuous} \tag{3.10}$$

$$q \in C(0,T], \quad q > 0 \quad \textit{on } (0,T] \quad \textit{and} \quad \int_0^T q(x)\,\mathrm{d}x < \infty \tag{3.11}$$

$$\left\{ \begin{array}{l} \textit{let } n \in \{n_0, n_0+1, \ldots\} \textit{ and associated with each } n \textit{ we} \\ \textit{have a constant } \rho_n \textit{ such that } \{\rho_n\} \textit{ is a nonincreasing} \\ \textit{sequence with } \lim_{n\to\infty} \rho_n = 0 \textit{ and such that for} \\ \frac{T}{2^{n+1}} \leqslant t \leqslant T \textit{ we have } q(t) f(t, \rho_n) \geqslant 0, \end{array} \right. \tag{3.12}$$

$$\begin{cases} \textit{there exists a function } \alpha \in C[0,T] \cap C^1(0,T] \\ \textit{with } \alpha(0) = 0, \alpha > 0 \textit{ on } (0,T] \textit{ such} \\ \textit{that } q(t) f\big(t, \alpha(t)\big) \geqslant \alpha'(t) \textit{ for } t \in (0,T) \end{cases} \tag{3.13}$$

and

$$\begin{cases} \textit{there exists a function } \beta \in C[0,T] \cap C^1(0,T] \textit{ with} \\ \beta(t) \geqslant \alpha(t) \textit{ and } \beta(t) \geqslant \rho_{n_0} \textit{ for } t \in [0,T] \textit{ with} \\ q(t) f\big(t, \beta(t)\big) \leqslant \beta'(t) \textit{ for } t \in (0,T) \textit{ and} \\ q(t) f\big(\frac{T}{2^{n_0+1}}, \beta(t)\big) \leqslant \beta'(t) \textit{ for } t \in \big(0, \frac{T}{2^{n_0+1}}\big). \end{cases} \tag{3.14}$$

Then (3.1) *has a solution* $y \in C[0,T] \cap C^1(0,T]$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0,T]$.

PROOF. For $n = n_0, n_0 + 1, \ldots$ let

$$e_n = \left[\frac{T}{2^{n+1}}, T\right] \quad \text{and} \quad \theta_n(t) = \max\left\{\frac{T}{2^{n+1}}, t\right\}, \quad 0 \leqslant t \leqslant T,$$

and

$$f_n(t,x) = \max\big\{f\big(\theta_n(t), x\big), f(t,x)\big\}.$$

Next we define inductively

$$g_{n_0}(t,x) = f_{n_0}(t,x)$$

and

$$g_n(t,x) = \min\big\{f_{n_0}(t,x), \ldots, f_n(t,x)\big\}, \quad n = n_0 + 1, n_0 + 2, \ldots.$$

Notice

$$f(t,x) \leqslant \cdots \leqslant g_{n+1}(t,x) \leqslant g_n(t,x) \leqslant \cdots \leqslant g_{n_0}(t,x)$$

for $(t,x) \in (0,T) \times (0,\infty)$ and

$$g_n(t,x) = f(t,x) \quad \text{for } (t,x) \in e_n \times (0,\infty).$$

Without loss of generality assume $\rho_{n_0} \leqslant \min_{t \in [\frac{T}{2}, T]} \alpha(t)$. Fix $n \in \{n_0, n_0 + 1, \ldots\}$. Let $t_n \in [0, \frac{T}{2}]$ be such that

$$\alpha(t_n) = \rho_n \quad \text{and} \quad \alpha(t) \leqslant \rho_n \quad \text{for } t \in [0, t_n].$$

Define

$$\alpha_n(t) = \begin{cases} \rho_n & \text{if } t \in [0, t_n], \\ \alpha(t) & \text{if } t \in (t_n, T]. \end{cases}$$

Notice $\alpha_{n+1}(t) \leqslant \alpha_n(t)$, $t \in [0,T]$, for each $n \in \{n_0, n_0+1, \ldots\}$ since $\{\rho_n\}$ is a nonincreasing sequence. We begin with the initial value problem

$$\begin{cases} y' = q(t) g_{n_0}^{\star}(t,y), & 0 < t < T, \\ y(0) = \rho_{n_0}; \end{cases} \tag{3.15}$$

here

$$g_{n_0}^{\star}(t,y) = \begin{cases} g_{n_0}\bigl(t, \alpha_{n_0}(t)\bigr), & y \leqslant \alpha_{n_0}(t), \\ g_{n_0}(t,y), & \alpha_{n_0}(t) \leqslant y \leqslant \beta(t), \\ g_{n_0}\bigl(t, \beta(t)\bigr), & y \geqslant \beta(t). \end{cases}$$

From Schauder's fixed point theorem we know that (3.15) has a solution $y_{n_0} \in C[0,T] \cap C^1(0,T]$. We first show

$$y_{n_0}(t) \geqslant \alpha_{n_0}(t), \quad t \in [0,T]. \tag{3.16}$$

Suppose (3.16) is not true. Then there exists $\tau_1 < \tau_2 \in [0,T]$ with

$$y_{n_0}(\tau_1) = \alpha_{n_0}(\tau_1), \qquad y_{n_0}(\tau_2) < \alpha_{n_0}(\tau_2)$$

and

$$y_{n_0}(t) < \alpha_{n_0}(t) \quad \text{for } t \in (\tau_1, \tau_2).$$

Of course

$$y_{n_0}(\tau_2) - \alpha_{n_0}(\tau_2) = \int_{\tau_1}^{\tau_2} (y_{n_0} - \alpha_{n_0})'(t)\, \mathrm{d}t. \tag{3.17}$$

We now claim

$$(y_{n_0} - \alpha_{n_0})'(t) \geqslant 0 \quad \text{for a.e. } t \in (\tau_1, \tau_2). \tag{3.18}$$

If (3.18) is true then (3.17) implies

$$y_{n_0}(\tau_2) - \alpha_{n_0}(\tau_2) \geqslant 0,$$

a contradiction. As a result if we show (3.18) is true then (3.16) will follow. To see (3.18) we will show

$$(y_{n_0} - \alpha_{n_0})'(t) \geqslant 0 \quad \text{for } t \in (\tau_1, \tau_2) \text{ provided } t \neq t_{n_0}.$$

Fix $t \in (\tau_1, \tau_2)$ and assume $t \neq t_{n_0}$. Then

$$\begin{aligned}(y_{n_0} - \alpha_{n_0})'(t) &= \bigl[q(t) g_{n_0}\bigl(t, \alpha_{n_0}(t)\bigr) - \alpha'_{n_0}(t)\bigr] \\ &= \begin{cases} q(t) g_{n_0}\bigl(t, \alpha(t)\bigr) - \alpha'(t) & \text{if } t \in (t_{n_0}, T), \\ q(t) g_{n_0}(t, \rho_{n_0}) & \text{if } t \in (0, t_{n_0}). \end{cases}\end{aligned}$$

Case (i). $t \geqslant \frac{T}{2^{n_0+1}}$.

Then since $g_{n_0}(t, x) = f(t, x)$ for $x \in (0, \infty)$ we have

$$\begin{aligned}(y_{n_0} - \alpha_{n_0})'(t) &= \begin{cases} q(t) f\bigl(t, \alpha(t)\bigr) - \alpha'(t) & \text{if } t \in (t_{n_0}, T), \\ q(t) f(t, \rho_{n_0}) & \text{if } t \in (0, t_{n_0}) \end{cases} \\ &\geqslant 0,\end{aligned}$$

from (3.12) and (3.13).

Case (ii). $t \in (0, \frac{T}{2^{n_0+1}})$.

Then since

$$g_{n_0}(t, x) = \max\left\{ f\left(\frac{T}{2^{n_0+1}}, x\right), f(t, x)\right\}$$

we have $g_{n_0}(t, x) \geqslant f(t, x)$ and $g_{n_0}(t, x) \geqslant f(\frac{T}{2^{n_0+1}}, x)$ for $x \in (0, \infty)$. Thus we have

$$\begin{aligned}(y_{n_0} - \alpha_{n_0})'(t) &\geqslant \begin{cases} q(t) f\bigl(t, \alpha(t)\bigr) - \alpha'(t) & \text{if } t \in (t_{n_0}, T), \\ q(t) f\bigl(\frac{T}{2^{n_0+1}}, \rho_{n_0}\bigr) & \text{if } t \in (0, t_{n_0}) \end{cases} \\ &\geqslant 0,\end{aligned}$$

from (3.12) and (3.13).

Consequently (3.18) (and so (3.16)) holds and now since $\alpha(t) \leqslant \alpha_{n_0}(t)$ for $t \in [0, T]$ we have

$$\alpha(t) \leqslant \alpha_{n_0}(t) \leqslant y_{n_0}(t) \quad \text{for } t \in [0, T]. \tag{3.19}$$

Next we show

$$y_{n_0}(t) \leqslant \beta(t) \quad \text{for } t \in [0, T]. \tag{3.20}$$

If (3.20) is not true then there exists $\tau_1 < \tau_2 \in [0, T]$ with

$$y_{n_0}(\tau_1) = \beta(\tau_1), \quad y_{n_0}(\tau_2) > \beta(\tau_2) \quad \text{and} \quad y_{n_0}(t) > \beta(t) \quad \text{for } t \in (\tau_1, \tau_2).$$

Notice also that

$$y_{n_0}(\tau_2) - y_{n_0}(\tau_1) = \int_{\tau_1}^{\tau_2} q(s) g_{n_0}\bigl(s, \beta(s)\bigr)\, \mathrm{d}s.$$

There are three cases to consider, namely (i) $\frac{T}{2^{n_0+1}} \leqslant \tau_1$; (ii) $\tau_1 < \tau_2 \leqslant \frac{T}{2^{n_0+1}}$; and (iii) $\tau_1 < \frac{T}{2^{n_0+1}} < \tau_2$.

Case (i). $\frac{T}{2^{n_0+1}} \leqslant \tau_1$.

Since $g_{n_0}(t,x) = f(t,x)$ for $(t,x) \in (\tau_1,\tau_2) \times (0,\infty)$ we have

$$y_{n_0}(\tau_2) - y_{n_0}(\tau_1) = \int_{\tau_1}^{\tau_2} q(s) f\bigl(s,\beta(s)\bigr)\,\mathrm{d}s \leqslant \int_{\tau_1}^{\tau_2} \beta'(s)\,\mathrm{d}s = \beta(\tau_2) - \beta(\tau_1),$$

a contradiction.

Case (ii). $\tau_1 < \tau_2 \leqslant \frac{T}{2^{n_0+1}}$.

Since

$$g_{n_0}(t,x) = \max\left\{ f\left(\frac{T}{2^{n_0+1}}, x\right), f(t,x)\right\}$$

for $(t,x) \in (\tau_1,\tau_2) \times (0,\infty)$ we have

$$\begin{aligned} y_{n_0}(\tau_2) - y_{n_0}(\tau_1) &= \int_{\tau_1}^{\tau_2} q(s) \max\left\{ f\left(\frac{T}{2^{n_0+1}}, \beta(s)\right), f\bigl(s,\beta(s)\bigr)\right\} \\ &\leqslant \int_{\tau_1}^{\tau_2} \beta'(s)\,\mathrm{d}s = \beta(\tau_2) - \beta(\tau_1), \end{aligned}$$

a contradiction.

Case (iii). $\tau_1 < \frac{T}{2^{n_0+1}} < \tau_2$.

Now

$$\begin{aligned} y_{n_0}\left(\frac{T}{2^{n_0+1}}\right) - y_{n_0}(\tau_1) &= \int_{\tau_1}^{\frac{T}{2^{n_0+1}}} q(s) \max\left\{ f\left(\frac{T}{2^{n_0+1}}, \beta(s)\right), f\bigl(s,\beta(s)\bigr)\right\} \\ &\leqslant \int_{\tau_1}^{\frac{T}{2^{n_0+1}}} \beta'(s)\,\mathrm{d}s = \beta\left(\frac{T}{2^{n_0+1}}\right) - \beta(\tau_1) \end{aligned}$$

and

$$y_{n_0}(\tau_2) - y_{n_0}\left(\frac{T}{2^{n_0+1}}\right) = \int_{\frac{T}{2^{n_0+1}}}^{\tau_2} q(s) f\bigl(s,\beta(s)\bigr)\,\mathrm{d}s \leqslant \beta(\tau_2) - \beta\left(\frac{T}{2^{n_0+1}}\right).$$

Combine to obtain

$$y_{n_0}(\tau_2) - y_{n_0}(\tau_1) \leqslant \beta(\tau_2) - \beta(\tau_1),$$

a contradiction.

Thus (3.20) holds, so we have

$$\alpha(t) \leqslant \alpha_{n_0}(t) \leqslant y_{n_0}(t) \leqslant \beta(t) \quad \text{for } t \in [0,T].$$

Next we consider

$$\begin{cases} y' = q(t)g^{\star}_{n_0+1}(t, y), & 0 < t < T, \\ y(0) = \rho_{n_0+1}; \end{cases} \tag{3.21}$$

here

$$g^{\star}_{n_0+1}(t, y) = \begin{cases} g_{n_0+1}\big(t, \alpha_{n_0+1}(t)\big), & y \leqslant \alpha_{n_0+1}(t), \\ g_{n_0+1}(t, y), & \alpha_{n_0+1}(t) \leqslant y \leqslant y_{n_0}(t), \\ g_{n_0+1}\big(t, y_{n_0}(t)\big), & y \geqslant y_{n_0}(t). \end{cases}$$

Now Schauder's fixed point theorem guarantees that (3.21) has a solution $y_{n_0+1} \in C[0, T] \cap C^1(0, T]$. We first show

$$y_{n_0+1}(t) \geqslant \alpha_{n_0+1}(t), \quad t \in [0, T]. \tag{3.22}$$

Suppose (3.22) is not true. Then there exists $\tau_1 < \tau_2 \in [0, T]$ with

$$y_{n_0+1}(\tau_1) = \alpha_{n_0+1}(\tau_1), \qquad y_{n_0+1}(\tau_2) < \alpha_{n_0+1}(\tau_2)$$

and

$$y_{n_0+1}(t) < \alpha_{n_0+1}(t) \quad \text{for } t \in (\tau_1, \tau_2).$$

If we show

$$(y_{n_0+1} - \alpha_{n_0+1})'(t) \geqslant 0 \quad \text{for a.e. } t \in (\tau_1, \tau_2), \tag{3.23}$$

then as before (3.22) is true. Fix $t \in (\tau_1, \tau_2)$ and assume $t \neq t_{n_0+1}$. Then

$$\begin{aligned} (y_{n_0+1} - \alpha_{n_0+1})'(t) &= \big[q(t)g_{n_0+1}\big(t, \alpha_{n_0+1}(t)\big) - \alpha'_{n_0+1}(t)\big] \\ &= \begin{cases} q(t)g_{n_0+1}\big(t, \alpha(t)\big) - \alpha'(t) & \text{if } t \in (t_{n_0+1}, T), \\ q(t)g_{n_0+1}(t, \rho_{n_0+1}) & \text{if } t \in (0, t_{n_0+1}). \end{cases} \end{aligned}$$

Case (i). $t \geqslant \frac{T}{2^{n_0+2}}$.

Then since $g_{n_0+1}(t, x) = f(t, x)$ for $x \in (0, \infty)$ we have

$$\begin{aligned} (y_{n_0+1} - \alpha_{n_0+1})'(t) &= \begin{cases} q(t)f\big(t, \alpha(t)\big) - \alpha'(t) & \text{if } t \in (t_{n_0+1}, T), \\ q(t)f(t, \rho_{n_0+1}) & \text{if } t \in (0, t_{n_0+1}) \end{cases} \\ &\geqslant 0, \end{aligned}$$

from (3.12) and (3.13).

Case (ii). $t \in (0, \frac{T}{2^{n_0+2}})$.

Then since $g_{n_0+1}(t,x)$ equals

$$\min\left\{\max\left\{f\left(\frac{T}{2^{n_0+1}}, x\right), f(t,x)\right\}, \max\left\{f\left(\frac{T}{2^{n_0+2}}, x\right), f(t,x)\right\}\right\}$$

we have

$$g_{n_0+1}(t,x) \geqslant f(t,x)$$

and

$$g_{n_0+1}(t,x) \geqslant \min\left\{f\left(\frac{T}{2^{n_0+1}}, x\right), f\left(\frac{T}{2^{n_0+2}}, x\right)\right\}$$

for $x \in (0,\infty)$. Thus we have

$$\begin{aligned}(y_{n_0+1} - \alpha_{n_0+1})'(t) &\leqslant \begin{cases} q(t) f\big(t, \alpha(t)\big) - \alpha'(t) & \text{if } t \in (t_{n_0+1}, T), \\ q(t) \min\big\{f\big(\frac{T}{2^{n_0+1}}, \rho_{n_0+1}\big), f\big(\frac{T}{2^{n_0+2}}, \rho_{n_0+1}\big)\big\} \\ \quad \text{if } t \in (0, t_{n_0+1}) \end{cases} \\ &\geqslant 0,\end{aligned}$$

from (3.12) and (3.13) (note $f(\frac{T}{2^{n_0+1}}, \rho_{n_0+1}) \geqslant 0$ since $f(t, \rho_{n_0+1}) \geqslant 0$ for $t \in [\frac{T}{2^{n_0+2}}, T]$ and $\frac{T}{2^{n_0+1}} \in (\frac{T}{2^{n_0+2}}, T)$).

Consequently (3.23) is true so

$$\alpha(t) \leqslant \alpha_{n_0+1}(t) \leqslant y_{n_0+1}(t) \quad \text{for } t \in [0,T]. \tag{3.24}$$

Next we show

$$y_{n_0+1}(t) \leqslant y_{n_0}(t) \quad \text{for } t \in [0,T]. \tag{3.25}$$

If (3.25) is not true then there exists $\tau_1 < \tau_2 \in [0,T]$ with

$$y_{n_0+1}(\tau_1) = y_{n_0}(\tau_1), \qquad y_{n_0+1}(\tau_2) > y_{n_0}(\tau_2)$$

and

$$y_{n_0+1}(t) > y_{n_0}(t) \quad \text{for } t \in (\tau_1, \tau_2).$$

Notice also since $g_{n_0}(t,x) \geqslant g_{n_0+1}(t,x)$ for $(t,x) \in (0,T) \times (0,\infty)$ that

$$y_{n_0+1}(\tau_2) - y_{n_0+1}(\tau_1) = \int_{\tau_1}^{\tau_2} q(s) g_{n_0+1}\big(s, y_{n_0}(s)\big)\, \mathrm{d}s$$

$$\begin{aligned}&\leqslant \int_{\tau_1}^{\tau_2} q(s) g_{n_0}\big(s, y_{n_0}(s)\big)\,\mathrm{d}s\\&= \int_{\tau_1}^{\tau_2} y'_{n_0}(s)\,\mathrm{d}s = y_{n_0}(\tau_2) - y_{n_0}(\tau_1),\end{aligned}$$

a contradiction.

Now proceed inductively to construct $y_{n_0+2}, y_{n_0+3}, \ldots$ as follows. Suppose we have y_k for some $k \in \{n_0+1, n_0+2, \ldots\}$ with $\alpha_k(t) \leqslant y_k(t) \leqslant y_{k-1}(t)$ for $t \in [0, T]$. Then consider the boundary value problem

$$\begin{cases} y' = q(t) g^\star_{k+1}(t, y), & 0 < t < T,\\ y(0) = \rho_{k+1}; \end{cases} \tag{3.26}$$

here

$$g^\star_{k+1}(t, y) = \begin{cases} g_{k+1}\big(t, \alpha_{k+1}(t)\big), & y \leqslant \alpha_{k+1}(t),\\ g_{k+1}(t, y), & \alpha_{k+1}(t) \leqslant y \leqslant y_k(t),\\ g_{k+1}\big(t, y_k(t)\big), & y \geqslant y_k(t). \end{cases}$$

Now Schauder's fixed point theorem guarantees that (3.26) has a solution $y_{k+1} \in C[0, T] \cap C^1(0, T]$, and essentially the same reasoning as above yields

$$\alpha(t) \leqslant \alpha_{k+1}(t) \leqslant y_{k+1}(t) \leqslant y_k(t) \quad \text{for } t \in [0, T]. \tag{3.27}$$

Thus for each $n \in \{n_0, n_0+1, \ldots\}$ we have

$$\alpha(t) \leqslant y_n(t) \leqslant y_{n-1}(t) \leqslant \cdots \leqslant y_{n_0}(t) \leqslant \beta(t) \quad \text{for } t \in [0, T]. \tag{3.28}$$

Lets look at the interval $[\frac{T}{2^{n_0+1}}, T]$. Let

$$R_{n_0} = \sup\left\{\big|q(x) f(x, y)\big|\colon\, x \in \left[\frac{T}{2^{n_0+1}}, T\right] \text{ and } \alpha(x) \leqslant y \leqslant y_{n_0}(x)\right\}.$$

We have immediately that

$$\begin{cases} \{y_n\}_{n=n_0+1}^{\infty} \text{ is a bounded, equicontinuous}\\ \text{family on } \big[\frac{T}{2^{n_0+1}}, T\big]. \end{cases} \tag{3.29}$$

The Arzela–Ascoli theorem guarantees the existence of a subsequence N_{n_0} of integers and a function $z_{n_0} \in C[\frac{T}{2^{n_0+1}}, T]$ with y_n converging uniformly to z_{n_0} on $[\frac{T}{2^{n_0+1}}, T]$ as $n \to \infty$ through N_{n_0}. Proceed inductively to obtain subsequences of integers

$$N_{n_0} \supseteq N_{n_0+1} \supseteq \cdots \supseteq N_k \supseteq \cdots$$

and functions

$$z_k \in C\left[\frac{T}{2^{k+1}}, T\right]$$

with

$$y_n \quad \text{converging uniformly to } z_k \text{ on } \left[\frac{T}{2^{k+1}}, T\right]$$

as $n \to \infty$ through N_k, and

$$z_k = z_{k-1} \quad \text{on } \left[\frac{T}{2^k}, T\right].$$

Define a function $y:[0,T] \to [0,\infty)$ by $y(x) = z_k(x)$ on $[\frac{T}{2^{k+1}}, T]$ and $y(0) = 0$. Notice y is well defined and $\alpha(t) \leqslant y(t) \leqslant y_{n_0}(t)(\leqslant \beta(t))$ for $t \in (0,T)$. Fix $t \in (0,T)$ and let $m \in \{n_0, n_0+1, \ldots\}$ be such that $\frac{T}{2^{m+1}} < t < T$. Let $N_m^\star = \{n \in N_m\colon n \geqslant m\}$. Now y_n, $n \in N_m^\star$, satisfies

$$\begin{aligned} y_n(t) &= y_n(T) - \int_t^T q(s) g_n^\star\big(s, y_n(s)\big)\,\mathrm{d}s \\ &= y_n(T) - \int_t^T q(s) f\big(s, y_n(s)\big)\,\mathrm{d}s. \end{aligned}$$

Let $n \to \infty$ through $N_m^\star$ to obtain

$$y(t) = y(T) - \int_t^T q(s) f\big(s, y(s)\big)\,\mathrm{d}s.$$

We can do this argument for each $t \in (0,T)$. It remains to show y is continuous at 0.

Let $\varepsilon > 0$ be given. Now since $\lim_{n\to\infty} y_n(0) = 0$ there exists $n_1 \in \{n_0, n_0+1, \ldots\}$ with $y_{n_1}(0) < \frac{\varepsilon}{2}$. Since $y_{n_1} \in C[0,T]$ there exists $\delta_{n_1} > 0$ with

$$y_{n_1}(t) < \frac{\varepsilon}{2} \quad \text{for } t \in [0, \delta_{n_1}].$$

Now for $n \geqslant n_1$ we have, since $\{y_n(t)\}$ is nonincreasing for each $t \in [0,T]$,

$$\alpha(t) \leqslant y_n(t) \leqslant y_{n_1}(t) < \frac{\varepsilon}{2} \quad \text{for } t \in [0, \delta_{n_1}].$$

Consequently

$$\alpha(t) \leqslant y(t) \leqslant \frac{\varepsilon}{2} < \varepsilon \quad \text{for } t \in (0, \delta_{n_1}],$$

and so y is continuous at 0. Thus $y \in C[0,T]$. □

Suppose (3.10)–(3.13) hold, and in addition assume the following conditions are satisfied:

$$q(t)f(t,y) \geqslant \alpha'(t) \quad \text{for } (t,y) \in (0,T) \times \bigl\{y \in (0,\infty)\colon\ y < \alpha(t)\bigr\} \tag{3.30}$$

and

$$\begin{cases} \text{there exists a function } \beta \in C[0,T] \cap C^1(0,T] \text{ with} \\ \beta(t) \geqslant \rho_{n_0} \text{ for } t \in [0,T] \text{ with } q(t)f\bigl(t,\beta(t)\bigr) \leqslant \beta'(t) \\ \text{for } t \in (0,T) \text{ and } q(t)f\bigl(\frac{T}{2^{n_0+1}},\beta(t)\bigr) \leqslant \beta'(t) \\ \text{for } t \in \bigl(0,\frac{T}{2^{n_0+1}}\bigr). \end{cases} \tag{3.31}$$

Then the result in Theorem 3.2 is again true. This follows immediately from Theorem 3.2 once we show (3.14) holds, i.e., once we show $\beta(t) \geqslant \alpha(t)$ for $t \in [0,T]$. Suppose it is false. Then there exists $\tau_1 < \tau_2 \in [0,T]$ with

$$\beta(\tau_1) = \alpha(\tau_1), \quad \beta(\tau_2) < \alpha(\tau_2) \quad \text{and} \quad \beta(t) < \alpha(t) \quad \text{for } t \in (\tau_1,\tau_2).$$

Now for $t \in (\tau_1,\tau_2)$, we have from (3.30) that

$$q(t)f\bigl(t,\beta(t)\bigr) \geqslant \alpha'(t),$$

and as a result

$$\begin{aligned} \beta(\tau_2) - \beta(\tau_1) &= \int_{\tau_1}^{\tau_2} \beta'(s)\,\mathrm{d}s \geqslant \int_{\tau_1}^{\tau_2} q(s)f\bigl(s,\beta(s)\bigr)\,\mathrm{d}s \\ &\geqslant \int_{\tau_1}^{\tau_2} \alpha'(s)\,\mathrm{d}s = \alpha(\tau_2) - \alpha(\tau_1), \end{aligned}$$

a contradiction. Thus we have

COROLLARY 3.3. *Let* $n_0 \in \{1,2,\ldots\}$ *be fixed and suppose* (3.10)–(3.13), (3.30) *and* (3.31) *hold. Then* (3.1) *has a solution* $y \in C[0,T] \cap C^1(0,T]$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0,T]$.

Next we discuss how to construct the lower solution α in (3.13) and in (3.30). Suppose the following condition is satisfied:

$$\begin{cases} \text{let } n \in \{n_0, n_0+1,\ldots\} \text{ and associated with each } n \text{ we} \\ \text{have a constant } \rho_n \text{ such that } \{\rho_n\} \text{ is a decreasing} \\ \text{sequence with } \lim_{n\to\infty} \rho_n = 0 \text{ and there exists a} \\ \text{constant } k_0 > 0 \text{ such that for } \frac{T}{2^{n+1}} \leqslant t \leqslant T \\ \text{and } 0 < y \leqslant \rho_n \text{ we have } q(t)f(t,y) \geqslant k_0. \end{cases} \tag{3.32}$$

Then an argument similar to the one before Theorem 2.14 guarantees that there exists a $\alpha \in C[0,T] \cap C^1(0,T]$, $\alpha(0) = 0$, $\alpha > 0$ for $t \in (0,T]$, $\alpha(t) \leqslant \rho_{n_0}$ for $t \in [0,T]$ with (3.13) and (3.30) holding. We combine this with Corollary 3.3 to obtain our next result.

THEOREM 3.4. *Let* $n_0 \in \{1,2,\dots\}$ *be fixed and suppose* (3.10), (3.11), (3.31), *and* (3.32) *hold. Then* (3.1) *has a solution* $y \in C[0,T] \cap C^1(0,T]$ *with* $y(t) > 0$ *for* $t \in (0,T]$.

Looking at Theorem 3.4 we see that the main difficulty when discussing examples is the construction of the β in (3.31). Our next result replaces (3.31) with a growth condition. We first present the result in its full generality.

THEOREM 3.5. *Let* $n_0 \in \{1,2,\dots\}$ *be fixed and suppose* (3.10)–(3.13) *hold. Also assume the following condition is satisfied*:

$$\begin{cases} |f(t,y)| \leqslant g(y) + h(y) \text{ on } [0,T] \times (0,\infty) \text{ with} \\ g > 0 \text{ continuous and nonincreasing on } (0,\infty) \\ \text{and } h \geqslant 0 \text{ continuous on } [0,\infty). \end{cases} \tag{3.33}$$

Also suppose there exists a constant $M > 0$ *with* $G^{-1}(M) > \sup_{t \in [0,T]} \alpha(t)$ *and with*

$$\int_0^T q(x)\,\mathrm{d}x < \int_0^M \frac{\mathrm{d}s}{\left[1 + \frac{h(G^{-1}(s))}{g(G^{-1}(s))}\right]} \tag{3.34}$$

holding; *here* $G(z) = \int_0^z \frac{\mathrm{d}u}{g(u)}$ (*note* G *is an increasing map from* $[0,\infty)$ *onto* $[0,\infty)$ *with* $G(0) = 0$). *Then* (3.1) *has a solution* $y \in C[0,T] \cap C^1(0,T]$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0,T]$.

PROOF. Choose $\varepsilon > 0$, $\varepsilon < M$ with

$$\int_0^T q(x)\,\mathrm{d}x < \int_\varepsilon^M \frac{\mathrm{d}s}{\left[1 + \frac{h(G^{-1}(s))}{g(G^{-1}(s))}\right]}. \tag{3.35}$$

Without loss of generality assume $G(\rho_{n_0}) < \varepsilon$. Let e_n, θ_n, f_n, g_n and α_n be as in Theorem 3.2. We consider the boundary value problem (3.15) with in this case $g^\star_{n_0}$ given by

$$g^\star_{n_0}(t,y) = \begin{cases} g_{n_0}\big(t, \alpha_{n_0}(t)\big), & y \leqslant \alpha_{n_0}(t), \\ g_{n_0}(t,y), & \alpha_{n_0}(t) \leqslant y \leqslant G^{-1}(M), \\ g_{n_0}\big(t, G^{-1}(M)\big), & y \geqslant G^{-1}(M). \end{cases}$$

Essentially the same reasoning as in Theorem 3.2 implies that (3.15) has a solution y_{n_0} with $y_{n_0}(t) \geqslant \alpha_{n_0}(t) \geqslant \alpha(t)$ for $t \in [0,T]$. Next we show

$$y_{n_0}(t) < G^{-1}(M) \quad \text{for } t \in [0,T]. \tag{3.36}$$

Suppose (3.36) is false. Then since $y_{n_0}(0) = \rho_{n_0}$ there exists $\tau_1 < \tau_2 \in [0, T]$ with

$$\rho_{n_0} \leqslant y_{n_0}(t) \leqslant G^{-1}(M) \quad \text{for } t \in (\tau_1, \tau_2),$$

with

$$y_{n_0}(\tau_1) = \rho_{n_0} \quad \text{and} \quad y_{n_0}(\tau_2) = G^{-1}(M).$$

Now for $t \in (\tau_1, \tau_2)$ we have from (3.33) that

$$g^\star_{n_0}\big(t, y_{n_0}(t)\big) \leqslant g\big(y_{n_0}(t)\big) + h\big(y_{n_0}(t)\big) = g\big(y_{n_0}(t)\big)\left\{1 + \frac{h(y_{n_0}(t))}{g(y_{n_0}(t))}\right\}.$$

Thus

$$\frac{y'_{n_0}(t)}{g(y_{n_0}(t))} \leqslant q(t)\left\{1 + \frac{h(y_{n_0}(t))}{g(y_{n_0}(t))}\right\} \quad \text{for } t \in (\tau_1, \tau_2).$$

Let

$$v_{n_0}(t) = \int_0^{y_{n_0}(t)} \frac{\mathrm{d}u}{g(u)} = G\big(y_{n_0}(t)\big)$$

and so

$$v'_{n_0}(t) \leqslant q(t)\left\{1 + \frac{h(G^{-1}(v_{n_0}(t)))}{g(G^{-1}(v_{n_0}(t)))}\right\} \quad \text{for } t \in (\tau_1, \tau_2).$$

Integrate from τ_1 to τ_2 to obtain

$$\begin{aligned}\int_\varepsilon^{v_{n_0}(\tau_2)} \frac{\mathrm{d}s}{\big[1 + \frac{h(G^{-1}(s))}{g(G^{-1}(s))}\big]} &\leqslant \int_{G(\rho_{n_0})}^{v_{n_0}(\tau_2)} \frac{\mathrm{d}s}{\big[1 + \frac{h(G^{-1}(s))}{g(G^{-1}(s))}\big]} \leqslant \int_0^T q(s)\,\mathrm{d}s \\ &< \int_\varepsilon^M \frac{\mathrm{d}s}{\big[1 + \frac{h(G^{-1}(s))}{g(G^{-1}(s))}\big]}.\end{aligned}$$

Consequently $v_{n_0}(\tau_2) < M$ so $y_{n_0}(\tau_2) < G^{-1}(M)$. This is a contradiction. Thus (3.36) holds and so

$$\alpha(t) \leqslant \alpha_{n_0}(t) \leqslant y_{n_0}(t) < G^{-1}(M) \quad \text{for } t \in [0, T]. \tag{3.37}$$

Essentially the same reasoning as in Theorem 3.2 (from (3.21) onwards) completes the proof. □

COROLLARY 3.6. *Let* $n_0 \in \{1, 2, \ldots\}$ *be fixed and suppose* (3.10)–(3.13), (3.30) *and* (3.33) *hold. In addition assume there is a constant* $M > 0$ *with*

$$\int_0^T q(x)\,\mathrm{d}x < \int_0^M \frac{\mathrm{d}s}{\left[1 + \frac{h(G^{-1}(s))}{g(G^{-1}(s))}\right]} \tag{3.38}$$

holding; *here* $G(z) = \int_0^z \frac{\mathrm{d}u}{g(u)}$. *Then* (3.1) *has a solution* $y \in C[0, T] \cap C^1(0, T]$ *with* $y(t) \geqslant \alpha(t)$ *for* $t \in [0, T]$.

PROOF. This follows immediately from Theorem 3.5 once we show

$$G^{-1}(M) > \alpha(t) \quad \text{for each } t \in [0, T].$$

Suppose this is false. Then since $\alpha(0) = 0$ there exists $\tau_1 < \tau_2 \in [0, T]$ with

$$0 \leqslant \alpha(t) \leqslant G^{-1}(M) \quad \text{for } t \in (\tau_1, \tau_2), \alpha(\tau_1) = 0 \text{ and } \alpha(\tau_2) = G^{-1}(M).$$

Notice (3.30) implies

$$\alpha'(t) \leqslant q(t) f\big(t, \alpha(t)\big) \quad \text{for } t \in (\tau_1, \tau_2),$$

so we have

$$\frac{\alpha'(t)}{g(\alpha(t))} \leqslant q(t)\left\{1 + \frac{h(\alpha(t))}{g(\alpha(t))}\right\} \quad \text{for } t \in (\tau_1, \tau_2).$$

Let

$$v(t) = \int_0^{\alpha(t)} \frac{\mathrm{d}u}{g(u)} = G\big(\alpha(t)\big),$$

so

$$v'(t) \leqslant q(t)\left\{1 + \frac{h(G^{-1}(v(t)))}{g(G^{-1}(v(t)))}\right\} \quad \text{for } t \in (\tau_1, \tau_2).$$

Integrate from τ_1 to τ_2 to obtain

$$\int_0^{v(\tau_2)} \frac{\mathrm{d}s}{\left[1 + \frac{h(G^{-1}(s))}{g(G^{-1}(s))}\right]} \leqslant \int_0^T q(s)\,\mathrm{d}s < \int_0^M \frac{\mathrm{d}s}{\left[1 + \frac{h(G^{-1}(s))}{g(G^{-1}(s))}\right]}.$$

Thus $v(\tau_2) < M$, so $\alpha(\tau_2) < G^{-1}(M)$, a contradiction. □

Combining Corollary 3.6 with the comments before Theorem 3.4 yields the following theorem.

THEOREM 3.7. *Let* $n_0 \in \{1, 2, \dots\}$ *be fixed and suppose* (3.10), (3.11), (3.32) *and* (3.33) *hold. In addition assume there is a constant* $M > 0$ *with* (3.38) *holding. Then* (3.1) *has a solution* $y \in C[0,T] \cap C^1(0,T]$ *with* $y(t) > 0$ *for* $t \in (0,T]$.

Next we present some examples which illustrate how easily the theory is applied in practice.

EXAMPLE 3.1. The initial value problem

$$\begin{cases} y' = t^\theta y^{-\alpha} + y^\beta + A, & 0 < t < T(<\infty), \\ y(0) = 0 \end{cases} \tag{3.39}$$

with $\theta > -1$, $\alpha > 0$, $\beta > 0$ and $A > 0$ has a solution $y \in C[0,T] \cap C^1(0,T]$ with $y(t) > 0$ for $t \in (0,T]$ if

$$\int_0^T q(s)\,\mathrm{d}s < \int_0^\infty \frac{\mathrm{d}s}{1 + B[(\alpha+1)s]^{\frac{\beta+\alpha}{\alpha+1}} + AC[(\alpha+1)s]^{\frac{\alpha}{\alpha+1}}}; \tag{3.40}$$

here

$$q(t) = \begin{cases} 1 & \text{if } \theta \geqslant 0, \\ t^\theta & \text{if } -1 < \theta < 0, \end{cases}$$

with

$$B = \begin{cases} T^{\frac{\theta(\beta-1)}{\alpha+1}} & \text{if } \theta \geqslant 0, \\ T^{-\theta} & \text{if } -1 < \theta < 0, \end{cases}$$

and

$$C = \begin{cases} T^{-\frac{\theta}{\alpha+1}} & \text{if } \theta \geqslant 0, \\ T^{-\theta} & \text{if } -1 < \theta < 0. \end{cases}$$

To see this we will apply Theorem 3.7. We will consider two cases, namely $\theta \geqslant 0$ and $-1 < \theta < 0$.

Case (i). $\theta \geqslant 0$.

We will apply Theorem 3.7 with

$$n_0 = 1, \qquad q = 1, \qquad g(y) = T^\theta y^{-\alpha}, \qquad h(y) = y^\beta + A,$$

together with

$$\rho_n = \left(\frac{T^\theta}{2^{(n+1)\theta}}\right)^{1/\alpha} \quad \text{and} \quad k_0 = 1.$$

Clearly (3.10) and (3.11) hold. Also for $n \in \{1, 2, \ldots\}$, $\frac{T}{2^{n+1}} \leqslant t \leqslant T$ and $0 < y \leqslant \rho_n$ we have

$$q(t) f(t, y) \geqslant t^{\theta} y^{-\alpha} \geqslant \left(\frac{T}{2^{n+1}}\right)^{\theta} \frac{1}{\rho_n^{\alpha}} = 1,$$

so (3.32) is satisfied. From (3.40) there exists $M > 0$ with

$$T < \int_0^M \frac{\mathrm{d}s}{1 + B[(\alpha+1)s]^{\frac{\beta+\alpha}{\alpha+1}} + AC[(\alpha+1)s]^{\frac{\alpha}{\alpha+1}}},$$

so now (3.38) holds with this M since

$$G(z) = \frac{1}{T^{\theta}} \frac{z^{\alpha+1}}{\alpha+1}, \quad \text{so} \quad G^{-1}(z) = \bigl[(\alpha+1)z\bigr]^{\frac{1}{\alpha+1}} T^{\frac{\theta}{\alpha+1}}.$$

Existence of a solution to (3.39) is now guaranteed from Theorem 3.7.

Case (ii). $-1 < \theta < 0$.

We will apply Theorem 3.7 with

$$n_0 = 1, \qquad q = t^{\theta}, \qquad g(y) = y^{-\alpha}, \qquad h(y) = T^{-\theta}\bigl[y^{\beta} + A\bigr],$$

together with

$$\rho_n = \left(\frac{T^{\theta}}{n}\right)^{1/\alpha} \quad \text{and} \quad k_0 = 1.$$

Clearly (3.10), (3.11) and (3.38) (as in Case (i)) hold. Also for $n \in \{1, 2, \ldots\}$, $\frac{T}{2^{n+1}} \leqslant t \leqslant T$ and $0 < y \leqslant \rho_n$ we have

$$q(t) f(t, y) \geqslant t^{\theta} y^{-\alpha} \geqslant \frac{T^{\theta}}{\rho_n^{\alpha}} = n \geqslant 1,$$

so (3.32) is satisfied. Existence of a solution to (3.39) is now guaranteed from Theorem 3.7.

EXAMPLE 3.2. The initial value problem

$$\begin{cases} y' = t^{\theta} y^{-\alpha} + y^{\beta} - A, & 0 < t < T(< \infty), \\ y(0) = 0 \end{cases}$$

with $\theta > -1$, $\alpha > 0$, $\beta > 0$ and $A > 0$ has a solution $y \in C[0, T] \cap C^1(0, T]$ with $y(t) > 0$ for $t \in (0, T]$ if (3.40) holds.

The proof is essentially the same as in Example 3.1 with

$$\rho_n = \left(\frac{T^{\theta}}{2^{(n+1)\theta}(A+1)}\right)^{1/\alpha} \quad \text{if } \theta \geqslant 0$$

and

$$\rho_n = \left(\frac{T^\theta}{n(A+1)}\right)^{1/\alpha} \quad \text{if } -1 < \theta < 0.$$

EXAMPLE 3.3. The initial value problem

$$\begin{cases} y' = t^\theta\big(y^{-\alpha} + y^\beta + A\big), & 0 < t < T(<\infty), \\ y(0) = 0 \end{cases}$$

with $\theta > -1, \alpha > 0, \beta > 0$ and $A > 0$ has a solution $y \in C[0,T] \cap C^1(0,T]$ with $y(t) > 0$ for $t \in (0,T]$ if

$$\frac{T^{\theta+1}}{\theta+1} < \int_0^\infty \frac{\mathrm{d}s}{1 + [(\alpha+1)s]^{\frac{\beta+\alpha}{\alpha+1}} + A[(\alpha+1)s]^{\frac{\alpha}{\alpha+1}}}.$$

Apply Theorem 3.7 with

$$q = t^\theta, \quad g(y) = y^{-\alpha} \quad \text{and} \quad h(y) = y^\beta + A.$$

References

[1] R.P. Agarwal and D. O'Regan, *Positive solutions to superlinear singular boundary value problems*, J. Comput. Appl. Math. **88** (1998), 129–147.

[2] R.P. Agarwal and D. O'Regan, *Twin solutions to singular Dirichlet problems*, J. Math. Anal. Appl. **240** (1999), 433–445.

[3] R.P. Agarwal and D. O'Regan, *Existence theory for single and multiple solutions to singular positone boundary value problems*, J. Differential Equations **175** (2001), 393–414.

[4] R.P. Agarwal, D. O'Regan, V. Lakshmikantham and S. Leela, *An upper and lower solution theory for singular Emden–Fowler equations*, Nonlinear Anal. (Real World Applications) **3** (2002), 275–291.

[5] R.P. Agarwal, D. O'Regan and V. Lakshmikantham, *Existence criteria for singular initial value problems with sign changing nonlinearities*, Math. Problems Engrg. **7** (2001), 503–524.

[6] R.P. Agarwal, D. O'Regan and P.J.Y. Wong, *Positive Solutions of Differential, Difference and Integral Equations*, Kluwer Academic, Dordrecht (1999).

[7] R. Aris, *The Mathematical Theory of Diffusion and Reaction of Permeable Catalysts*, Clarendon, Oxford (1975).

[8] J.V. Baxley, *A singular nonlinear boundary value problem: membrane response of a spherical cap*, SIAM J. Appl. Math. **48** (1988), 497–505.

[9] L.E. Bobisud, *Behaviour of solutions for a Robin problem*, J. Differential Equations **85** (1990), 91–104.

[10] L.E. Bobisud, J.E. Calvert and W.D. Royalty, *Some existence results for singular boundary value problems*, Differential Integral Equations **6** (1993), 553–571.

[11] A. Callegari and A. Nachman, *Some singular nonlinear differential equations arising in boundary layer theory*, J. Math. Anal. Appl. **64** (1978), 96–105.

[12] A. Callegari and A. Nachman, *A nonlinear singular boundary value problem in the theory of pseudoplastic fluids*, SIAM J. Appl. Math. **38** (1980), 275–282.

[13] K. Deimling, *Nonlinear Functional Analysis*, Springer-Verlag, New York (1985).

[14] P.W. Eloe and J. Henderson, *Singular nonlinear boundary value problems for higher order ordinary differential equations*, Nonlinear Anal. **17** (1991), 1–10.

[15] J.A. Gatica, V. Oliker and P. Waltman, *Singular nonlinear boundary value problems for second order differential equations*, J. Differential Equations **79** (1989), 62–78.

[16] P. Habets and F. Zanolin, *Upper and lower solutions for a generalized Emden–Fowler equation*, J. Math. Anal. Appl. **181** (1994), 684–700.

[17] C.D. Luning and W.L. Perry, *Positive solutions of negative exponent generalized Emden–Fowler boundary value problems*, SIAM J. Math. Anal. **12** (1981), 874–879.

[18] M.N. Nhashama, *A generalized upper and lower solution method and multiplicity results for nonlinear first order ordinary differential equations*, J. Math. Anal. Appl. **140** (1989), 381–395.

[19] D. O'Regan, *Existence Theory for Nonlinear Ordinary Differential Equations*, Kluwer Academic, Dordrecht (1997).

[20] S. Taliaferro, *A nonlinear singular boundary value problem*, Nonlinear Anal. **3** (1979), 897–904.

[21] A. Tineo, *Existence theorems for a singular two point Dirichlet problem*, Nonlinear Anal. **19** (1992), 323–333.

CHAPTER 2

The Lower and Upper Solutions Method for Boundary Value Problems

C. De Coster

Université du Littoral, 50 rue F. Buisson, BP 699, 62228 Calais Cédex, France
E-mail: colette.de.coster@lmpa.univ.littoral.fr

P. Habets

Université Catholique de Louvain, Institut de Mathématique Pure et Appliquée, Chemin du cyclotron 2,
1348 Louvain-la-Neuve, Belgium

Contents

Introduction . . . 71
1. Well ordered lower and upper solutions . . . 73
    1.1. A derivative independent periodic problem . . . 73
    1.2. A priori bounds on the derivatives . . . 79
    1.3. Derivative dependent periodic problems . . . 83
    1.4. Derivative dependent Dirichlet problem . . . 89
    1.5. A derivative independent Dirichlet problem . . . 90
    1.6. Historical and bibliographical notes . . . 92
2. Relation with degree theory . . . 94
    2.1. The periodic problem . . . 94
    2.2. The Dirichlet problem . . . 101
    2.3. Non well-ordered lower and upper solutions . . . 108
    2.4. Historical and bibliographical notes . . . 114
3. Variational methods . . . 115
    3.1. The minimization method . . . 115
    3.2. The minimax method . . . 121
    3.3. Historical and bibliographical notes . . . 134
4. Monotone methods . . . 135
    4.1. Abstract results . . . 135
    4.2. Well-ordered lower and upper solutions . . . 137
    4.3. Lower and upper solutions in reversed order . . . 145
    4.4. A mixed approximation scheme . . . 152

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 1
Edited by A. Cañada, P. Drábek and A. Fonda

© 2004 Elsevier B.V. All rights reserved

4.5. Historical and bibliographical notes 154
References 157

Introduction

The premises of the lower and upper solutions method can be traced back to Picard. In 1890 for partial differential equations [79] and in 1893 for ordinary differential equations [80], he introduced monotone iterations from a lower solution. This is the starting point of the use of lower and upper solutions in connection with monotone methods.

Independently, some of the basic ideas of the method appeared in the study of first order Cauchy problems made in 1915 by Perron [78] and in its extension to systems worked by Müller [72] in 1926. These authors deduced existence of solutions together with their localization between lower and upper solutions, i.e., ordered functions that satisfy differential inequalities. A good account of this theory can be found in Szarski [100] or Walter [101]. This approach is however limited to the Cauchy problem.

The major breakthrough was due to Scorza Dragoni in 1931. In two successive papers [94] and [95], this author introduced lower and upper solutions for the boundary value problem

$$u'' = f(t, u, u'), \quad u(a) = A,\ u(b) = B,$$

i.e., he considered functions α, $\beta \in C^2([a, b])$ such that $\alpha \leqslant \beta$ and

$$\begin{aligned} &\alpha'' \geqslant f(t, \alpha, \alpha'), \quad \alpha(a) \leqslant A,\ \alpha(b) \leqslant B, \\ &\beta'' \leqslant f(t, \beta, \beta'), \quad \beta(a) \geqslant A,\ \beta(b) \geqslant B. \end{aligned}$$

As for the Cauchy problem, he proved existence of a solution u and its localization between the lower and the upper solutions

$$\alpha \leqslant u \leqslant \beta.$$

Section 1 describes the present evolution of these basic ideas.

In 1972, Amann [5] associated a degree to a pair of strict lower and upper solutions. The introduction of degree theory was essential to deal with a larger class of problems such as multiplicity results. An outline of this approach is given in Section 2.

Another important step was due independently to Chang [17,18] and de Figueiredo and Solimini [38]. In 1983 and 1984 respectively, they pointed out that between lower and upper solutions the related functional has a critical point which is a minimum. This was the starting point of results relating lower and upper solutions with the variational method. Section 3 presents basic results in this direction.

Recent results that extend the old idea of Picard to use lower and upper solutions with monotone methods are discussed in Section 4.

Throughout the paper we consider two basic problems, the periodic problem

$$u'' = f(t, u, u'), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{0.1}$$

and the Dirichlet problem

$$u'' = f(t, u, u'), \quad u(a) = 0,\ u(b) = 0. \tag{0.2}$$

Although the method applies to a larger class of boundary value problems, we restrict attention to these two, in order to keep our work within a reasonable length. We also choose to describe only basic results and to select a small number of applications. A more thorough description of the method will appear in [33].

The type of nonlinearities $f : D \subset [a,b] \times \mathbb{R}^n \to \mathbb{R}$ ($n = 1$ or 2) we consider are *Carathéodory functions*, which means they satisfy the *Carathéodory conditions*:

(a) for a.e. $t \in [a,b]$, the function $f(t,\cdot)$ with domain $\{z \in \mathbb{R}^n \mid (t,z) \in D\}$ is continuous;

(b) for all $z \in \mathbb{R}^n$, the function $f(\cdot,z)$ with domain $\{t \in [a,b] \mid (t,z) \in D\}$ is measurable.

If further, for some $p \in [1,\infty]$, the Carathéodory function f satisfies

(c) for all $r > 0$, there exists $h \in L^p(a,b)$ such that for all $(t,z) \in D$ with $|z| \leqslant r$, $|f(t,z)| \leqslant h(t)$,

we say that f is an L^p*-Carathéodory function* or that it satisfies an L^p*-Carathéodory condition*. The lower and upper solution method was first developed for continuous nonlinearities. The generalization to L^p-Carathéodory function is by no means trivial and brings a better understanding of the fundamentals of the method. This is why we adopted this framework as long as it does not imply an overwhelming technicality.

In this paper, we use the following notations:

$\mathcal{C}([a,b])$ is the set of continuous functions $u : [a,b] \to \mathbb{R}$;

$\mathcal{C}_0([a,b])$ is the set of functions $u \in \mathcal{C}([a,b])$, so that $u(a) = 0$, $u(b) = 0$;

$\mathcal{C}^1([a,b])$ is the set of differentiable functions $u : [a,b] \to \mathbb{R}$ so that $u' \in \mathcal{C}([a,b])$;

$\mathcal{C}^1_0([a,b])$ is the set of functions $u \in \mathcal{C}^1([a,b])$, so that $u(a) = 0$, $u(b) = 0$;

$L^2(a,b)$ is the set of measurable functions $u : [a,b] \to \mathbb{R}$ such that

$$\|u\|_{L^2} = \left[\int_a^b |u(t)|^2 \, \mathrm{d}t\right]^{1/2} \in \mathbb{R};$$

$H^1_0(a,b)$ is the set of functions $u \in \mathcal{C}_0([a,b])$, with a weak derivative $u' \in L^2(a,b)$;

$W^{2,1}(a,b)$ is the set of functions $u \in \mathcal{C}^1([a,b])$, with a weak second derivative $u'' \in L^1(a,b)$;

given α and $\beta \in \mathcal{C}([a,b])$, we write $\alpha \leqslant \beta$ if $\alpha(t) \leqslant \beta(t)$ for all $t \in [a,b]$;

given α and $\beta \in \mathcal{C}([a,b])$, we define $[\alpha,\beta] = \{u \in \mathcal{C}([a,b]) \mid \alpha \leqslant u \leqslant \beta\}$;

$\mathbb{N}_0 = \mathbb{N} \setminus \{0\}$, $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$.

Given $u \in \mathcal{C}([a,b])$, we define the Dini derivatives

$$D_+u(t) = \liminf_{h\to 0^+} \frac{u(t+h)-u(t)}{h}, \qquad D^+u(t) = \limsup_{h\to 0^+} \frac{u(t+h)-u(t)}{h},$$

$$D_-u(t) = \liminf_{h\to 0^-} \frac{u(t+h)-u(t)}{h}, \qquad D^-u(t) = \limsup_{h\to 0^-} \frac{u(t+h)-u(t)}{h}.$$

Considering a function $u : [a,b] \to \mathbb{R}$, its *periodic extension on* $\mathbb{R}$ is the function $u : \mathbb{R} \to \mathbb{R}$ defined from $u(t) \equiv u(t+b-a)$.

1. Well ordered lower and upper solutions

1.1. *A derivative independent periodic problem*

Consider the derivative independent periodic problem

$$u'' = f(t,u), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{1.1}$$

where f is an L^1-Carathéodory function. Solutions of (1.1) are in $W^{2,1}(a,b)$ so that it is natural to look for lower and upper solutions which are in this space or at least which are piecewise $W^{2,1}$. This motivates the definitions we present here. To simplify the notations, we extend $f(t,u)$ by periodicity, i.e., $f(t,u) = f(t+b-a,u)$ for all $(t,u) \in \mathbb{R}^2$.

DEFINITIONS 1.1. A function $\alpha \in \mathcal{C}([a,b])$ such that $\alpha(a) = \alpha(b)$ is a *lower solution* of (1.1) if its periodic extension on $\mathbb{R}$ is such that for any $t_0 \in \mathbb{R}$ either $D_-\alpha(t_0) < D^+\alpha(t_0)$, or there exists an open interval I_0 such that $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$ and, for a.e. $t \in I_0$,

$$\alpha''(t) \geqslant f\bigl(t,\alpha(t)\bigr).$$

A function $\beta \in \mathcal{C}([a,b])$ such that $\beta(a) = \beta(b)$ is an *upper solution* of (1.1) if its periodic extension on $\mathbb{R}$ is such that for any $t_0 \in \mathbb{R}$ either $D^-\beta(t_0) > D_+\beta(t_0)$, or there exists an open interval I_0 such that $t_0 \in I_0$, $\beta \in W^{2,1}(I_0)$ and, for a.e. $t \in I_0$,

$$\beta''(t) \leqslant f\bigl(t,\beta(t)\bigr).$$

Notice that the condition $D_-\alpha(t_0) < D^+\alpha(t_0)$ cannot hold for all $t_0 \in [a,b]$. In practical problems it only holds at a finite number of points. Further the "natural" lower and upper solutions are often more regular than needed in Definitions 1.1. For example, a lower solution α will often be produced as the solution of some auxiliary problem so that it will be in $W^{2,1}(a,b)$ and satisfy for a.e. $t \in [a,b]$

$$\alpha''(t) \geqslant f\bigl(t,\alpha(t)\bigr), \quad \alpha(a) = \alpha(b),\ \alpha'(a) \geqslant \alpha'(b).$$

The following theorem is the basic existence result of the lower and upper solutions method for solutions of the periodic problem (1.1).

THEOREM 1.1. *Let α and β be lower and upper solutions of* (1.1) *such that $\alpha \leqslant \beta$,*

$$E = \bigl\{(t,u) \in [a,b] \times \mathbb{R} \mid \alpha(t) \leqslant u \leqslant \beta(t)\bigr\} \tag{1.2}$$

and $f : E \to \mathbb{R}$ be an L^1-Carathéodory function. Then the problem (1.1) *has at least one solution $u \in W^{2,1}(a,b)$ such that for all $t \in [a,b]$*

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

PROOF. We consider the modified problem

$$u'' - u = f\big(t, \gamma(t, u)\big) - \gamma(t, u), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{1.3}$$

where

$$\gamma(t, u) = \max\big\{\alpha(t), \min\{u, \beta(t)\}\big\}. \tag{1.4}$$

Claim 1. *The problem* (1.3) *has at least one solution.* Let us write (1.3) as an integral equation

$$u(t) = \int_a^b G(t, s)\big[f\big(s, \gamma\big(s, u(s)\big)\big) - \gamma\big(s, u(s)\big)\big]\,\mathrm{d}s,$$

where $G(t, s)$ is the Green's function corresponding to the problem

$$u'' - u = f(t), \quad u(a) = u(b),\ u'(a) = u'(b). \tag{1.5}$$

The operator

$$T : \mathcal{C}\big([a, b]\big) \to \mathcal{C}\big([a, b]\big),$$

defined by

$$(Tu)(t) = \int_a^b G(t, s)\big[f\big(s, \gamma\big(s, u(s)\big)\big) - \gamma\big(s, u(s)\big)\big]\,\mathrm{d}s,$$

is completely continuous and bounded. By Schauder's theorem, T has a fixed point which is a solution of (1.3).

Claim 2. *All solutions* u *of* (1.3) *satisfy on* $[a, b]$

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

Let us assume on the contrary that, for some $t_0 \in [a, b]$

$$\min_t\big(u(t) - \alpha(t)\big) = u(t_0) - \alpha(t_0) < 0.$$

Extending the functions by periodicity, we have then

$$u'(t_0) - D_-\alpha(t_0) \leqslant 0 \leqslant u'(t_0) - D^+\alpha(t_0)$$

and by definition of a lower solution $u'(t_0) - \alpha'(t_0) = 0$. Further, there exists an open interval I_0, with $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$ and for almost every $t \in I_0$

$$\alpha''(t) \geqslant f\big(t, \alpha(t)\big).$$

Hence, for $t \geqslant t_0$, near enough t_0,

$$u'(t) - \alpha'(t) = \int_{t_0}^{t} \bigl(u''(s) - \alpha''(s)\bigr)\,\mathrm{d}s$$
$$\leqslant \int_{t_0}^{t} \bigl[f\bigl(s, \alpha(s)\bigr) + u(s) - \alpha(s) - f\bigl(s, \alpha(s)\bigr)\bigr]\,\mathrm{d}s < 0.$$

This proves $u(t_0) - \alpha(t_0)$ is not a minimum of $u - \alpha$ which is a contradiction. A similar argument holds to prove $u \leqslant \beta$. □

Theorem 1.1 furnishes two kinds of information. It is an existence result but it gives also a localization of the solution. In the following example, such a localization provides an asymptotic estimate on the solution.

EXAMPLE 1.1. Consider the problem

$$\epsilon u'' = \bigl(u - |t|\bigr)^3, \quad u(-1) = u(1),\ u'(-1) = u'(1),$$

where $\epsilon > 0$ is a parameter. Let $k = \epsilon^p$ with $p \in]0, 1/4[$,

$$\alpha(t) = 1 - \sqrt{(|t| - 1)^2 + k^2} \quad \text{and} \quad \beta(t) = \sqrt{t^2 + k^2}.$$

If ϵ is small enough, these functions are lower and upper solutions and we deduce from Theorem 1.1 the existence of a solution u such that $\alpha(t) \leqslant u(t) \leqslant \beta(t)$ on $[-1, 1]$. This implies an asymptotic estimate

$$u(t) = |t| + \mathrm{O}\bigl(\epsilon^p\bigr).$$

Notice also that in this example $\beta'(-1) \neq \beta'(1)$ and α is not differentiable at $t = 0$.

Another illustration of Theorem 1.1 is the following.

EXAMPLE 1.2. Consider the problem

$$u'' = \frac{1}{\sqrt{t}} u^2 + q(t), \quad u(0) = u(2\pi),\ u'(0) = u'(2\pi),$$

where $q \in L^1(0, 2\pi)$ is such that

$$\bar{q} + (2\pi)^{3/2} \|\tilde{q}\|_{L^1}^2 \leqslant 0,$$

with $\bar{q} = \frac{1}{2\pi} \int_0^{2\pi} q(t)\,\mathrm{d}t$ and $\tilde{q}(t) = q(t) - \bar{q}$.

Let us prove this problem has a solution u. We define w to be a solution of

$$w'' = \tilde{q}(t), \quad w(0) = w(2\pi),\ w'(0) = w'(2\pi),$$

and $\alpha(t) = w(t) - w(0)$. Hence, we have $\|\alpha'\|_\infty \leqslant \|\tilde{q}\|_{L^1}$ and for all $t \in [0, 2\pi]$, $|\alpha(t)| \leqslant \|\tilde{q}\|_{L^1} t$. We compute then

$$\alpha'' - \frac{1}{\sqrt{t}}\alpha^2 - q(t) \geqslant -\frac{1}{\sqrt{t}}\|\tilde{q}\|_{L^1}^2 t^2 - \bar{q} \geqslant -\|\tilde{q}\|_{L^1}^2 (2\pi)^{3/2} - \bar{q} \geqslant 0,$$

which proves $\alpha(t)$ is a lower solution. If c is a large enough positive constant then $\beta(t) = \alpha(t) + c \geqslant \alpha(t)$ is an upper solution and existence of a solution follows from Theorem 1.1.

REMARK. Theorem 1.1 depends strongly on the ordering $\alpha \leqslant \beta$. In case this ordering is not satisfied, the result does not hold as such. Consider for example the following problem.

EXAMPLE 1.3. From Fredholm alternative, it is clear that the problem

$$u'' + u = \sin t, \quad u(0) = u(2\pi),\ u'(0) = u'(2\pi)$$

has no solution. However

$$\alpha(t) = 1 \quad \text{and} \quad \beta(t) = -1$$

are lower and upper solutions.

REMARK. Another remark of the same type is that the result is no longer true if we allow for α angles with opening from below. This is the case in the following example.

EXAMPLE 1.4. Consider the problem

$$u'' = 1, \quad u(-1) = u(1),\ u'(-1) = u'(1).$$

It has no solution although $\alpha(t) = t^2 - 1$ is almost a lower solution (i.e., $\alpha''(t) = 2 > 1$, $\alpha(-1) = \alpha(1)$), $\beta(t) = 1$ is an upper solution ($\beta''(t) \leqslant 1$, $\beta(-1) = \beta(1)$, $\beta'(-1) = \beta'(1)$) and $\alpha(t) < \beta(t)$. Clearly, Theorem 1.1 does not apply here since $D_-\alpha(1) > D^+\alpha(1) = D^+\alpha(-1)$ which means that α is not a lower solution.

In applications it is often useful to use the maximum of lower solutions and the minimum of upper solutions. Although it is probably true, it is not obvious with our definition that such functions are lower and upper solutions. However, we can prove the existence of solutions between such maximum and minimum.

THEOREM 1.2. *Let α_i $(i = 1, \ldots, n)$ be lower solutions and β_j $(j = 1, \ldots, m)$ be upper solutions of* (1.1), *$\alpha := \max_{1\leqslant i\leqslant n} \alpha_i$ and $\beta := \min_{1\leqslant j\leqslant m} \beta_j$ be such that $\alpha \leqslant \beta$. Define E*

from (1.2) *and let* $f : E \to \mathbb{R}$ *be an* L^1*-Carathéodory function. Then the problem* (1.1) *has at least one solution* $u \in W^{2,1}(a,b)$ *such that for all* $t \in [a,b]$

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

PROOF. Consider the modified problem

$$u'' - u = \bar{f}(t,u) - \gamma(t,u), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{1.6}$$

where $\gamma(t,u)$ is defined in (1.4) and

$$\bar{f}(t,u) = \begin{cases} \min_{1\leqslant i\leqslant n} f\bigl(t, \max\{\alpha_i(t), u\}\bigr), & \text{if } u \leqslant \alpha(t), \\ f(t,u), & \text{if } \alpha(t) < u < \beta(t), \\ \max_{1\leqslant j\leqslant m} f\bigl(t, \min\{\beta_j(t), u\}\bigr), & \text{if } \beta(t) \leqslant u. \end{cases}$$

First, we prove as in Theorem 1.1 that problem (1.6) has a solution. Let us show next that

$$\alpha(t) \leqslant u(t) \leqslant \beta(t)$$

on $[a,b]$. Extend α and u by periodicity and assume by contradiction that $\min_t(u(t) - \alpha(t)) < 0$. It follows that, for some t_0 and $i \in \{1,\ldots,n\}$, we have $\min_t(u(t) - \alpha(t)) = u(t_0) - \alpha_i(t_0) = \min_t(u(t) - \alpha_i(t)) < 0$. A contradiction follows now as in the proof of Theorem 1.1. Finally, $u \leqslant \beta$ follows from the same argument. □

The next result concerns existence of minimal and maximal solutions between lower and upper solutions.

THEOREM 1.3. *Let* α *and* β *be lower and upper solutions of* (1.1) *such that* $\alpha \leqslant \beta$. *Define E from* (1.2) *and let* $f : E \to \mathbb{R}$ *satisfy an* L^1*-Carathéodory condition. Then the problem* (1.1) *has a* minimal and a maximal solution in $[\alpha,\beta]$, *i.e., solutions* $u_{\min}$ *and* $u_{\max}$ *such that*

$$\alpha \leqslant u_{\min} \leqslant u_{\max} \leqslant \beta$$

and any other solution u of (1.1) *with* $\alpha \leqslant u \leqslant \beta$ *satisfies*

$$u_{\min} \leqslant u \leqslant u_{\max}.$$

PROOF. Notice first that solutions of (1.1) are fixed points of the operator

$$T : \mathcal{C}\bigl([a,b]\bigr) \to \mathcal{C}\bigl([a,b]\bigr)$$

defined by

$$(Tu)(t) = \int_a^b G(t,s)\bigl[f\bigl(s,u(s)\bigr) - u(s)\bigr]\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function of (1.5). Define then the set

$$S = \{u \mid u = Tu,\ \alpha \leqslant u \leqslant \beta\}.$$

From Theorem 1.1, $S \neq \emptyset$. Further S is compact as T is completely continuous. Consider next the family of sets

$$F_x = \{u \in S \mid u \leqslant x\},$$

where $x \in S$. This family has the finite intersection property as follows from Theorem 1.2. Hence, it is known (see [60, Theorem 5.1]) that there exists

$$u_{\min} \in \bigcap_{x \in S} F_x,$$

which is a minimal solution.

Similarly, we prove existence of a maximal solution. □

The structure of the set of solutions is richer if f is nondecreasing with respect to u. In such a case, we have a continuum of solutions as follows from the following theorem.

THEOREM 1.4. *Assume the hypotheses of Theorem* 1.3 *hold and f is nondecreasing with respect to u. Then the set of solutions u of* (1.1) *with $u_{\min} \leqslant u \leqslant u_{\max}$ is such that for any $t_0 \in [a,b]$ and $u^* \in [u_{\min}(t_0), u_{\max}(t_0)]$ there exists one of them with $u(t_0) = u^*$.*

PROOF. Let $t_0 \in [a,b]$ and u^* be such that $u_{\min}(t_0) \leqslant u^* \leqslant u_{\max}(t_0)$. Choose $\epsilon > 0$ large enough so that $u_{\max} - \epsilon \leqslant u_{\min} + \epsilon$ and define

$$\alpha_1(t) = \max\bigl\{u_{\min}(t), u_{\max}(t) - \epsilon\bigr\}, \qquad \beta_1(t) = \min\bigl\{u_{\max}(t), u_{\min}(t) + \epsilon\bigr\}.$$

Observe that $u_{\max} - \epsilon$ and $u_{\min} + \epsilon$ are respectively lower and upper solutions of (1.1). By Theorem 1.2, the problem (1.1) has a solution u_1 such that, for all $t \in [a,b]$,

$$u_{\min}(t) \leqslant u_1(t) \leqslant u_{\min}(t) + \epsilon, \qquad u_{\max}(t) - \epsilon \leqslant u_1(t) \leqslant u_{\max}(t).$$

In case $u^* \in [u_{\min}(t_0), u_1(t_0)]$ we define

$$\alpha_2(t) = \max\bigl\{u_{\min}(t), u_1(t) - \epsilon/2\bigr\}, \qquad \beta_2(t) = \min\bigl\{u_1(t), u_{\min}(t) + \epsilon/2\bigr\}$$

and obtain from Theorem 1.2 a solution u_2 such that on $[a,b]$

$$u_{\min}(t) \leqslant u_2(t) \leqslant u_{\min}(t) + \epsilon/2, \qquad u_1(t) - \epsilon/2 \leqslant u_2(t) \leqslant u_1(t).$$

If $u^* \in \,]u_1(t_0), u_{\max}(t_0)]$, we proceed in a similar way. This defines a sequence of solutions $(u_k)_k$ that satisfies $|u_k(t_0) - u^*| \leqslant \epsilon/2^{k-1}$. Next, from Arzelà–Ascoli theorem, there is a subsequence $(u_{k_n})_n$ such that, for some $u \in \mathcal{C}([a,b])$, u_{k_n} converges to u in $\mathcal{C}([a,b])$. It follows then that u is a solution of (1.1). Further, we have $u(t_0) = \lim_{n\to\infty} u_{k_n}(t_0) = u^*$. □

Existence of a continuum of solutions depends strongly on the nondecreasingness of f. Such a continuum does not exist in the following example.

EXAMPLE 1.5. Consider the problem

$$u'' = u^3 - u^2,$$
$$u(0) = u(T),\ u'(0) = u'(T).$$

Notice first that $\alpha(t) = -2$ and $\beta(t) = 2$ are lower and upper solutions. Also, it is straightforward from a phase plane analysis that this problem has only two solutions, $u_1 = 0$ and $u_2 = 1$, which are between α and β.

Notice also that solutions between lower and upper solutions are not necessarily ordered as shown in the example that follows.

EXAMPLE 1.6. The piecewise linear problem

$$u'' = \min\{u+2, \max\{-u, u-2\}\},$$
$$u(0) = u(2\pi),\ u'(0) = u'(2\pi),$$

is such that $u_1(t) = \sin t$ and $u_2(t) = -\sin t$ are nonordered solutions which lie between the lower solution $\alpha(t) = -3$ and the upper one $\beta(t) = 3$. In this problem, it follows from the phase plane analysis that the minimal and maximal solutions are respectively $u_{\min}(t) = -2$ and $u_{\max}(t) = 2$.

1.2. *A priori bounds on the derivatives*

Consider the problem

$$u'' = f(t,u,u'), \quad u(a) = u(b),\ u'(a) = u'(b). \tag{1.7}$$

Here the Nemitskii operator reads $N(u) := f(\cdot, u, u')$ and therefore the fixed point problem associated to (1.7) is defined now on $\mathcal{C}^1([a,b])$. Lower and upper solutions will give a priori bounds on u. In order to apply the Schauder Fixed Point theorem or degree theory we shall also need a priori bounds on the derivative u'. In some cases, the special structure of the nonlinearity f gives this information. In others, this follows from a Nagumo condition. The following example shows that in any case some condition is necessary since the existence of solutions does not follow from the existence of ordered lower and upper solutions.

EXAMPLE 1.7. Consider the problem

$$u'' = \big(1 + u'^2\big)^2\big(u - p(t)\big), \quad u(0) = u(T),\ u'(0) = u'(T), \tag{1.8}$$

where p is a continuous function such that $p(t) = 2$ on $[0, T/3[$, $p(t) \in [-2, 2]$ on $[T/3, 2T/3[$ and $p(t) = -2$ on $[2T/3, T]$. For such a problem we define lower and upper solutions from Definitions 1.2 below. It follows that $\alpha = -3$ is a lower solution and $\beta = 3$ an upper one. However, it can be proved using elementary methods that if $T > 0$ is large enough, (1.8) has no solution (see [33,50]).

To illustrate how a priori bounds on the derivative can follow from the structure of the equation, consider the periodic problem for a Rayleigh equation

$$u'' + g(u') + h(t, u) = 0, \quad u(a) = u(b),\ u'(a) = u'(b). \tag{1.9}$$

PROPOSITION 1.5. *Let $h : [a, b] \times [-r, r] \to \mathbb{R}$ be a Carathéodory function such that for some $h_0 \in L^2(a, b)$, for a.e. $t \in [a, b]$ and all $u \in [-r, r]$, we have $|h(t, u)| \leqslant h_0(t)$. Then there exists $R > 0$ such that for every function $g \in \mathcal{C}(\mathbb{R})$, any solution u of* (1.9) *with $\|u\|_\infty \leqslant r$ satisfies $\|u'\|_\infty < R$.*

PROOF. Define $R := \sqrt{b - a}\,\|h_0\|_{L^2} + 1 > 0$. Let then $g \in \mathcal{C}(\mathbb{R})$ be given and u be a solution of (1.9) with $\|u\|_\infty \leqslant r$. Multiplying (1.9) by u'' and integrating we obtain

$$\|u''\|_{L^2}^2 = -\int_a^b g\big(u'(t)\big)u''(t)\,\mathrm{d}t - \int_a^b h\big(t, u(t)\big)u''(t)\,\mathrm{d}t \leqslant \|h_0\|_{L^2}\|u''\|_{L^2}.$$

Now it is easy to see that for some $t_0 \in [a, b]$, $u'(t_0) = 0$ so that for all $t \in [a, b]$

$$\big|u'(t)\big| = \left|\int_{t_0}^t u''(s)\,\mathrm{d}s\right| \leqslant \sqrt{b - a}\,\|h_0\|_{L^2} < R.$$

□

In case the equation does not have any special structure, a priori bounds on the derivative can still be obtained for nonlinearities which do not grow too quickly with respect to the derivative. For continuous functions, Nagumo conditions describe such a control. A typical result is the following.

PROPOSITION 1.6. *Let $E \subset [a, b] \times [-r, r] \times \mathbb{R}$ and let $\bar{\varphi} : \mathbb{R}^+ \to \mathbb{R}_0^+$ be a continuous function that satisfies*

$$\int_0^\infty \frac{s\,\mathrm{d}s}{\bar{\varphi}(s)} = +\infty. \tag{1.10}$$

Then there exists $R > 0$ such that for every continuous function $f : E \to \mathbb{R}$ that satisfies

$$\forall (t, u, v) \in E, \quad \big|f(t, u, v)\big| \leqslant \bar{\varphi}\big(|v|\big) \tag{1.11}$$

and every solution u of

$$u'' = f(t, u, u') \tag{1.12}$$

such that $\|u\|_\infty \leqslant r$, *we have* $\|u'\|_\infty < R$.

REMARK. A function $f : E \to \mathbb{R}$ is said to satisfy a *Nagumo condition* if (1.11) holds.

PROOF. Define $R > 0$ to be such that

$$\int_{2r/(b-a)}^{R} \frac{s\,\mathrm{d}s}{\bar{\varphi}(s)} > 2r \tag{1.13}$$

and let u be a solution of (1.12) such that $(t, u(t), u'(t)) \in E$ on $[a, b]$. Observe that there exists $\tau \in [a, b]$ with $|u'(\tau)| \leqslant 2r/(b-a)$. Assume now there exists an interval $I = [t_0, t_1]$ (or $[t_1, t_0]$) such that $u'(t_0) = 2r/(b-a)$, $u'(t_1) = R$ and $u'(t) \in [2r/(b-a), R]$ on I. We have then

$$\begin{aligned}\int_{u'(t_0)}^{u'(t_1)} \frac{s\,\mathrm{d}s}{\bar{\varphi}(s)} &= \int_{t_0}^{t_1} \frac{u'(t)u''(t)}{\bar{\varphi}(u'(t))}\,\mathrm{d}t = \int_{t_0}^{t_1} \frac{u'(t)f(t, u(t), u'(t))}{\bar{\varphi}(u'(t))}\,\mathrm{d}t \\ &\leqslant \bigl|u(t_1) - u(t_0)\bigr| \leqslant 2r,\end{aligned}$$

which contradicts (1.13).

In the same way, we prove that, for any $t \in [a, b]$, $u'(t) > -R$ and the result follows. □

REMARKS. In the above proof we do not use the divergence of the integral in (1.10) but rather the fact that for some $R > 0$,

$$\int_{\frac{2r}{b-a}}^{R} \frac{s\,\mathrm{d}s}{\bar{\varphi}(s)} > 2r.$$

In case $\varphi(s) = 1 + s^p$, condition (1.10) implies that $p \leqslant 2$. However this condition still holds if $\varphi(s) = s^2 \ln(s^2 + 1) + 1$.

A fundamental generalization of this result concerns one-sided Nagumo conditions. This applies to problems where some a priori bound on the derivative of solutions is known at the points a and b. For the periodic problem we can extend the solution by periodicity and consider an interval $[\bar{a}, \bar{a} + b - a]$ so that $u'(\bar{a}) = u'(\bar{a} + b - a) = 0$.

PROPOSITION 1.7. *Let* $E \subset [\bar{a}, \bar{b}] \times [-r, r] \times \mathbb{R}$, $k \geqslant 0$ *and let* $\bar{\varphi} : \mathbb{R}^+ \to \mathbb{R}_0^+$ *be a continuous function that satisfies* (1.10). *Then there exists* $R > 0$ *such that for every continuous function* $f : E \to \mathbb{R}$ *that satisfies*

$$\forall (t, u, v) \in E, \quad f(t, u, v) \leqslant \bar{\varphi}\bigl(|v|\bigr) \tag{1.14}$$

and for every solution u of (1.12) *on* $[\bar a, \bar b]$ *such that* $\|u\|_\infty \leqslant r$, $u'(\bar a) \leqslant k$ *and* $u'(\bar b) \geqslant -k$, *we have* $\|u'\|_\infty < R$.

PROOF. The proof follows the argument of the proof of Proposition 1.6. □

REMARK. Condition (1.14) is called a *one-sided Nagumo condition*.

The above result still holds for other such conditions. We can consider any of the following:

(a) $f(t,u,v) \geqslant -\bar\varphi(|v|)$ for all $(t,u,v) \in E$, $u'(\bar a) \geqslant -k$ and $u'(\bar b) \leqslant k$;
(b) $\operatorname{sgn}(v) f(t,u,v) \leqslant \bar\varphi(|v|)$ for all $(t,u,v) \in E$ and $|u'(\bar a)| \leqslant k$;
(c) $\operatorname{sgn}(v) f(t,u,v) \geqslant -\bar\varphi(|v|)$ for all $(t,u,v) \in E$ and $|u'(\bar b)| \leqslant k$.

Such assumptions apply for problems of the type

$$u'' \pm (2+\sin t)(u')^m + h(t,u) = 0, \quad u(a) = u(b),\ u'(a) = u'(b),$$

with $m \geqslant 0$.

The Nagumo condition implies the function f at hand is L^∞-Carathéodory. Hence this condition has to be extended so as to deal with L^p-Carathéodory functions which are not L^∞-Carathéodory.

PROPOSITION 1.8. *Let* $E \subset [a,b] \times [-r,r] \times \mathbb{R}$, p, $q \in [1,\infty]$ *with* $\frac{1}{p} + \frac{1}{q} = 1$ *and* $\bar\psi \in L^p(a,b)$. *Let also* $\bar\varphi : \mathbb{R}^+ \to \mathbb{R}_0^+$ *be a continuous function that satisfies*

$$\int_0^{+\infty} \frac{s^{1/q}}{\bar\varphi(s)}\,\mathrm{d}s = +\infty. \tag{1.15}$$

Then, there exists $R > 0$ *so that for every Carathéodory function* $f : E \to \mathbb{R}$ *such that*

$$\begin{aligned} &\text{for a.e. } t \in [a,b] \text{ and all } (u,v) \in \mathbb{R}^2, \text{ with } (t,u,v) \in E,\\ &\bigl|f(t,u,v)\bigr| \leqslant \bar\psi(t)\bar\varphi\bigl(|v|\bigr), \end{aligned} \tag{1.16}$$

and for every solution u of (1.12) *such that* $\|u\|_\infty \leqslant r$, *we have*

$$\|u'\|_\infty < R.$$

PROOF. Define $R > 0$ to be such that

$$\int_{2r/(b-a)}^{R} \frac{s^{1/q}}{\bar\varphi(s)}\,\mathrm{d}s > \|\bar\psi\|_{L^p}(2r)^{1/q}.$$

Let u be a solution of (1.12) and $t \in [a,b]$ be such that $u'(t) \geqslant R$. We can choose, as in the proof of Proposition 1.6, $t_0 < t_1$ (or $t_0 > t_1$) such that $u'(t_0) = 2r/(b-a)$, $u'(t_1) = R$ and

$u'(t) \in [2r/(b-a), R]$ on $[t_0, t_1]$ (or $[t_1, t_0]$). Next, we write

$$\begin{aligned}\int_{u'(t_0)}^{u'(t_1)} \frac{s^{1/q}}{\bar{\varphi}(s)}\,\mathrm{d}s &= \int_{t_0}^{t_1} \frac{(u')^{1/q}(t)u''(t)}{\bar{\varphi}(u'(t))}\,\mathrm{d}t\\ &= \int_{t_0}^{t_1} \frac{(u')^{1/q}(t)f(t,u(t),u'(t))}{\bar{\varphi}(u'(t))}\,\mathrm{d}t\\ &\leqslant \left|\int_{t_0}^{t_1} \bar{\psi}(t)(u')^{1/q}(t)\,\mathrm{d}t\right| \leqslant \|\bar{\psi}\|_{L^p}\left|\int_{t_0}^{t_1} u'(t)\,\mathrm{d}t\right|^{1/q}\\ &\leqslant \|\bar{\psi}\|_{L^p}(2r)^{1/q}.\end{aligned}$$

We obtain a contradiction and deduce that $u'(t) < R$ on $[a, b]$. In the same way we prove that $u'(t) > -R$ on $[a, b]$. □

Assumption (1.16) is a Nagumo condition. Similar one-sided conditions can also be worked out for the Carathéodory case.

PROPOSITION 1.9. *Let $E \subset [\bar{a}, \bar{b}] \times [-r, r] \times \mathbb{R}$, $k \geqslant 0$, $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$ and $\bar{\psi} \in L^p(a, b)$. Let also $\bar{\varphi} : \mathbb{R}^+ \to \mathbb{R}_0^+$ be a continuous function that satisfies* (1.15). *Then, there exists $R > 0$ so that for every L^p-Carathéodory function $f : E \to \mathbb{R}$ such that*

$$\begin{aligned}&\text{for a.e. } t \in [a, b] \text{ and all } (u, v) \in \mathbb{R}^2, \text{ with } (t, u, v) \in E,\\ &f(t, u, v) \leqslant \bar{\psi}(t)\bar{\varphi}(|v|),\end{aligned} \tag{1.17}$$

and for every solution u of (1.12) *such that $\|u\|_\infty \leqslant r$, $u'(\bar{a}) \leqslant k$ and $u'(\bar{b}) \geqslant -k$, we have*

$$\|u'\|_\infty < R.$$

PROOF. The proof follows the line of the proof of Proposition 1.8. □

REMARK. Similar results hold for the other one-sided Nagumo conditions as in the remark following Proposition 1.7.

1.3. *Derivative dependent periodic problems*

Definitions of lower and upper solutions for the periodic problem (1.7) are straightforward extensions of Definitions 1.1.

DEFINITIONS 1.2. A function $\alpha \in \mathcal{C}([a, b])$ such that $\alpha(a) = \alpha(b)$ is a *lower solution* of (1.7) if its periodic extension on $\mathbb{R}$ is such that for any $t_0 \in \mathbb{R}$ either $D_-\alpha(t_0) < D^+\alpha(t_0)$, or there exists an open interval I_0 such that $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$ and, for a.e. $t \in I_0$,

$$\alpha''(t) \geqslant f\big(t, \alpha(t), \alpha'(t)\big).$$

A function $\beta \in \mathcal{C}([a,b])$ such that $\beta(a)=\beta(b)$ is an *upper solution* of (1.7) if its periodic extension on $\mathbb{R}$ is such that for any $t_0 \in \mathbb{R}$ either $D^-\beta(t_0) > D_+\beta(t_0)$, or there exists an open interval I_0 such that $t_0 \in I_0$, $\beta \in W^{2,1}(I_0)$ and, for a.e. $t \in I_0$,

$$\beta''(t) \leqslant f\big(t,\beta(t),\beta'(t)\big).$$

As a first result we consider periodic solutions of the Rayleigh equation.

THEOREM 1.10. *Let α, $\beta \in \mathcal{C}([a,b])$ be lower and upper solutions of the problem* (1.9) *such that $\alpha \leqslant \beta$. Let E be defined by* (1.2), *$g \in \mathcal{C}(\mathbb{R})$ and let $h: E \to \mathbb{R}$ be an L^2-Carathéodory function. Then problem* (1.9) *has at least one solution $u \in W^{2,2}(a,b)$ such that for all $t \in [a,b]$*

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

PROOF. Consider the family of modified problems

$$\begin{aligned} &u'' - C(t)u = -\big[\lambda g(u') + h\big(t,\gamma(t,u)\big) + C(t)\gamma(t,u)\big],\\ &u(a)=u(b),\ u'(a)=u'(b), \end{aligned} \tag{1.18}$$

where $\gamma(t,u)$ is defined from (1.4), $C \in L^1(a,b)$ is chosen such that $C(t) > |g(0)| + 1 + |h(t,u)|$ for $(t,u) \in E$ and $\lambda \in [0,1]$.

Claim 1. *Define $\rho = \max\{\|\alpha\|_\infty, \|\beta\|_\infty\} + 1$. Then every solution u of* (1.18) *satisfies $\|u\|_\infty < \rho$.* Let us assume on the contrary that for some $t_0 \in \mathbb{R}$

$$\min_t u(t) = u(t_0) \leqslant -\rho.$$

Hence $u'(t_0)=0$ and we compute for $t \geqslant t_0$ close enough to t_0

$$\begin{aligned} u'(t) &= \int_{t_0}^t u''(s)\,\mathrm{d}s\\ &\leqslant \int_{t_0}^t \big[C(s)\big(u(s)-\alpha(s)\big) - \lambda g\big(u'(s)\big) - h\big(s,\alpha(s)\big)\big]\,\mathrm{d}s\\ &\leqslant -\int_{t_0}^t \big[C(s) - \big|g\big(u'(s)\big)\big| - h\big(s,\alpha(s)\big)\big]\,\mathrm{d}s < 0. \end{aligned}$$

This proves that $u(t_0)$ is not a minimum of u which is a contradiction. A similar argument holds to prove that $u \leqslant \rho$.

Claim 2. *There exists $R > 0$ such that every solution u of* (1.18) *with $\|u\|_\infty < \rho$ satisfies $\|u'\|_\infty < R$.* The proof follows the argument in Proposition 1.5.

Claim 3. *There exists a solution u of* (1.18) *with $\lambda = 1$.* Define the operator $T_\lambda : \mathcal{C}^1([a,b]) \to \mathcal{C}^1([a,b])$ by

$$T_\lambda(u) = -\int_a^b G(t,s)\big[\lambda g\big(u'(s)\big) + h\big(s,\gamma\big(s,u(s)\big)\big) + C(s)\gamma\big(s,u(s)\big)\big]\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function of

$$u'' - C(t)u = f(t),$$
$$u(a) = u(b),\ u'(a) = u'(b).$$

Observe that there exists $R_0 > 0$ such that $T_0(\mathcal{C}^1([a,b])) \subset B(0, R_0)$. Hence,

$$\deg\bigl(I - T_0, B(0, R_0)\bigr) = 1$$

and by the properties of the degree, we prove easily that (1.18) with $\lambda = 1$ has a solution.

Claim 4. *The solution u of* (1.18) *with $\lambda = 1$ is such that $\alpha(t) \leqslant u(t) \leqslant \beta(t)$ on* $[a,b]$. Extend α and u by periodicity. Let $t_0 \in [a,b[$ be such that

$$u(t_0) - \alpha(t_0) = \min_t \bigl(u(t) - \alpha(t)\bigr) < 0.$$

From the definition of lower solution there exists an open interval I_0 such that $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$ and, for a.e. $t \in I_0$,

$$\alpha''(t) + g\bigl(\alpha'(t)\bigr) + h\bigl(t, \alpha(t)\bigr) \geqslant 0.$$

Further, for $t \geqslant t_0$ near enough t_0, we compute

$$\begin{aligned} u'(t) - \alpha'(t) &= \int_{t_0}^{t} \bigl(u''(s) - \alpha''(s)\bigr)\,\mathrm{d}s \\ &\leqslant \int_{t_0}^{t} \bigl[g\bigl(\alpha'(s)\bigr) - g\bigl(u'(s)\bigr) + C(s)\bigl(u(s) - \alpha(s)\bigr)\bigr]\,\mathrm{d}s < 0, \end{aligned}$$

which follows as $g(\alpha'(s)) - g(u'(s))$ is small and $C(s)(\alpha(s) - u(s)) > \alpha(s) - u(s) \geqslant k$ for some $k > 0$. This contradicts the minimality of $u - \alpha$ at t_0.

In a similar way we prove that $u \leqslant \beta$. Hence u is also a solution of (1.18). □

A similar result is easy to work out for the general problem (1.7) in case a one-sided Nagumo condition is satisfied. Here we work the case of a continuous function f. The key of the proof of this generalization is to use an appropriate modified problem.

THEOREM 1.11. *Let α, $\beta \in \mathcal{C}([a,b])$ be lower and upper solutions of the problem* (1.7) *such that $\alpha \leqslant \beta$. Let*

$$E = \bigl\{(t, u, v) \in [a,b] \times \mathbb{R}^2 \mid \alpha(t) \leqslant u \leqslant \beta(t)\bigr\}, \tag{1.19}$$

$\varphi : \mathbb{R}^+ \to \mathbb{R}$ *be a positive continuous function satisfying* (1.10) *and $f : E \to \mathbb{R}$ a continuous function which satisfies the one-sided Nagumo condition* (1.14) (*with $\bar{\varphi} = \varphi$*). *Then the problem* (1.7) *has at least one solution $u \in \mathcal{C}^2([a,b])$ such that for all $t \in [a,b]$*

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

PROOF. Consider the modified problem

$$\begin{aligned} &u'' = \lambda f\big(t, \gamma(t,u), u'\big) + \varphi\big(|u'|\big)\big(u - \lambda\gamma(t,u)\big),\\ &u(a) = u(b),\ u'(a) = u'(b), \end{aligned} \tag{1.20}$$

where $\gamma(t,u)$ is defined by (1.4). Choose $r > 0$ such that

$$\begin{aligned} &-r < \alpha(t) \leqslant \beta(t) < r,\\ &f\big(t, \alpha(t), 0\big) + \varphi(0)\big(-r - \alpha(t)\big) < 0,\\ &f\big(t, \beta(t), 0\big) + \varphi(0)\big(r - \beta(t)\big) > 0. \end{aligned}$$

Claim 1. *Every solution u of* (1.20) *with* $\lambda \in [0,1]$ *is such that* $-r < u(t) < r$. Assume there exists t_0 such that $u(t_0) = \min_t u(t) \leqslant -r$. This leads to a contradiction since then

$$0 \leqslant u''(t_0) = \lambda\big[f\big(t_0, \alpha(t_0), 0\big) + \varphi(0)\big(u(t_0) - \alpha(t_0)\big)\big] + (1-\lambda)\varphi(0)u(t_0) < 0.$$

Similarly, we prove that $u(t) < r$.

Claim 2. *There exists $R > 0$ such that every solution u of* (1.20) *with* $\lambda \in [0,1]$ *satisfies* $\|u'\|_\infty < R$. The claim follows choosing $R > 0$ from Proposition 1.7, where $\bar{\varphi}(v) = (1 + 2r)\varphi(v)$, $k = 0$ and $\bar{a}$ is such that $u'(\bar{a}) = u'(\bar{a} + b - a) = 0$.

Claim 3. *Existence of solutions of* (1.20) *for* $\lambda = 1$. Let us define the operators

$$\begin{aligned} &L : \operatorname{Dom} L \subset \mathcal{C}^1\big([a,b]\big) \to \mathcal{C}\big([a,b]\big) : u \to u'' - u,\\ &N_\lambda : \mathcal{C}^1\big([a,b]\big) \to \mathcal{C}\big([a,b]\big) : u \to \lambda f\big(t, \gamma(t,u), u'\big) + \varphi\big(|u'|\big)\big(u - \lambda\gamma(t,u)\big) - u, \end{aligned}$$

where $\operatorname{Dom} L = \{u \in \mathcal{C}^2([a,b]) \mid u(a) = u(b),\ u'(a) = u'(b)\}$. Observe that L has a compact inverse. Hence, we can define the completely continuous operator

$$T_\lambda(u) = L^{-1}N_\lambda(u).$$

From degree theory, we have that

$$\deg(T_0, \Omega) = \deg(T_1, \Omega),$$

where

$$\Omega = \big\{u \in \mathcal{C}^1([a,b]) \mid \|u\|_\infty < r,\ \|u'\|_\infty < R\big\}.$$

Using the Odd Mapping theorem (see [69]), we compute that

$$\deg(T_0, \Omega) \neq 0$$

and the problem (1.20) with $\lambda = 1$ has a solution u.

Claim 4. *The solution u of* (1.20) *with $\lambda = 1$ is such that $\alpha \leqslant u \leqslant \beta$.* This claim follows as the corresponding argument in the proof of Theorem 1.10. As a consequence, u satisfies (1.7). □

The following theorem considers the case of L^p-Carathéodory nonlinearities.

THEOREM 1.12. *Let α and $\beta \in \mathcal{C}([a,b])$ be lower and upper solutions of* (1.7) *such that $\alpha \leqslant \beta$. Define $A \subset [a,b]$* (*respectively $B \subset [a,b]$*) *to be the set of points where α* (*respectively β*) *is differentiable. Let p, $q \in [1,\infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, $\psi \in L^p(a,b)$ and $\varphi \in \mathcal{C}(\mathbb{R}^+, \mathbb{R}_0^+)$ be such that* (1.15) *holds* (*with $\bar{\varphi} = \varphi$*). *Let E be defined from* (1.19) *and suppose $f : E \to \mathbb{R}$ is an L^p-Carathéodory function that satisfies the one-sided Nagumo condition* (1.17) (*with $\bar{\varphi} = \varphi$ and $\bar{\psi} = \psi$*). *Assume there exists $N \in L^1(a,b)$, $N > 0$, such that for a.e. $t \in A$* (*respectively for a.e. $t \in B$*)

$$f\big(t,\alpha(t),\alpha'(t)\big) \geqslant -N(t) \quad \big(\text{respectively } f\big(t,\beta(t),\beta'(t)\big) \leqslant N(t)\big). \tag{1.21}$$

Then the problem (1.7) *has at least one solution $u \in W^{2,p}(a,b)$ such that for all $t \in [a,b]$*

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

PROOF. The proof proceeds in several steps.

Step 1. *The modified problem.* Let $R > 0$ be large enough so that

$$\int_0^R \frac{s^{1/q}}{\varphi(s)}\,\mathrm{d}s > \|\psi\|_{L^p}\big(\max_t \beta(t) - \min_t \alpha(t)\big)^{1/q}.$$

Increasing N if necessary, we can assume $N(t) \geqslant |f(t,u,v)|$ if $t \in [a,b]$, $\alpha(t) \leqslant u \leqslant \beta(t)$ and $|v| \leqslant R$. Define then

$$\begin{aligned}
\bar{f}(t,u,v) &= \max\big\{\min\big\{f\big(t,\gamma(t,u),v\big), N(t)\big\}, -N(t)\big\},\\
\omega_1(t,\delta) &= \chi_A(t)\max_{|v|\leqslant\delta}\big|\bar{f}\big(t,\alpha(t),\alpha'(t)+v\big) - \bar{f}\big(t,\alpha(t),\alpha'(t)\big)\big|,\\
\omega_2(t,\delta) &= \chi_B(t)\max_{|v|\leqslant\delta}\big|\bar{f}\big(t,\beta(t),\beta'(t)+v\big) - \bar{f}\big(t,\beta(t),\beta'(t)\big)\big|,
\end{aligned}$$

where γ is defined from (1.4), χ_A and χ_B are the characteristic functions of the sets A and B. It is clear that ω_i are L^1-Carathéodory functions, nondecreasing in δ, such that $\omega_i(t,0) = 0$ and $|\omega_i(t,\delta)| \leqslant 2N(t)$.

We consider now the modified problem

$$u'' - u = \bar{f}(t,u,u') - \omega(t,u), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{1.22}$$

where

$$\omega(t,u) = \begin{cases} \beta(t) - \omega_2\big(t,u-\beta(t)\big), & \text{if } u > \beta(t),\\ u, & \text{if } \alpha(t) \leqslant u \leqslant \beta(t),\\ \alpha(t) + \omega_1\big(t,\alpha(t)-u\big), & \text{if } u < \alpha(t). \end{cases}$$

Step 2. *Existence of a solution of* (1.22). This claim follows from Schauder's theorem.

Step 3. *The solution u of* (1.22) *is such that* $\alpha \leqslant u \leqslant \beta$. Let us assume on the contrary that for some $t_0 \in \mathbb{R}$

$$\min_t \bigl(u(t) - \alpha(t)\bigr) = u(t_0) - \alpha(t_0) < 0.$$

Then, as in Theorem 1.1, there exists an open interval I_0 with $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$ and, for a.e. $t \in I_0$

$$\alpha''(t) \geqslant f\bigl(t, \alpha(t), \alpha'(t)\bigr).$$

Further $u'(t_0) - \alpha'(t_0) = 0$ and for $t \geqslant t_0$ near enough t_0

$$\bigl|u'(t) - \alpha'(t)\bigr| \leqslant \alpha(t) - u(t).$$

As ω_1 is nondecreasing and $\bar{f}(t, \alpha(t), \alpha'(t)) \leqslant f(t, \alpha(t), \alpha'(t))$,

$$\begin{aligned} u'(t) - \alpha'(t) &= \int_{t_0}^{t} \bigl(u''(s) - \alpha''(s)\bigr)\, \mathrm{d}s \\ &\leqslant \int_{t_0}^{t} \bigl[\bar{f}\bigl(s, \alpha(s), u'(s)\bigr) - \bar{f}\bigl(s, \alpha(s), \alpha'(s)\bigr) \\ &\qquad + u(s) - \alpha(s) - \omega_1\bigl(s, \alpha(s) - u(s)\bigr)\bigr]\, \mathrm{d}s < 0. \end{aligned}$$

This proves $u(t_0) - \alpha(t_0)$ is not a minimum of $u - \alpha$ which is a contradiction.

A similar argument holds to prove $u \leqslant \beta$.

Step 4. *The solution u of* (1.22) *is such that* $\|u'\|_\infty < R$. Observe that, for all $(t, u, v) \in E$,

$$\max\bigl\{\min\bigl\{f(t, u, v), N(t)\bigr\}, -N(t)\bigr\} \leqslant \psi(t)\varphi\bigl(|v|\bigr).$$

From Proposition 1.9, every solution $u \in [\alpha, \beta]$ of (1.22) satisfies

$$\|u'\|_\infty < R.$$

Hence $|f(t, u(t), u'(t))| \leqslant N(t)$ and the function u is a solution of (1.7). ☐

REMARK. Notice that the condition (1.21) is satisfied in case f does not depend on u' or if α, $\beta \in W^{1,\infty}(a, b)$.

EXAMPLE 1.8. Existence of a solution to problem

$$u'' = \frac{1}{\sqrt{t}}|u'|^a + u + t, \quad u(0) = u(1),\ u'(0) = u'(1),$$

where $1 \leqslant a < 3/2$, follows from Theorem 1.12. Notice that here Theorem 1.11 does not apply.

1.4. *Derivative dependent Dirichlet problem*

Dirichlet boundary value problems

$$u'' = f(t, u, u'), \quad u(a) = 0, \ u(b) = 0, \tag{1.23}$$

can be studied in a similar way. To this end, we adapt accordingly the definitions of lower and upper solutions.

DEFINITIONS 1.3. A function $\alpha \in \mathcal{C}([a,b])$ is a *lower solution* of (1.23) if

(a) for any $t_0 \in]a,b[$, either $D_-\alpha(t_0) < D^+\alpha(t_0)$, or there exists an open interval $I_0 \subset]a,b[$ such that $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$ and, for a.e. $t \in I_0$,

$$\alpha''(t) \geqslant f\bigl(t, \alpha(t), \alpha'(t)\bigr);$$

(b) $\alpha(a) \leqslant 0, \alpha(b) \leqslant 0$.

A function $\beta \in \mathcal{C}([a,b])$ is an *upper solution* of (1.23) if

(a) for any $t_0 \in]a,b[$, either $D^-\beta(t_0) > D_+\beta(t_0)$, or there exists an open interval $I_0 \subset]a,b[$ such that $t_0 \in I_0$, $\beta \in W^{2,1}(I_0)$ and, for a.e. $t \in I_0$,

$$\beta''(t) \leqslant f\bigl(t, \beta(t), \beta'(t)\bigr);$$

(b) $\beta(a) \geqslant 0$, $\beta(b) \geqslant 0$.

A typical result is then the following.

THEOREM 1.13. *Assume α and $\beta \in \mathcal{C}([a,b])$ are lower and upper solutions of problem* (1.23) *such that $\alpha \leqslant \beta$. Define $A \subset [a,b]$* (*respectively $B \subset [a,b]$*) *to be the set of points where α* (*respectively β*) *is derivable. Let p, $q \in [1,\infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, $\varphi \in \mathcal{C}(\mathbb{R}^+, \mathbb{R}_0^+)$ and $\psi \in L^p(a,b)$ be such that* (1.15) *holds* (*with $\bar{\varphi} = \varphi$*). *Let E be defined in* (1.19) *and suppose $f : E \to \mathbb{R}$ is an L^p-Carathéodory function that satisfies the Nagumo condition* (1.16) (*with $\bar{\varphi} = \varphi$ and $\bar{\psi} = \psi$*). *Assume there exists $N \in L^1(a,b)$, $N > 0$ such that, for a.e. $t \in A$* (*respectively for a.e. $t \in B$*),

$$f\bigl(t, \alpha(t), \alpha'(t)\bigr) \geqslant -N(t) \quad \bigl(\textit{respectively } f\bigl(t, \beta(t), \beta'(t)\bigr) \leqslant N(t)\bigr).$$

Then the problem (1.23) *has at least one solution $u \in W^{2,p}(a,b)$ such that for all $t \in [a,b]$*

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

PROOF. The proof of this result follows the lines of the proof of Theorem 1.12 and is left to the reader as an exercise. □

Notice that a similar result using one-sided Nagumo conditions can be worked out provided some a priori bound on the derivative of solutions is known at the points a and b (see Proposition 1.7). These a priori bounds are valid, for example, in case the lower or the upper solution satisfies the boundary conditions.

1.5. *A derivative independent Dirichlet problem*

Dirichlet boundary value problems

$$u'' = f(t, u), \quad u(a) = 0,\ u(b) = 0, \tag{1.24}$$

can be studied for more general nonlinearities than L^1-Carathéodory functions. For example, the problem

$$t(1-t)u'' = 1, \quad u(0) = 0,\ u(1) = 0,$$

has the solution $u(t) = t \ln t + (1-t)\ln(1-t)$ which is in $\mathcal{C}_0([0,1]) \cap W^{2,1}_{loc}(]0,1[)$ but not even in $\mathcal{C}^1([0,1])$. This generalizes to the linear problem

$$u'' = h(t), \quad u(a) = 0,\ u(b) = 0. \tag{1.25}$$

In case

$$h \in \mathcal{A} := \bigl\{h \mid (s-a)(b-s)h(s) \in L^1(a,b)\bigr\},$$

problem (1.25) has one and only one solution in

$$W^{2,\mathcal{A}}(a,b) := \bigl\{u \in W^{1,1}(a,b) \mid u'' \in \mathcal{A}\bigr\} \subset \mathcal{C}\bigl([a,b]\bigr) \cap \mathcal{C}^1\bigl(]a,b[\bigr),$$

which reads

$$u(t) = \int_a^b G(t,s)h(s)\,\mathrm{d}s,$$

where $G(t,s)$ is the corresponding Green's function. Further, we have

$$\|u\|_\infty \leqslant \frac{1}{b-a}\|h\|_{\mathcal{A}},$$

where $\|h\|_{\mathcal{A}} = \int_a^b (s-a)(b-s)\bigl|h(s)\bigr|\,\mathrm{d}s$ (see [53,33]).

In order to deal with nonlinear problems, we consider *$\mathcal{A}$-Carathéodory functions* $f : E \subset [a,b] \times \mathbb{R} \to \mathbb{R}$. These are Carathéodory functions such that for any $r > 0$

there exists $h \in \mathcal{A}$ so that for a.e. $t \in [a,b]$ and all $u \in \mathbb{R}$ with $(t,u) \in E$ and $|u| \leqslant r$, $|f(t,u)| \leqslant h(t)$.

Notice that if $g : \mathbb{R} \to \mathbb{R}$ is continuous, the function

$$f(t,u) = \frac{g(u)}{t(1-t)}$$

is an $\mathcal{A}$-Carathéodory function. Observe also that $L^1(a,b) \subset \mathcal{A}$, so that this definition generalizes the classical L^1-Carathéodory conditions on f.

We can now state the main result for (1.24).

THEOREM 1.14. *Assume that* α *and* β *are lower and upper solutions of* (1.24) *such that* $\alpha \leqslant \beta$. *Let* E *be defined from* (1.2) *and* $f : E \to \mathbb{R}$ *be an* $\mathcal{A}$*-Carathéodory function. Then the problem* (1.24) *has at least one solution* $u \in W^{2,\mathcal{A}}(a,b)$ *such that for all* $t \in [a,b]$

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

PROOF. We consider the modified problem

$$u'' - u = f\big(t, \gamma(t,u)\big) - \gamma(t,u), \quad u(a) = 0,\ u(b) = 0, \tag{1.26}$$

where γ is defined by (1.4).

Claim 1. *The problem* (1.26) *has at least one solution* $u \in W^{2,\mathcal{A}}(a,b)$. Define the operator $T : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ given by

$$(Tu)(t) = \int_a^b G(t,s)\big[f\big(s, \gamma\big(s, u(s)\big)\big) - \gamma\big(s, u(s)\big)\big]\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function associated with

$$u'' - u = f(t), \quad u(a) = 0,\ u(b) = 0.$$

We can prove that T is completely continuous and bounded (see [53,33]). Hence by Schauder Fixed Point theorem, T has a fixed point u which is a solution of (1.26) in $W^{2,\mathcal{A}}(a,b)$.

Claim 2. *Any solution* u *of* (1.26) *satisfies* $\alpha \leqslant u \leqslant \beta$. Observe first that $\alpha(a) \leqslant u(a) \leqslant \beta(a)$ and $\alpha(b) \leqslant u(b) \leqslant \beta(b)$. Next, we argue as in Theorem 1.1 to obtain the result.

Conclusion. From Claim 2, the function u, solution of (1.26), solves (1.24). □

In the definition of $\mathcal{A}$-Carathéodory functions, the condition $h \in \mathcal{A}$ is used to insure

$$\left\| \int_a^b G(\cdot, s) h(s)\,\mathrm{d}s \right\|_\infty < +\infty \quad \text{and} \quad \int_a^b \left| \frac{\partial G}{\partial t}(\cdot, s) \right| h(s)\,\mathrm{d}s \in L^1(a,b).$$

EXAMPLE 1.9. Consider the boundary value problem

$$u'' + |u|^{1/2} - \frac{1}{t} = 0, \quad u(0) = 0,\ u(\pi) = 0.$$

It is easy to see that $\beta(t) = 0$ is an upper solution and $\alpha(t) = t \ln \frac{t}{\pi} - t$ is a lower solution. Hence we have a solution u such that for all $t \in [0, \pi]$

$$t \ln \frac{t}{\pi} - t \leqslant u(t) \leqslant 0.$$

Notice that, in this example, the function $f(t, u)$ is not L^1-Carathéodory.

If we want more regularity, we need more restrictive conditions on f. For example, we have the following theorem.

THEOREM 1.15. *Assume that α and β are lower and upper solutions of* (1.24) *such that $\alpha \leqslant \beta$. Let E be defined from* (1.2) *and let $f : E \to \mathbb{R}$ satisfy a Carathéodory condition. Assume that there exists a measurable function h such that $\int_a^b (s - a)h(s)\,\mathrm{d}s < \infty$ and for a.e. $t \in [a, b]$ and all $u \in \mathbb{R}$ with $(t, u) \in E$,*

$$\bigl|f(t, u)\bigr| \leqslant h(t).$$

Then the problem (1.24) *has at least one solution $u \in W^{2,\mathcal{A}}(a, b) \cap \mathcal{C}^1(]a, b])$ such that for all $t \in [a, b]$*

$$\alpha(t) \leqslant u(t) \leqslant \beta(t).$$

PROOF. Existence of a solution $u \in W^{2,\mathcal{A}}(a, b)$ follows from Theorem 1.14. Further

$$u(t) = \int_a^b G(t, s) f\bigl(s, u(s)\bigr)\,\mathrm{d}s,$$

where $G(t, s)$ is the Green's function corresponding to (1.25). It is now standard to see that $u \in \mathcal{C}^1(]a, b])$. □

1.6. *Historical and bibliographical notes*

The method of lower and upper solutions applied to boundary value problems is due to Scorza Dragoni in 1931. In a first paper [94], this author considers the boundary value problem

$$u'' = f(t, u, u'), \quad u(a) = A,\ u(b) = B,$$

where f is bounded for u in bounded sets. The basic assumption is the existence of functions α, $\beta \in \mathcal{C}^2([a,b])$ such that

$$\alpha'' = f(t, \alpha, \alpha'), \quad \alpha(a) \leqslant A,\ \alpha(b) \leqslant B,$$
$$\beta'' = f(t, \beta, \beta'), \quad \beta(a) \geqslant A,\ \beta(b) \geqslant B.$$

The same year [95], he improves his result assuming that the functions α and β satisfy differential inequalities. In these papers, the method of proof already uses an auxiliary problem, modified for $u \notin [\alpha(t), \beta(t)]$, whose solution is proved, from a maximum principle type argument, to be such that $\alpha \leqslant u \leqslant \beta$. The basic ideas are already set.

The introduction of angles in lower and upper solutions has been used by Picard [80] in 1893. This idea was rediscovered by Nagumo [74] in 1954 who worked with maximum of lower solutions. A first order condition, similar to $D_{-}\alpha(t_0) < D^{+}\alpha(t_0)$, was used by Knobloch [64] in 1963.

In 1938, Scorza Dragoni [96,97] already considers L^1-Carathéodory functions f. Such an assumption is also made by Epheser [40] in 1955. More recently, we can quote Jackson [58], Kiguradze [62] (see also [63]), Gudkov and Lepin [46], Habets and Laloy [47], Hess [55] (see also Stampacchia [99]), Fabry and Habets [41], Adje [1], Habets and Sanchez [51], De Coster and Habets [31,32]. These papers present a variety of definitions of lower and upper solutions which are strongly related and there is no obvious reason to choose one rather than the other. Our choice, Definitions 1.1 and 1.2 (see [29,31,32]), tends to be general enough for applications and simple enough to model the geometric intuition built into the concept.

Theorems 1.1 and 1.2 are variants of the basic results of the theory. This last result can be found in [35] for the parabolic problem.

Theorem 1.3 deals with extremal solutions. Existence of extremal solutions for the Cauchy problem associated with first order ODEs were already studied by Peano [77] in 1885 and Perron [78] in 1915. Using a monotonicity assumption, Satō [89] worked in 1954 elliptic PDEs in relation with lower and upper solutions. The monotonicity assumption was deleted in 1960 for parabolic problems by Mlak [71] and in 1961 for elliptic PDEs by Akô [2]. The proof used here is inspired by the recent paper of Angel Cid [22]. For a classical proof, we refer to Akô [2] or Schmitt [92] in case of $\mathcal{C}^2$-solutions. Such proofs use maxima of lower solutions and minima of upper ones as in Theorem 1.2.

As far as the structure of the solution set is concerned, i.e., Theorem 1.4, we refer to [90] and [2].

A priori bounds on the derivative of solutions were already worked out by Bernstein [11] in 1904. Nagumo [73], in 1937, generalized these ideas introducing the so-called Nagumo condition which is both simple and very general. This is why it has been widely used since then. Basically, it is our Proposition 1.6. In 1967, Kiguradze [61] (see also Epheser [40] in 1955) observed that, for some boundary value problems, the Nagumo condition can be restricted to be one-sided conditions. Proposition 1.7 introduced such a condition. In the same paper, Kiguradze extended the Nagumo condition so as to deal with $W^{2,1}$-solutions as we did in Propositions 1.8 and 1.9. It seems that we have to wait for Knobloch [64] in 1963 to consider the periodic problem. In 1954, Nagumo [74] pointed out that the existence of well ordered lower and upper solutions is not sufficient to ensure existence of solutions of a Dirichlet problem. Example 1.5 presented here for the periodic problem is adapted from Habets and Pouso [50].

Theorem 1.10 extends a result of Habets and Torres [52] avoiding the assumption α, $\beta \in W^{1,\infty}(a,b)$. Theorem 1.11 is due to Mawhin [70] and Theorem 1.12 is the counterpart for the periodic problem of a result of De Coster [28].

Rosenblatt [88] already noticed in 1933 that Dirichlet problems can be studied for more singular nonlinearities than L^p-Carathéodory functions. In 1953, Prodi [86] used lower and upper solutions for such singular problems. Theorems 1.14 and 1.15 follow Habets and Zanolin [53] and [54], where the Dirichlet boundary value problem is investigated with another definition of lower and upper solutions.

2. Relation with degree theory

2.1. *The periodic problem*

In order to use degree theory, we need to reinforce the notion of lower and upper solution.

DEFINITIONS 2.1. A lower solution α of

$$u'' = f(t,u,u'), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{2.1}$$

is said to be a *strict lower solution* if every solution u of (2.1) with $u \geqslant \alpha$ is such that $u(t) > \alpha(t)$ on $[a,b]$.

Similarly, an upper solution β of (2.1) is said to be a *strict upper solution* if every solution u of (2.1) with $u \leqslant \beta$ is such that $u(t) < \beta(t)$ on $[a,b]$.

The classical way to obtain lower and upper solutions in the case of a continuous function f is described in the following propositions.

PROPOSITION 2.1. *Let $f : [a,b] \times \mathbb{R}^2 \to \mathbb{R}$ be continuous and $\alpha \in C^2([a,b])$ be such that*
(a) *for all $t \in [a,b]$, $\alpha''(t) > f(t,\alpha(t),\alpha'(t))$;*
(b) $\alpha(a) = \alpha(b)$, $\alpha'(a) \geqslant \alpha'(b)$.
Then α is a strict lower solution of (2.1).

PROOF. Let u be a solution of (2.1) such that $u \geqslant \alpha$ and assume, by contradiction, that

$$\min_t \big(u(t) - \alpha(t)\big) = u(t_0) - \alpha(t_0) = 0.$$

We have $u'(t_0) - \alpha'(t_0) = 0$; in case $t_0 = a$ or b, this follows from assumption (b). Hence, we obtain the contradiction

$$0 \leqslant u''(t_0) - \alpha''(t_0) = f\big(t_0, \alpha(t_0), \alpha'(t_0)\big) - \alpha''(t_0) < 0.$$

□

Using the same argument we obtain the corresponding result for upper solutions.

PROPOSITION 2.2. *Let* $f : [a,b] \times \mathbb{R}^2 \to \mathbb{R}$ *be continuous and* $\beta \in \mathcal{C}^2([a,b])$ *be such that*

(a) *for all* $t \in [a,b]$, $\beta''(t) < f(t, \beta(t), \beta'(t))$;

(b) $\beta(a) = \beta(b)$, $\beta'(a) \leqslant \beta'(b)$.

Then β *is a strict upper solution of* (2.1).

If f is not continuous but L^1-Carathéodory, these last results do not hold anymore. Even the stronger condition

$$\text{for a.e. } t \in [a,b], \quad \alpha''(t) \geqslant f\big(t, \alpha(t), \alpha'(t)\big) + 1$$

does not prevent solutions u of (2.1) to be tangent to the curve $u = \alpha(t)$ from above. This is, for example, the case for the bounded function

$$f(t,u) := \begin{cases} -1, & u \leqslant -1, \\ -\frac{u^2+\sin t}{1+\sin t}, & -1 < u \leqslant \sin t, \\ -\sin t, & \sin t < u, \end{cases}$$

if we consider $\alpha(t) \equiv -1$, $u(t) \equiv \sin t$, $a = 0$ and $b = 2\pi$. This remark motivates the following proposition.

PROPOSITION 2.3. *Let* $f : [a,b] \times \mathbb{R}^2 \to \mathbb{R}$ *be an* L^1*-Carathéodory function. Let* $\alpha \in \mathcal{C}([a,b])$ *be such that* $\alpha(a) = \alpha(b)$ *and consider its periodic extension on* $\mathbb{R}$. *Assume that* α *is not a solution of* (2.1) *and for any* $t_0 \in \mathbb{R}$, *either*

(a) $D_-\alpha(t_0) < D^+\alpha(t_0)$ *or*

(b) *there exist an open interval* I_0 *and an* $\epsilon_0 > 0$ *such that* $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$ *and for a.e.* $t \in I_0$, *all* $u \in [\alpha(t), \alpha(t) + \epsilon_0]$ *and all* $v \in [\alpha'(t) - \epsilon_0, \alpha'(t) + \epsilon_0]$,

$$\alpha''(t) \geqslant f(t, u, v).$$

Then α *is a strict lower solution of* (2.1).

PROOF. The function α is a lower solution since clearly it satisfies Definition 1.2. Let u be a solution of (2.1) such that $u \geqslant \alpha$. As α is not a solution, there exists t^* such that $u(t^*) > \alpha(t^*)$. Extend u and α by periodicity and assume, by contradiction, that

$$t_0 = \inf\big\{t > t^* \mid u(t) = \alpha(t)\big\}$$

exists. As $\alpha - u$ is maximum at t_0, we have $D_-\alpha(t_0) - u'(t_0) \geqslant D^+\alpha(t_0) - u'(t_0)$. Hence, assumption (b) applies. This implies that $\alpha'(t_0) - u'(t_0) = 0$ and there exist I_0 and $\epsilon_0 > 0$ according to (b). It follows we can choose $t_1 \in I_0$ with $t_1 < t_0$ such that $\alpha'(t_1) - u'(t_1) > 0$ and for every $t \in \,]t_1, t_0[$

$$u(t) \leqslant \alpha(t) + \epsilon_0, \quad \big|\alpha'(t) - u'(t)\big| < \epsilon_0.$$

Hence, for almost every $t \in]t_1, t_0[$, we can write

$$\alpha''(t) \geqslant f\big(t, u(t), u'(t)\big),$$

which leads to the contradiction

$$0 > (\alpha' - u')(t_0) - (\alpha' - u')(t_1) = \int_{t_1}^{t_0} \big[\alpha''(t) - f\big(t, u(t), u'(t)\big)\big]\, \mathrm{d}t \geqslant 0.$$

□

In the same way we can prove the following result on strict upper solutions.

PROPOSITION 2.4. *Let $f : [a,b] \times \mathbb{R}^2 \to \mathbb{R}$ be an L^1-Carathéodory function. Let $\beta \in \mathcal{C}([a,b])$ be such that $\beta(a) = \beta(b)$ and consider its periodic extension on $\mathbb{R}$. Assume that β is not a solution of* (2.1) *and for any $t_0 \in \mathbb{R}$, either*

(a) *$D^-\beta(t_0) > D_+\beta(t_0)$ or*

(b) *there exist an open interval I_0 and an $\epsilon_0 > 0$ such that $t_0 \in I_0$, $\beta \in W^{2,1}(I_0)$ and for a.e. $t \in I_0$, all $u \in [\beta(t) - \epsilon_0, \beta(t)]$ and all $v \in [\beta'(t) - \epsilon_0, \beta'(t) + \epsilon_0]$,*

$$\beta''(t) \leqslant f(t, u, v).$$

Then β is a strict upper solution of (2.1).

REMARK. Notice that Propositions 2.3 and 2.4 apply with nonstrict inequalities $\alpha''(t) \geqslant f(t, \alpha(t), \alpha'(t))$ and $\beta''(t) \leqslant f(t, \beta(t), \beta'(t))$. Hence, even if f is continuous, these propositions generalize Propositions 2.1 and 2.2.

The relation between degree theory and lower and upper solutions is described in the following result which completes Theorem 1.12.

THEOREM 2.5. *Let α and $\beta \in \mathcal{C}([a,b])$ be strict lower and upper solutions of* (2.1) *such that $\alpha \leqslant \beta$. Define $A \subset [a,b]$* (*respectively $B \subset [a,b]$*) *to be the set of points where α* (*respectively β*) *is differentiable. Let p, $q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, $\psi \in L^p(a,b)$ and $\varphi \in \mathcal{C}(\mathbb{R}^+, \mathbb{R}_0^+)$ be such that* (1.15) *holds* (*with $\bar{\varphi} = \varphi$*). *Let E be defined from* (1.19) *and suppose $f : E \to \mathbb{R}$ is an L^p-Carathéodory function that satisfies the one-sided Nagumo condition* (1.17) (*with $\bar{\varphi} = \varphi$ and $\bar{\psi} = \psi$*). *Assume there exists $N \in L^1(a,b)$, $N > 0$, such that for a.e. $t \in A$* (*respectively for a.e. $t \in B$*) (1.21) *is satisfied. Then*

$$\deg(I - T, \Omega) = 1, \tag{2.2}$$

where $T : \mathcal{C}^1([a,b]) \to \mathcal{C}^1([a,b])$ is defined by

$$(Tu)(t) := \int_a^b G(t,s)\big[f\big(s, u(s), u'(s)\big) - u(s)\big]\, \mathrm{d}s, \tag{2.3}$$

$G(t,s)$ is the Green's function corresponding to (1.5) *and Ω is given by*

$$\Omega = \{u \in \mathcal{C}^1([a,b]) \mid \forall t \in [a,b],\ \alpha(t) < u(t) < \beta(t),\ |u'(t)| < R\}, \tag{2.4}$$

with $R > 0$ large enough. In particular, the problem (2.1) *has at least one solution $u \in W^{2,p}(a,b)$ such that for all $t \in [a,b]$*

$$\alpha(t) < u(t) < \beta(t).$$

PROOF. Define $R > 0$ as in the proof of Theorem 1.12 and consider the modified problem (1.22). This problem is equivalent to the fixed point problem

$$u = \overline{T}u,$$

where $\overline{T} : \mathcal{C}^1([a,b]) \to \mathcal{C}^1([a,b])$ is defined by

$$(\overline{T}u)(t) = \int_a^b G(t,s)\big[\bar{f}\big(s, \gamma(s,u(s)), u'(s)\big) - \omega\big(s,u(s)\big)\big]\,\mathrm{d}s.$$

Observe that $\overline{T}$ is completely continuous. Further there exists $\overline{R}$ large enough so that $\Omega \subset B(0,\overline{R})$ and $\overline{T}(\mathcal{C}^1([a,b])) \subset B(0,\overline{R})$. Hence we have, by the properties of the degree,

$$\deg\big(I - \overline{T}, B(0,\overline{R})\big) = 1.$$

We know that every fixed point of $\overline{T}$ is a solution of (1.22). Arguing as in the proof of Theorem 1.12, we see that $\alpha \leqslant u \leqslant \beta$ and $\|u'\|_\infty < R$. As α and β are strict, $\alpha < u < \beta$. Hence, every fixed point of $\overline{T}$ is in Ω and by the excision property we obtain

$$\deg(I - T, \Omega) = \deg(I - \overline{T}, \Omega) = \deg\big(I - \overline{T}, B(0,\overline{R})\big) = 1.$$

Existence of a solution u such that for all $t \in [a,b]$,

$$\alpha(t) < u(t) < \beta(t)$$

follows now from the properties of the degree. □

In case we consider a derivative independent problem

$$u'' = f(t,u), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{2.5}$$

we can simplify this theorem defining T on $\mathcal{C}([a,b])$ rather than on $\mathcal{C}^1([a,b])$.

THEOREM 2.6. *Let α and $\beta \in \mathcal{C}([a,b])$ be strict lower and upper solutions of* (2.5) *such that $\alpha \leqslant \beta$. Let E be defined from* (1.2) *and suppose $f : E \to \mathbb{R}$ is an L^1-Carathéodory function. Then*

$$\deg(I - T, \Omega) = 1,$$

where $T : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ *is defined by*

$$(Tu)(t) := \int_a^b G(t,s)\big[f\big(s,u(s)\big) - u(s)\big]\,\mathrm{d}s,$$

$G(t,s)$ *is the Green's function corresponding to* (1.5) *and* Ω *is given by*

$$\Omega = \big\{u \in \mathcal{C}\big([a,b]\big) \mid \forall t \in [a,b],\ \alpha(t) < u(t) < \beta(t)\big\}.$$

In particular, the problem (2.5) *has at least one solution* $u \in W^{2,1}(a,b)$ *such that for all* $t \in [a,b]$

$$\alpha(t) < u(t) < \beta(t).$$

The proof of this theorem repeats the argument used in the proof of Theorem 2.5.

These theorems can be used to obtain multiplicity results. We present here such a theorem where, for simplicity, we assume the nonlinearity not to depend on the derivative. Such a restriction is by no means essential.

THEOREM 2.7 (The Three Solutions theorem). *Let* α_1, β_1 *and* α_2, $\beta_2 \in \mathcal{C}([a,b])$ *be two pairs of lower and upper solutions of* (2.5) *such that for all* $t \in [a,b]$,

$$\alpha_1(t) \leqslant \beta_1(t), \qquad \alpha_2(t) \leqslant \beta_2(t), \qquad \alpha_1(t) \leqslant \beta_2(t)$$

and for some $t_0 \in [a,b]$,

$$\alpha_2(t_0) > \beta_1(t_0).$$

Assume further β_1 *and* α_2 *are strict upper and lower solutions. Let* E *be defined from* (1.2) (*with* $\alpha = \min\{\alpha_1, \alpha_2\}$ *and* $\beta = \max\{\beta_1, \beta_2\}$) *and suppose* $f : E \to \mathbb{R}$ *is an* L^1-*Carathéodory function. Then, the problem* (2.5) *has at least three solutions* $u_1, u_2, u_3 \in W^{2,1}(a,b)$ *such that for all* $t \in [a,b]$,

$$\alpha_1(t) \leqslant u_1(t) < \beta_1(t), \qquad \alpha_2(t) < u_2(t) \leqslant \beta_2(t), \qquad u_1(t) \leqslant u_3(t) \leqslant u_2(t)$$

and for some t_1, $t_2 \in [a,b]$,

$$u_3(t_1) > \beta_1(t_1), \qquad u_3(t_2) < \alpha_2(t_2).$$

PROOF. Consider the modified problem

$$u'' - u = f\big(t, \gamma(t,u)\big) - \gamma(t,u), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{2.6}$$

where $\gamma(t,u) = \max\{\min\{\beta_2(t), u\}, \alpha_1(t)\}$. Let us choose k so that $\beta_1 \leqslant \beta_2 + k$ and $\alpha_1 - k \leqslant \alpha_2$, and define $T : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ by

$$(Tu)(t) = \int_a^b G(t,s)\big[f\big(s, \gamma\big(s,u(s)\big)\big) - \gamma\big(s,u(s)\big)\big]\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function corresponding to (1.5).

Step 1. *Computation of* $d(I-T,\Omega_{1,1})$, *where*

$$\Omega_{1,1}=\big\{u\in\mathcal{C}\big([a,b]\big)\mid\forall t\in[a,b],\ \alpha_1(t)-k<u(t)<\beta_1(t)\big\}.$$

Define the alternative modified problem

$$u''-u=\bar{f}(t,u)-\bar{\gamma}(t,u),\quad u(a)=u(b),\ u'(a)=u'(b),$$

where $\bar{\gamma}(t,u)=\max\{\min\{\beta_1(t),\beta_2(t),u\},\alpha_1(t)\}$ and

$$\bar{f}(t,u)=\begin{cases}f\big(t,\alpha_1(t)\big), & \text{if } u\leqslant\alpha_1(t),\\ f(t,u), & \text{if } \alpha_1(t)<u<\min\big\{\beta_1(t),\beta_2(t)\big\},\\ \max_{i=1,2}\big\{f\big(t,\min\big\{\beta_i(t),u\big\}\big)\big\}, & \text{if } \min\big\{\beta_1(t),\beta_2(t)\big\}\leqslant u.\end{cases}$$

Define next $\overline{T}:\mathcal{C}([a,b])\to\mathcal{C}([a,b])$ by

$$(\overline{T}u)(t)=\int_a^b G(t,s)\big[f\big(s,\bar{\gamma}\big(s,u(s)\big)\big)-\bar{\gamma}\big(s,u(s)\big)\big]\,\mathrm{d}s.$$

For any $\lambda\in[0,1]$, we consider then the homotopy $\mathcal{T}_\lambda=\lambda\overline{T}+(1-\lambda)T$.

Claim 1. *If* $\lambda\in[0,1]$ *and* u *is a fixed point of* $\mathcal{T}_\lambda$, *we have* $u\geqslant\alpha_1$. This follows from the usual maximum principle argument as in Claim 2 of the proof of Theorem 1.1.

Claim 2. *If* $\lambda\in[0,1]$ *and* u *is a fixed point of* $\mathcal{T}_\lambda$, *we have* $u\leqslant\beta_2$. Notice that u solves

$$\begin{aligned}&u''-u=\lambda\big[\bar{f}(t,u)-\bar{\gamma}(t,u)\big]+(1-\lambda)\big[f\big(t,\gamma(t,u)\big)-\gamma(t,u)\big],\\&u(a)=u(b),\ u'(a)=u'(b).\end{aligned}$$

Assume now that for some $t_0\in[a,b]$, $u(t_0)-\beta_2(t_0)>0$. Hence for t near enough t_0, we can write

$$u''(t)-u(t)\geqslant f\big(t,\beta_2(t)\big)-\beta_2(t)\geqslant\beta_2''(t)-\beta_2(t),$$

i.e.,

$$u''(t)-\beta_2''(t)\geqslant u(t)-\beta_2(t)>0,$$

which contradicts the fact that t_0 maximizes $u-\beta_2$.

Claim 3. *If* $\lambda\in[0,1]$ *and* $u\in\overline{\Omega}_{1,1}$ *is a fixed point of* $\mathcal{T}_\lambda$, *we have* $u<\beta_1$. Assume there exists $t_0\in[a,b]$ such that $u(t_0)=\beta_1(t_0)$. We deduce from Claims 1 and 2 that $\alpha_1(t)\leqslant u(t)\leqslant\beta_2(t)$ for all $t\in[a,b]$ so that u solves (2.5). As further β_1 is a strict upper solution, the claim follows.

Claim 4. $\deg(I-T,\Omega_{1,1})=1$. Observe that $\overline{T}$ is completely continuous. Further there exists R large enough so that $\Omega_{1,1}\subset B(0,R)$ and $\overline{T}(\mathcal{C}^1([a,b]))\subset B(0,R)$. On the other hand, we know from the usual maximum principle argument (see the proof of Theorem 1.2) that fixed points u of $\overline{T}$ are such that $\alpha_1 \leqslant u \leqslant \min\{\beta_1(t), \beta_2(t)\}$. Hence we deduce from Claim 3 that $u < \beta_1$. Now we can deduce from the properties of the degree,

$$\deg(I - \overline{T}, \Omega_{1,1}) = \deg\bigl(I - \overline{T}, B(0, R)\bigr) = 1$$

and using the above claims

$$\deg(I - T, \Omega_{1,1}) = \deg(I - \mathcal{T}_\lambda, \Omega_{1,1}) = \deg(I - \overline{T}, \Omega_{1,1}) = 1.$$

Step 2. $\deg(I - T, \Omega_{2,2}) = 1$*, where*

$$\Omega_{2,2} = \bigl\{u \in \mathcal{C}\bigl([a,b]\bigr) \mid \forall t \in [a,b],\ \alpha_2(t) < u(t) < \beta_2(t) + k\bigr\}.$$

The proof of this result parallels the proof of Step 1.

Step 3. *There exist three solutions* $\bar{u}_i$ $(i = 1, 2, 3)$ *of* (2.6) *such that*

$$\alpha_1 - k < \bar{u}_1 < \beta_1, \qquad \alpha_2 < \bar{u}_2 < \beta_2 + k, \qquad \alpha_1 - k < \bar{u}_3 < \beta_2 + k$$

and there exist t_1*,* $t_2 \in [a, b]$ *with*

$$\bar{u}_3(t_1) > \beta_1(t_1), \qquad \bar{u}_3(t_2) < \alpha_2(t_2).$$

The two first solutions are obtained from the fact that

$$\deg(I - T, \Omega_{1,1}) = 1 \quad \text{and} \quad \deg(I - T, \Omega_{2,2}) = 1.$$

Define

$$\Omega_{1,2} = \bigl\{u \in \mathcal{C}\bigl([a,b]\bigr) \mid \forall t \in [a,b],\ \alpha_1(t) - k < u(t) < \beta_2(t) + k\bigr\}.$$

We have

$$\begin{aligned} 1 &= \deg(I - T, \Omega_{1,2}) \\ &= \deg(I - T, \Omega_{1,1}) + \deg(I - T, \Omega_{2,2}) + \deg\bigl(I - T, \Omega_{1,2} \backslash (\overline{\Omega}_{1,1} \cup \overline{\Omega}_{2,2})\bigr) \end{aligned}$$

which implies

$$\deg\bigl(I - T, \Omega_{1,2} \setminus (\overline{\Omega}_{1,1} \cup \overline{\Omega}_{2,2})\bigr) = -1$$

and the existence of $\bar{u}_3 \in \Omega_{1,2} \setminus (\overline{\Omega}_{1,1} \cup \overline{\Omega}_{2,2})$ follows.

Step 4. *There exist solutions* u_i $(i = 1, 2, 3)$*, of* (2.5) *such that*

$$\alpha_1 \leqslant u_1 < \beta_1, \qquad \alpha_2 < u_2 \leqslant \beta_2, \qquad u_1 \leqslant u_3 \leqslant u_2$$

and there exist t_1, $t_2 \in [a, b]$, with

$$u_3(t_1) > \beta_1(t_1), \qquad u_3(t_2) < \alpha_2(t_2).$$

We know that solutions u of (2.6) are such that

$$\alpha_1 \leqslant u \leqslant \beta_2,$$

i.e., they are solutions of (2.5). Next, from Theorem 1.3, we know there exist extremal solutions $u_{\min}$ and $u_{\max}$ of (2.5) in $[\alpha_1, \beta_2]$. The claim follows then with $u_1 = u_{\min}$, $u_2 = u_{\max}$ and $u_3 = \bar{u}_3$. □

Observe that in this theorem $u_1 \leqslant \min\{\beta_1, \beta_2\}$ and $u_2 \geqslant \max\{\alpha_1, \alpha_2\}$.

EXAMPLE 2.1. Consider the problem

$$u'' + \sin u = h(t), \quad u(0) = u(2\pi), \ u'(0) = u'(2\pi). \tag{2.7}$$

Let $h \in \mathcal{C}([0, 2\pi])$ and write $\bar{h} = \frac{1}{2\pi} \int_0^{2\pi} h(s)\, \mathrm{d}s$ and $\tilde{h} = h - \bar{h}$. Assume $\|\tilde{h}\|_{L^1} \leqslant 3$ and $|\bar{h}| < \cos(\frac{\pi}{6} \|\tilde{h}\|_{L^1})$. Let w be the solution of

$$w'' = \tilde{h}(t), \quad w(0) = w(2\pi), \ w'(0) = w'(2\pi), \ \int_0^{2\pi} w(s)\, \mathrm{d}s = 0,$$

and

$$\alpha_1 = -\frac{3\pi}{2} + w, \qquad \beta_1 = -\frac{\pi}{2} + w, \qquad \alpha_2 = \frac{\pi}{2} + w, \qquad \beta_2 = \frac{3\pi}{2} + w.$$

Using Theorem 2.7 and the estimate $\|w\|_\infty \leqslant \frac{\pi}{6} \|\tilde{h}\|_{L^1}$ (see [44] or [33]) we find three solutions of (2.7)

$$u_1 \in \,]\alpha_1, \beta_1[, \quad u_2 \in \,]\alpha_2, \beta_2[\quad \text{and} \quad u_3 \in \,]\alpha_1, \beta_2[,$$

with $u_3(t_1) \geqslant \beta_1(t_1)$ and $u_3(t_2) \leqslant \alpha_2(t_2)$ for some t_1 and $t_2 \in [0, 2\pi]$. Notice then that u_1 might be $u_2 - 2\pi$ but $u_3 \neq u_1 \bmod 2\pi$. Hence, this problem has at least two geometrically different solutions.

2.2. *The Dirichlet problem*

In this section, we consider the Dirichlet problem

$$u'' = f(t, u), \quad u(a) = 0, \ u(b) = 0, \tag{2.8}$$

where f is an L^p-Carathéodory function. To focus on the main ideas and avoid technical difficulties we restrict our analysis to derivative independent nonlinearities. Problem (2.8) is equivalent to the fixed point problem

$$u(t) = (Tu)(t) := \int_a^b G(t,s) f\big(s, u(s)\big)\,\mathrm{d}s, \tag{2.9}$$

where $G(t,s)$ is the Green's function corresponding to (1.25).

In this section, we consider the degree of $I - T$ for an open set Ω of functions u that lie between the lower and upper solutions α and β. If we allow α and β to satisfy the boundary conditions, the set

$$\Omega = \big\{u \in \mathcal{C}_0\big([a,b]\big) \mid \forall t \in \left]a,b\right[,\ \alpha(t) < u(t) < \beta(t)\big\},$$

might not be open in $\mathcal{C}_0([a,b])$. A way out is to impose some additional conditions on the functions u at these boundary points. To this end, for $u, v \in \mathcal{C}([a,b])$, we write $u \succ v$ or $v \prec u$ if there exists $\epsilon > 0$ such that for any $t \in [a,b]$

$$u(t) - v(t) \geqslant \epsilon e(t),$$

where $e(t) := \sin(\pi \frac{t-a}{b-a})$. We can then work with the space $\mathcal{C}_0^1([a,b])$ and use the set

$$\Omega = \big\{u \in \mathcal{C}_0^1\big([a,b]\big) \mid \alpha \prec u \prec \beta\big\}. \tag{2.10}$$

This set is open in $\mathcal{C}_0^1([a,b])$.

A possible alternative used by Amann (see [6]) is to work with the space

$$\mathcal{C}_e = \big\{u \in \mathcal{C}\big([a,b]\big) \mid \exists \lambda > 0,\ \forall t \in [a,b],\ \big|u(t)\big| \leqslant \lambda e(t)\big\}.$$

In our case, this approach does not seem to be simpler and as the solutions are anyhow in $\mathcal{C}_0^1([a,b])$, we choose to work in the more usual space $\mathcal{C}_0^1([a,b])$.

DEFINITIONS 2.2. A lower solution α of (2.8) is said to be a *strict lower solution* if every solution u of (2.8) with $\alpha \leqslant u$ is such that $\alpha \prec u$.

An upper solution β of (2.8) is said to be a *strict upper solution* if every solution u of (2.8) with $u \leqslant \beta$ is such that $u \prec \beta$.

A first result concerns lower and upper solutions which are $\mathcal{C}^2$.

PROPOSITION 2.8. *Let $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ be continuous and $\alpha \in \mathcal{C}([a,b]) \cap \mathcal{C}^2(]a,b[)$ be such that*

(a) *for all $t \in \left]a,b\right[$, $\alpha''(t) > f(t, \alpha(t))$;*
(b) *for $t_0 \in \{a,b\}$, either $\alpha(t_0) < 0$*
or $\alpha(t_0) = 0$, $\alpha \in \mathcal{C}^2(]a,b[\cup \{t_0\})$ and $\alpha''(t_0) > f(t_0, \alpha(t_0))$.

Then α is a strict lower solution of (2.8).

PROOF. From the assumptions, α is a lower solution of (2.8). Let then u be a solution of (2.8) such that $\alpha \leqslant u$ and assume, by contradiction, that for any $n \in \mathbb{N}_0$, there exists $t_n \in [a,b]$ such that

$$u(t_n) - \alpha(t_n) < \frac{1}{n} \sin\left(\pi \frac{t_n - a}{b - a}\right). \tag{2.11}$$

It follows there exists a subsequence of $(t_n)_n$ that converges to a point t_0 such that $u(t_0) = \alpha(t_0)$. If $t_0 \in]a,b[$, we have $u'(t_0) = \alpha'(t_0)$. On the other hand, if $t_0 = a$, we know that $\alpha \in \mathcal{C}^2([a,b[)$ and we deduce from (2.11) that

$$\frac{u(t_n) - u(a)}{t_n - a} \leqslant \frac{\alpha(t_n) + \frac{1}{n} \sin(\pi \frac{t_n - a}{b - a}) - \alpha(a)}{t_n - a}.$$

This implies $u'(a) \leqslant \alpha'(a)$. As further $u - \alpha$ is minimum at $t = a$, we also have $u'(a) \geqslant \alpha'(a)$. Hence, $u'(a) = \alpha'(a)$. A similar reasoning applies if $t_0 = b$ so that in all cases $u'(t_0) - \alpha'(t_0) = 0$. Finally, we obtain the contradiction $0 \leqslant u''(t_0) - \alpha''(t_0) = f(t_0, \alpha(t_0)) - \alpha''(t_0) < 0$. □

In a similar way, we can write

PROPOSITION 2.9. *Let* $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ *be continuous and* $\beta \in \mathcal{C}([a,b]) \cap \mathcal{C}^2(]a,b[)$ *be such that*

(a) *for all* $t \in]a,b[$, $\beta''(t) < f(t, \beta(t))$;
(b) *for* $t_0 \in \{a,b\}$, *either* $\beta(t_0) > 0$
or $\beta(t_0) = 0$, $\beta \in \mathcal{C}^2(]a,b[\cup \{t_0\})$ *and* $\beta''(t_0) < f(t_0, \beta(t_0))$.

Then β *is a strict upper solution of* (2.8).

Notice that an upper solution β such that

$$\beta''(t) < f(t, \beta(t)) \quad \text{on }]a,b[, \qquad \beta(a) \geqslant 0, \qquad \beta(b) \geqslant 0$$

is not necessarily strict. Consider for example the problem (2.8) defined on $[a,b] = [0, 2\pi]$ with

$$f(t,u) = \begin{cases} 0, & \text{if } u \leqslant 0, \\ 7u/t^2, & \text{if } 0 < u \leqslant t^3, \\ 7t, & \text{if } u > t^3. \end{cases}$$

The function $u(t) = 0$ is a solution and $\beta(t) = t^3$ is an upper solution that satisfies $\beta''(t) < f(t, \beta(t))$ on $]0, 2\pi]$, $\beta(0) = u(0)$ and $\beta'(0) = u'(0)$.

In the Carathéodory case, we can use the following propositions.

PROPOSITION 2.10. *Let* $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ *be an* L^1*-Carathéodory function. Assume that* $\alpha \in \mathcal{C}([a,b])$ *is not a solution of* (2.8) *and that*

(a) *for any* $t_0 \in]a,b[$, *either* $D_-\alpha(t_0) < D^+\alpha(t_0)$
or there exist an open interval $I_0 \subset [a,b]$ *and* $\epsilon_0 > 0$ *such that* $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$ *and for a.e.* $t \in I_0$ *and all* $u \in [\alpha(t), \alpha(t) + \epsilon_0 \sin(\pi \frac{t-a}{b-a})]$,

$$\alpha''(t) \geqslant f(t,u);$$

(b) *either* $\alpha(a) < 0$
or $\alpha(a) = 0$ *and there exists* $\epsilon_0 > 0$ *such that* $\alpha \in W^{2,1}(a, a+\epsilon_0)$ *and for a.e.* $t \in [a, a+\epsilon_0]$ *and all* $u \in [\alpha(t), \alpha(t) + \epsilon_0 \sin(\pi \frac{t-a}{b-a})]$,

$$\alpha''(t) \geqslant f(t,u);$$

(c) *either* $\alpha(b) < 0$
or $\alpha(b) = 0$ *and there exists* $\epsilon_0 > 0$ *such that* $\alpha \in W^{2,1}(b-\epsilon_0, b)$ *and for a.e.* $t \in [b-\epsilon_0, b]$ *and all* $u \in [\alpha(t), \alpha(t) + \epsilon_0 \sin(\pi \frac{t-a}{b-a})]$,

$$\alpha''(t) \geqslant f(t,u).$$

Then α *is a strict lower solution of* (2.8).

PROOF. Notice first that α satisfies Definition 1.3 and therefore is a lower solution.

Let u be a solution of (2.8) such that $u \geqslant \alpha$. Arguing by contradiction as in Proposition 2.8 there exists a sequence $(t_n)_n$ that satisfies (2.11) and converges to a point t_0 such that $u(t_0) = \alpha(t_0)$ and $u'(t_0) = \alpha'(t_0)$.

As α is not a solution, we can find t^* such that $u(t^*) > \alpha(t^*)$. Assume $t_0 < t^*$ and define $t_1 = \max\{t < t^* \mid u(t) = \alpha(t)\}$. Notice then that $u(t_1) = \alpha(t_1)$ and $u'(t_1) = \alpha'(t_1)$ and fix $\epsilon_0 > 0$ according to the assumptions. Next, for $t \geqslant t_1$ near enough t_1

$$u(t) \in \left[\alpha(t), \alpha(t) + \epsilon_0 \sin\left(\pi \frac{t-a}{b-a}\right)\right]$$

and we compute

$$u'(t) - \alpha'(t) = \int_{t_1}^{t} \bigl[f\bigl(s, u(s)\bigr) - \alpha''(s)\bigr]\,\mathrm{d}s \leqslant 0,$$

which contradicts the definition of t_1. A similar argument holds if $t_0 > t^*$. □

Strict upper solutions can be obtained from a similar proposition.

PROPOSITION 2.11. *Let* $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ *be an* L^1*-Carathéodory function. Assume that* $\beta \in \mathcal{C}([a,b])$ *is not a solution of* (2.8) *and that*

(a) *for any* $t_0 \in]a,b[$, *either* $D^-\beta(t_0) > D_+\beta(t_0)$ *or there exist an open interval* $I_0 \subset [a,b]$ *and* $\epsilon_0 > 0$ *such that* $t_0 \in I_0$, $\beta \in W^{2,1}(I_0)$ *and for a.e.* $t \in I_0$ *and all* $u \in [\beta(t) - \epsilon_0 \sin(\pi \frac{t-a}{b-a}), \beta(t)]$,

$$\beta''(t) \leqslant f(t,u);$$

(b) *either* $\beta(a) > 0$ *or* $\beta(a) = 0$ *and there exists* $\epsilon_0 > 0$ *such that* $\beta \in W^{2,1}(a, a+\epsilon_0)$ *and for a.e.* $t \in [a, a+\epsilon_0]$ *and all* $u \in [\beta(t) - \epsilon_0 \sin(\pi \frac{t-a}{b-a}), \beta(t)]$,

$$\beta''(t) \leqslant f(t,u);$$

(c) *either* $\beta(b) > 0$ *or* $\beta(b) = 0$ *and there exists* $\epsilon_0 > 0$ *such that* $\beta \in W^{2,1}(b-\epsilon_0, b)$ *and for a.e.* $t \in [b-\epsilon_0, b]$ *and all* $u \in [\beta(t) - \epsilon_0 \sin(\pi \frac{t-a}{b-a}), \beta(t)]$,

$$\beta''(t) \leqslant f(t,u).$$

Then β *is a strict upper solution of* (2.8).

We can study cases where f satisfies a one-sided Lipschitz condition in u.

PROPOSITION 2.12. *Let* $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ *be an* L^1*-Carathéodory function such that, for some* $k \in L^1(a,b;\mathbb{R}^+)$, *for a.e.* $t \in [a,b]$, *all* $u_1, u_2 \in \mathbb{R}$,

$$u_1 \leqslant u_2 \Rightarrow f(t,u_2) - f(t,u_1) \leqslant k(t)(u_2 - u_1);$$

Let α (*respectively* β) *be a lower* (*respectively upper*) *solution of* (2.8) *which is not a solution and assume*

(a) *either* $\alpha(a) < 0$ (*respectively* $\beta(a) > 0$) *or* $\alpha(a) = 0$ (*respectively* $\beta(a) = 0$) *and there exists an interval* $I_0 = [a,c[\subset [a,b]$ *such that* $\alpha \in W^{2,1}(I_0)$ (*respectively* $\beta \in W^{2,1}(I_0)$) *and, for a.e.* $t \in I_0$,

$$\alpha''(t) \geqslant f\big(t,\alpha(t)\big) \quad \big(\textit{respectively } \beta''(t) \leqslant f\big(t,\beta(t)\big)\big);$$

(b) *either* $\alpha(b) < 0$ (*respectively* $\beta(b) > 0$) *or* $\alpha(b) = 0$ (*respectively* $\beta(b) = 0$) *and there exists an interval* $I_0 =]c,b] \subset [a,b]$ *such that* $\alpha \in W^{2,1}(I_0)$ (*respectively* $\beta \in W^{2,1}(I_0)$) *and, for a.e.* $t \in I_0$,

$$\alpha''(t) \geqslant f\big(t,\alpha(t)\big) \quad \big(\textit{respectively } \beta''(t) \leqslant f\big(t,\beta(t)\big)\big).$$

Then α (*respectively* β) *is a strict lower solution* (*respectively a strict upper solution*) *of* (2.8).

PROOF. We prove the proposition for a lower solution α.

Let u be a solution of (2.8) such that $u \geqslant \alpha$ and assume by contradiction, as in Proposition 2.8, that for all $n \in \mathbb{N}_0$ there exists $t_n \in [a, b]$ such that (2.11) holds and that the sequence $(t_n)_n$ converges to some $t^* \in [a, b]$ which satisfies $u(t^*) = \alpha(t^*)$ and $u'(t^*) = \alpha'(t^*)$. Define $t_0 = \min\{t \in [a, b] \mid u(t) = \alpha(t) \text{ and } u'(t) = \alpha'(t)\}$. If $t_0 \neq a$, we can find an interval I_0 such that for a.e. $t \in I_0$

$$\alpha''(t) \geqslant f\bigl(t, \alpha(t)\bigr)$$

and $t_1 \in I_0$ with $t_1 < t_0$. Notice that $u(t_1) > \alpha(t_1)$. On $[t_1, t_0[$ the function $w = u - \alpha$ verifies

$$-w'' + k(t)w \geqslant -f\bigl(t, u(t)\bigr) + f\bigl(t, \alpha(t)\bigr) + k(t)\bigl(u(t) - \alpha(t)\bigr) \geqslant 0,$$
$$w(t_1) > 0,\ w(t_0) = 0,\ w'(t_0) = 0.$$

Define then v to be the solution of

$$v'' = k(t)v, \quad v(t_1) = 0,\ v'(t_1) = 1.$$

As v is positive on $]t_1, t_0[$, we have the contradiction

$$0 < w(t_1) = \bigl(w'(t)v(t) - w(t)v'(t)\bigr)\Big|_{t_1}^{t_0} = \int_{t_1}^{t_0} \bigl(w''(s) - k(s)w(s)\bigr)v(s)\,\mathrm{d}s \leqslant 0.$$

If $t_0 = a$ we define t^0 to be the maximum of the points $t \in [a, b]$ such that for all $s \in [a, t]$, $u(s) = \alpha(s)$ and $u'(s) = \alpha'(s)$. As α is not a solution, $t^0 < b$ and a similar argument applies to the right of t^0. □

As in the periodic case, we associate with lower and upper solutions some sets Ω which are such that if T is defined from (2.9), the degree of $I - T$ on such a set is 1.

THEOREM 2.13. *Let α and $\beta \in C([a, b])$ be strict lower and upper solutions of the problem* (2.8) *such that $\alpha \prec \beta$. Define E from* (1.2) *and assume $f : E \to \mathbb{R}$ is an L^1-Carathéodory function. Then, for $R > 0$ large enough,*

$$\deg(I - T, \Omega) = 1,$$

where $\Omega = \{u \in C_0^1([a, b]) \mid \alpha \prec u \prec \beta,\ \|u\|_{C^1} < R\}$ and $T : C_0^1([a, b]) \to C_0^1([a, b])$ is defined by (2.9). *In particular, the problem* (2.8) *has at least one solution $u \in W^{2,1}(a, b)$ such that*

$$\alpha \prec u \prec \beta.$$

PROOF. We consider the modified problem

$$u'' = f\bigl(t, \gamma(t,u)\bigr), \quad u(a) = 0,\ u(b) = 0, \tag{2.12}$$

with $\gamma(t,u)$ defined from (1.4). This problem is equivalent to the fixed point problem

$$u = \overline{T}u,$$

where $\overline{T} : \mathcal{C}_0^1([a,b]) \to \mathcal{C}_0^1([a,b])$ is defined by

$$(\overline{T}u)(t) := \int_a^b G(t,s) f\bigl(s, \gamma\bigl(s, u(s)\bigr)\bigr)\,\mathrm{d}s$$

and $G(t,s)$ is the Green's function corresponding to (1.25). Notice that $\overline{T}$ is completely continuous and for R large enough, $\overline{T}(\mathcal{C}_0^1([a,b])) \subset B(0,R)$. Hence by the properties of the degree

$$\deg\bigl(I - \overline{T}, B(0,R)\bigr) = 1.$$

We know that any fixed point u of $\overline{T}$ is a solution of (2.12). Arguing as in the proof of Theorem 1.1 we prove that $\alpha \leqslant u \leqslant \beta$ and as α and β are strict $\alpha \prec u \prec \beta$. Hence, every fixed point of $\overline{T}$ is in $\Omega \subset B(0,R)$ and using the excision property we obtain

$$\deg(I - T, \Omega) = \deg(I - \overline{T}, \Omega) = 1.$$

Existence of the solution u follows now from the properties of the degree. □

It is easy to deal with $\mathcal{A}$-Carathéodory functions if we reinforce the notion of strict lower and upper solutions imposing that these functions do not satisfy the boundary conditions.

THEOREM 2.14. *Let α and $\beta \in \mathcal{C}([a,b])$ be strict lower and upper solutions of the problem* (2.8) *such that*

$$\alpha(a) < 0 < \beta(a), \quad \alpha(b) < 0 < \beta(b) \quad \text{and} \quad \forall t \in \,]a,b[,\ \alpha(t) < \beta(t).$$

Define E from (1.2) *and assume $f : E \to \mathbb{R}$ is an $\mathcal{A}$-Carathéodory function. Then,*

$$\deg(I - T, \Omega) = 1,$$

where

$$\Omega = \bigl\{u \in \mathcal{C}_0\bigl([a,b]\bigr) \mid \forall t \in [a,b],\ \alpha(t) < u(t) < \beta(t)\bigr\}$$

and $T : \mathcal{C}_0([a,b]) \to \mathcal{C}_0([a,b])$ is defined by (2.9). *In particular, the problem* (2.8) *has at least one solution $u \in W^{2,\mathcal{A}}(a,b)$ such that, for all $t \in [a,b]$,*

$$\alpha(t) < u(t) < \beta(t).$$

The proof of this result is similar to the proof of Theorem 2.13.

2.3. *Non well-ordered lower and upper solutions*

We already noticed in Example 1.3 that the method of lower and upper solutions depends strongly on the ordering $\alpha \leqslant \beta$. On the other hand, lower and upper solutions α, β satisfying the reversed ordering condition $\beta \leqslant \alpha$ arise naturally in situations where the corresponding problem has a solution. As a very simple example, we can consider the linear problem

$$u'' + \frac{2}{3}u = \sin t, \quad u(0) = u(2\pi),\ u'(0) = u'(2\pi).$$

The functions $\alpha = \frac{3}{2}$ and $\beta = -\frac{3}{2}$ are lower and upper solutions such that $\alpha \geqslant \beta$. Notice, however, that the unique solution $u(t) = -3\sin t$ does not lie between the lower and the upper solution. As we shall see, the reason for this example to work is that the "nonlinearity" $f(t,u) = \sin t - \frac{2}{3}u$ "lies" between the two first eigenvalues of the problem. As a first approach we consider the following result which concerns a nonresonance problem using a bounded perturbation of the linear problem at the first eigenvalue. To simplify, we consider a derivative independent problem.

THEOREM 2.15. *Let α and $\beta \in \mathcal{C}([a,b])$ be lower and upper solutions of* (2.5) *such that $\alpha \nleqslant \beta$. Assume $f:[a,b]\times\mathbb{R}\to\mathbb{R}$ is an L^1-Carathéodory function and for some $h \in L^1(a,b)$ either*

$$f(t,u) \leqslant h(t) \quad \text{on } [a,b]\times\mathbb{R}$$

or

$$f(t,u) \geqslant h(t) \quad \text{on } [a,b]\times\mathbb{R}.$$

Then, there exists a solution u of (2.5) *in*

$$\mathcal{S} := \bigl\{u \in \mathcal{C}\bigl([a,b]\bigr) \mid \exists t_1, t_2 \in [a,b],\ u(t_1) \geqslant \beta(t_1),\ u(t_2) \leqslant \alpha(t_2)\bigr\}. \tag{2.13}$$

PROOF. For each $r > 0$, we define

$$f_r(t,u) = \begin{cases} f(t,u), & \text{if } |u| < r,\\ \bigl(1 + r - |u|\bigr)f(t,u) + \bigl(|u| - r\bigr)\frac{u}{r}, & \text{if } r \leqslant |u| < r+1,\\ \frac{u}{r}, & \text{if } r+1 \leqslant |u|, \end{cases}$$

and consider the problem

$$u'' = f_r(t,u), \quad u(a) = u(b),\ u'(a) = u'(b). \tag{2.14}$$

Claim. There exists $k > 0$ *such that for any* $r \geqslant 2(b-a)^2$, *solutions* u *of* (2.14)*, which are in* $\mathcal{S}$*, are such that* $\|u\|_\infty \leqslant k$. Consider the case $f(t,u) \leqslant h(t)$ on $[a,b] \times \mathbb{R}$. Let $u \in \mathcal{S}$ be a solution of (2.14) and let t_0, t_1 and $t_2 \in [a,b]$ be such that

$$u'(t_0) = 0, \quad u(t_1) \geqslant \beta(t_1) \geqslant -\|\beta\|_\infty \quad \text{and} \quad u(t_2) \leqslant \alpha(t_2) \leqslant \|\alpha\|_\infty.$$

Extending u by periodicity, we can write for $t \in [t_0, t_0 + b - a]$, and therefore for all $t \in \mathbb{R}$,

$$u'(t) = -\int_t^{t_0+b-a} f_r\big(s, u(s)\big)\,\mathrm{d}s \geqslant -\|h\|_{L^1} - \frac{\|u\|_\infty}{2(b-a)}.$$

It follows that for $t \in [t_1, t_1 + b - a]$

$$u(t) = u(t_1) + \int_{t_1}^{t} u'(s)\,\mathrm{d}s \geqslant -\|\beta\|_\infty - \|h\|_{L^1}(b-a) - \frac{\|u\|_\infty}{2}$$

and for $t \in [t_2 - b + a, t_2]$

$$u(t) = u(t_2) - \int_t^{t_2} u'(s)\,\mathrm{d}s \leqslant \|\alpha\|_\infty + \|h\|_{L^1}(b-a) + \frac{\|u\|_\infty}{2}.$$

Hence, we have

$$\|u\|_\infty \leqslant 2\big(\|\alpha\|_\infty + \|\beta\|_\infty + \|h\|_{L^1}(b-a)\big) =: k.$$

A similar argument holds if $f(t,u) \geqslant h(t)$.

Conclusion. Consider the problem (2.14), with $r > \max\{k, 2(b-a)^2\}$. It is easy to see that $\alpha_1 = -r - 2$ and $\beta_2 = r + 2$ are lower and upper solutions. Recall that $r \geqslant \|\alpha\|_\infty + \|\beta\|_\infty$ so that $\alpha_1 < \alpha < \beta_2$ and $\alpha_1 < \beta < \beta_2$.

Assume β is not a strict upper solution. There exists then a solution u of (2.14) such that $u \leqslant \beta$ and for some $t_1 \in [a,b]$, $u(t_1) = \beta(t_1)$. As further $\alpha \nleqslant \beta$, there exists $t_2 \in [a,b]$ such that $\alpha(t_2) > \beta(t_2)$. It follows that $\alpha(t_2) > u(t_2)$, $u \in \mathcal{S}$, and we deduce from the claim that $\|u\|_\infty \leqslant k$. Hence, u is a solution of (2.5) in $\mathcal{S}$.

We come to the same conclusion if α is not a strict lower solution.

Suppose now that $\beta_1 = \beta$ and $\alpha_2 = \alpha$ are strict upper and lower solutions. We deduce then from Theorem 2.7 the existence of three solutions of (2.14) one of them, u, being such that for some t_1, $t_2 \in [a,b]$, $u(t_1) > \beta(t_1)$ and $u(t_2) < \alpha(t_2)$. Hence, $u \in \mathcal{S}$ and from the claim $\|u\|_\infty < k$. This implies that u solves (2.5) and proves the theorem. □

Such a result can be used if we assume some asymptotic control on the quotient $f(t,u)/u$ as $|u|$ goes to infinity. We can generalize further assuming different behaviours as u goes to plus or minus infinity. This is worked out in the following theorem.

THEOREM 2.16. *Let* α *and* $\beta \in \mathcal{C}([a,b])$ *be lower and upper solutions of* (2.5) *such that* $\alpha \not\leqslant \beta$. *Let* $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ *be an* L^1*-Carathéodory function such that for some functions* $a_\pm \leqslant 0$, $b_\pm \geqslant 0$ *in* $L^1(a,b)$,

$$a_\pm(t) \leqslant \liminf_{u \to \pm\infty} \frac{f(t,u)}{u} \leqslant \limsup_{u \to \pm\infty} \frac{f(t,u)}{u} \leqslant b_\pm(t),$$

uniformly in $t \in [a,b]$. *Assume further that for any* $p, q \in L^1(a,b)$, *with* $a_+ \leqslant p \leqslant b_+$ *and* $a_- \leqslant q \leqslant b_-$, *the nontrivial solutions of*

$$u'' = p(t)u^+ - q(t)u^-, \quad u(a) = u(b),\ u'(a) = u'(b), \tag{2.15}$$

where $u^+(t) = \max\{u(t), 0\}$ *and* $u^-(t) = \max\{-u(t), 0\}$, *do not have zeros. Then the problem* (2.5) *has at least one solution* $u \in \mathcal{S}$, *where* $\mathcal{S}$ *is defined from* (2.13).

PROOF. *Step* 1. *Claim. There exists* $\epsilon > 0$ *so that for any* $p, q \in L^1(a,b)$, *with* $a_+ - \epsilon \leqslant p \leqslant b_+ + \epsilon$ *and* $a_- - \epsilon \leqslant q \leqslant b_- + \epsilon$, *the nontrivial solutions of* (2.15) *do not have zeros.* If the claim were wrong, there would exist sequences $(p_n)_n$, $(q_n)_n \subset L^1(a,b)$, $(t_n)_n \subset [a,b]$ and $(u_n)_n \subset W^{2,1}(a,b)$ so that $a_+ - 1/n \leqslant p_n \leqslant b_+ + 1/n$, $a_- - 1/n \leqslant q_n \leqslant b_- + 1/n$ and u_n is a solution of (2.15) (with $p = p_n$ and $q = q_n$) such that $\|u_n\|_{\mathcal{C}^1} = 1$ and $u_n(t_n) = 0$. Going to subsequences we can assume, using the Dunford–Pettis theorem (see [12]),

$$p_n \rightharpoonup p, \qquad q_n \rightharpoonup q \quad \text{in } L^1(a,b), \qquad u_n \to u \quad \text{in } \mathcal{C}^1\big([a,b]\big), \quad t_n \to t_0.$$

It follows that $a_+ \leqslant p \leqslant b_+$, $a_- \leqslant q \leqslant b_-$, u is a solution of (2.15) and $u(t_0) = 0$, which contradicts the assumptions.

Step 2. *The modified problem.* Let us choose $R > 0$ large enough so that

$$a_+ - \epsilon \leqslant g_+(t,u) = \frac{f(t,u)}{u} \leqslant b_+ + \epsilon, \quad \text{for } u \geqslant R,$$
$$a_- - \epsilon \leqslant g_-(t,u) = \frac{f(t,u)}{u} \leqslant b_- + \epsilon, \quad \text{for } u \leqslant -R$$

and extend these functions on $[a,b] \times \mathbb{R}$ so that these inequalities remain valid. As f is L^1-Carathéodory, there exists $\ell \in L^1(a,b)$ such that

$$f(t,u) = g_+(t,u)u^+ - g_-(t,u)u^- + h(t,u)$$

and

$$\big|h(t,u)\big| \leqslant \ell(t).$$

Next, for each $r \geqslant 1$, we define

$$g_r^\pm(t,u) = \begin{cases} g_\pm(t,u), & \text{if } |u| < r, \\ \big(1 + r - |u|\big)g_\pm(t,u), & \text{if } r \leqslant |u| < r+1, \\ 0, & \text{if } r+1 \leqslant |u|, \end{cases}$$

$$h_r(t,u) = \begin{cases} h(t,u), & \text{if } |u| < r, \\ \big(1 + r - |u|\big)h(t,u), & \text{if } r \leqslant |u| < r+1, \\ 0, & \text{if } r+1 \leqslant |u|, \end{cases}$$

and consider the modified problem

$$\begin{aligned} &u'' + g_r^+(t,u)u^+ - g_r^-(t,u)u^- + h_r(t,u) = 0, \\ &u(a) = u(b),\ u'(a) = u'(b). \end{aligned} \tag{2.16}$$

Step 3. *Claim. There exists* $k > 0$ *such that, for any* $r > k$*, solutions* u *of* (2.16)*, which are in* $\mathcal{S}$*, are such that* $\|u\|_\infty < k$. Assume by contradiction, there exist sequences $(r_n)_n$ and $(u_n)_n \subset \mathcal{S}$, where $r_n \geqslant n$ and u_n is a solution of (2.16) (with $r = r_n$) such that $\|u_n\|_\infty \geqslant n$.

As $u_n \in \mathcal{S}$, there exist sequences $(t_{1n})_n$ and $(t_{2n})_n \subset [a,b]$ such that $u_n(t_{1n}) \geqslant \beta(t_{1n})$ and $u_n(t_{2n}) \leqslant \alpha(t_{2n})$.

Consider now the functions $v_n = u_n/\|u_n\|_\infty$ which solve the problems

$$\begin{aligned} &v_n'' = g_{r_n}^+(t,u_n)v_n^+ - g_{r_n}^-(t,u_n)v_n^- + \frac{h_{r_n}(t,u_n)}{\|u_n\|_\infty}, \\ &v_n(a) = v_n(b),\ v_n'(a) = v_n'(b). \end{aligned}$$

Going to subsequence, we can assume as above

$$\begin{aligned} &g_n^+(\cdot,u_n) \rightharpoonup p, \quad g_n^-(\cdot,u_n) \rightharpoonup q \quad \text{in } L^1(a,b), \\ &\frac{h_{r_n}(t,u_n)}{\|u_n\|_\infty} \to 0 \quad \text{in } L^1(a,b), \\ &v_n \to v \quad \text{in } \mathcal{C}^1\big([a,b]\big), \qquad t_{1n} \to t_1, \qquad t_{2n} \to t_2. \end{aligned}$$

It follows that v satisfies (2.15) and by assumption has no zeros. Hence, we come to a contradiction since $v(t_1) \geqslant 0$ and $v(t_2) \leqslant 0$ which implies v has a zero.

Conclusion, We deduce from Theorem 2.15 that (2.16) with $r > \max\{k,\ \|\alpha\|_\infty,\ \|\beta\|_\infty\}$ has a solution $u \in \mathcal{S}$ and conclude from Step 3 that u solves (2.5). ☐

In the previous theorem we control asymptotically the nonlinearity using the functions $a_\pm, b_\pm$. Next we impose some admissibility condition on the box $[a_+, b_+] \times [a_-, b_-]$ which is to assume that for any functions $(p,q) \in [a_+, b_+] \times [a_-, b_-]$, the nontrivial solutions of problem (2.15) do not have zeros. Such a condition implies the nonlinearity does not interfere with the second eigenvalue $\lambda_2 = 4(\frac{\pi}{b-a})^2$ of the periodic problem, i.e., $(-\lambda_2, -\lambda_2) \notin [a_+, b_+] \times [a_-, b_-]$. This remark can be made up considering the second curve of the Fučík spectrum. The Fučík spectrum is the set $\mathcal{F}$ of points $(\mu, \nu) \in \mathbb{R}^2$ such that the problem

$$\begin{aligned} &u'' + \mu u^+ - \nu u^- = 0, \\ &u(a) = u(b),\ u'(a) = u'(b), \end{aligned}$$

has nontrivial solutions. From explicit computations of the solution it is easy to see that

$$\mathcal{F}=\bigcup_{n=1}^{\infty}\mathcal{F}_n,$$

where

$$\mathcal{F}_1=\big\{(\mu,0)\mid\mu\in\mathbb{R}\big\}\cup\big\{(0,\nu)\mid\nu\in\mathbb{R}\big\}$$

and

$$\mathcal{F}_n=\left\{(\mu,\nu)\;\middle|\;\frac{1}{\sqrt{\mu}}+\frac{1}{\sqrt{\nu}}=\frac{b-a}{\pi(n-1)}\right\},\quad n=2,3,\ldots.$$

The following proposition relates the admissibility of the box $[a_+,b_+]\times[a_-,b_-]$ with the Fučík spectrum.

PROPOSITION 2.17. *Let* $(\mu,\nu)\in\mathcal{F}_2$ *and* p, $q\in L^1(a,b)$. *Assume that for some set* $I\subset[a,b]$ *of positive measure*

$$\begin{aligned}&p(t)\geqslant-\mu,\quad q(t)\geqslant-\nu,\quad \text{for a.e. } t\in[a,b],\\&p(t)>-\mu,\quad q(t)>-\nu,\quad \text{for a.e. } t\in I.\end{aligned}$$

Then, the nontrivial solutions of problem (2.15) *have no zeros.*

PROOF. Assume there exists a nontrivial solution u which has a zero. Extend u by periodicity and let t_0 and t_1 be consecutive zeros such that u is positive on $]t_0,t_1[$. Define $v(t)=\sin(\sqrt{\mu}(t-t_0))$ and compute

$$(uv'-vu')\big|_{t_0}^{t_1}=-\int_{t_0}^{t_1}\big(p(t)+\mu\big)u(t)v(t)\,\mathrm{d}t.$$

If $t_1-t_0<\frac{\pi}{\sqrt{\mu}}$, we come to a contradiction

$$0<-v(t_1)u'(t_1)\leqslant 0.$$

Hence $t_1-t_0\geqslant\frac{\pi}{\sqrt{\mu}}$ and we only have equality in case $p(t)=-\mu$ on $[t_0,t_1]$. Similarly, we prove the distance between two consecutive zeros t_1 and t_2 with u negative on $]t_1,t_2[$ is such that $t_2-t_1\geqslant\frac{\pi}{\sqrt{\nu}}$ with equality if and only if $q(t)=-\nu$ on $[t_1,t_2]$. It follows that

$$b-a\geqslant t_2-t_0\geqslant\frac{\pi}{\sqrt{\mu}}+\frac{\pi}{\sqrt{\nu}}=b-a.$$

This implies $t_2-t_0=b-a$, $p(t)=-\mu$ on $[t_0,t_1]$ and $q(t)=-\nu$ on $[t_1,t_2]$ which contradicts the assumptions. □

Existence of several solutions can be obtained using non-well-ordered lower and upper solutions. The following result complements Theorem 2.16.

THEOREM 2.18. *Let α_1 and $\alpha_2 \in \mathcal{C}([a,b])$ be lower solutions of* (2.5) *and $\beta \in \mathcal{C}([a,b])$ be a strict upper solution such that $\alpha_2 \nleqslant \beta$, $\alpha_1 \leqslant \alpha_2$ and $\alpha_1 \leqslant \beta$. Assume $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ is an L^1-Carathéodory function such that for some function $b_+ \in L^1(a,b)$,*

$$\limsup_{u \to +\infty} \frac{f(t,u)}{u} \leqslant b_+(t),$$

uniformly in $t \in [a,b]$. Then the problem (2.5) *has at least two solutions u_1 and u_2 such that*

$$\alpha_1 \leqslant u_1 < \beta, \quad u_2 \in \mathcal{S} \quad \text{and} \quad u_1 \leqslant u_2,$$

where $\mathcal{S}$ is defined in (2.13) *with $\alpha = \alpha_2$.*

PROOF. For any $r > \max\{\|\alpha_1\|_\infty, \|\alpha_2\|_\infty, \|\beta\|_\infty\}$, we consider the modified problem

$$u'' = f_r(t,u), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{2.17}$$

where

$$f_r(t,u) = \begin{cases} f\big(t, \alpha_1(t)\big) + u - \alpha_1(t), & \text{if } u \leqslant \alpha_1(t), \\ f(t,u), & \text{if } \alpha_1(t) < u \leqslant r, \\ (1 + r - u) f(t,u), & \text{if } r < u \leqslant r + 1, \\ 0, & \text{if } r + 1 < u. \end{cases}$$

Claim 1. *Every solution of* (2.17) *is such that $u \geqslant \alpha_1$.* This follows from the usual maximum principle argument as it is used, for example, in the proof of Theorem 1.1.

Claim 2. *There exists $k > 0$ so that for any $r > \max\{\|\alpha_1\|_\infty, \|\alpha_2\|_\infty, \|\beta\|_\infty\}$ and any solution $u \in \mathcal{S}$ of* (2.17)*, we have $\|u\|_\infty < k$.* As $u \in \mathcal{S}$, there exist t_0 and t_1 such that

$$u(t_0) = \min_{t \in [a,b]} u(t) \leqslant u(t_1) \leqslant \alpha_2(t_1) \leqslant \|\alpha_2\|_\infty.$$

Further, we deduce from the asymptotic character of f that there exists $\hat{b}_+$ and $h \in L^1(a,b)$ such that, for a.e. $t \in [a,b]$ and all $u \geqslant \alpha_1(t)$,

$$f_r(t,u) \leqslant \hat{b}_+(t) u + h(t).$$

Hence, we have for $t \in [t_0, t_0 + b - a]$

$$\begin{aligned} u(t) &= u(t_0) + \int_{t_0}^{t} f_r\big(s, u(s)\big)(t - s)\, \mathrm{d}s \\ &\leqslant \|\alpha_2\|_\infty + \|h\|_{L^1}(b - a) + (b - a) \int_{t_0}^{t} \hat{b}_+(s) u(s)\, \mathrm{d}s \end{aligned}$$

and the claim follows from Gronwall's lemma.

Conclusion. Consider the problem (2.17) for

$$r > \max\{\|\alpha_1\|_\infty, \|\alpha_2\|_\infty, \|\beta\|_\infty, k\},$$

where k is given in Claim 2.

It follows from Theorem 1.3 that there exists a solution u_1 of (2.17) which is minimal in $[\alpha_1, \beta]$. Hence, $\alpha_1(t) \leqslant u_1(t) \leqslant r$ which implies this solution solves also (2.5).

Next, we can apply Theorem 2.15 to obtain a solution $u_2 \in \mathcal{S}$ of (2.17). From Claim 1, $u_2 \geqslant \alpha_1$, and from Claim 2, $u_2 \leqslant k < r$. Therefore u_2 is a solution of (2.5).

Finally, notice that if $u_2 \not\geqslant u_1$, u_1 and u_2 are upper solutions of (2.17) and we deduce from Theorem 1.2 the existence of a solution u_3 with $\alpha_1 \leqslant u_3 \leqslant \min\{u_1, u_2\}$ which contradicts u_1 to be minimal. ☐

REMARK. We can drop the assumption $\alpha_1 \leqslant \alpha_2$, but this needs additional work as in the proof of Theorem 1.2.

Dirichlet problem (1.24) can be investigated along the same lines. For example, we can consider the problem of interaction with Fučík spectrum and write a result similar to Theorem 2.16.

THEOREM 2.19. *Assume α and $\beta \in \mathcal{C}^1([a,b])$ are lower and upper solutions of* (1.24) *such that $\alpha \not\leqslant \beta$. Let $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ be an L^1-Carathéodory function such that for some functions $a_\pm \leqslant -\lambda_1$, $b_\pm \geqslant -\lambda_1$ in $L^1(a,b)$, where $\lambda_1 = (\frac{\pi}{b-a})^2$*

$$a_\pm(t) \leqslant \liminf_{u\to\pm\infty} \frac{f(t,u)}{u} \leqslant \limsup_{u\to\pm\infty} \frac{f(t,u)}{u} \leqslant b_\pm(t),$$

uniformly in $t \in [a,b]$. Assume further that for any p, $q \in L^1(a,b)$, with $a_+ \leqslant p \leqslant b_+$ and $a_- \leqslant q \leqslant b_-$, the nontrivial solutions of

$$u'' = p(t)u^+ - q(t)u^-, \quad u(a) = 0,\ u(b) = 0,$$

where $u^+(t) = \max\{u(t), 0\}$ and $u^-(t) = \max\{-u(t), 0\}$, do not have interior zeros. Then the problem (1.24) *has at least one solution $u \in \mathcal{S}$, where $\mathcal{S} \subset \mathcal{C}_0^1([a,b])$ is the closure in the $\mathcal{C}^1$-topology of the set*

$$\{u \in \mathcal{C}_0^1([a,b]) \mid \exists t_1, t_2 \in [a,b],\ u(t_1) > \beta(t_1),\ u(t_2) < \alpha(t_2)\}.$$

2.4. *Historical and bibliographical notes*

In 1972, Amann [5] proved a degree result for boundary value problems with strict lower and upper solutions. He considered the associated fixed point problem $u = Tu$ together with the set Ω of functions u that lie between strict lower and upper solutions and proved

$\deg(I - T, \Omega) = 1$. Until recently, such results were only obtained for continuous nonlinearities.

A study of the Carathéodory case for a Dirichlet problem with f independent of u' can be found in De Coster [27], De Coster, Grossinho and Habets [29], Habets and Omari [48] and De Coster and Habets [30]. The derivative dependent case for a Rayleigh equation is worked out in [52]. Our approach follows these papers. Theorem 2.14 which considers $\mathcal{A}$-Carathéodory functions can be found in [42].

For the periodic problem we present here the counterpart of similar results for the Dirichlet problem. Theorem 2.5 extends [87], avoiding the assumption $\alpha, \beta \in W^{1,\infty}(a, b)$.

The abstract idea used in the Three Solutions theorem (Theorem 2.7), goes back to Kolesov [66] in 1970 and Amann [4] in 1971. The first one who proved such a result with degree theory seems to be Amann [5] in 1972 (see also [6]). Extensions were also given by Shivaji [98] and Bongsoo Ko [65]. Our result, Theorem 2.7, improves all these in the special case of ODE and extends them to the Carathéodory case. It extends [32] by relaxing the order relations between α_i and β_i.

In 1972, Sattinger [91] presented as an open problem the question of existence of a solution for the problem

$$-\Delta u = f(x, u), \quad \text{in } \Omega, \qquad u = 0, \quad \text{on } \partial\Omega, \tag{2.18}$$

in presence of lower and upper solutions which does not satisfy the ordering relation $\alpha \leqslant \beta$. A first important contribution to this question was given by Amann, Ambrosetti and Mancini [8] in 1978. They consider (2.18) assuming the nonlinearity $f(x, u)$ is a bounded pertubation of $\lambda_1 u$, where λ_1 is the first eigenvalue of the Laplacian. In 1994, Gossez and Omari [45] assumed some asymptotic control on the nonlinearity so that $\frac{f(t,u)}{u}$ remains, within small perturbations, between the two first eigenvalues of the linearized problem. They prove then existence of a solution in presence of a lower and an upper solution without any order relation. More recently, Habets and Omari [48] extended this work providing a general nonresonance condition with respect to the second curve of the Fučík spectrum as in Theorems 2.16 and 2.19. They obtain an existence and localization result in presence of a lower and an upper solution satisfying the reversed order $\alpha > \beta$. Such a condition did not appear in previous works. In [34], this last result was extended by cancelling the reversed order condition as in Theorem 2.19. In [34] the results were obtained for a general elliptic problem. The parabolic case can be found in [35] and the results for the periodic ODE in [36].

3. Variational methods

3.1. *The minimization method*

Another approach in working with lower and upper solutions is to relate them with variational methods. Consider for example the problem

$$u'' = f(t, u), \quad u(a) = 0, \ u(b) = 0, \tag{3.1}$$

where f is an L^1-Carathéodory function. It is well known that the related functional

$$\phi : H_0^1(a,b) \to \mathbb{R}, \qquad u \mapsto \int_a^b \left[\frac{u'^2(t)}{2} + F\big(t, u(t)\big)\right] \mathrm{d}t,$$

with $F(t,u) = \int_0^u f(t,s)\,\mathrm{d}s$, is of class $\mathcal{C}^1$ and its critical points are the solutions of (3.1). A first link between the two methods is that existence of a well ordered pair of lower and upper solutions α and β, implies the functional ϕ has a minimum on the convex but noncompact set $[\alpha, \beta]$. This minimum solves (3.1).

THEOREM 3.1. *Let α and β be lower and upper solutions of* (3.1) *with $\alpha \leqslant \beta$ on $[a,b]$ and E be defined from* (1.2). *Assume $f : E \to \mathbb{R}$ is an L^1-Carathéodory function. Then the functional ϕ is minimum on $[\alpha, \beta]$, i.e., there exists u with $\alpha \leqslant u \leqslant \beta$ so that*

$$\phi(u) = \min_{\substack{v \in H_0^1(a,b)\\ \alpha \leqslant v \leqslant \beta}} \phi(v).$$

Further, u is a solution of (3.1).

PROOF. Consider the modified problem

$$u'' = f\big(t, \gamma(t,u)\big), \quad u(a) = 0,\ u(b) = 0, \tag{3.2}$$

where $\gamma(t,u)$ is defined from (1.4), and define the functional

$$\bar{\phi} : H_0^1(a,b) \to \mathbb{R}, \qquad u \mapsto \int_a^b \left[\frac{u'^2(t)}{2} + \overline{F}\big(t, u(t)\big)\right] \mathrm{d}t,$$

where $\overline{F}(t,u) = \int_0^u f\big(t, \gamma(t,s)\big)\,\mathrm{d}s$.

Claim 1. *$\bar{\phi}$ has a global minimum u which is a solution of* (3.2). It is easy to verify that $\bar{\phi}$ is of class $\mathcal{C}^1$ and its critical points are precisely the solutions of (3.2). Moreover $\bar{\phi}$ is weakly lower semicontinuous and coercive. Hence the claim follows.

Claim 2. *$\alpha \leqslant u \leqslant \beta$.* Assume $\min_t(u(t) - \alpha(t)) < 0$ and define $t_0 = \max\{t \in [a,b] \mid u(t) - \alpha(t) = \min_s(u(s) - \alpha(s))\}$. We proceed now as in the proof of Theorem 1.1 and obtain that for any $t \geqslant t_0$, near enough t_0,

$$u'(t) - \alpha'(t) = \int_{t_0}^t \big(u''(s) - \alpha''(s)\big)\,\mathrm{d}s = \int_{t_0}^t \big(f\big(s, \alpha(s)\big) - \alpha''(s)\big)\,\mathrm{d}s \leqslant 0.$$

This contradicts the definition of t_0.

Conclusion. Notice that if u is such that $\alpha \leqslant u \leqslant \beta$, the difference $\bar{\phi}(u) - \phi(u)$ is a constant independent of u. Hence, both functionals are minimized together between α and β so that the theorem follows from the previous claims. □

EXAMPLE 3.1. Consider the problem

$$u'' = \lambda f(t,u), \quad u(a) = 0,\ u(b) = 0, \tag{3.3}$$

where $f : [a,b] \times \mathbb{R} \to \mathbb{R}$ is an L^1-Carathéodory function such that $f(t,0) = 0$, $f(t,R) \geqslant 0$ for some $R > 0$ and there exists $\mu \in H_0^1(a,b)$, $0 \leqslant \mu \leqslant R$ that satisfies $\int_a^b F(t,\mu(t))\,\mathrm{d}t < 0$ with $F(t,u) = \int_0^u f(t,s)\,\mathrm{d}s$. Then, we can prove the existence of $\Lambda \geqslant 0$ such that for all $\lambda \geqslant \Lambda$, (3.3) has, beside the trivial solution, at least one nontrivial nonnegative solution.

We just have to observe that $\alpha = 0$ is a lower solution, $\beta = R$ is an upper solution and $\phi(\mu) < 0$ for λ large enough. Hence, for such values of λ, there exists $u \in [0,R]$ which solves (3.3) and minimizes ϕ on $[0,R]$, i.e.,

$$\phi(u) = \min_{\substack{v \in H_0^1(a,b)\\ 0 \leqslant v \leqslant R}} \phi(v) \leqslant \phi(\mu) < 0 = \phi(0).$$

This last inequality implies $u \not\equiv 0$.

The method applies to other boundary value problems such as the periodic problem

$$u'' = f(t,u), \quad u(a) = u(b),\ u'(a) = u'(b). \tag{3.4}$$

Here, the associated functional reads

$$\phi : H^1_{\mathrm{per}}(a,b) \to \mathbb{R}, \qquad u \mapsto \int_a^b \left[\frac{u'^2(t)}{2} + F\bigl(t,u(t)\bigr)\right] \mathrm{d}t, \tag{3.5}$$

with $F(t,u) = \int_0^u f(t,s)\,\mathrm{d}s$ and $H^1_{\mathrm{per}}(a,b) = \{u \in H^1(a,b) \mid u(a) = u(b)\}$. For this problem, we can write an equivalent of Theorem 3.1.

THEOREM 3.2. *Let α and β be lower and upper solutions of* (3.4) *with $\alpha \leqslant \beta$ on $[a,b]$ and E be defined from* (1.2). *Assume $f : E \to \mathbb{R}$ is an L^1-Carathéodory function. Then the functional ϕ defined by* (3.5) *is minimum on $[\alpha,\beta]$, i.e., there exists u with $\alpha \leqslant u \leqslant \beta$ so that*

$$\phi(u) = \min_{\substack{v \in H^1_{\mathrm{per}}(a,b)\\ \alpha \leqslant v \leqslant \beta}} \phi(v).$$

Further, u is a solution of (3.4).

PROOF. The proof of this result follows the argument of the proof of Theorem 3.1. □

As an application of Theorem 3.1, consider the problem

$$u'' + \mu(t)g(u) + h(t) = 0, \quad u(0) = 0,\ u(\pi) = 0. \tag{3.6}$$

THEOREM 3.3. *Let* μ, $h \in L^\infty(0,\pi)$ *and suppose* $\mu_0 = \operatorname{essinf}\mu(t) > 0$. *Assume* $g : \mathbb{R} \to \mathbb{R}$ *is a continuous function,* $G(u) = \int_0^u g(s)\,\mathrm{d}s$,

$$-\infty < \liminf_{u\to\pm\infty} \frac{G(u)}{u^2} \leqslant 0 \quad \textit{and} \quad \limsup_{u\to\pm\infty} \frac{G(u)}{u^2} = +\infty.$$

Then the problem (3.6) *has two infinite sequences of solutions* $(u_n)_n$ *and* $(v_n)_n$ *satisfying*

$$\cdots \leqslant v_{n+1} \leqslant v_n \leqslant \cdots \leqslant v_1 \leqslant u_1 \leqslant \cdots \leqslant u_n \leqslant u_{n+1} \leqslant \cdots$$

and

$$\lim_{n\to\infty}\bigl(\max_t u_n(t)\bigr) = +\infty, \qquad \lim_{n\to\infty}\bigl(\min_t v_n(t)\bigr) = -\infty.$$

PROOF. *Step* 1. *Claim. For every* $M \geqslant 0$ *there exists* β, *an upper solution of* (3.6), *with* $\beta(t) \geqslant M$ *on* $[0,\pi]$. First observe that, if g is unbounded from below on $[0,+\infty[$, we have a sequence of constant upper solutions $\beta_n \to +\infty$. In the opposite case, we can assume there exists $K \geqslant 0$ such that $g(u) \geqslant -K$ for $u \geqslant 0$.

Given $M > 0$, we can choose d so that

$$\|\mu\|_\infty\left(\frac{G(d)}{d^2} + \frac{K}{d}\right) + \frac{\|h\|_\infty}{d} \leqslant \frac{1}{8\pi^2} \quad \text{and} \quad d > 2M.$$

We define then β to be the solution of the Cauchy problem

$$u'' + \|\mu\|_\infty\bigl(g(u) + K\bigr) + \|h\|_\infty = 0, \quad u(0) = d,\ u'(0) = 0. \tag{3.7}$$

Assume there exists $t_0 \in\,]0,\pi]$ such that $\beta(t) > M$ on $[0,t_0[$ and $\beta(t_0) = M$. Notice that on $[0,t_0]$, $\beta'(t) \leqslant 0$ and $\|\mu\|_\infty(G(\beta(t)) + K\beta(t)) + \|h\|_\infty\beta(t) \geqslant 0$. From the conservation of energy for (3.7), we have

$$\begin{aligned}\frac{\beta'^2(t)}{2} &\leqslant \frac{\beta'^2(t)}{2} + \|\mu\|_\infty\bigl(G\bigl(\beta(t)\bigr) + K\beta(t)\bigr) + \|h\|_\infty\beta(t)\\ &= \|\mu\|_\infty\bigl(G(d) + Kd\bigr) + \|h\|_\infty d \leqslant \frac{d^2}{8\pi^2},\end{aligned}$$

i.e.,

$$0 \leqslant -\beta'(t) \leqslant \frac{d}{2\pi}.$$

It follows that for any $t \in [0,t_0]$,

$$d - \beta(t) \leqslant \frac{d}{2\pi}t_0 \leqslant \frac{d}{2},$$

which leads to the contradiction $\beta(t_0) \geqslant \frac{d}{2} > M$. Hence, the claim follows.

In a similar way, we prove the following.

Claim. For every $M \geqslant 0$ *there exists a lower solution* α *of* (3.6) *such that* $\alpha(t) \leqslant -M$ *on* $[0, \pi]$.

Step 2. *Claim. There exist a sequence of positive real numbers* $(s_n)_n$ *with* $s_n \to +\infty$ *and* $z \succ 0$ *such that* $\phi(s_n z) \to -\infty$. Define z to be a C^1-function such that $0 < z(t) \leqslant 1$ on $[0, \pi]$, $z(0) = 0$, $z(\pi) = 0$, $z'(0) > 0$, $z'(\pi) < 0$ and $z(t) = 1$ on $[\epsilon, \pi - \epsilon]$ for some $\epsilon > 0$. Choose $(s_n)_n$ a sequence of positive real numbers with $s_n \to +\infty$ and $\frac{G(s_n)}{s_n^2} \to +\infty$. Recall that the assumptions imply $G(u) \geqslant -K(u^2 + 1)$, for some $K > 0$. We compute then

$$\begin{aligned}
\phi(s_n z) &= \int_0^\pi \left[\frac{s_n^2 z'^2(t)}{2} - \mu(t) G\bigl(s_n z(t)\bigr) - h(t) s_n z(t) \right] \mathrm{d}t \\
&= \int_{[0,\pi]\setminus[\epsilon,\pi-\epsilon]} \frac{s_n^2 z'^2(t)}{2} \,\mathrm{d}t - G(s_n) \int_\epsilon^{\pi-\epsilon} \mu(t) \,\mathrm{d}t \\
&\quad - \int_{[0,\pi]\setminus[\epsilon,\pi-\epsilon]} \mu(t) G\bigl(s_n z(t)\bigr) \,\mathrm{d}t - s_n \int_0^\pi h(t) z(t) \,\mathrm{d}t \\
&\leqslant s_n^2 \|z'\|_\infty^2 \epsilon - G(s_n) \mu_0 (\pi - 2\epsilon) + K\bigl(s_n^2 + 1\bigr) \|\mu\|_{L^1} + s_n \|h\|_{L^1} \|z\|_\infty.
\end{aligned}$$

It follows that $\phi(s_n z) \to -\infty$.

Step 3. *Claim. There exist a sequence of negative real numbers* $(t_n)_n$ *with* $t_n \to -\infty$ *and* $z \succ 0$ *such that* $\phi(t_n z) \to -\infty$. The argument is similar to Step 2.

Step 4. *Conclusion.* By Step 1, we have α_1, β_1 lower and upper solutions of (3.6) with $\alpha_1 \leqslant \beta_1$. Hence, we obtain from Theorem 3.1 a solution u_1 of (3.6) such that $\alpha_1 \leqslant u_1 \leqslant \beta_1$.

From Step 2, we have z and s_1 such that $s_1 z \geqslant u_1$ and $\phi(s_1 z) < \phi(u_1)$. Moreover, Step 1 provides the existence of an upper solution β_2 with $u_1 \leqslant s_1 z \leqslant \beta_2$. Now, by Theorem 3.1, we have a solution u_2 of (3.6) satisfying

$$u_1 \leqslant u_2 \leqslant \beta_2$$

and

$$\phi(u_2) = \min_{\substack{v \in H_0^1(a,b) \\ u_1 \leqslant v \leqslant \beta_2}} \phi(v) \leqslant \phi(s_1 z) < \phi(u_1).$$

It follows that $u_2 \neq u_1$.

Iterating this argument and reproducing it in the negative part, we prove the result. □

REMARK 3.1. The condition on $G(u)/u^2$ cannot be replaced by analogous conditions on $g(u)/u$ (see [39]).

As a next problem consider the following prescribed mean curvature problem

$$\left(\frac{u'}{\sqrt{1 + u'^2}} \right)' = \lambda f(t, u), \quad u(0) = 0,\ u(1) = 0. \tag{3.8}$$

This equation is equivalent to

$$u'' = \lambda\big(1 + u'^2\big)^{3/2} f(t,u), \quad u(0) = 0,\ u(1) = 0.$$

As in Example 3.1, the following result provides a nontrivial nonnegative solution. Notice however that the equation does not satisfy a Nagumo condition which will force us to modify not only the dependence in u but also in the derivative u'.

PROPOSITION 3.4. *Let $f : [0,1] \times \mathbb{R} \to \mathbb{R}$ be a continuous function such that $f(t,0) \leqslant 0$ and define $F(t,u) = \int_0^u f(t,s)\,\mathrm{d}s$. Assume that for some $[a,b] \subset\,]0,1[$, $a \neq b$,*

$$\lim_{u\to 0^+}\left[\max_{t\in[a,b]} \frac{F(t,u)}{u^2}\right] = -\infty. \tag{3.9}$$

Then there exists $\lambda^* > 0$ such that, for each $\lambda \in\,]0,\lambda^*[$, problem (3.8) *has a nontrivial, nonnegative solution.*

PROOF. *The modified problem.* Define

$$p(s) = \begin{cases} (1+s)^{-1/2}, & \text{if } 0 \leqslant s < 1,\\ \frac{1}{8\sqrt{2}}\big[(s-2)^2 + 7\big], & \text{if } 1 \leqslant s < 2,\\ \frac{7}{8\sqrt{2}}, & \text{if } 2 \leqslant s, \end{cases}$$

$\alpha(t) = 0$ and $\beta(t) = t(1-t)$. Consider then the functional $\phi : H_0^1(0,1) \to \mathbb{R}$, defined by

$$\phi(u) = \int_0^1 \left[\frac{1}{2} P\big(u'^2\big) + \lambda \overline{F}\big(t, u(t)\big)\right] \mathrm{d}t,$$

where $P(v) = \int_0^v p(s)\,\mathrm{d}s$, $\overline{F}(t,u) = \int_0^u f(t,\gamma(t,s))\,\mathrm{d}s$ and γ is defined from (1.4). Critical points of ϕ solve

$$\big(p\big(u'^2\big)u'\big)' = \lambda f\big(t, \gamma(t,u)\big), \quad u(0) = 0,\ u(1) = 0,$$

which can also be written

$$u'' = \lambda\big(p\big(u'^2\big) + 2p'\big(u'^2\big)u'^2\big)^{-1} f\big(t,\gamma(t,u)\big), \quad u(0) = 0,\ u(1) = 0. \tag{3.10}$$

Notice at last that $p(s) + 2p'(s)s \geqslant \frac{19}{40\sqrt{2}}$.

Claim 1. *Existence for small values of λ of a solution $u_{\min} \in [0,\beta]$ of* (3.10) *such that*

$$\phi(u_{\min}) = \min_{\substack{v\in H_0^1(0,1)\\ 0\leqslant v\leqslant \beta}} \phi(v).$$

Notice that $\alpha = 0$ is a lower solution. Further, for $\lambda > 0$ small enough, β is an upper solution of (3.10). The claim follows now from an argument similar to the proof of Theorem 3.1.

Claim 2. $\phi(u_{\min}) < \phi(0)$. Let $\zeta \in \mathcal{C}_0^1([0,1])$ be such that $\zeta(t) = 0$ on $[0,a] \cup [b,1]$ and $\zeta(t) \in]0,1]$ on $]a,b[$. From (3.9), we deduce the existence of a sequence $(c_n)_n \subset \mathbb{R}^+$ such that

$$\lim_{n\to\infty} c_n = 0 \quad \text{and} \quad \text{for all } x \in]0,c_n], \quad \max_{t\in[a,b]} \frac{F(t,x)}{x^2} \leqslant \max_{t\in[a,b]} \frac{F(t,c_n)}{c_n^2}.$$

Hence, for n large enough

$$\begin{aligned}\phi(c_n\zeta) &= \int_0^1 \sqrt{1 + c_n^2|\zeta'|^2}\,\mathrm{d}t + \int_a^b \lambda F(\cdot, c_n\zeta)\,\mathrm{d}t - 1\\ &\leqslant c_n^2\biggl(\int_0^1 \frac{1}{c_n^2}\bigl(\sqrt{1+c_n^2|\zeta'|^2} - 1\bigr)\,\mathrm{d}t + \lambda \max_{t\in[a,b]} \frac{F(t,c_n)}{c_n^2}\int_a^b \zeta^2\,\mathrm{d}t\biggr)\\ &< 0 = \phi(0).\end{aligned}$$

The claim follows.

Claim 3. *For $\lambda > 0$ small enough, $u_{\min}$ is a solution of* (3.8). There exists $K > 0$ so that solutions of (3.10) are such that $\|u''\|_\infty \leqslant K\lambda$. Hence for $\lambda > 0$ small enough, $\|u'_{\min}\|_\infty \leqslant 1$ and the claim follows. □

3.2. *The minimax method*

The minus gradient flow

One of the main techniques in variational methods uses the deformation of paths or surfaces along the minus gradient (or pseudo-gradient) flow. In this section, we study the dynamical system associated with such a flow.

We shall define the minus gradient flow using the following assumptions:

(H) Let $f : [a,b] \times \mathbb{R} \to \mathbb{R}$, $(t,u) \mapsto f(t,u)$ be an L^1-Carathéodory function, locally Lipschitz in u. Let also $m \in L^1(a,b)$ be such that $m > 0$ a.e. in $[a,b]$ and $f(t,u) - m(t)\,u$ is decreasing in u.

Now, let us define on $H_0^1(a,b)$ the scalar product

$$(u,v)_{H_0^1} = \int_a^b \bigl[u'(t)v'(t) + m(t)u(t)v(t)\bigr]\,\mathrm{d}t$$

and let

$$F(t,u) = \int_0^u f(t,s)\,\mathrm{d}s.$$

It is then easy to see that the functional

$$\phi : H_0^1(a,b) \to \mathbb{R}, \qquad u \mapsto \int_a^b \left[\frac{u'^2(t)}{2} + F\big(t, u(t)\big)\right] \mathrm{d}t$$

is of class $\mathcal{C}^1$ and

$$\nabla\phi(u) = u - KN(u), \tag{3.11}$$

where

$$\begin{aligned} &N : H_0^1(a,b) \to L^1(a,b), && u \mapsto f(\cdot\,,u) - m(\cdot)u, \\ &K : L^1(a,b) \to H_0^1(a,b), && h \mapsto Kh \end{aligned} \tag{3.12}$$

and Kh is defined to be the unique solution of

$$u'' - m(t)u = h(t), \quad u(a) = 0,\ u(b) = 0.$$

Let us notice at last that if assumptions (H) are satisfied, the function

$$\nabla\phi : \mathcal{C}_0^1\big([a,b]\big) \to \mathcal{C}_0^1\big([a,b]\big),$$

defined from (3.11), is a locally Lipschitzian function. Next, we define for any $r \in \mathbb{R}$ a $\mathcal{C}^1$-function $\psi_r : \mathbb{R} \to [0,1]$ such that $\psi_r(s) = 1$ if $s \geqslant r$ and $\psi_r(s) = 0$ if $s \leqslant r - 1$.

We consider then the Cauchy problem

$$\frac{\mathrm{d}}{\mathrm{d}t}u = -\psi_r\big(\phi(u)\big)\nabla\phi(u) = -\psi_r\big(\phi(u)\big)\big(u - KN(u)\big), \quad u(0) = u_0, \tag{3.13}$$

where $u_0 \in \mathcal{C}_0^1([a,b])$. From the theory of ordinary differential equations, we know that the solution $u(\cdot\,; u_0)$ of (3.13) exists, is unique and is defined in the future on a maximal interval $[0, \omega(u_0)[$. We also know that for any $t \in [0, \omega(u_0)[$, the function $u(t;\cdot) : \mathcal{C}_0^1([a,b]) \to \mathcal{C}_0^1([a,b])$ is continuous. We call the *minus gradient flow* the local semi-dynamical system defined on $\mathcal{C}_0^1([a,b])$ by $u(t; u_0)$.

We could have defined the minus gradient flow in $X = \mathcal{C}_0([a,b])$ or $H_0^1(a,b)$. However, such choices are not suitable in our context. We have to work with sets such as $\{u \in X \mid u \prec \beta\}$ and $\{u \in X \mid u \succ \alpha\}$, where α and β are lower and upper solutions that satisfy the boundary conditions. With the $\mathcal{C}_0([a,b])$ or the $H_0^1(a,b)$-topology, these sets have empty interior which creates major difficulties.

A first result shows that the solutions of (3.13) are defined for all $t \geqslant 0$.

PROPOSITION 3.5. *Let assumptions* (H) *be satisfied and* $u(t; u_0)$ *be the minus gradient flow defined for some* $r \in \mathbb{R}$. *Then for any* $u_0 \in \mathcal{C}_0^1([a,b])$, *we have* $\omega(u_0) = +\infty$.

PROOF. Notice that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\phi\big(u(t;u_0)\big) &= \left(\nabla\phi\big(u(t;u_0)\big), \frac{\mathrm{d}}{\mathrm{d}t}u(t;u_0)\right)_{H_0^1}\\ &= -\psi_r\big(\phi\big(u(t;u_0)\big)\big)\big\|\nabla\phi\big(u(t;u_0)\big)\big\|^2_{H_0^1},\end{aligned}$$

which implies that for all $t \in [0, \omega(u_0)[$

$$\phi\big(u(t;u_0)\big) \leqslant \phi(u_0). \tag{3.14}$$

Observe also that $\phi(u(t;u_0)) \geqslant \min\{r-1, \phi(u_0)\} =: C$. Finally we have for any $0 \leqslant t_1 < t_2 < \omega(u_0)$

$$\begin{aligned}&\big\|u(t_2;u_0) - u(t_1;u_0)\big\|_{H_0^1}\\ &\quad\leqslant \int_{t_1}^{t_2} \psi_r\big(\phi\big(u(s;u_0)\big)\big)\big\|\nabla\phi\big(u(s;u_0)\big)\big\|_{H_0^1}\,\mathrm{d}s\\ &\quad\leqslant \left[\int_{t_1}^{t_2} \psi_r\big(\phi\big(u(s;u_0)\big)\big)\big\|\nabla\phi\big(u(s;u_0)\big)\big\|^2_{H_0^1}\,\mathrm{d}s\right]^{\frac{1}{2}}\left[\int_{t_1}^{t_2} \psi_r\big(\phi\big(u(s;u_0)\big)\big)\,\mathrm{d}s\right]^{\frac{1}{2}}\\ &\quad\leqslant \left[-\int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}s}\phi\big(u(s;u_0)\big)\,\mathrm{d}s\right]^{\frac{1}{2}}\sqrt{t_2-t_1} \leqslant \big[\phi(u_0) - C\big]^{\frac{1}{2}}\sqrt{t_2-t_1}.\end{aligned}$$

Hence, if $\omega(u_0) < +\infty$, there exists $u^* \in H_0^1(a,b)$ such that $u(t;u_0) \overset{H_0^1}{\to} u^*$ as $t \to \omega(u_0)$. It follows that the function $u(\,\cdot\,;u_0) : [0,\omega(u_0)] \to \mathcal{C}([a,b])$, where $u(\omega(u_0);u_0) = u^*$, is continuous and $KNu(\,\cdot\,;u_0) \in \mathcal{C}([0,\omega(u_0)], \mathcal{C}_0^1([a,b]))$. Let $a(t) = \psi_r(\phi(u(t;u_0)))$. For all $t \in [0,\omega(u_0)]$, we can write

$$u(t;u_0) = u_0 \mathrm{e}^{-\int_0^t a(r)\,\mathrm{d}r} + \int_0^t \mathrm{e}^{-\int_s^t a(r)\,\mathrm{d}r} a(s) KNu(s;u_0)\,\mathrm{d}s \in \mathcal{C}_0^1\big([a,b]\big).$$

Hence, $u(\,\cdot\,;u_0) : [0,\omega(u_0)] \to \mathcal{C}_0^1([a,b])$ is continuous, which implies

$$u(t;u_0) \overset{\mathcal{C}_0^1}{\to} u^* \quad \text{as } t \to \omega(u_0).$$

This contradicts the maximality of $\omega(u_0)$. □

Invariant sets

An important property of the cones

$$C_\alpha = \big\{u \in \mathcal{C}_0^1\big([a,b]\big) \mid u \succ \alpha\big\} \tag{3.15}$$

and

$$C^\beta = \{u \in \mathcal{C}_0^1([a,b]) \mid u \prec \beta\}, \tag{3.16}$$

which are associated to lower and upper solutions α and β for (3.1), is that they are positively invariant. To make this precise, let us introduce the following definitions.

DEFINITION 3.1. Let $u(t; u_0)$ be the minus gradient flow defined for some $r \in \mathbb{R}$. A nonempty set $M \subset \mathcal{C}_0^1([a,b])$ is called a *positively invariant set* if

$$\forall u_0 \in M,\ \forall t \geqslant 0, \quad u(t; u_0) \in M.$$

As a first example, notice that (3.14) implies that the set

$$\phi^c = \{u \in \mathcal{C}_0^1([a,b]) \mid \phi(u) < c\}$$

is positively invariant. Also, unions and intersections of positively invariant sets are positively invariant.

To investigate the positive invariance of the cones C_α and C^β defined by (3.15) and (3.16), we need the following lemma.

LEMMA 3.6. *Let assumptions* (H) *be satisfied. Assume* $\alpha \in W^{2,1}(a,b)$ *is a lower solution of* (3.1)*. Then for all* $u \geqslant \alpha$*, we have* $KNu \geqslant \alpha$*, where* K *and* N *are defined by* (3.12).

PROOF. Let $u \geqslant \alpha$, set $w = KNu - \alpha$ and observe that w satisfies

$$\begin{aligned} w'' - m(t)w &= (KNu)''(t) - m(t)(KNu)(t) - \big(\alpha''(t) - m(t)\alpha(t)\big) \\ &\leqslant f\big(t, u(t)\big) - m(t)u(t) - \big(f\big(t, \alpha(t)\big) - m(t)\alpha(t)\big) \leqslant 0, \\ w(a) &\geqslant 0,\ w(b) \geqslant 0. \end{aligned}$$

It follows that $w \geqslant 0$. □

PROPOSITION 3.7. *Let assumptions* (H) *be satisfied and* $u(t; u_0)$ *be the minus gradient flow defined for some* $r \in \mathbb{R}$*. If* $\alpha \in W^{2,1}(a,b)$ *is a lower solution of* (3.1)*, the set* C_α *defined by* (3.15) *is positively invariant. Similarly, if* $\beta \in W^{2,1}(a,b)$ *is an upper solution of* (3.1)*, the set* C^β *defined by* (3.16) *is positively invariant.*

PROOF. If the claim is wrong, we can find $u_0 \in C_\alpha$ and $t_1 > 0$ so that for all $t \in [0, t_1[$, $u(t; u_0) \in C_\alpha$ and $u(t_1; u_0) \in \partial C_\alpha$.

Let $w(t) = u(t; u_0) - \alpha$, define $a(t) = \psi_r(\phi(u(t; u_0)))$ and observe that for all $t \in [0, t_1[$

$$\frac{\mathrm{d}}{\mathrm{d}t} w(t) = \frac{\mathrm{d}}{\mathrm{d}t} u(t; u_0) = -a(t)\big(u(t, u_0) - KN\big(u(t; u_0)\big)\big) = -a(t)w(t) + h(t),$$

where from Lemma 3.6 and for $t \in [0, t_1[$, $h(t) = a(t)(KN(u(t; u_0)) - \alpha) \geqslant 0$. As $w(0) = u_0 - \alpha \succ 0$ we have

$$w(t_1) = w(0)\mathrm{e}^{-\int_0^{t_1} a(r)\,\mathrm{d}r} + \int_0^{t_1} \mathrm{e}^{-\int_s^{t_1} a(r)\,\mathrm{d}r} h(s)\,\mathrm{d}s \succ 0,$$

which contradicts $u(t_1; u_0) \in \partial C_\alpha$. □

Non-well-ordered lower and upper solutions

The first result of this section provides Palais–Smale type sequences from non-well-ordered lower and upper solutions. As usual, this gives a solution of (3.1) with the help of a Palais–Smale type condition.

PROPOSITION 3.8. *Let assumptions* (H) *be satisfied. Suppose α and $\beta \in W^{2,1}(a, b)$ are lower and upper solutions of* (3.1) *and $\alpha \nleqslant \beta$. Define C_α and C^β from* (3.15) *and* (3.16),

$$\begin{aligned} \Gamma &= \bigl\{\gamma \in \mathcal{C}\bigl([0, 1], \mathcal{C}_0^1([a, b])\bigr) \mid \gamma(0) \in C^\beta,\ \gamma(1) \in C_\alpha\bigr\}, \\ T_\gamma &= \bigl\{s \in [0, 1] \mid \gamma(s) \in \mathcal{C}_0^1([a, b]) \setminus \bigl(C^\beta \cup C_\alpha\bigr)\bigr\}, \end{aligned} \tag{3.17}$$

and assume

$$c := \inf_{\gamma \in \Gamma} \max_{s \in T_\gamma} \phi\bigl(\gamma(s)\bigr) \in \mathbb{R}.$$

Finally, let $u(t; u_0)$ be the minus gradient flow defined with $r = c - 1$. Then, for any $\delta \in]0, 1[$ there exists $u_0 \in \mathcal{C}_0^1([a, b])$ such that

$$\forall t > 0, \quad u(t; u_0) \in \phi^{-1}\bigl([c - \delta, c + \delta]\bigr) \setminus \bigl(C^\beta \cup C_\alpha\bigr)$$

and there exists an increasing unbounded sequence $(t_n)_n \subset \mathbb{R}^+$ such that

$$\nabla\phi\bigl(u(t_n; u_0)\bigr) \overset{H_0^1}{\to} 0 \quad \textit{as } n \to \infty.$$

PROOF. Let us fix $\delta \in]0, 1[$ and define $E = \phi^{c-\delta} \cup C_\alpha \cup C^\beta$. Observe that E is positively invariant. Define $A(E) = \{u_0 \in \mathcal{C}_0^1([a, b]) \mid \exists t \geqslant 0,\ u(t; u_0) \in E\}$. Obviously, this set is open and positively invariant.

Consider a path $\gamma \in \Gamma$ so that $c \leqslant \max_{s \in T_\gamma} \phi(\gamma(s)) \leqslant c + \delta$.

Claim. There exists $u_0 \in \gamma(T_\gamma) \setminus A(E)$. Assume by contradiction that for every $s \in T_\gamma$, $\gamma(s) \in A(E)$, i.e., that for every $s \in [0, 1]$, $\gamma(s) \in A(E)$.

Let us prove first that in such a case there exists $T \geqslant 0$ such that for all $s \in [0, 1]$, $u(T; \gamma(s)) \in E$. For any $s \in [0, 1]$, we can find $t_s \geqslant 0$ such that $u(t_s; \gamma(s)) \in E$. Assume next that for every $n \in \mathbb{N}$, there exists $s_n \in [0, 1]$ such that $u(n, \gamma(s_n)) \notin E$. Going to a subsequence, we can assume $s_n \to s^* \in [0, 1]$ and, using the contradiction assumption, there exists $t_{s^*} \geqslant 0$ such that $u(t_{s^*}; \gamma(s^*)) \in E$. As E is open, for all n large enough $u(t_{s^*}; \gamma(s_n)) \in E$ which leads to a contradiction as E is positively invariant and $u(n; \gamma(s_n)) \notin E$.

Notice now that $u(T; \gamma(\cdot))$ is in Γ and such that $\phi(u(T; \gamma(\cdot))) \leqslant c - \delta$ on T_γ which contradicts the definition of c.

Conclusion. By construction of $A(E)$ and as $\phi(u(\cdot\,; u_0))$ is decreasing, we have for all $t > 0$, $u(t; u_0) \in \phi^{-1}([c-\delta, c+\delta]) \setminus (C^\beta \cup C_\alpha)$. Hence, u satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}u = -\nabla\phi(u)$$

and there exists an increasing unbounded sequence $(t_n)_n$ which verifies

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi\bigl(u(t_n; u_0)\bigr) = -\bigl\|\nabla\phi\bigl(u(t_n; u_0)\bigr)\bigr\|^2_{H^1_0} \to 0.$$

□

In order to obtain existence of solutions of (3.1), we need to prove that the sequence $(u(t_n; u_0))_n$ converges towards such a solution. This holds true in case we assume a Palais–Smale condition.

THE PALAIS–SMALE CONDITION. *For every $(u_n)_n \subset H^1_0(a,b)$ such that*

$$\phi(u_n) \text{ is bounded} \quad \text{and} \quad \nabla\phi(u_n) \overset{H^1_0}{\to} 0,$$

there exists a subsequence that converges in $H^1_0(a,b)$ to some function u such that $\nabla\phi(u) = 0$.

It is known that, under our assumptions, the Palais–Smale condition is easy to verify if $\phi(u)$ is coercive or more generally if the Palais–Smale sequences $(u_n)_n$ are bounded in $H^1_0(a,b)$.

Our next result is an existence result that uses the Palais–Smale condition.

THEOREM 3.9. *Let assumptions* (H) *be satisfied. Suppose α and $\beta \in W^{2,1}(a,b)$ are lower and upper solutions of* (3.1) *and $\alpha \not\leqslant \beta$. Define Γ and T_γ as in* (3.17) *and assume*

$$c := \inf_{\gamma\in\Gamma}\max_{s\in T_\gamma}\phi\bigl(\gamma(s)\bigr) \in \mathbb{R}.$$

Finally, assume that the Palais–Smale condition is satisfied. Then there exists $v \in \mathcal{C}^1_0([a,b]) \setminus (C^\beta \cup C_\alpha)$ a solution of (3.1) *such that $\phi(v) = c$.*

PROOF. Let $u(t; u_0)$ be the minus gradient flow defined with $r = c - 1$.

Part 1. *For any $k \in \mathbb{N}$, there exists $v_k \in H^1_0(a,b)$ such that*

$$c - \frac{1}{k} \leqslant \phi(v_k) \leqslant c + \frac{1}{k} \quad \text{and} \quad \nabla\phi(v_k) = 0.$$

Let us fix $k \in \mathbb{N}$. From Proposition 3.8, there exists $u_k \in \mathcal{C}_0^1([a,b])$ such that

$$\forall t > 0, \quad u(t; u_k) \in \phi^{-1}\left(\left[c - \frac{1}{k}, c + \frac{1}{k}\right]\right) \setminus \left(C^\beta \cup C_\alpha\right),$$

and there exists an increasing unbounded sequence $(t_n)_n \subset \mathbb{R}^+$ such that

$$\nabla\phi\bigl(u(t_n; u_k)\bigr) \overset{H_0^1}{\to} 0 \quad \text{as } n \to \infty.$$

Notice that $\phi(u(t; u_k)) > c - 1$ so that $u(t; u_k)$ solves

$$\frac{\mathrm{d}}{\mathrm{d}t} u = -\nabla\phi(u), \quad u(0) = u_k.$$

Using the Palais–Smale condition, we can find a subsequence that we still write $(u(t_n; u_k))_n$ and $v_k \in H_0^1(a,b)$ such that

$$u(t_n; u_k) \overset{H_0^1}{\to} v_k \quad \text{as } n \to \infty,$$
$$c - \frac{1}{k} \leqslant \phi(v_k) \leqslant c + \frac{1}{k} \quad \text{and} \quad \nabla\phi(v_k) = 0.$$

Part 2. $v_k \in \mathcal{C}_0^1([a,b]) \setminus (C^\beta \cup C_\alpha)$.

Claim 1. *There exists* $R > 0$ *such that for all* $s \in [0, +\infty[$,

$$\bigl\|u(s; u_k)\bigr\|_{H_0^1} \geqslant R \quad \textit{implies} \quad \bigl\|\nabla\phi\bigl(u(s; u_k)\bigr)\bigr\|_{H_0^1} \geqslant \frac{1}{R}.$$

If not, there exists a sequence $(s_m)_m \subset [0, +\infty[$ such that

$$\bigl\|u(s_m; u_k)\bigr\|_{H_0^1} \geqslant m \tag{3.18}$$

and $\|\nabla\phi(u(s_m; u_k))\|_{H_0^1} \leqslant \frac{1}{m}$. As $\phi(u(t; u_k))$ is bounded, by the Palais–Smale condition, there exists a subsequence $(s_{m_j})_j$ so that $u(s_{m_j}; u_k)$ converges in $H_0^1(a,b)$ which contradicts (3.18).

Claim 2. $\|u(t; u_k)\|_{H_0^1}$ *is bounded on* $\mathbb{R}^+$. Assume that for some $t > 0$, $\|u(t; u_k)\|_{H_0^1} > R_0 = \max\{\|u_k\|_{H_0^1}, R\}$. Then there exists $t_1 \in [0, t]$ so that $\|u(t_1; u_k)\|_{H_0^1} = R_0$ and for any $s \in [t_1, t]$, $\|u(s; u_k)\|_{H_0^1} \geqslant R_0 \geqslant R$. It follows that

$$\bigl|\phi\bigl(u(t; u_k)\bigr) - \phi\bigl(u(t_1; u_k)\bigr)\bigr| = \int_{t_1}^{t} \bigl\|\nabla\phi\bigl(u(s; u_k)\bigr)\bigr\|_{H_0^1}^2 \,\mathrm{d}s \geqslant \frac{1}{R^2}(t - t_1).$$

On the other hand, we have

$$\bigl\|u(t; u_k) - u(t_1; u_k)\bigr\|_{H_0^1}$$

$$
\begin{aligned}
&= \left\| \int_{t_1}^{t} \nabla\phi\big(u(s;u_k)\big)\,\mathrm{d}s \right\|_{H_0^1} \\
&\leqslant \left[\int_{t_1}^{t} \big\|\nabla\phi\big(u(s;u_k)\big)\big\|_{H_0^1}^2\,\mathrm{d}s \right]^{\frac{1}{2}} (t-t_1)^{\frac{1}{2}} \\
&= \big|\phi\big(u(t;u_k)\big) - \phi\big(u(t_1;u_k)\big)\big|^{\frac{1}{2}} (t-t_1)^{\frac{1}{2}} \\
&\leqslant \big|\phi\big(u(t;u_k)\big) - \phi\big(u(t_1;u_k)\big)\big| R.
\end{aligned}
$$

As $\phi(u(t;u_k))$ is bounded, the claim follows.

Claim 3. $v_k \in \mathcal{C}_0^1([a,b]) \setminus (C^\beta \cup C_\alpha)$. To prove this claim let us show that for some subsequence $u(t_n;u_k) \overset{\mathcal{C}_0^1}{\to} v_k$. Consider the sequence $(w_n)_n \subset \mathcal{C}_0^1([a,b])$, defined by

$$
w_n(r) = \int_0^{t_n} \mathrm{e}^{-(t_n-s)} \big(KNu(s;u_k)\big)(r)\,\mathrm{d}s,
$$

with K and N defined from (3.12). As $\|u(t;u_k)\|_{H_0^1}$ is bounded, there exists $h \in L^1(a,b)$ so that

$$
\begin{aligned}
\big|w_n''(r)\big| &= \left| \int_0^{t_n} \mathrm{e}^{-(t_n-s)} \big[f\big(\cdot, u(s;u_k)\big) - m(\cdot)\big(u(s;u_k) - KNu(s;u_k)\big)\big](r)\,\mathrm{d}s \right| \\
&\leqslant \int_0^{t_n} \mathrm{e}^{-(t_n-s)} h(r)\,\mathrm{d}s \leqslant h(r).
\end{aligned}
$$

Using the Arzelà–Ascoli theorem, we can find a subsequence $(w_{n_i})_i$ converging in $\mathcal{C}_0^1([a,b])$. The same holds true for

$$
u(t_n;u_k) = u_k \mathrm{e}^{-t_n} + \int_0^{t_n} \mathrm{e}^{-(t_n-s)} KNu(s;u_k)\,\mathrm{d}s,
$$

i.e.,

$$
u(t_{n_i};u_k) \overset{\mathcal{C}_0^1}{\to} v_k
$$

and the claim follows.

Conclusion. From the Palais–Smale condition, a subsequence of $(v_k)_k$ converges in $H_0^1(a,b)$ to some function v. As $v_k = KNv_k$, the convergence also holds in $\mathcal{C}_0^1([a,b])$, i.e.,

$$
v \in \mathcal{C}_0^1\big([a,b]\big) \setminus \big(C^\beta \cup C_\alpha\big), \quad \phi(v) = c \quad \text{and} \quad \nabla\phi(v) = 0.
$$

□

A four solutions theorem

This section deals with a problem (3.1) which has the trivial solution $u = 0$. We consider assumptions which imply existence of lower and upper solutions α_i, β_i so that

$$\alpha_1 \leqslant \beta_1 \leqslant 0 \leqslant \alpha_2 \leqslant \beta_2.$$

The Three Solution theorem (see Theorem 2.7 for the periodic case) will provide three solutions, two one-sign ones $u_1 \in [\alpha_1, \beta_1]$ and $u_2 \in [\alpha_2, \beta_2]$ and a third one that can be the zero solution. The difficulty is to obtain a third nontrivial solution. Here such a result is obtained assuming the slope $\frac{f(t,u)}{u}$ crosses the two first eigenvalues.

THEOREM 3.10. *Let assumptions* (H) *be satisfied and assume*:

(i) *there exist* $\lambda > \lambda_2 = \frac{4\pi^2}{(b-a)^2}$ *and* $\delta > 0$ *such that for a.e.* $t \in [a, b]$ *and all* $u \in [-\delta, \delta]$,

$$\frac{f(t,u)}{u} \leqslant -\lambda;$$

(ii) *there exist* $\mu < \lambda_1 = \frac{\pi^2}{(b-a)^2}$ *and* $R > 0$ *such that for a.e.* $t \in [a, b]$ *and all* $u \in \mathbb{R}$ *with* $|u| \geqslant R$,

$$\frac{f(t,u)}{u} \geqslant -\mu.$$

Then the problem (3.1) *has at least three nontrivial solutions* u_i *such that* $u_1 \prec 0$, $u_2 \succ 0$ *and* u_3 *changes sign.*

PROOF. *Claim. There exists* $\alpha_1 \prec 0$ *which is a lower solution of* (3.1). Let $h \in L^1(a, b)$ be such that $h \geqslant 0$ and for a.e. $t \in [a, b]$ and all $u \leqslant 0$,

$$f(t,u) < -\mu u + h(t).$$

Define then α_1 to be the solution of

$$u'' = -\mu u + h(t), \quad u(a) = 0,\ u(b) = 0.$$

As $\mu < \lambda_1$ and $h \geqslant 0$ we have $\alpha_1 \prec 0$ and

$$\alpha_1''(t) = -\mu\alpha_1(t) + h(t) > f\big(t, \alpha_1(t)\big), \quad \alpha_1(a) = 0,\ \alpha_1(b) = 0,$$

i.e., α_1 is a lower solution.

Claim. There exists $\beta_2 \succ 0$ *which is an upper solution of* (3.1). We construct β_2 as we did for α_1.

The modified problem. Consider the modified problem

$$u'' = \bar{f}(t,u), \quad u(a) = 0,\ u(b) = 0, \tag{3.19}$$

where

$$\bar{f}(t,u) = \begin{cases} f\big(t,\alpha_1(t)\big), & \text{if } u < \alpha_1(t), \\ f(t,u), & \text{if } \alpha_1(t) \leqslant u \leqslant \beta_2(t), \\ f\big(t,\beta_2(t)\big), & \text{if } u > \beta_2(t), \end{cases}$$

and the corresponding functional

$$\bar{\phi}(u) = \int_a^b \left[\frac{u'^2(t)}{2} + \overline{F}\big(t,u(t)\big) \right] \mathrm{d}t,$$

with $\overline{F}(t,u) = \int_0^u \bar{f}(t,s)\,\mathrm{d}s$. As usual, it is easy to see that every solution of (3.19) satisfies $\alpha_1 \leqslant u \leqslant \beta_2$ and is a solution of (3.1).

Existence of the solutions u_1 and u_2. Define $\varphi_1(t) = \sin(\pi \frac{t-a}{b-a})$ and let us fix $\epsilon > 0$ small enough so that $\epsilon < \min\{\delta/4, \lambda - \lambda_2\}$, $-4\epsilon\varphi_1 \succ \alpha_1$ and $4\epsilon\varphi_1 \prec \beta_2$. It is easy to see that $\beta_1 = -\epsilon\varphi_1$ and $\alpha_2 = \epsilon\varphi_1$ are respectively upper and lower solutions of (3.1) but are not solutions. This follows from

$$\begin{aligned} \beta_1''(t) &= \epsilon\lambda_1\varphi_1(t) < f\big(t,-\epsilon\varphi_1(t)\big) = f\big(t,\beta_1(t)\big), \\ \alpha_2''(t) &= -\epsilon\lambda_1\varphi_1(t) > f\big(t,\epsilon\varphi_1(t)\big) = f\big(t,\alpha_2(t)\big). \end{aligned}$$

Using Assumption (H) and Proposition 2.12, they are strict upper and lower solutions. Theorem 3.1 applies then with $\alpha = \alpha_i$ and $\beta = \beta_i$, which implies the existence of solutions $u_1 \in C^{\beta_1}$ and $u_2 \in C_{\alpha_2}$.

Existence of a third nontrivial solution. Observe that as $\bar{\phi}$ is coercive, it satisfies the Palais–Smale condition. Hence we can apply Theorem 3.9 with $\alpha = \alpha_2$ and $\beta = \beta_1$. This proves the existence of a solution $u_3 \in \mathcal{C}_0^1([a,b]) \setminus (C^{\beta_1} \cup C_{\alpha_2})$, i.e. $u_3 \neq u_1$ and $u_3 \neq u_2$. The main problem is to prove that u_3 is not the trivial solution. To this aim, we prove that $c = \bar{\phi}(u_3) < 0 = \bar{\phi}(0)$.

Define $\gamma \in \Gamma$ (with $\alpha = \alpha_2$, $\beta = \beta_1$) in the following way

$$\gamma(s) = \begin{cases} 2\epsilon\big((2s-1)\varphi_1 + 2s\varphi_2\big), & \text{if } s \in \big[0, \frac{1}{2}\big], \\ 2\epsilon\big((2s-1)\varphi_1 + 2(1-s)\varphi_2\big), & \text{if } s \in \big[\frac{1}{2}, 1\big], \end{cases}$$

where $\varphi_2(t) = \sin(2\pi \frac{t-a}{b-a})$. Observe that

$$\begin{aligned} &\gamma(0) = -2\epsilon\varphi_1 \prec \beta_1, \qquad \gamma(1) = 2\epsilon\varphi_1 \succ \alpha_2, \\ &\alpha_1 \prec -4\varepsilon\varphi_1 \leqslant \gamma(s) \leqslant 4\varepsilon\varphi_1 \prec \beta_2 \quad \text{for all } s \in [0,1]. \end{aligned}$$

Moreover, for $s \in [0, \frac{1}{2}]$,

$$\bar{\phi}\big(\gamma(s)\big) = \int_a^b \big[2\epsilon^2\big((2s-1)^2(\varphi_1')^2(t) + 4s^2(\varphi_2')^2(t)\big)$$

$$
\begin{aligned}
&+ \overline{F}\big(t, 2\epsilon\big((2s-1)\varphi_1(t) + 2s\varphi_2(t)\big)\big)\big]\,\mathrm{d}t \\
&\leqslant \int_a^b \big[2\epsilon^2\big((2s-1)^2\lambda_1\varphi_1^2(t) + 4s^2\lambda_2\varphi_2^2(t)\big) \\
&\quad - \lambda 2\epsilon^2\big((2s-1)\varphi_1(t) + 2s\varphi_2(t)\big)^2\big]\,\mathrm{d}t \\
&\leqslant \epsilon^2(b-a)\big[(2s-1)^2(\lambda_1-\lambda) + 4s^2(\lambda_2-\lambda)\big] \\
&\leqslant \epsilon^2(b-a)\big[(2s-1)^2 + 4s^2\big](\lambda_2-\lambda) \\
&\leqslant -\frac{\epsilon^3}{2}(b-a).
\end{aligned}
$$

In the same way, we compute for $s \in [\frac{1}{2}, 1]$,

$$
\bar{\phi}\big(\gamma(s)\big) \leqslant -\frac{\epsilon^3}{2}(b-a).
$$

Hence $c \leqslant -\frac{\epsilon^3}{2}(b-a) < 0$. This implies the third solution u_3 is nontrivial.

Claim. The function u_3 changes sign. Assume $u_3 \geqslant 0$ and define $\eta = \max\{\tau \geqslant 0 \mid u_3 - \tau\varphi_1 \geqslant 0\}$. Observe first that

$$
\begin{aligned}
&u_3'' - m(t)u_3 = f(tu_3) - m(t)u_3 \leqslant f(t,0) = 0, \\
&u_3(a) = 0, \qquad u_3(b) = 0.
\end{aligned}
$$

As $u_3 \neq 0$, we deduce from the maximum principle that $u_3 \succ 0$ which implies that $\eta > 0$. Let us assume now that $\eta < \delta$. We can find then $t_0 \in [a,b]$ such that $u_3(t_0) - \eta\varphi_1(t_0) = 0$, $u_3'(t_0) - \eta\varphi_1'(t_0) = 0$ and for t close enough to t_0

$$
\begin{aligned}
(t-t_0)(u_3 - \eta\varphi_1)'(t) &= (t-t_0)\int_{t_0}^t \big(f\big(s, u_3(s)\big) + \lambda_1\eta\varphi_1(s)\big)\,\mathrm{d}s \\
&\leqslant -(\lambda-\lambda_1)(t-t_0)\eta\int_{t_0}^t \varphi_1(s)\,\mathrm{d}s < 0.
\end{aligned}
$$

This contradicts the minimality of $u_3 - \eta\varphi_1$ for $t = t_0$. It follows that $u_3 \geqslant \delta\varphi_1 \succ \alpha_2$ which contradicts the localization of u_3.

We prove in a similar way that u_3 cannot be negative. Therefore u_3 changes sign. □

A five solutions theorem

An additional solution can be obtained by combining variational methods and degree theory. Here we impose that the slope $-\frac{f(t,u)}{u}$ lies between two consecutive eigenvalues for small values of u.

THEOREM 3.11. *Assume that $f \in C^1([a,b] \times \mathbb{R})$ satisfies assumption* (H) *together with*

(i) *there exist* p, q, $k \geqslant 2$ ($k \in \mathbb{N}$) *and* $\delta > 0$ *such that for a.e.* $t \in [a,b]$ *and all* $u \in [-\delta, \delta]$,

$$\lambda_k := \frac{k^2\pi^2}{(b-a)^2} < p \leqslant -\frac{f(t,u)}{u} \leqslant q < \frac{(k+1)^2\pi^2}{(b-a)^2} =: \lambda_{k+1};$$

(ii) *there exist* $\mu < \lambda_1 = \frac{\pi^2}{(b-a)^2}$ *and* $R > 0$ *such that for a.e.* $t \in [a,b]$ *and all* $u \in \mathbb{R}$ *with* $|u| \geqslant R$,

$$\frac{f(t,u)}{u} \geqslant -\mu.$$

Then the problem (3.1) *has at least four nontrivial solutions* u_i *such that* $u_1 \prec 0$, $u_2 \succ 0$ *and* u_3, u_4 *change sign.*

PROOF. As in the proof of Theorem 3.10, we choose strict lower solutions α_i and strict upper solutions β_i such that

$$\alpha_1 \prec -\delta\varphi_1 \prec \beta_1 = -\varepsilon\varphi_1 \quad \text{and} \quad \alpha_2 = \varepsilon\varphi_1 \prec \delta\varphi_1 \prec \beta_2,$$

where $\varepsilon \in]0, \delta[$. As in Theorem 1.3, we can prove the problem (3.1) has two solutions $u_1 \prec \beta_1$ and $u_2 \succ \alpha_2$ such that u_1 is the maximum solution in $[\alpha_1, \beta_1]$ and u_2 is the minimum solution in $[\alpha_2, \beta_2]$. Moreover we can prove as in the proof of Theorem 3.10 that

$$u_1 \prec -\delta\varphi_1 \quad \text{and} \quad u_2 \succ \delta\varphi_1.$$

Consider now the modified problem

$$u'' = \bar{f}(t,u), \quad u(a) = 0,\ u(b) = 0, \tag{3.20}$$

where

$$\bar{f}(t,u) = \begin{cases} f\big(t, u_1(t)\big), & \text{if } u < u_1(t), \\ f(t,u), & \text{if } u_1(t) \leqslant u < u_2(t), \\ f\big(t, u_2(t)\big), & \text{if } u_2(t) \leqslant u, \end{cases}$$

and the corresponding functional

$$\bar{\phi}(u) = \int_a^b \left[\frac{u'^2}{2} + \overline{F}\big(t, u(t)\big)\right] \mathrm{d}t,$$

with $\overline{F}(t,u(t)) = \int_0^u \bar{f}(t,s)\,\mathrm{d}s$. As usual, it is easy to see that every solution of (3.20) satisfies $u_1 \leqslant u \leqslant u_2$ and is a solution of (3.1). As in the proof of Theorem 3.10, we see that the problem (3.20) and hence (3.1) has a third solution $u_3 \neq 0$, which changes sign and is such that

$$\bar{\phi}(u_3) = \inf_{\gamma \in \Gamma} \max_{s \in T_\gamma} \bar{\phi}\bigl(\gamma(s)\bigr) < 0,$$

where Γ and T_γ are defined in (3.17) with ϕ replaced by $\bar{\phi}$. Assume by contradiction that the only solutions of (3.20) are u_1, u_2, u_3 and 0. As $u_1 - 1$ is a strict lower solution of (3.20) and u_1 is the only solution in $C_{u_1-1} \cap C^{\beta_1}$, by Theorem 3.1, u_1 minimizes $\bar{\phi}$ on a $\mathcal{C}_0^1([a,b])$-neighbourhood of this point. It is also a minimizer on some $H_0^1(a,b)$-neighbourhood as follows from Theorem 8 in [38]. Similarly, u_2 is a local minimizer of $\bar{\phi}$ in $H_0^1(a,b)$. By [7], there exists $r > 0$ such that

$$\deg\bigl(I - KN, B(u_1,r)\bigr) = 1 \quad \text{and} \quad \deg\bigl(I - KN, B(u_2,r)\bigr) = 1,$$

where K, N are defined from (3.12) with f replaced by $\bar{f}$ and $m = 0$.

Suppose $\bar{\phi}(u_1) \geqslant \bar{\phi}(u_2)$; a similar argument holds if $\bar{\phi}(u_2) > \bar{\phi}(u_1)$. We can prove (see [37]) that there exists $\gamma > 0$ such that

$$\inf\bigl\{\bar{\phi}(u) \mid \|u - u_1\|_{H_0^1} = \gamma\bigr\} > \bar{\phi}(u_1).$$

Hence, by the Mountain Pass theorem [57], u_3 is of mountain pass type and there exists $r > 0$ such that

$$\deg\bigl(I - KN, B(u_3,r)\bigr) = -1.$$

Moreover, as $KN(H_0^1(a,b)) \subset B(0,R)$ for some $R > 0$,

$$\deg\bigl(I - KN, B(0,R)\bigr) = 1.$$

Let us prove next that for $r > 0$ small enough

$$\bigl|\deg\bigl(I - KN, B(0,r)\bigr)\bigr| = 1.$$

Consider the homotopy

$$u'' = s\bar{f}(t,u) - (1-s)\frac{p+q}{2}u, \quad u(a) = 0,\ u(b) = 0. \tag{3.21}$$

Notice that for a.e. $t \in [a,b]$ and all $u \in [\max(-\delta, u_1(t)), \min(u_2(t), \delta)]$

$$\lambda_k < p \leqslant -s\frac{\bar{f}(t,u)}{u} + (1-s)\frac{p+q}{2} \leqslant q < \lambda_{k+1}.$$

Hence, for every $\epsilon > 0$, we can find $r > 0$ small enough such that if $u \in \partial B(0,r)$ is a solution of (3.21), we have

$$-u'' = A(t)u, \quad u(a) = 0,\ u(b) = 0,$$

with

$$A(t) := -s\frac{\bar{f}(t,u)}{u} + (1-s)\frac{p+q}{2} \in [p,q] \subset]\lambda_k, \lambda_{k+1}[\quad \text{for } t \in [a+\epsilon, b-\epsilon]$$

and $A(t) \in [0,q]$ for $t \in [a,b]$. By eigenvalue comparison, we conclude that $u \equiv 0$. Hence, using properties of the degree, we can write

$$\left|\deg\bigl(I - KN, B(0,r)\bigr)\right| = \left|\deg\left(I - K\left(\frac{p+q}{2}I\right), B(0,r)\right)\right| = 1.$$

We come to the contradiction

$$\begin{aligned}
&\deg\bigl(I - KN, B(0,R)\bigr)\\
&\quad = \deg\bigl(I - KN, B(u_1,r)\bigr) + \deg\bigl(I - KN, B(u_2,r)\bigr)\\
&\qquad + \deg\bigl(I - KN, B(u_3,r)\bigr) + \deg\bigl(I - KN, B(0,r)\bigr)\\
&\quad = 2 - 1 \pm 1 \neq 1.
\end{aligned}$$

This proves existence of an additional nontrivial solution u_4 of (3.20). Recall that such a solution lies in $[u_1, u_2]$ and from the definition of u_1 and u_2, $u_4 \notin C^{\beta_1} \cup C_{\alpha_2}$. Arguing as in Theorem 3.10, we prove then that this solution changes sign. □

3.3. *Historical and bibliographical notes*

As we mentioned in the introduction, existence of a minimum of the related functional between a lower and an upper solution was noticed independently by Chang [17,18] and de Figueiredo and Solimini [38]. Theorems 3.1 and 3.2 provide such a result respectively for the Dirichlet and the periodic problem.

De Figueiredo and Solimini [38] noticed that under certain conditions, the minimum obtained between the lower and the upper solution is valid in the H_0^1-topology. More recently Brezis and Nirenberg [13] pointed out the interest of this result and extended it to nonlinearities with critical growth.

Application to the existence of sequences of solutions as in Theorem 3.3 is due to Omari and Zanolin [76]. Proposition 3.4, which considers a prescribed mean curvature problem, is adapted from Habets and Omari [49].

The idea to combine invariant sets in $\mathcal{C}^1$ with variational methods is worked out in Section 3.2. This gives a new point of view on the relation with variational methods. This goes back to Chang [17,18]. Developments of this idea can be found in [10,19,26,56,67,68]. The study of the Dirichlet problem with nonordered lower and upper solutions as in Theorem 3.9 presents an alternative to the result of [24]. The Four Solutions theorem and the Five Solutions theorem (Theorems 3.10 and 3.11) are known results and can be found, with another proof, in [56].

4. Monotone methods

4.1. *Abstract results*

Let Z be a Banach space. An *order cone* $K \subset Z$ is a closed set such that

$$\text{for all } u \text{ and } v \in K,\ u + v \in K,$$

$$\text{for all } t \in \mathbb{R}^+ \text{ and } u \in K,\ tu \in K,$$

$$\text{if } u \in K \text{ and } -u \in K \text{ then } u = 0.$$

Such a cone K induces an *order* on Z:

$$u \leqslant v \quad \text{if and only if} \quad v - u \in K.$$

We write equivalently $u \leqslant v$ or $v \geqslant u$. The cone is said to be *normal* if there exists $c > 0$ such that $0 \leqslant u \leqslant v$ implies $\|u\| \leqslant c\|v\|$.

The following theorem gives conditions for an increasing sequence $(\alpha_n)_n$ to converge to a fixed point of an operator T.

THEOREM 4.1. *Let $X \subset Z$ be continuously included Banach spaces so that Z has a normal order cone. Let α and $\beta \in X$, $\alpha \leqslant \beta$,*

$$\mathcal{E} = \{u \in X \mid \alpha \leqslant u \leqslant \beta\} \tag{4.1}$$

and let $T : \mathcal{E} \to X$ be completely continuous in X. Assume the sequence $(\alpha_n)_n$ defined by

$$\alpha_0 = \alpha, \qquad \alpha_n = T\alpha_{n-1}, \tag{4.2}$$

is bounded in X and for all $n \in \mathbb{N}$

$$\alpha_n \leqslant \alpha_{n+1} \leqslant \beta.$$

Then the sequence $(\alpha_n)_n$ converges monotonically in X to a fixed point u of T such that

$$\alpha \leqslant u \leqslant \beta.$$

PROOF. *Claim. The sequence $(\alpha_n)_n$ converges in X.* The sequence $(\alpha_n)_n$ is increasing and included in $\mathcal{E}$. As the set $A = \{\alpha_n \mid n \in \mathbb{N}\}$ is bounded in X, $T(A)$ is relatively compact in X. Hence, any sequence $(\alpha_{n_k})_k \subset (\alpha_n)_n$ has a converging subsequence in X and therefore in Z. As the order cone is normal and the sequence is monotone, the sequence itself converges in Z, i.e., there exists $u \in Z$ so that

$$\alpha \leqslant u \leqslant \beta \quad \text{and} \quad \alpha_n \xrightarrow{Z} u.$$

It follows that all such subsequences converging in X have the same limit u, which implies

$$\alpha_n \overset{X}{\to} u.$$

Next, we deduce from the continuity of T that u is a fixed point of T. □

A similar result holds to prove the convergence of decreasing sequences $(\beta_n)_n$.

THEOREM 4.2. *Let $X \subset Z$ be continuously included Banach spaces so that Z has a normal order cone. Let α and $\beta \in X$, $\alpha \leqslant \beta$, $\mathcal{E}$ be defined by* (4.1) *and $T : \mathcal{E} \to X$ be completely continuous in X. Assume the sequence $(\beta_n)_n$ defined by*

$$\beta_0 = \beta, \qquad \beta_n = T\beta_{n-1}, \tag{4.3}$$

is bounded in X and for all $n \in \mathbb{N}$

$$\beta_n \geqslant \beta_{n+1} \geqslant \alpha.$$

Then the sequence $(\beta_n)_n$ converges monotonically in X to a fixed point v of T such that

$$\alpha \leqslant v \leqslant \beta.$$

As a corollary we can write the following result which deals with maps T that are *monotone increasing*, i.e., $u \leqslant v$ implies $Tu \leqslant Tv$.

THEOREM 4.3. *Let $X \subset Z$ be continuously included Banach spaces so that Z has a normal order cone. Let α and $\beta \in X$, $\alpha \leqslant \beta$, $\mathcal{E}$ be defined by* (4.1) *and let $T : \mathcal{E} \to X$ be continuous and monotone increasing. Assume $T(\mathcal{E})$ is relatively compact in X and*

$$\alpha \leqslant T\alpha \quad \text{and} \quad T\beta \leqslant \beta.$$

Then, the sequence $(\alpha_n)_n$ and $(\beta_n)_n$ defined by (4.2) *and* (4.3) *converge monotonically in X to fixed points $u_{\min}$ and $u_{\max}$ of T such that*

$$\alpha \leqslant u_{\min} \leqslant u_{\max} \leqslant \beta.$$

Further, any fixed point $u \in \mathcal{E}$ of T verifies

$$u_{\min} \leqslant u \leqslant u_{\max}.$$

PROOF. *Claim* 1. *The sequence $(\alpha_n)_n$ converges in X to a fixed point $u_{\min}$ of T such that $\alpha \leqslant u_{\min} \leqslant \beta$.* As T is monotone increasing, we prove by induction that for any $n \in \mathbb{N}$, $\alpha_n \leqslant \alpha_{n+1} \leqslant \beta$. Hence, $(\alpha_n)_n \subset \mathcal{E}$ and since $T(\mathcal{E})$ is relatively compact in X the sequence $(\alpha_n)_n$ is bounded in X. The claim follows now from Theorem 4.1. Recall that T is completely continuous as $T(\mathcal{E})$ is relatively compact.

Claim 2. *The sequence $(\beta_n)_n$ converges in X to a fixed point $u_{\max}$ of T such that $u_{\min} \leqslant u_{\max} \leqslant \beta$.* Using Theorem 4.2 with $\alpha = u_{\min}$, we prove, as for Claim 1, existence of a fixed point $u_{\max}$ such that $u_{\min} \leqslant u_{\max} \leqslant \beta$.

Claim 3. *Any fixed point $u \in \mathcal{E}$ of T verifies $u_{\min} \leqslant u \leqslant u_{\max}$.* Since $\alpha \leqslant u \leqslant \beta$, we deduce by induction $\alpha_n = T\alpha_{n-1} \leqslant Tu = u \leqslant T\beta_{n-1} = \beta_n$. The claim follows now by going to the limit. □

4.2. *Well-ordered lower and upper solutions*

The periodic problem

Consider the periodic boundary value problem

$$u'' = f(t,u), \quad u(a) = u(b),\ u'(a) = u'(b), \tag{4.4}$$

where f is a continuous function.

Our aim is to build an approximation scheme, easy to compute, that converges to solutions of (4.4). To this end, given continuous functions α and β, and $M > 0$, we consider the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ defined by

$$\begin{aligned}
&\alpha_0 = \alpha,\\
&\alpha_n'' - M\alpha_n = f(t,\alpha_{n-1}) - M\alpha_{n-1},\\
&\alpha_n(a) = \alpha_n(b),\ \alpha_n'(a) = \alpha_n'(b)
\end{aligned} \tag{4.5}$$

and

$$\begin{aligned}
&\beta_0 = \beta,\\
&\beta_n'' - M\beta_n = f(t,\beta_{n-1}) - M\beta_{n-1},\\
&\beta_n(a) = \beta_n(b),\ \beta_n'(a) = \beta_n'(b).
\end{aligned} \tag{4.6}$$

The approximations α_n and β_n are "easy to compute", in the sense that for every n, the problems (4.5) and (4.6) are linear and have unique solutions which read explicitly

$$\begin{aligned}
\alpha_n(t) &= \int_a^b G(t,s)\big(f\big(s,\alpha_{n-1}(s)\big) - M\alpha_{n-1}(s)\big)\,\mathrm{d}s,\\
\beta_n(t) &= \int_a^b G(t,s)\big(f\big(s,\beta_{n-1}(s)\big) - M\beta_{n-1}(s)\big)\,\mathrm{d}s,
\end{aligned}$$

where $G(t,s)$ is the Green's function of the problem

$$\begin{aligned}
&u'' - Mu = f(t),\\
&u(a) = u(b),\ u'(a) = u'(b).
\end{aligned} \tag{4.7}$$

Clearly this does not avoid numerical difficulties such as those related to stiff systems. This will be the case if we have to pick M very large.

The following theorem proves the convergence of the α_n and β_n.

THEOREM 4.4. *Let α and $\beta \in \mathcal{C}^2([a,b])$, $\alpha \leqslant \beta$ and E be defined from* (1.2). *Assume $f : E \to \mathbb{R}$ is a continuous function, there exists $M > 0$ such that for all (t, u_1), $(t, u_2) \in E$,*

$$u_1 \leqslant u_2 \quad \textit{implies} \quad f(t,u_2) - f(t,u_1) \leqslant M(u_2 - u_1)$$

and for all $t \in [a,b]$

$$\alpha''(t) \geqslant f\big(t, \alpha(t)\big), \quad \alpha(a) = \alpha(b),\ \alpha'(a) \geqslant \alpha'(b),$$
$$\beta''(t) \leqslant f\big(t, \beta(t)\big), \quad \beta(a) = \beta(b),\ \beta'(a) \leqslant \beta'(b).$$

Then the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ defined by (4.5) *and* (4.6) *converge monotonically in $\mathcal{C}^1([a,b])$ to solutions $u_{\min}$ and $u_{\max}$ of* (4.4) *such that*

$$\alpha \leqslant u_{\min} \leqslant u_{\max} \leqslant \beta.$$

Further, any solution u of (4.4) *with graph in E verifies*

$$u_{\min} \leqslant u \leqslant u_{\max}.$$

PROOF. Let $X = \mathcal{C}^1([a,b])$, $Z = \mathcal{C}([a,b])$, $K = \{u \in Z \mid u(t) \geqslant 0 \text{ on } [a,b]\}$ be the order cone in Z and $\mathcal{E}$ be defined from (4.1). Define the operator $T : \mathcal{E} \to X$ by

$$Tu(t) = \int_a^b G(t,s)\big(f\big(s, u(s)\big) - Mu(s)\big)\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function of (4.7). This operator is continuous in X and monotone increasing. Further, $T(\mathcal{E})$ is relatively compact in X, $\alpha \leqslant T\alpha$ and $\beta \geqslant T\beta$. The proof follows now from Theorem 4.3. □

REMARK. Notice that the assumption α and $\beta \in \mathcal{C}^2([a,b])$ is not restrictive. If these functions are lower and upper solutions with angles, the first iterates α_1 and β_1 satisfy the assumptions of the theorem and are such that $\alpha \leqslant \alpha_1 \leqslant \beta_1 \leqslant \beta$.

Next, we consider a derivative dependent problem

$$u'' = f(t, u, u'), \quad u(a) = u(b),\ u'(a) = u'(b). \tag{4.8}$$

As above, given α, $\beta \in \mathcal{C}^1([a,b])$ and $L > 0$, we consider the approximation schemes

$$\begin{aligned} &\alpha_0 = \alpha, \\ &\alpha_n'' - L\alpha_n = f(t, \alpha_{n-1}, \alpha_{n-1}') - L\alpha_{n-1}, \\ &\alpha_n(a) = \alpha_n(b),\ \alpha_n'(a) = \alpha_n'(b) \end{aligned} \tag{4.9}$$

and

$$\begin{aligned}&\beta_0 = \beta,\\ &\beta_n'' - L\beta_n = f(t, \beta_{n-1}, \beta_{n-1}') - L\beta_{n-1},\\ &\beta_n(a) = \beta_n(b),\ \beta_n'(a) = \beta_n'(b).\end{aligned} \tag{4.10}$$

Such problems lead to a major difficulty. A straightforward application of the previous ideas would be to assume that for any u_1, u_2, v_1 and v_2,

$$u_1 \leqslant u_2 \quad \text{implies} \quad f(t, u_2, v_2) - f(t, u_1, v_1) \leqslant L(u_2 - u_1).$$

This would mean that f does not depend on derivatives.

The next theorem works out the difficulty. It is however weaker than Theorem 4.4 since it does not give maximal and minimal solutions. Its proof relies on the following maximum principle (see [6] or [85]).

PROPOSITION 4.5 (Maximum Principle). *Let p, $q \in L^1(a, b)$ be such that the first eigenvalue λ_1 of*

$$-u'' + pu' + qu + \lambda u = 0, \quad u(a) = u(b),\ u'(a) = u'(b),$$

satisfies $\lambda_1 < 0$. Assume $u \in W^{2,1}(a, b)$ is a nontrivial function such that

$$-u'' + pu' + qu \geqslant 0, \quad u(a) = u(b),\ u'(a) \leqslant u'(b).$$

Then $u > 0$ on $[a, b]$.

In general, we shall use the case where $p = 0$ and $q > 0$ are constants.

THEOREM 4.6. *Let α and $\beta \in \mathcal{C}^2([a, b])$ and E be defined from* (1.19). *Assume $f : E \to \mathbb{R}$ is a continuous function, there exists $M \geqslant 0$ such that for all (t, u_1, v), $(t, u_2, v) \in E$,*

$$u_1 \leqslant u_2 \quad \textit{implies} \quad f(t, u_2, v) - f(t, u_1, v) \leqslant M(u_2 - u_1), \tag{4.11}$$

there exists $N \geqslant 0$ such that for all (t, u, v_1), $(t, u, v_2) \in E$,

$$\big|f(t, u, v_2) - f(t, u, v_1)\big| \leqslant N|v_2 - v_1| \tag{4.12}$$

and for all $t \in [a, b]$

$$\begin{aligned}&\alpha''(t) \geqslant f\big(t, \alpha(t), \alpha'(t)\big), \quad \alpha(a) = \alpha(b),\ \alpha'(a) \geqslant \alpha'(b),\\ &\beta''(t) \leqslant f\big(t, \beta(t), \beta'(t)\big), \quad \beta(a) = \beta(b),\ \beta'(a) \leqslant \beta'(b).\end{aligned}$$

Finally, let $L > 0$ *be such that*

$$L \geqslant M + \frac{N^2}{2} + \frac{N}{2}\sqrt{N^2+4M} \tag{4.13}$$

and for all $t \in [a,b]$

$$f\big(t,\alpha(t),\alpha'(t)\big) - f\big(t,\beta(t),\beta'(t)\big) + L\big(\beta(t)-\alpha(t)\big) \geqslant 0.$$

Then, the sequences $(\alpha_n)_n$ *and* $(\beta_n)_n$ *defined by* (4.9) *and* (4.10) *converge monotonically in* $\mathcal{C}^1([a,b])$ *to solutions* u *and* v *of* (4.8) *such that*

$$\alpha \leqslant u \leqslant v \leqslant \beta.$$

REMARKS. (a) The function $w = \beta - \alpha \geqslant 0$ satisfies

$$-w'' + N|w'| + (M+1)w = h(t) \geqslant 0, \quad w(a) = w(b),\ w'(b) \geqslant w'(a).$$

Hence, using the maximum principle we can prove that, if $\alpha \neq \beta$, our assumptions imply $\alpha < \beta$ on $[a,b]$. Also if u is a solution of (4.8) such that $\alpha \gneqq u \gneqq \beta$, we have $\alpha < u < \beta$.

(b) It is clear from Remark (a) that the assumptions on L are satisfied if L is large enough so that the theorem applies for any values of M and N which satisfy the assumptions (4.11) and (4.12).

(c) The conditions on L are immediately satisfied with $L = M$ if the function f does not depend on the derivative u' (i.e., $N = 0$).

(d) If α or β is a solution of (4.8), we have $\alpha_n = \alpha$ for all $n \in \mathbb{N}$ or $\beta_n = \beta$ for all $n \in \mathbb{N}$.

PROOF OF THEOREM 4.6. The proof uses Theorems 4.1 and 4.2 with $X = \mathcal{C}^1([a,b])$, $Z = \mathcal{C}([a,b])$ and $K = \{u \in Z \mid u(t) \geqslant 0 \text{ on } [a,b]\}$ as the order cone in Z. Let $\mathcal{E}$ be defined from (4.1). The operator $T : \mathcal{E} \to X$, defined by

$$Tu(t) = \int_a^b G(t,s)\big(f\big(s,u(s),u'(s)\big) - Lu(s)\big)\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function of (4.7) with $M = L$, is completely continuous in X. With these notations, the approximation schemes (4.9) and (4.10) are equivalent to (4.2) and (4.3).

A: *Claim. Let* $L > 0$ *satisfy* (4.13). *Then the functions* α_n *defined recursively by* (4.9) *are such that for all* $n \in \mathbb{N}$,

(a) α_n *is a lower solution, i.e.,*

$$\begin{aligned} &\alpha_n''(t) \geqslant f\big(t,\alpha_n(t),\alpha_n'(t)\big), \\ &\alpha_n(a) = \alpha_n(b),\ \alpha_n'(a) \geqslant \alpha_n'(b), \end{aligned} \tag{4.14}$$

(b) $\alpha_{n+1} \geqslant \alpha_n$.

The proof is by induction.

Initial step: $n=0$. The condition (4.14) for $n=0$ is an assumption. Next, $w=\alpha_1-\alpha_0$ is a solution of

$$\begin{aligned}&-w''+Lw=\alpha_0''(t)-f\big(t,\alpha_0(t),\alpha_0'(t)\big)\geqslant 0,\\&w(a)=w(b),\ w'(a)\leqslant w'(b).\end{aligned}$$

Hence, we deduce (b) from the maximum principle.

Inductive step – 1st part: *assume* (a) *and* (b) *hold for some n and let us prove that*

$$\begin{aligned}&\alpha_{n+1}''(t)\geqslant f\big(t,\alpha_{n+1}(t),\alpha_{n+1}'(t)\big),\\&\alpha_{n+1}(a)=\alpha_{n+1}(b),\ \alpha_{n+1}'(a)\geqslant\alpha_{n+1}'(b).\end{aligned}$$

Let $w=\alpha_{n+1}-\alpha_n$. We have

$$\begin{aligned}&-\alpha_{n+1}''+f(t,\alpha_{n+1},\alpha_{n+1}')\\&\quad=f(t,\alpha_{n+1},\alpha_{n+1}')-f(t,\alpha_n,\alpha_n')-L(\alpha_{n+1}-\alpha_n)\\&\quad\leqslant M(\alpha_{n+1}-\alpha_n)+N|\alpha_{n+1}'-\alpha_n'|-L(\alpha_{n+1}-\alpha_n)\\&\quad=(M-L)w+N|w'|.\end{aligned}$$

On the other hand, w satisfies

$$-w''+Lw=h(t),\quad w(a)=w(b),\ w'(b)-w'(a)=A,\tag{4.15}$$

with $h(t):=\alpha_n''(t)-f(t,\alpha_n(t),\alpha_n'(t))\geqslant 0$ and $A\geqslant 0$. Its solution w reads

$$\begin{aligned}w(t)=k\bigg[&\int_a^t h(s)\cosh\sqrt{L}\Big(\frac{b-a}{2}+s-t\Big)\,\mathrm{d}s\\&+\int_t^b h(s)\cosh\sqrt{L}\Big(\frac{b-a}{2}+t-s\Big)\,\mathrm{d}s+A\cosh\sqrt{L}\Big(t-\frac{a+b}{2}\Big)\bigg]\end{aligned}$$

where

$$k=\Big(2\sqrt{L}\sinh\sqrt{L}\frac{b-a}{2}\Big)^{-1}.$$

Hence, to prove α_{n+1} is a lower solution, we only have to verify

$$\begin{aligned}\int_a^t\bigg[&(M-L)\cosh\sqrt{L}\Big(\frac{b-a}{2}+s-t\Big)\\&+N\sqrt{L}\Big|\sinh\sqrt{L}\Big(\frac{b-a}{2}+s-t\Big)\Big|\bigg]h(s)\,\mathrm{d}s\leqslant 0,\end{aligned}$$

$$\int_t^b \left[(M-L)\cosh\sqrt{L}\left(\frac{b-a}{2}+t-s\right) + N\sqrt{L}\left|\sinh\sqrt{L}\left(\frac{b-a}{2}+t-s\right)\right|\right] h(s)\,\mathrm{d}s \leqslant 0,$$

and

$$(M-L)\cosh\sqrt{L}\left(t-\frac{a+b}{2}\right) + N\sqrt{L}\left|\sinh\sqrt{L}\left(t-\frac{a+b}{2}\right)\right| \leqslant 0.$$

Since h is non-positive and

$$(M-L)\cosh x + N\sqrt{L}\,|\sinh x| \leqslant (M-L+N\sqrt{L})|\sinh x|$$

for all $x \in \mathbb{R}$, we obtain $(M-L)w + N|w'| \leqslant 0$ if $M-L+N\sqrt{L} \leqslant 0$, which follows from (4.13).

Inductive step – 2nd part: *assume* (a) *and* (b) *hold for some n and let us prove that* $\alpha_{n+2} \geqslant \alpha_{n+1}$. The function $w = \alpha_{n+2} - \alpha_{n+1}$ satisfies (4.15), where

$$h(t) := \alpha''_{n+1}(t) - f\big(t, \alpha_{n+1}(t), \alpha'_{n+1}(t)\big) \quad \text{and} \quad A = 0.$$

From the previous step $h(t) \geqslant 0$ and the claim follows from the maximum principle.

B: *Claim. Let* $L > 0$ *satisfy* (4.13). *Then the functions* β_n *defined recursively by* (4.10) *are such that for all* $n \in \mathbb{N}$,

(a) β_n is an upper solution, i.e.,

$$\beta''_n(t) \leqslant f\big(t, \beta_n(t), \beta'_n(t)\big),$$
$$\beta_n(a) = \beta_n(b), \ \beta'_n(a) \leqslant \beta'_n(b),$$

(b) $\beta_{n+1} \leqslant \beta_n$.

The proof of this claim parallels the proof of Claim A.

C: *Claim.* $\alpha_n \leqslant \beta_n$. Define, for all $i \in \mathbb{N}$, $w_i = \beta_i - \alpha_i$ and

$$h_i(t) := f\big(t, \alpha_i(t), \alpha'_i(t)\big) - f\big(t, \beta_i(t), \beta'_i(t)\big) + L\big(\beta_i(t) - \alpha_i(t)\big).$$

The proof of the claim is by induction.

Initial step: $\alpha_1 \leqslant \beta_1$. The function w_1 is a solution of (4.15) with $h = h_0 \geqslant 0$ and $A = 0$. Using the maximum principle, we deduce that $w_1 \geqslant 0$, i.e., $\alpha_1 \leqslant \beta_1$.

Inductive step: *Let* $n \geqslant 2$. *If* $h_{n-2} \geqslant 0$ *and* $\alpha_{n-1} \leqslant \beta_{n-1}$, *then* $h_{n-1} \geqslant 0$ *and* $\alpha_n \leqslant \beta_n$. First, let us prove that, for all $t \in [a,b]$, the function h_{n-1} is nonnegative. Indeed, we have

$$\begin{aligned} h_{n-1} &= f(\cdot, \alpha_{n-1}, \alpha'_{n-1}) - f(\cdot, \beta_{n-1}, \beta'_{n-1}) + L(\beta_{n-1} - \alpha_{n-1}) \\ &\geqslant -M(\beta_{n-1} - \alpha_{n-1}) - N|\beta'_{n-1} - \alpha'_{n-1}| + L(\beta_{n-1} - \alpha_{n-1}) \\ &= (L-M)w_{n-1} - N|w'_{n-1}|. \end{aligned}$$

Recall that w_{n-1} is a solution of (4.15) with $h(t) = h_{n-2}(t) \geqslant 0$ and $A = 0$. Hence, we can proceed as in the proof of Claim A to show that $h_{n-1} \geqslant 0$. It follows then from the maximum principle that w_n is nonnegative, i.e., $\alpha_n \leqslant \beta_n$.

D: *Claim. There exists* $R > 0$ *such that any solution* u *of*

$$u'' \leqslant f(t, u, u'), \quad u(a) = u(b),\ u'(a) = u'(b),$$

with $\alpha \leqslant u \leqslant \beta$ *satisfies* $\|u'\|_\infty < R$. We deduce from the assumptions that

$$u'' = f(t, u, u') + h(t)$$

where $h(t) \leqslant 0$ and $f(t, u, u') + h(t) \leqslant \max_E |f(t, u, 0)| + N|u'|$. The proof follows now using Proposition 1.7.

E: *Claim. There exists* $R > 0$ *such that any solution* u *of*

$$u'' \geqslant f(t, u, u'), \quad u(a) = u(b),\ u'(a) = u'(b),$$

with $\alpha \leqslant u \leqslant \beta$ *satisfies* $\|u'\|_\infty < R$. The proof repeats the argument of Claim D (See the remark that follows Proposition 1.7).

F: *Conclusion.* We deduce from Theorems 4.1 and 4.2 that the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ converge monotonically in $C^1([a, b])$ to functions u and v which are solutions of (4.8) such that $\alpha \leqslant u \leqslant \beta$ and $\alpha \leqslant v \leqslant \beta$. Further, as $\alpha_n \leqslant \beta_n$ for any n, we have $u \leqslant v$. ☐

REMARK. Notice that with a little additional work we can reduce problems to cases where Theorem 4.6 applies. For example, it does not apply to problem

$$\begin{aligned} &u'' - u + (u')^2 = \sin t, \\ &u(0) = u(2\pi),\ u'(0) = u'(2\pi), \end{aligned}$$

as (4.12) is not satisfied. However, we can work it out as follows. Notice first that this problem satisfies a Nagumo condition. Next, we know that lower and upper solutions, α and $\beta \in [-1, 1]$, of problems that satisfy such a Nagumo condition have a priori bounded derivatives: $\|\alpha'\|_\infty$ and $\|\beta'\|_\infty \leqslant R$. We can modify then the equation for $|u'| \geqslant R$ so that the same Nagumo condition is satisfied for the modified problem together with (4.12). It follows then that the approximations defined from (4.9) and (4.10) satisfy the same bounds and that convergence of the approximations holds.

The Dirichlet problem

As in the periodic case, we can work with the Dirichlet problem

$$u'' = f(t, u), \quad u(a) = 0,\ u(b) = 0, \tag{4.16}$$

where f is a continuous function.

The following result paraphrases Theorem 4.4.

THEOREM 4.7. *Let* α *and* $\beta \in C^2([a,b])$*,* $\alpha \leqslant \beta$ *and let* E *be defined from* (1.2)*. Assume* $f : E \to \mathbb{R}$ *is a continuous function, there exists* $M \geqslant 0$ *such that for all* (t, u_1)*,* $(t, u_2) \in E$*,*

$$u_1 \leqslant u_2 \quad \textit{implies} \quad f(t,u_2) - f(t,u_1) \leqslant M(u_2 - u_1)$$

and for all $t \in [a,b]$

$$\begin{aligned} &\alpha''(t) \geqslant f\big(t,\alpha(t)\big), \quad \alpha(a) \leqslant 0,\ \alpha(b) \leqslant 0,\\ &\beta''(t) \leqslant f\big(t,\beta(t)\big), \quad \beta(a) \geqslant 0,\ \beta(b) \geqslant 0. \end{aligned}$$

Then the sequences $(\alpha_n)_n$ *and* $(\beta_n)_n$ *defined by*

$$\begin{aligned} &\alpha_0 = \alpha,\\ &\alpha_n'' - M\alpha_n = f(t,\alpha_{n-1}) - M\alpha_{n-1},\\ &\alpha_n(a) = 0,\ \alpha_n(b) = 0, \end{aligned}$$

and

$$\begin{aligned} &\beta_0 = \beta,\\ &\beta_n'' - M\beta_n = f(t,\beta_{n-1}) - M\beta_{n-1},\\ &\beta_n(a) = 0,\ \beta_n(b) = 0, \end{aligned}$$

converge uniformly and monotonically to solutions $u_{\min}$ *and* $u_{\max}$ *of* (4.16) *such that*

$$\alpha \leqslant u_{\min} \leqslant u_{\max} \leqslant \beta.$$

Further, any solution u *of* (4.16) *with graph in* E *verifies*

$$u_{\min} \leqslant u \leqslant u_{\max}.$$

PROOF. The proof of this theorem repeats the argument of Theorem 4.4. □

In case of derivative dependent equations

$$\begin{aligned} &u'' = f(t,u,u'),\\ &u(a) = 0,\ u(b) = 0, \end{aligned} \tag{4.17}$$

approximation schemes similar to (4.9), (4.10) do not work. Here we have to work out a generalization as in the following theorem. Notice that we use lower and upper solutions that verify the boundary conditions.

THEOREM 4.8. *Let* α *and* $\beta \in C^2([a,b])$*,* $\alpha \leqslant \beta$ *and let* E *be defined from* (1.19)*. Assume* $f : E \to \mathbb{R}$ *is a continuous function, there exists* $M \geqslant 0$ *such that for all* (t,u_1,v)*,* $(t,u_2,v) \in E$*,*

$$u_1 \leqslant u_2 \quad \textit{implies} \quad f(t, u_2, v) - f(t, u_1, v) \leqslant M(u_2 - u_1),$$

there exists $N \geqslant 0$ *such that for all* (t, u, v_1), $(t, u, v_2) \in E$,

$$\bigl|f(t, u, v_2) - f(t, u, v_1)\bigr| \leqslant N|v_2 - v_1|$$

and for all $t \in [a, b]$

$$\begin{aligned} &\alpha''(t) \geqslant f\bigl(t, \alpha(t), \alpha'(t)\bigr), \quad \alpha(a) = 0,\ \alpha(b) = 0, \\ &\beta''(t) \leqslant f\bigl(t, \beta(t), \beta'(t)\bigr), \quad \beta(a) = 0,\ \beta(b) = 0. \end{aligned}$$

Finally, let $K_0 \in \mathcal{C}([a, b])$ *be such that* $K_0(a) > 0$ *and for all* $t \in [a, b]$, $K_0(t) = -K_0(b + a - t)$. *Then, for* L *large enough, the sequences* $(\alpha_n)_n$ *and* $(\beta_n)_n$ *defined by*

$$\begin{aligned} &\alpha''_{n+1} - \sqrt[3]{L}\, K_0(t)\alpha'_{n+1} - L\alpha_{n+1} = f(t, \alpha_n, \alpha'_n) - \sqrt[3]{L}\, K_0(t)\alpha'_n - L\alpha_n, \\ &\alpha_{n+1}(a) = 0,\ \alpha_{n+1}(b) = 0, \\ &\beta''_{n+1} - \sqrt[3]{L}\, K_0(t)\beta'_{n+1} - L\beta_{n+1} = f(t, \beta_n, \beta'_n) - \sqrt[3]{L}\, K_0(t)\beta'_n - L\beta_n, \\ &\beta_{n+1}(a) = 0,\ \beta_{n+1}(b) = 0, \end{aligned}$$

converge monotonically in $\mathcal{C}^1([a, b])$ *to solutions* u *and* v *of* (4.17) *such that, for all* $t \in [a, b]$, *we have*

$$\alpha(t) \leqslant u(t) \leqslant v(t) \leqslant \beta(t).$$

PROOF. The above result can be proved adapting the proof of Theorem 4.6. □

4.3. *Lower and upper solutions in reversed order*

Consider the periodic boundary value problem

$$\begin{aligned} &u'' = f(t, u), \\ &u(a) = u(b),\ u'(a) = u'(b), \end{aligned} \tag{4.18}$$

where f is a continuous function.

In Section 4.2, we have built an approximation scheme for solutions of (4.18) based on the maximum principle. Here, we consider a similar approach based on the antimaximum principle. Given continuous functions α and β, and $M > 0$, we consider the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ defined by

$$\begin{aligned} &\alpha_0 = \alpha, \\ &\alpha''_n + M\alpha_n = f(t, \alpha_{n-1}) + M\alpha_{n-1}, \\ &\alpha_n(a) = \alpha_n(b),\ \alpha'_n(a) = \alpha'_n(b) \end{aligned} \tag{4.19}$$

and

$$\begin{aligned}&\beta_0=\beta,\\&\beta_n''+M\beta_n=f(t,\beta_{n-1})+M\beta_{n-1},\\&\beta_n(a)=\beta_n(b),\ \beta_n'(a)=\beta_n'(b).\end{aligned}\tag{4.20}$$

If M is not an eigenvalue of the periodic problem, i.e., $M\neq(\frac{2n\pi}{b-a})^2$ with $n\in\mathbb{N}$, the functions α_n and β_n, solutions of (4.19) and (4.20), can be written explicitly

$$\alpha_n(t)=\int_a^b G(t,s)\bigl(f\bigl(s,\alpha_{n-1}(s)\bigr)+M\alpha_{n-1}(s)\bigr)\,\mathrm{d}s,$$

$$\beta_n(t)=\int_a^b G(t,s)\bigl(f\bigl(s,\beta_{n-1}(s)\bigr)+M\beta_{n-1}(s)\bigr)\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function of the problem

$$\begin{aligned}&u''+Mu=f(t),\\&u(a)=u(b),\ u'(a)=u'(b).\end{aligned}\tag{4.21}$$

The next theorem indicates a framework to obtain convergence of the α_n and β_n to extremal solutions of (4.18). To prove this result we need the following antimaximum principle.

PROPOSITION 4.9 (Antimaximum Principle). *Let* $q\in\,]0,\frac{\pi^2}{(b-a)^2}]$. *Suppose* $u\in\mathcal{C}^2([a,b])$ *is a solution of*

$$u''+qu=f(t),\quad u(a)-u(b)=0,\ u'(a)-u'(b)=A,\tag{4.22}$$

where $A\geqslant 0$ *and* $f\in\mathcal{C}([a,b])$, $f(t)\geqslant 0$. *Then* $u\geqslant 0$ *on* $[a,b]$.

PROOF. *Claim* 1. *If* $f\not\equiv 0$, *solutions* u *of* (4.22) *are one-signed.* Let t_0 be a zero of u. Extend u by periodicity, define $v(t)=\sin\sqrt{q}(t-t_0)$ and compute

$$\begin{aligned}u'(t_0)\sin\sqrt{q}(b-a)&\geqslant(u'v-v'u)|_{t_0}^{b}+(u'v-v'u)|_{b}^{t_0+b-a}\\&=\int_{t_0}^{t_0+b-a}f(s)\sin\sqrt{q}(s-t_0)\,\mathrm{d}s>0.\end{aligned}$$

If $q=\frac{\pi^2}{(b-a)^2}$, this is contradictory. If $q\neq\frac{\pi^2}{(b-a)^2}$, this implies $u'(t_0)>0$ and u cannot be a periodic function.

Claim 2. *If* $f \not\equiv 0$, *any one-signed solution of* (4.22) *is positive.* Direct integration of (4.22) gives

$$q \int_a^b u(s)\,\mathrm{d}s = \int_a^b f(s)\,\mathrm{d}s + A > 0.$$

Hence, we have $u(t) > 0$.

Claim 3. *If* $f(t) \equiv 0$, $u(t) \geqslant 0$. We deduce from a direct integration that

$$u(t) = A \frac{\cos \sqrt{q}(\frac{a+b}{2} - t)}{2\sqrt{q} \sin \sqrt{q}(\frac{b-a}{2})} \geqslant 0.$$

□

THEOREM 4.10. *Let* α *and* $\beta \in \mathcal{C}^2([a,b])$, $\beta \leqslant \alpha$ *and*

$$E := \bigl\{(t,u) \in [a,b] \times \mathbb{R} \mid \beta(t) \leqslant u \leqslant \alpha(t)\bigr\}.$$

Assume $f : E \to \mathbb{R}$ *is a continuous function, there exists* $M \in \left]0, \frac{\pi^2}{(b-a)^2}\right]$ *such that for all* (t,u_1), $(t,u_2) \in E$,

$$u_1 \leqslant u_2 \quad \textit{implies} \quad f(t,u_2) - f(t,u_1) \geqslant -M(u_2 - u_1)$$

and for all $t \in [a,b]$

$$\alpha''(t) \geqslant f\bigl(t,\alpha(t)\bigr), \quad \alpha(a) = \alpha(b),\ \alpha'(a) \geqslant \alpha'(b),$$
$$\beta''(t) \leqslant f\bigl(t,\beta(t)\bigr), \quad \beta(a) = \beta(b),\ \beta'(a) \leqslant \beta'(b).$$

Then the sequences $(\alpha_n)_n$ *and* $(\beta_n)_n$ *defined by* (4.19) *and* (4.20) *converge monotonically in* $\mathcal{C}^1([a,b])$ *to solutions* $u_{\max}$ *and* $u_{\min}$ *of* (4.18) *such that*

$$\beta \leqslant u_{\min} \leqslant u_{\max} \leqslant \alpha.$$

Further, any solution u *of* (4.18) *with graph in* E *verifies*

$$u_{\min} \leqslant u \leqslant u_{\max}.$$

PROOF. Let $X = \mathcal{C}^1([a,b])$, $Z = \mathcal{C}([a,b])$, $K = \{u \in Z \mid u(t) \geqslant 0 \text{ on } [a,b]\}$ be the order cone in Z and

$$\mathcal{E} = \{u \in X \mid \beta \leqslant u \leqslant \alpha\}. \tag{4.23}$$

The operator $T : \mathcal{E} \to X$, defined by

$$Tu(t) = \int_a^b G(t,s)\bigl(f\bigl(s,u(s)\bigr) + Mu(s)\bigr)\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function of (4.21), is continuous in X and monotone increasing (see Proposition 4.9). Further, $T(\mathcal{E})$ is relatively compact in X, $\beta \leqslant T\beta$ and $\alpha \geqslant T\alpha$. The proof follows now from Theorem 4.3, where α and β have to be interchanged. □

Next, we consider the derivative dependent problem

$$\begin{aligned} &u'' = f(t,u,u'),\\ &u(a) = u(b),\ u'(a) = u'(b). \end{aligned} \tag{4.24}$$

As above, given α, $\beta \in \mathcal{C}^1([a,b])$ and $L > 0$, we consider the approximation schemes

$$\begin{aligned} &\alpha_0 = \alpha,\\ &\alpha_n'' + L\alpha_n = f(t,\alpha_{n-1},\alpha_{n-1}') + L\alpha_{n-1},\\ &\alpha_n(a) = \alpha_n(b),\ \alpha_n'(a) = \alpha_n'(b) \end{aligned} \tag{4.25}$$

and

$$\begin{aligned} &\beta_0 = \beta,\\ &\beta_n'' + L\beta_n = f(t,\beta_{n-1},\beta_{n-1}') + L\beta_{n-1},\\ &\beta_n(a) = \beta_n(b),\ \beta_n'(a) = \beta_n'(b). \end{aligned} \tag{4.26}$$

The following result is a counterpart of Theorem 4.6.

THEOREM 4.11. *Let α and $\beta \in \mathcal{C}^2([a,b])$, $\beta \leqslant \alpha$ and*

$$E := \bigl\{(t,u,v) \in [a,b] \times \mathbb{R}^2 \mid \beta(t) \leqslant u \leqslant \alpha(t)\bigr\}.$$

Assume $f : E \to \mathbb{R}$ is a continuous function, there exists $M \in \,]0, \frac{\pi^2}{(b-a)^2}[$ such that for all (t,u_1,v), $(t,u_2,v) \in E$,

$$u_1 \leqslant u_2 \quad \textit{implies} \quad f(t,u_2,v) - f(t,u_1,v) \geqslant -M(u_2 - u_1),$$

there exists $N \geqslant 0$ such that for all (t,u,v_1), $(t,u,v_2) \in E$,

$$\bigl|f(t,u,v_2) - f(t,u,v_1)\bigr| \leqslant N|v_2 - v_1|$$

and for all $t \in [a,b]$

$$\begin{aligned} &\alpha''(t) \geqslant f\bigl(t,\alpha(t),\alpha'(t)\bigr), \quad \alpha(a) = \alpha(b),\ \alpha'(a) \geqslant \alpha'(b),\\ &\beta''(t) \leqslant f\bigl(t,\beta(t),\beta'(t)\bigr), \quad \beta(a) = \beta(b),\ \beta'(a) \leqslant \beta'(b). \end{aligned}$$

Finally, let $L \in \left]M, \frac{\pi^2}{(b-a)^2}\right[$ *be such that*

$$(L - M)\cos\sqrt{L}\left(\frac{b-a}{2}\right) - N\sqrt{L}\sin\sqrt{L}\left(\frac{b-a}{2}\right) \geqslant 0 \tag{4.27}$$

and

$$f\bigl(t, \alpha(t), \alpha'(t)\bigr) - f\bigl(t, \beta(t), \beta'(t)\bigr) + L\bigl(\alpha(t) - \beta(t)\bigr) \geqslant 0.$$

Then, the sequences $(\alpha_n)_n$ *and* $(\beta_n)_n$ *defined by* (4.25) *and* (4.26) *converge monotonically in* $\mathcal{C}^1([a,b])$ *to solutions* u *and* v *of* (4.24) *such that*

$$\beta \leqslant v \leqslant u \leqslant \alpha.$$

PROOF. The proof uses Theorems 4.1 and 4.2 with $X = \mathcal{C}^1([a,b])$, $Z = \mathcal{C}([a,b])$ and $K = \{u \in Z \mid u(t) \geqslant 0 \text{ on } [a,b]\}$ as the order cone in Z. Let $\mathcal{E}$ be defined from (4.23). The operator $T : \mathcal{E} \to X$, defined by

$$Tu(t) = \int_a^b G(t,s)\bigl(f\bigl(s, u(s), u'(s)\bigr) + Lu(s)\bigr)\,\mathrm{d}s,$$

where $G(t,s)$ is the Green's function of (4.21) with M replaced by L, is completely continuous in X. With these notations, the approximation schemes (4.25) and (4.26) are equivalent to (4.3) and (4.2).

A: *Claim. Let* $L > 0$ *satisfy* (4.27). *Then the functions* α_n *defined recursively by* (4.25) *are such that for all* $n \in \mathbb{N}$,

(a) α_n is a lower solution, i.e.,

$$\begin{aligned} &\alpha_n''(t) \geqslant f\bigl(t, \alpha_n(t), \alpha_n'(t)\bigr), \\ &\alpha_n(a) = \alpha_n(b),\ \alpha_n'(a) \geqslant \alpha_n'(b), \end{aligned} \tag{4.28}$$

(b) $\alpha_{n+1} \leqslant \alpha_n$.

The proof is by induction.

Initial step: $n = 0$. The condition (4.28) for $n = 0$ is an assumption. Next, $w = \alpha_0 - \alpha_1$ is a solution of

$$\begin{aligned} &w'' + Lw = \alpha_0''(t) - f\bigl(t, \alpha_0(t), \alpha_0'(t)\bigr) \geqslant 0, \\ &w(a) = w(b),\ w'(a) \geqslant w'(b). \end{aligned}$$

Hence, we deduce (b) from the antimaximum principle (Proposition 4.9).

Inductive step – 1st part: *assume* (a) *and* (b) *hold for some* n *and let us prove that*

$$\begin{aligned} &\alpha_{n+1}''(t) \geqslant f\bigl(t, \alpha_{n+1}(t), \alpha_{n+1}'(t)\bigr), \\ &\alpha_{n+1}(a) = \alpha_{n+1}(b),\ \alpha_{n+1}'(a) \geqslant \alpha_{n+1}'(b). \end{aligned}$$

Let $w = \alpha_n - \alpha_{n+1} \geqslant 0$. We have

$$\begin{aligned}
&-\alpha''_{n+1} + f(t, \alpha_{n+1}, \alpha'_{n+1})\\
&\quad = -f(t, \alpha_n, \alpha'_n) + f(t, \alpha_{n+1}, \alpha'_{n+1}) - L(\alpha_n - \alpha_{n+1})\\
&\quad \leqslant M(\alpha_n - \alpha_{n+1}) + N|\alpha'_{n+1} - \alpha'_n| - L(\alpha_n - \alpha_{n+1})\\
&\quad = (M - L)w + N|w'|.
\end{aligned}$$

On the other hand, w satisfies

$$w'' + Lw = h(t), \quad w(a) = w(b),\ w'(a) - w'(b) = C, \tag{4.29}$$

with $h(t) := \alpha''_n(t) - f(t, \alpha_n(t), \alpha'_n(t)) \geqslant 0$ and $C \geqslant 0$. Observe that

$$\begin{aligned}
w(t) = \frac{1}{2\sqrt{L}\sin\sqrt{L}(\frac{b-a}{2})}\bigg[&C\cos\sqrt{L}\bigg(\frac{a+b}{2} - t\bigg)\\
&+ \int_a^t h(s)\cos\sqrt{L}\bigg(\frac{b-a}{2} + s - t\bigg)\,\mathrm{d}s\\
&+ \int_t^b h(s)\cos\sqrt{L}\bigg(\frac{b-a}{2} + t - s\bigg)\,\mathrm{d}s\bigg].
\end{aligned}$$

Hence, using (4.27) and denoting $D = 2\sqrt{L}\sin\sqrt{L}(\frac{b-a}{2})$, we compute

$$\begin{aligned}
&(M - L)w(t) + N\big|w'(t)\big|\\
&\quad \leqslant \frac{1}{D}\bigg[C\bigg[(M - L)\cos\sqrt{L}\bigg(\frac{a+b}{2} - t\bigg) + N\sqrt{L}\bigg|\sin\sqrt{L}\bigg(\frac{a+b}{2} - t\bigg)\bigg|\bigg]\\
&\qquad + \int_a^t h(s)\bigg[(M - L)\cos\sqrt{L}\bigg(\frac{b-a}{2} + s - t\bigg)\\
&\qquad\qquad + N\sqrt{L}\bigg|\sin\sqrt{L}\bigg(\frac{b-a}{2} + s - t\bigg)\bigg|\bigg]\,\mathrm{d}s\\
&\qquad + \int_t^b h(s)\bigg[(M - L)\cos\sqrt{L}\bigg(\frac{b-a}{2} + t - s\bigg)\\
&\qquad\qquad + N\sqrt{L}\bigg|\sin\sqrt{L}\bigg(\frac{b-a}{2} + t - s\bigg)\bigg|\bigg]\,\mathrm{d}s\bigg]\\
&\quad \leqslant 0.
\end{aligned}$$

Hence α_{n+1} is a lower solution.

Inductive step – 2nd part: *assume* (a) *and* (b) *hold for some* n *and let us prove that* $\alpha_{n+2} \leqslant \alpha_{n+1}$. The function $w = \alpha_{n+1} - \alpha_{n+2}$ satisfies (4.29), where

$$h(t) := \alpha''_{n+1}(t) - f\bigl(t, \alpha_{n+1}(t), \alpha'_{n+1}(t)\bigr) \quad \text{and} \quad C = 0.$$

From the previous step $h(t) \geqslant 0$ and the claim follows from the antimaximum principle (Proposition 4.9).

B: *Claim. Let* $L > 0$ *satisfy* (4.27). *Then the functions* β_n *defined recursively by* (4.26) *are such that for all* $n \in \mathbb{N}$,

(a) β_n *is an upper solution, i.e.*,

$$\begin{aligned} &\beta''_n(t) \leqslant f\bigl(t, \beta_n(t), \beta'_n(t)\bigr),\\ &\beta_n(a) = \beta_n(b),\ \beta'_n(a) \leqslant \beta'_n(b), \end{aligned}$$

(b) $\beta_{n+1} \geqslant \beta_n$.

The proof of this claim parallels the proof of Claim A.

C: *Claim.* $\alpha_n \geqslant \beta_n$. Define, for all $i \in \mathbb{N}$, $w_i = \alpha_i - \beta_i$ and

$$h_i(t) := f\bigl(t, \alpha_i(t), \alpha'_i(t)\bigr) - f\bigl(t, \beta_i(t), \beta'_i(t)\bigr) + L\bigl(\alpha_i(t) - \beta_i(t)\bigr).$$

The proof of the claim is by induction.

Initial step: $\alpha_1 \geqslant \beta_1$. The function w_1 is a solution of (4.29) with $h = h_0 \geqslant 0$ and $C = 0$. Using the antimaximum principle (Proposition 4.9), we deduce that $w_1 \geqslant 0$, i.e., $\alpha_1 \geqslant \beta_1$.

Inductive step: *Let* $n \geqslant 2$. *If* $h_{n-2} \geqslant 0$ *and* $\alpha_{n-1} \geqslant \beta_{n-1}$, *then* $h_{n-1} \geqslant 0$ *and* $\alpha_n \geqslant \beta_n$. First, let us prove that, for all $t \in [a, b]$, the function h_{n-1} is nonnegative. Indeed, we have

$$\begin{aligned} h_{n-1} &= f(\cdot, \alpha_{n-1}, \alpha'_{n-1}) - f(\cdot, \beta_{n-1}, \beta'_{n-1}) + L(\alpha_{n-1} - \beta_{n-1})\\ &\geqslant -M(\alpha_{n-1} - \beta_{n-1}) - N|\alpha'_{n-1} - \beta'_{n-1}| + L(\alpha_{n-1} - \beta_{n-1})\\ &= (L - M)w_{n-1} - N|w'_{n-1}|. \end{aligned}$$

Recall that w_{n-1} is a solution of (4.29) with $h(t) = h_{n-2}(t) \geqslant 0$ and $C = 0$. Hence, we can proceed as in the proof of Claim A to show that $h_{n-1} \geqslant 0$. It follows then from the antimaximum principle (Proposition 4.9) that w_n is nonnegative, i.e., $\alpha_n \leqslant \beta_n$.

D: *Claim. There exists* $R > 0$ *such that any solution* u *of*

$$u'' \leqslant f(t, u, u'), \quad u(a) = u(b),\ u'(a) = u'(b),$$

with $\beta \leqslant u \leqslant \alpha$ *satisfies* $\|u'\|_\infty < R$. We deduce from the assumptions that

$$u'' = f(t, u, u') + h(t),$$

where $h(t) \leqslant 0$ and $f(t, u, u') + h(t) \leqslant \max_E |f(t, u, 0)| + N|u'|$. The proof follows now using Proposition 1.7.

E: *Claim. There exists* $R > 0$ *such that any solution* u *of*

$$u'' \geqslant f(t, u, u'), \quad u(a) = u(b),\ u'(a) = u'(b),$$

with $\beta \leqslant u \leqslant \alpha$ *satisfies* $\|u'\|_\infty < R$. The proof repeats the argument of Claim D but uses the remark following Proposition 1.7.

F: *Conclusion.* We deduce from Theorems 4.2 and 4.1, where α and β have to be interchanged, that the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ converge monotonically in $\mathcal{C}^1([a,b])$ to functions u and v such that $\beta \leqslant v \leqslant \alpha$ and $\beta \leqslant u \leqslant \alpha$. Further Claim C implies $v \leqslant u$. □

4.4. *A mixed approximation scheme*

In this section, we consider a derivative independent Dirichlet problem

$$u'' = f(t, u, u), \quad u(a) = 0,\ u(b) = 0. \tag{4.30}$$

Here we write the nonlinearity so that $f(t, u, v)$ is nonincreasing in u and nondecreasing in v. We work out an approximation scheme which provides only bounds on the solutions. However, we shall give assumptions which imply these bounds to be equal, so that they are solutions.

Let us introduce the following definition.

DEFINITION 4.1. Functions α and $\beta \in \mathcal{C}([a,b])$ are *coupled lower and upper quasi-solutions* of (4.30) if

(a) for any $t \in [a,b]$, $\alpha(t) \leqslant \beta(t)$;

(b) for any $t_0 \in]a,b[$, either $D_-\alpha(t_0) < D^+\alpha(t_0)$ or there exists an open interval $I_0 \subset]a,b[$ such that $t_0 \in I_0$, $\alpha \in W^{2,1}(I_0)$, and for a.e. $t \in I_0$

$$\alpha''(t) \geqslant f\big(t, \alpha(t), \beta(t)\big);$$

(c) for any $t_0 \in]a,b[$, either $D^-\beta(t_0) > D_+\beta(t_0)$ or there exists an open interval $I_0 \subset]a,b[$ such that $t_0 \in I_0$, $\beta \in W^{2,1}(I_0)$, and for a.e. $t \in I_0$

$$\beta''(t) \leqslant f\big(t, \beta(t), \alpha(t)\big);$$

(d) $\alpha(a) \leqslant 0 \leqslant \beta(a)$, $\alpha(b) \leqslant 0 \leqslant \beta(b)$.

Consider the following auxiliary problem

$$\begin{aligned} u'' &= f(t, u, v), \quad u(a) = 0,\ u(b) = 0,\\ v'' &= f(t, v, u), \quad v(a) = 0,\ v(b) = 0. \end{aligned} \tag{4.31}$$

PROPOSITION 4.12. *Let* α_0, $\beta_0 \in \mathcal{C}([a,b])$,

$$E := \{(t,u,v) \mid t \in [a,b],\ u,v \in [\alpha_0(t), \beta_0(t)]\}$$

and $f : E \to \mathbb{R}$ *be an* L^1*-Carathéodory function such that* $f(t,u,v)$ *is nonincreasing in* u *and nondecreasing in* v. *Assume* α_0 *and* β_0 *are coupled lower and upper quasi-solutions of* (4.30). *Then, the sequences* $(\alpha_n)_n$ *and* $(\beta_n)_n$, *defined for* $n \geqslant 1$ *by*

$$\alpha_n'' = f(t, \alpha_{n-1}, \beta_{n-1}), \quad \alpha_n(a) = 0,\ \alpha_n(b) = 0,$$
$$\beta_n'' = f(t, \beta_{n-1}, \alpha_{n-1}), \quad \beta_n(a) = 0,\ \beta_n(b) = 0,$$

converge in $\mathcal{C}^1([a,b])$ *to functions* $u_{\min}$ *and* $u_{\max}$. *The pair* $(u_{\min}, u_{\max})$ *is a solution of* (4.31) *such that*

$$\alpha_0 \leqslant u_{\min} \leqslant u_{\max} \leqslant \beta_0.$$

Moreover, any solution (u,v) *of* (4.31) *with* $\alpha_0 \leqslant u \leqslant \beta_0$, $\alpha_0 \leqslant v \leqslant \beta_0$ *is such that*

$$u_{\min} \leqslant u \leqslant u_{\max}, \qquad u_{\min} \leqslant v \leqslant u_{\max}.$$

PROOF. Let $X = \mathcal{C}^1([a,b]) \times \mathcal{C}^1([a,b])$, $Z = \mathcal{C}([a,b]) \times \mathcal{C}([a,b])$, $K = \{(u,v) \in Z \mid u \geqslant 0, v \leqslant 0\}$ and $\mathcal{E} = \{(u,v) \in X \mid \alpha_0 \leqslant u \leqslant \beta_0,\ \alpha_0 \leqslant v \leqslant \beta_0\}$. We define $T : \mathcal{E} \to X$, $(u,v) \mapsto T(u,v)$, where $T(u,v)$ is the solution (x,y) of

$$x'' = f(t,u,v), \quad x(a) = 0,\ x(b) = 0,$$
$$y'' = f(t,v,u), \quad y(a) = 0,\ y(b) = 0.$$

Next, we verify that T is continuous, monotone increasing, $T(\mathcal{E})$ is relatively compact in X and

$$(\alpha, \beta) \leqslant T(\alpha, \beta), \qquad (\beta, \alpha) \geqslant T(\beta, \alpha).$$

Now, Theorem 4.3 applies with $\alpha = (\alpha_0, \beta_0)$ and $\beta = (\beta_0, \alpha_0)$, and the claims follow. □

Notice that if u is a solution of the given problem (4.30), then (u,u) is a solution of the auxiliary problem (4.31), whence $u_{\min}$ and $u_{\max}$ are bounds on solutions of (4.30). The following proposition proves, under appropriate assumptions, convergence of the sequences defined in Proposition 4.12 to the unique solution of the given problem (4.30).

THEOREM 4.13. *Suppose the assumptions of Proposition* 4.12 *hold. Assume moreover*

(i) *there exists* $\epsilon > 0$ *such that* $\alpha_0 \geqslant \epsilon\beta_0$;

(ii) *for every* $s \in [\epsilon, 1[$, *almost every* $t \in [a,b]$ *and every* u, $v \in [\alpha_0, \beta_0]$ *with* $sv \leqslant u \leqslant v$,

$$sf\left(t, \frac{u}{s}, sv\right) > f(t,u,v).$$

Then, the functions $u_{\min}$ *and* $u_{\max}$ *defined in Proposition* 4.12 *are equal, i.e., are solutions of* (4.30).

PROOF. From assumption (i), we deduce

$$\epsilon u_{\max} \leqslant u_{\min} \leqslant u_{\max}.$$

Let $s_0 = \sup\{s \mid s u_{\max} \leqslant u_{\min}\}$. It is obvious that $s_0 \in [\epsilon, 1]$ and that $s_0 u_{\max} \leqslant u_{\min}$. From the definition of s_0, we deduce the existence of $t_0 \in [a, b]$ such that

$$u_{\min}(t_0) - s_0 u_{\max}(t_0) = 0, \qquad u'_{\min}(t_0) - s_0 u'_{\max}(t_0) = 0.$$

If $t_0 \neq b$, we also have $t_1 > t_0$ such that $u'_{\min}(t_1) - s_0 u'_{\max}(t_1) = 0$.

Assume now that $s_0 < 1$. Hence, we can write

$$\begin{aligned} s_0 u''_{\max} &= s_0 f(\,\cdot\,, u_{\max}, u_{\min}) \geqslant s_0 f\left(\,\cdot\,, \frac{1}{s_0} u_{\min}, s_0 u_{\max}\right) \\ &> f(\,\cdot\,, u_{\min}, u_{\max}) = u''_{\min}, \end{aligned}$$

which leads to the contradiction

$$0 = (u'_{\min} - s_0 u'_{\max})|_{t_0}^{t_1} = \int_{t_0}^{t_1} \big(u''_{\min}(t) - s_0 u''_{\max}(t)\big)\,\mathrm{d}t < 0.$$

A similar argument holds if $t_0 = b$. Hence $s_0 = 1$ and $u_{\max} = u_{\min}$. □

4.5. *Historical and bibliographical notes*

The idea of associating to a pair of well-ordered lower and upper solutions, monotone sequences of lower and upper solutions converging to solutions is far older than the theory presented here. It goes back at least to Picard whose contribution can be found in two "mémoires", the first one [79] in 1890 and the second one [80] in 1893. In [80], the author considers the problem

$$u'' + f(t, u) = 0, \quad u(a) = 0, \; u(b) = 0, \tag{4.32}$$

in case $f(t, \cdot)$ is increasing. He exhibits then a function α_0 such that

$$\alpha_0'' + f(t, \alpha_0) > 0, \quad \text{on }]a, b[, \qquad \alpha_0(a) = 0, \qquad \alpha_0(b) = 0,$$

i.e., a lower solution, and considers the iterations

$$\alpha_n'' + f(t, \alpha_{n-1}) = 0, \quad \alpha_n(a) = 0, \; \alpha_n(b) = 0.$$

At last he observes that the sequence $(\alpha_n)_n$ is increasing and converges to a solution u of (4.32) greater than α_0. This gives an approximation process to compute the solutions of (4.32) that we will call *the first monotone approximation scheme*. Theorems 4.4 and 4.7 are very close to this approach.

In [79], Picard introduces a *second monotone approximation scheme*. He studies an elliptic Dirichlet problem in the opposite situation where $f(t,\cdot)$ is decreasing. His main result gives bounds on the solutions of (4.32). Under appropriate assumptions on f, he defines $\alpha_0 = 0$ and considers, for $n \geqslant 1$, the solutions α_n and β_n of the following problems

$$\begin{aligned} &\beta_n'' + f(t,\alpha_{n-1}) = 0, \quad \beta_n(a) = 0,\ \beta_n(b) = 0,\\ &\alpha_n'' + f(t,\beta_n) = 0, \quad \alpha_n(a) = 0,\ \alpha_n(b) = 0. \end{aligned}$$

Here $\alpha_n \leqslant \beta_n$, $(\alpha_n)_n$ is increasing and $(\beta_n)_n$ is decreasing. Hence the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ converge pointwise respectively to functions u and v with $u \leqslant v$. The author proves in 1898 in case of ODE [83] and in 1900 for the PDE [84], that the convergence is uniform and that the limit functions u and v satisfy

$$\begin{aligned} &u'' + f(t,v) = 0, \quad u(a) = 0,\ u(b) = 0,\\ &v'' + f(t,u) = 0, \quad v(a) = 0,\ v(b) = 0. \end{aligned}$$

Moreover the solution z of (4.32), which is unique under the given assumptions, is such that

$$u \leqslant z \leqslant v.$$

At last, Picard provides in 1894 [81] (see also [82]) an example of a problem (4.32) such that $u \neq v$ which shows that this second approximation scheme does not necessarily converge to a solution of (4.32). This method is described in Proposition 4.12 and Theorem 4.13 which come from [20].

Following Chaplygin [16], the Russian school studies the monotone iterative methods extensively. In 1954, Babkin [9] considers the problem (4.32) under assumptions on $f(t,u)$ which imply the uniqueness of the solutions of (4.32). In his approach, he considers two approximation sequences. Starting from lower and upper solutions α_0 and $\beta_0 \geqslant \alpha_0$, these approximations are obtained (for $n \geqslant 1$) as solutions of the linear problem

$$\begin{aligned} &-\alpha_n'' + K\alpha_n = f(t,\alpha_{n-1}) + K\alpha_{n-1}, \quad \alpha_n(a) = 0,\ \alpha_n(b) = 0,\\ &-\beta_n'' + K\beta_n = f(t,\beta_{n-1}) + K\beta_{n-1}, \quad \beta_n(a) = 0,\ \beta_n(b) = 0. \end{aligned}$$

The main assumption to prove that the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ are monotone and convergent is to choose $K > 0$ such that the function $f(t,u) + Ku$ is increasing in u so that the corresponding differential operator satisfies a maximum principle such as Proposition 4.5 (see [6] or [85]). The observation that the limits of the sequences $(\alpha_n)_n$ and $(\beta_n)_n$ are extremal solutions was noticed in 1962 by Courant and Hilbert [25].

One important step is due to Kantorovich [59] in 1939. He observes that the first approximation scheme, used for the Cauchy problem associated with ODE as well as for other boundary value problem, has a common structure related to positive operators. He then develops an abstract formulation of the method. In 1959, Collatz and Schröder [23] give an abstract formulation of the second monotone approximation scheme. These two abstract formulations have been unified in 1960 by Schröder [93] who shows that the second scheme can be reduced to the first one. Our first section takes into account such an abstract formulation.

The study of the monotone iterative methods for nonlinearities depending on the derivative was initiated in 1964 by Gendzhoyan [43] who considers the problem

$$u'' + f(t, u, u') = 0, \quad u(a) = 0,\ u(b) = 0.$$

Starting from lower and upper solutions α_0 and $\beta_0 \geqslant \alpha_0$, he defines sequences of approximations $(\alpha_n)_n$, $(\beta_n)_n$ as solutions of

$$\begin{aligned}
&-\alpha_n'' + l(t)\alpha_n' + k(t)\alpha_n = f(t, \alpha_{n-1}, \alpha_{n-1}') + l(t)\alpha_{n-1}' + k(t)\alpha_{n-1},\\
&\alpha_n(a) = 0,\ \alpha_n(b) = 0,\\
&-\beta_n'' + l(t)\beta_n' + k(t)\beta_n = f(t, \beta_{n-1}, \beta_{n-1}') + l(t)\beta_{n-1}' + k(t)\beta_{n-1},\\
&\beta_n(a) = 0,\ \beta_n(b) = 0,
\end{aligned}$$

where $k(t)$ and $l(t)$ are functions related to the assumptions on f. Here also, the convergence is monotone and gives approximations of the solution together with some error bounds. Our Theorems 4.6 and 4.8 simplify considerably this approach. These can be found in [21].

All the above-quoted papers consider the usual order $\alpha \leqslant \beta$ for the lower and upper solutions. The monotone iterative method was also developed in case lower and upper solutions appear in the reversed order, i.e., $\alpha \geqslant \beta$. We can first quote the paper of Omari and Trombetta [75] in 1992. They consider problems such as

$$-u'' + cu' + f(t, u) = 0, \quad u(a) = u(b),\ u'(a) = u'(b).$$

The key assumptions are that the function $f(t, u) - \lambda u$ is nondecreasing in u for some $\lambda < 0$ and that this λ is such that the operator $-u'' + cu' + \lambda u$ is inverse negative on the space of periodic functions, i.e., that an antimaximum principle holds (see Proposition 4.9 [75]). Theorem 4.10 is a particular case of the results in [75]. The Neumann problem was considered by Cabada and Sanchez [15]. We also refer to [14] for other results in this direction. Recently Cherpion, De Coster and Habets [21] worked out an approach of derivative dependent problems very much along the lines of Theorems 4.6 and 4.8. This result is Theorem 4.11.

References

[1] A. Adje, *Sur et sous-solutions généralisées et problèmes aux limites du second ordre*, Bull. Soc. Math. Belg. Ser. B **42** (1990), 347–368.

[2] K. Akô, *On the Dirichlet problem for quasi-linear elliptic differential equations of the second order*, J. Math. Soc. Jpn. **13** (1961), 45–62.

[3] K. Akô, *Subfunctions for ordinary differential equations II*, Funkcialaj Ekvacioj **10** (1967), 145–162.

[4] H. Amann, *On the existence of positive solutions of nonlinear elliptic boundary value problems*, Indiana Univ. Math. J. **21** (1971), 125–146.

[5] H. Amann, *On the number of solutions of nonlinear equations in ordered Banach spaces*, J. Funct. Anal. **11** (1972), 346–384.

[6] H. Amann, *Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces*, SIAM Rev. **18** (1976), 620–709.

[7] H. Amann, *A note on degree theory for gradient mappings*, Proc. Amer. Math. Soc. **85** (1982), 591–595.

[8] H. Amann, A. Ambrosetti and G. Mancini, *Elliptic equations with noninvertible Fredholm linear part and bounded nonlinearities*. Math. Z. **158** (1978), 179–194.

[9] B.N. Babkin, *Solution of a boundary value problem for an ordinary differential equation of second order by Caplygin's method*, Prikl. Mat. Mekh. Akad. Nauk SSSR **18** (1954), 239–242.

[10] T. Bartsch and Z.-Q. Wang, *On the existence of sign-changing solutions for semilinear Dirichlet problems*, Topol. Methods Nonlinear Anal. **7** (1996), 115–131.

[11] S. Bernstein, *Sur certaines équations différentielles ordinaires du second ordre*, C. R. Acad. Sci. Paris **138** (1904), 950–951.

[12] H. Brezis, *Analyse fonctionnelle, Théorie et applications*, Masson, Paris (1983).

[13] H. Brezis and L. Nirenberg, *H^1 versus $\mathcal{C}^1$ local minimizers*, C. R. Acad. Sci. Paris **317** (1993), 465–472.

[14] A. Cabada, P. Habets and R. Pouso, *Optimal existence conditions for Φ-Laplacian equations with upper and lower solutions in the reversed order*, J. Differential Equations **166** (2000), 385–401.

[15] A. Cabada and L. Sanchez, *A positive operator approach to the Neumann problem for a second order ordinary differential equation*, J. Math. Anal. Appl. **204** (1996), 774–785.

[16] S.A. Chaplygin, *Complete Works III*, Akademia Nauk SSSR, Leningrad (1935).

[17] K.C. Chang, *A variant mountain pass lemma*, Scientia Sinica (Ser. A) **26** (1983), 1241–1255.

[18] K.C. Chang, *Variational methods and sub- and super-solutions*, Scientia Sinica (Ser. A) **26** (1983), 1256–1265.

[19] K.C. Chang, *Infinite Dimensional Morse Theory and Multiple Solution Problems*, Birkhäuser, Boston (1993).

[20] M. Cherpion, C. De Coster and P. Habets, *Monotone iterative methods for boundary value problems*, Differential Integral Equations **12** (1999), 309–338.

[21] M. Cherpion, C. De Coster and P. Habets, *A constructive monotone iterative method for second order BVP in presence of lower and upper solutions*, Appl. Math. Comput. **123** (2001) 75–91.

[22] J. Ángel Cid, *On extremal fixed points in Schauder's theorem with applications to differential equations*, Bull. Belgian Math. Soc. **11** (2004), 15–20.

[23] L. Collatz and J. Schröder, *Einschliessen der Lösungen von Randwertaufgaben*, Numer. Math. **1** (1959), 61–72.

[24] M. Conti, L. Merizzi and S. Terracini, *Remarks on variational methods and lower–upper solutions*, NoDEA Nonlinear Differential Equations Appl. **6** (1999), 371–393.

[25] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol 2: Partial Differential Equations*, Interscience, New York–London (1962).

[26] E.N. Dancer and Y. Du, *A note on multiple solutions of some semilinear elliptic problems*, J. Math. Anal. Appl. **211** (1997), 626–640.

[27] C. De Coster, *Pairs of positive solutions for the one-dimensional p-Laplacian*, Nonlinear Anal. TMA **23** (1994), 669–681.

[28] C. De Coster, *Lower and upper solutions for singular derivative dependent Dirichlet problem*, Math. Inequalities Appl. **4** (2001), 377–396.

[29] C. De Coster, M.R. Grossinho and P. Habets, *On pairs of positive solutions for a singular boundary value problem*, Appl. Anal. **59** (1995), 241–256.

[30] C. De Coster and P. Habets, *A two parameters Ambrosetti–Prodi problem*, Portugaliae Math. **53** (1996), 279–303.

[31] C. De Coster and P. Habets, *Upper and lower solutions in the theory of ODE boundary value problems: classical and recent results*, Nonlinear Analysis and Boundary Value Problems for Ordinary Differential Equations, F. Zanolin, ed., CISM Courses and Lectures, Vol. 371, Springer-Verlag, New York (1996), 1–79.

[32] C. De Coster and P. Habets, *An overview of the method of lower and upper solutions for ODEs*, Nonlinear Analysis and its Applications to Differential Equations, M.R. Grossinho, M. Ramos, C. Rebelo and L. Sanchez, eds., Progress in Nonlinear Differential Equations and their Applications, Vol. 43, Birkhauser, Boston (2001), 3–22.

[33] C. De Coster and P. Habets, *On the Method of Lower and Upper Solutions in the Study of Boundary Value Problems*, to appear.

[34] C. De Coster and M. Henrard, *Existence and localization of solution for elliptic problem in presence of lower and upper solutions without any order*, J. Differential Equations **145** (1998), 420–452.

[35] C. De Coster and P. Omari, *Unstable periodic solutions of a parabolic problem in presence of non-well-ordered lower and upper solutions*, J. Functional Anal. **175** (2000), 52–88.

[36] C. De Coster and M. Tarallo, *Foliations, associated reductions and lower-upper solutions*, Calculus Variations PDE **15** (2002), 25–44.

[37] D.G. de Figueiredo, *Lectures on the Ekeland Variational Principle with Applications and Detours*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics, Vol. 81, Springer-Verlag, Berlin (1989).

[38] D.G. de Figueiredo and S. Solimini, *A variational approach to superlinear elliptic problems*, Comm. Partial Differential Equations **9** (1984), 699–717.

[39] T. Ding, R. Iannacci and F. Zanolin, *Existence and multiplicity results for periodic solutions of semilinear Duffing equations*, J. Differential Equations **6** (1993), 364–409.

[40] H. Epheser, *Über die Existenz der Lösungen von Randwertaufgaben mit gewöhnlichen, nichtlinearen Differentialgleichungen zweiter Ordnung*, Math. Z. **61** (1955), 435–454.

[41] C. Fabry and P. Habets, *Upper and lower solutions for second-order boundary value problems with nonlinear boundary conditions*, Nonlinear Anal. TMA **10** (1986), 985–1007.

[42] M. Gaudenzi, P. Habets and F. Zanolin, *Positive solutions of superlinear boundary value problems with singular indefinite weight*, Bull. Belgian Math. Soc. **9** (2002), 607–619.

[43] G.V. Gendzhoyan, *On two-sided Chaplygin approximations to the solution of the two point boundary value problem*, Izv. Arm. SSR Fiz. Mat. Nauk **17** (1964), 21–27.

[44] J.-P. Gossez, *Some nonlinear differential equations with resonance at the first eigenvalue*, Confer. Sem. Mat. Univ. Bari **167** (1979), 355–389.

[45] J.-P. Gossez and P. Omari, *Non-ordered lower and upper solutions in semilinear elliptic problems*, Comm. PDE **19** (1994), 1163–1184.

[46] V.V. Gudkov and A.J. Lepin, *On necessary and sufficient conditions for the solvability of certain boundary-value problems for a second-order ordinary differential equation*, Dokl. Akad. Nauk SSSR **210** (1973), 800–803.

[47] P. Habets and M. Laloy, *Perturbations singulières de problèmes aux limites: Les sur- et sous-solutions*, Sém. Math. Appl. Mécanique UCL **76** (1974).

[48] P. Habets and P. Omari, *Existence and localization of solutions of second order elliptic problems using lower and upper solutions in the reversed order*, Topol. Methods Nonlinear Anal. **8** (1996), 25–56.

[49] P. Habets and P. Omari, *Positive solutions of an indefinite prescribed mean curvature problem on a general domain*, Adv. Nonlinear Stud. **4** (2004), 1–13.

[50] P. Habets and R. Pouso, *Examples of non existence of solution in presence of upper and lower solutions*, ANZIAM J. **44** (2003), 591–594.

[51] P. Habets and L. Sanchez, *Periodic solutions of some Liénard equations with singularities*, Proc. Amer. Math. Soc. **109** (1990), 1035–1044.

[52] P. Habets and P. Torres, *Some multiplicity for periodic solutions of a Rayleigh differential equation*, Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal. **8** (2001), 335–351.

[53] P. Habets and F. Zanolin, *Upper and lower solutions for a generalized Emden–Fowler equation*, J. Math. Anal. Appl. **181** (1994), 684–700.

[54] P. Habets and F. Zanolin, *Positive solutions for a class of singular boundary value problem*, Boll. Un. Mat. Ital. (7) **9A** (1995), 273–286.

[55] P. Hess, *On the solvability of nonlinear elliptic boundary value problems*, Indiana Univ. Math. J. **25** (1976), 461–466.

[56] H. Hofer, *Variational and topological methods in partially ordered Hilbert spaces*, Math. Ann. **261** (1982), 493–514.

[57] H. Hofer, *A note on the topological degree at a critical point of mountainpass-type*, Proc. Amer. Math. Soc. **90** (1984), 309–315.

[58] L.K. Jackson, *Subfunctions and second-order ordinary differential inequalities*, Adv. Math. **2** (1967), 307–363.

[59] L. Kantorovich, *The method of successive approximations for functional equations*, Acta Math. **71** (1939), 63–97.

[60] J.L. Kelley, *General Topology*, Graduate Texts in Mathematics, Vol. 27, Springer-Verlag, New York (1955).

[61] I.T. Kiguradze, *A priori estimates for derivatives of bounded functions satisfying second-order differential inequalities*, Differentsial'nye Uravneniya **3** (1967), 1043–1052.

[62] I.T. Kiguradze, *Some singular boundary value problems for ordinary nonlinear second-order differential equations*, Differentsial'nye Uravneniya **4** (1968), 1753–1773.

[63] I.T. Kiguradze and B.L. Shekhter, *Singular boundary-value problems for ordinary second-order differential equations*, Itogi Nauki i Tekhniki, Seriya Sovremennye Problemy Matematiki, Noveishie Dostizheniya **30** (1987), 105–201, translated in J. Soviet Math. **43** (1988), 2340–2417.

[64] H.W. Knobloch, *Eine neue Methode zur Approximation periodischer Lösungen nichtlinearer Differentialgleichungen zweiter Ordnung*, Math. Z. **82** (1963), 177–197.

[65] Bongsoo Ko, *The third solution of semilinear elliptic boundary value problems and applications to singular perturbation problems*, J. Differential Equations **101** (1993), 1–14.

[66] Y.S. Kolesov, *Periodic solutions of quasilinear parabolic equations of second order*, Trans. Moscow Math. Soc. **21** (1970), 114–146.

[67] S.J. Li and T. Wang, *Mountain pass theorem in order intervals and multiple solutions for semilinear elliptic Dirichlet problems*, J. Anal. Math. **81** (2000), 373–396.

[68] Zhaoli Liu and Jingxian Sun, *Invariant sets of descending flow in critical point theory with applications to nonlinear differential equations*, J. Differential Equations **172** (2001), 257–299.

[69] N.G. Lloyd, *Degree Theory*, Cambridge Univ. Press, Cambridge (1978).

[70] J. Mawhin, *Points fixes, points critiques et problèmes aux limites*, Sém. de Math. Supérieures, Univ. Montréal (1985).

[71] W. Mlak, *Parabolic differential inequalities and Chaplighin's method*, Ann. Polon. Math. **8** (1960), 139–152.

[72] M. Müller, *Über das Fundamentaltheorem in der Theorie der gewöhnlichen Differentialgleichungen*, Math. Z. **26** (1926), 619–649.

[73] M. Nagumo, *Über die Differentialgleichung* $y'' = f(t, y, y')$, Proc. Phys-Math. Soc. Jpn. **19** (1937), 861–866.

[74] M. Nagumo, *On principally linear elliptic differential equations of the second order*, Osaka Math. J. **6** (1954), 207–229.

[75] P. Omari and M. Trombetta, *Remarks on the lower and upper solutions method for second- and third-order periodic boundary value problem*, Appl. Math. Comput. **50** (1992), 1–21.

[76] P. Omari and F. Zanolin, *Infinitely many solutions of a quasilinear elliptic problem with an oscillatory potential*, Comm. Partial Differential Equations **21** (1996), 721–733.

[77] G. Peano, *Sull'integrabilità delle equazioni differenziali di primo ordine*, Atti Acad. Torino **21** (1885), 677–685.

[78] O. Perron, *Ein neuer Existenzbeweis für die Integrale der Differentialgleichung* $y' = f(x, y)$, Math. Ann. **76** (1915), 471–484.

[79] E. Picard, *Mémoire sur la théorie des équations aux dérivées partielles et la méthode des approximations successives*, J. Math. **6** (1890), 145–210.

[80] E. Picard, *Sur l'application des méthodes d'approximations successives à l'étude de certaines équations différentielles ordinaires*, J. Math. **9** (1893), 217–271.

[81] E. Picard, *Sur un exemple d'approximations successives divergentes*, Comptes Rendus **118** (1894), 899–902.

[82] E. Picard, *Traité d'analyse, III*, Gauthier-Villars, Paris (1896).

[83] E. Picard, *Sur certains exemples singuliers d'approximations successives*, Comptes Rendus **126** (1898), 497–500.

[84] E. Picard, *Sur un exemple d'approximations successives divergentes*, Bull. Soc. Math. **28** (1900), 137–143.

[85] M.H. Protter and H.F. Weinberger, *Maximum Principles in Differential Equations*, Prentice–Hall, Englewood Cliffs (1967).

[86] G. Prodi, *Teoremi di esistenza per equazioni alle derivate parziali non lineari di tipo parabolico, Nota I e II*, Rend. Ist. Lombardo **86** (1953), 1–47.

[87] I. Rachůnková and M. Tvrdý, *Nonlinear systems of differential inequalities and solvability of certain boundary value problems*, J. Inequal. Appl. **6** (2001), 199–226.

[88] A. Rosenblatt, *Sur les théorèmes de M. Picard dans la théorie des problèmes aux limites des équations différentielles ordinaires non linéaires*, Bull. Sci. Math. **57** (1933), 100–106.

[89] T. Satō, *Sur l'équation aux dérivées partielles* $\Delta z = f(x, y, z, p, q)$, Compositio Math. **12** (1954), 157–177.

[90] T. Satō, *Sur l'équation aux dérivées partielles* $\Delta z = f(x, y, z, p, q)$ *II*, Compositio Math. **14** (1959), 152–171.

[91] D.H. Sattinger, *Monotone methods in nonlinear elliptic and parabolic boundary value problems*, Indiana Univ. Math. J. **21** (1972), 979–1000.

[92] K. Schmitt, *Boundary value problems for quasilinear second order elliptic equations*, Nonlinear Anal. TMA **2** (1978), 263–309.

[93] J. Schröder, *Anwendungen von Fixpunktsätzen bei der numerischen Behandlung nichtlinearer Gleichungen in halbgeordneten Räumen*, Arch. Rat. Mech. Anal. **4** (1959/1960), 177–192.

[94] G. Scorza Dragoni, *Il problema dei valori ai limiti studiato in grande per gli integrali di una equazione differenziale del secondo ordine*, Giornale Mat. (Battaglini) **69** (1931), 77–112.

[95] G. Scorza Dragoni, *Il problema dei valori ai limiti studiato in grande per le equazioni differenziali del secondo ordine*, Math. Ann. **105** (1931), 133–143.

[96] G. Scorza Dragoni, *Elementi uniti di transformazioni funzionali e problemi di valori ai limiti*, Rend. Semin. Mat. R. Univ. Roma **2** (1938), 255–275.

[97] G. Scorza Dragoni, *Intorno a un criterio di esistenza per un problema di valori ai limiti*, Rend. Semin. R. Accad. Naz. Lincei **28** (1938), 317–325.

[98] R. Shivaji, *A remark on the existence of three solutions via sub-super solutions*, Nonlinear Analysis and Applications, V. Lakshmikantham, ed., Dekker, New York–Basel (1987), 561–566.

[99] G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier **15** (1965), 189–258.

[100] J. Szarski, *Differential Inequalities*, Monografie Matematyczne, Vol. 43, Warszawa (1965).

[101] W. Walter, *Differential and Integral Inequalities*, Springer-Verlag, New York (1970).

CHAPTER 3

Half-Linear Differential Equations

O. Došlý

Masaryk University, Department of Mathematics, Janackovo nam. 2A, 662 95 Brno, Czech Republic
E-mail: dosly@math.muni.cz

Contents

Preface 165
3A. Basic Theory 167
  1. Existence and uniqueness 167
    1.1. First order half-linear system 167
    1.2. Half-linear trigonometric functions 168
    1.3. Half-linear Prüfer transformation 169
    1.4. Existence and uniqueness 170
  2. Sturmian theory 171
    2.1. Riccati equation 171
    2.2. Picone's identity 172
    2.3. Energy functional 174
    2.4. Roundabout theorem 174
    2.5. Sturmian separation and comparison theorems 176
    2.6. Transformation of independent variable 177
    2.7. Reciprocity principle 178
    2.8. Leighton–Wintner oscillation criterion 178
  3. Differences between linear and half-linear equations 180
    3.1. Wronskian identity 180
    3.2. Transformation formula 181
    3.3. Fredholm alternative 182
  4. Some elementary half-linear equations 183
    4.1. Equations with constant coefficients 183
    4.2. Euler-type half-linear differential equation 188
    4.3. Kneser-type oscillation and nonoscillation criteria 194
3B. Nonoscillatory Equations 197
  5. Nonoscillation criteria 197
    5.1. Variational principle and Wirtinger's inequality 197
    5.2. Nonoscillation criteria via Wirtinger inequality 198
    5.3. Riccati inequality 200
    5.4. Half-linear Hartman–Wintner theorem 203

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 1
Edited by A. Cañada, P. Drábek and A. Fonda
© 2004 Elsevier B.V. All rights reserved

- 5.5. Riccati integral equation and Hille–Wintner comparison theorem . . . 206
- 5.6. Hille–Nehari criteria . . . 208
- 5.7. Modified Hille–Nehari's criteria . . . 209
- 5.8. Comparison theorem with respect to p . . . 214

6. Asymptotic of nonoscillatory solutions . . . 215
    - 6.1. Integral conditions and classification of solutions . . . 215
    - 6.2. The case c negative . . . 218
    - 6.3. Uniqueness in $\mathbb{M}^-$. . . . 224
    - 6.4. The case c positive . . . 226
7. Principal solution . . . 229
    - 7.1. Principal solution of linear equations . . . 229
    - 7.2. Mirzov's construction of the principal solution . . . 231
    - 7.3. Construction of Elbert and Kusano . . . 232
    - 7.4. Comparison theorem for eventually minimal solutions of Riccati equations . . . 234
    - 7.5. Sturmian property of the principal solution . . . 235
    - 7.6. Integral characterization of the principal solution . . . 236
    - 7.7. Another integral characterization . . . 240
    - 7.8. Limit characterization of the principal solution . . . 241
8. Conjugacy and disconjugacy of half-linear equations . . . 241
    - 8.1. Leighton's conjugacy criterion . . . 241
    - 8.2. Singular Leighton's theorem . . . 242
    - 8.3. Lyapunov inequality . . . 246
    - 8.4. Vallée Poussin-type inequality . . . 248
    - 8.5. Focal point criteria . . . 249
    - 8.6. Lyapunov-type focal points and conjugacy criteria . . . 252

3C. Oscillatory Equations . . . 255

9. Oscillation criteria . . . 255
    - 9.1. General observations . . . 255
    - 9.2. Coles-type criteria . . . 256
    - 9.3. Generalized Hartman–Wintner's criterion . . . 261
    - 9.4. Generalized Kamenev criterion . . . 263
    - 9.5. Another refinement of the Hartman–Wintner theorem . . . 264
    - 9.6. Half-linear Willet's criteria . . . 266
    - 9.7. Equations with periodic coefficient . . . 268
    - 9.8. Equations with almost periodic coefficient . . . 270
    - 9.9. Generalized H-function averaging technique . . . 272
10. Various oscillation problems . . . 275
    - 10.1. Asymptotic formula for distance of zeros of oscillatory solutions . . . 275
    - 10.2. Half-linear Milloux and Armellini–Tonelli–Sansone theorems . . . 279
    - 10.3. Strongly and conditionally oscillatory equation . . . 283
    - 10.4. Oscillation of forced half-linear differential equations . . . 284
    - 10.5. Oscillation of retarded half-linear equations . . . 286
11. Half-linear Sturm–Liouville problem . . . 293
    - 11.1. Basic Sturm–Liouville problem . . . 293
    - 11.2. Regular problem with indefinite weight . . . 295
    - 11.3. Singular Sturm–Liouville problem . . . 300
12. Perturbation principle . . . 302
    - 12.1. General idea . . . 302
    - 12.2. Leighton–Wintner type oscillation criterion . . . 303
    - 12.3. Hille–Nehari-type oscillation criterion . . . 304
    - 12.4. Hille–Nehari-type nonoscillation criterion . . . 308
    - 12.5. Perturbed Euler equation . . . 310

3D. Related Equations and Problems . . . 313
- 13. Half-linear boundary value problems . . . 313
  - 13.1. Basic boundary value problem . . . 313
  - 13.2. Variational characterization of eigenvalues . . . 314
  - 13.3. Nonresonance problem . . . 315
  - 13.4. Fredholm alternative for the scalar p-Laplacian . . . 317
  - 13.5. Homotopic deformation along p and Leray–Schauder degree . . . 319
  - 13.6. Resonance problem . . . 321
- 14. Quasilinear and related differential equations . . . 322
  - 14.1. Equation (14.1) with constant coefficients . . . 323
  - 14.2. Emden–Fowler type equation . . . 326
  - 14.3. More about quasilinear equations . . . 327
- 15. Partial differential equations with $\boldsymbol{p}$-Laplacian . . . 329
  - 15.1. Dirichlet BVP with p-Laplacian . . . 330
  - 15.2. Picone's identity for equations with p-Laplacian . . . 333
  - 15.3. Second eigenvalue of p-Laplacian . . . 335
  - 15.4. Equations involving pseudolaplacian . . . 336
- 16. Half-linear difference equations . . . 337
  - 16.1. Roundabout theorem for half-linear difference equations . . . 338
  - 16.2. Discrete Leighton–Wintner criterion . . . 341
  - 16.3. Riccati inequality . . . 343
  - 16.4. Hille–Nehari nonoscillation criterion . . . 345
  - 16.5. Half-linear dynamic equations on time scales . . . 348
- References . . . 349

This Page Intentionally Left Blank

Preface

In this part of the book we deal with the half-linear second order differential equation

$$\big(r(t)\Phi(x')\big)' + c(t)\Phi(x) = 0, \quad \Phi(x) := |x|^{p-1}\operatorname{sgn} x,\ p > 1, \tag{HL}$$

where r, c are continuous functions and $r(t) > 0$. The investigation of solutions of (HL) has attracted considerable attention in the last two decades, and it was shown that solutions of this equation behave in many aspects like those of the Sturm–Liouville equation

$$\big(r(t)x'\big)' + c(t)x = 0 \tag{SL}$$

which is the special case of (HL) when $p = 2$. The aim of this part is to present the substantial results of this investigation. We show similarities in qualitative behavior of solutions of (HL) and (SL), we also point out phenomena where properties of solutions of (HL) and (SL) (considerably) differ. Note that the term half-linear equations is motivated by the fact that the solution space of (HL) has just one half of the properties which characterize linearity, namely homogeneity (but not additivity).

The investigation of qualitative properties of nonlinear second order differential equations has a long history. Recall here only the papers of Emden [95], Fowler [102], Thomas [211], and the book of Sansone [206] containing the survey of the results achieved in the first half of the last century. In the fifties and the later decades the number of papers devoted to nonlinear second order differential equations increased rapidly, so we mention here only treatments directly associated with (HL). Even if some ideas concerning the properties of solutions of (HL) can already be found in the papers of Bihari [20,21], Elbert and Mirzov with their papers [85,176] are the ones usually regarded as pioneers of the qualitative theory of (HL). In later years, in particular in the nineties, the striking similarity between oscillatory properties of (HL) and (SL) was revealed. On the other hand, in some aspects, e.g., the Fredholm-type alternative for solutions of boundary value problems associated with (HL), it turned out that the situation is completely different in the linear and half-linear case, and that the absence of additivity of the solution space of (HL) brings completely new phenomena.

This part of the book is divided into four chapters. In the first one we present a brief survey of the basic properties of solutions of (HL). A particular attention is devoted to the existence, uniqueness, Sturmian theory and to some elementary half-linear differential equations. Then we turn our attention to the oscillation theory of half-linear equations. First we deal with nonoscillatory equations and nonoscillation criteria (Chapter 3B), and in Chapter 3C we deal with their oscillation counterparts. We also present some related results concerning asymptotic behavior of nonoscillatory solutions and properties of some distinguished solutions of (HL). In the last chapter we deal with boundary value problems associated with (HL) (in this part of the qualitative theory of (HL) we can see the biggest difference between linear and half-linear second order differential equations), and with equations related to (HL), in particular, with partial differential equations with p-Laplacian, quasilinear equations and half-linear difference equations.

Comparing our treatment of half-linear differential equations with Chapter 3C of the recent book [3] (this chapter is devoted to the oscillation theory of (HL), there are some common points, but the most part of our presentation differ from that of [3]. More precisely, the treatment analogous to Sections 1.3, 1.4, 2.2, 2.3, 3.3, 3.4 and the whole Chapter 3D are missing in [3]. On the other hand, [3] devotes more space to particular (non)oscillation criteria, to forced half-linear equations and to equations with deviated argument. Moreover, the parts which overlap here and in [3] are presented from a different point of view.

I would like to thank Prof. Mariella Cecchi and Prof. Mauro Marini from the University of Florence, who prepared the section devoted to the asymptotics of nonoscillatory solutions of (HL) (Section 6), and to Assoc. Prof. Jaromír Kuben for his help with the LaTeX elaboration of the text. My special thanks belong to my colleague Dr. Pavel Řehák, who read the whole text, for numerous helpful suggestions and comments to the final version of this text.

The preparation of this part of the book was supported by the Grant 201/01/0079 of the Grant Agency of the Czech Republic and by the Research Project J07/98/143100001 of the Czech Government.

December 2003 Ondřej Došlý

CHAPTER 3A

Basic Theory

In this chapter we deal with the basic properties of solutions of the half-linear second order differential equation

$$\big(r(t)\Phi(x')\big)' + c(t)\Phi(x) = 0, \quad \Phi(x) := |x|^{p-1}\operatorname{sgn} x,\ p > 1. \tag{0.1}$$

We suppose that the functions r, c are continuous and $r(t) > 0$ in the interval under consideration. In the first section we deal with the existence and the unique solvability of (0.1). Then we present the basic oscillatory properties of (0.1), in particular, we show that the linear Sturmian oscillation theory extends almost verbatim to half-linear equations. In Section 3 we show basic differences between second order linear and half-linear equations. The last section of this chapter deals with some special half-linear equations.

1. Existence and uniqueness

1.1. *First order half-linear system*

Consider the Sturm–Liouville *linear* differential equation

$$\big(r(t)x'\big)' + c(t)x = 0 \tag{1.1}$$

which is a special case $p = 2$ in (0.1). Then, given $t_0, x_0, x_1 \in \mathbb{R}$, there exists the unique solution of (1.1) satisfying the initial conditions $x(t_0) = x_0$, $x'(t_0) = x_1$, which is extensible over the whole interval where the functions r, c are continuous and $r(t) > 0$. This follows, e.g., from the fact that (1.1) can be written as the 2-dimensional first order linear system

$$x' = \frac{1}{r(t)}u, \qquad u' = -c(t)x$$

and the linearity (hence Lipschitz property) of this system implies the above mentioned statement concerning the existence and unique solvability of (1.1). On the other hand, if we rewrite (0.1) into the first order system (substituting $u = r\Phi(x')$), we get the system

$$x' = r^{1-q}(t)\Phi^{-1}(u), \qquad u' = -c(t)\Phi(x), \tag{1.2}$$

where q is the conjugate number of p, i.e., $\frac{1}{p} + \frac{1}{q} = 1$, and Φ^{-1} is the inverse function of Φ. The right hand-side of (1.2) is no longer Lipschitzian in x, u, hence the standard existence and uniqueness theorems do not apply directly to this system. Moreover, it is known that the so-called Emden–Fowler differential equation

$$x'' + p(t)|x|^{\alpha-2}x = 0, \quad \alpha > 1, \tag{1.3}$$

(which looks similarly to (0.1)) admits the so-called singular solutions (see Section 14.2 and, e.g., the books [127,178]), i.e., solutions which violate uniqueness and continuability of solutions of (1.3).

1.2. *Half-linear trigonometric functions*

In proving the existence and uniqueness result for (0.1), the fundamental role is played by the generalized Prüfer transformation introduced in [85]. Consider a special half-linear equation of the form (0.1)

$$\bigl(\Phi(x')\bigr)' + (p-1)\Phi(x) = 0 \tag{1.4}$$

and denote by $S = S(t)$ its solution given by the initial conditions $S(0) = 0$, $S'(0) = 1$. We will show that the behavior of this solution is very similar to that of the classical sine function. Multiplying (1.4) (with x replaced by S) by S' and using the fact that $(\Phi(S'))' = (p-1)|S'|^{p-2}S''$, we get the identity $[|S'|^p + |S|^p]' = 0$. Substituting here $t = 0$ and using the initial condition for S we have the generalized Pythagorian identity

$$|S(t)|^p + |S'(t)|^p \equiv 1. \tag{1.5}$$

The function S is positive in some right neighborhood of $t = 0$ and using (1.5) $S' = \sqrt[p]{1 - S^p}$, i.e., $\frac{dS}{\sqrt[p]{1-S^p}} = dt$ in this neighborhood, hence

$$t = \int_0^{S(t)} \bigl(1 - s^p\bigr)^{-\frac{1}{p}}\, \mathrm{d}s. \tag{1.6}$$

Following the analogy with the case $p = 2$, we denote

$$\frac{\pi_p}{2} = \int_0^1 \bigl(1 - s^p\bigr)^{-\frac{1}{p}}\, \mathrm{d}s = \frac{1}{p}\int_0^1 (1-u)^{-\frac{1}{p}} u^{-\frac{1}{q}}\, \mathrm{d}u = \frac{1}{p} B\left(\frac{1}{p}, \frac{1}{q}\right),$$

where

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, \mathrm{d}t$$

is the Euler beta function. Using the formulas

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \qquad \Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}$$

with the Euler gamma function $\Gamma(x)=\int_0^\infty t^{x-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$, we have

$$\pi_p=\frac{2\pi}{p\sin\frac{\pi}{p}}. \tag{1.7}$$

The formula (1.6) defines uniquely the function $S=S(t)$ on $[0,\pi_p/2]$ with $S(\pi_p/2)=1$ and hence by (1.5) $S'(\pi_p/2)=0$. Now we define the generalized sine function S_p on the whole real line as the odd $2\pi_p$ periodic continuation of the function

$$S_p(t)=\begin{cases} S(t), & 0\leqslant t\leqslant \frac{\pi_p}{2},\\ S(\pi_p-t), & \frac{\pi_p}{2}\leqslant t\leqslant \pi_p. \end{cases}$$

The function S_p reduces to the classical function sin in case $p=2$ and in some literature (e.g., [73]) this function is denoted by $\sin_p t$. We will skip the index p by S_p if no ambiguity may occur.

In addition, we introduce the half-linear tangent and cotangent functions $\tan_p$ and $\cot_p$ by

$$\tan_p t=\frac{S_p(t)}{S_p'(t)},\qquad \cot_p t=\frac{S_p'(t)}{S_p(t)}.$$

The function $\tan_p$ is periodic with the period π_p and has discontinuities at $\pi_p/2+k\pi_p$, $k\in\mathbb{Z}$. The function $\cot_p$ is also π_p periodic, with discontinuities at $t=k\pi_p$, $k\in\mathbb{Z}$. By (1.4) and (1.5) we have

$$\begin{aligned}(\tan_p t)'&=\frac{1}{|S_p'(t)|^p}=1+|\tan_p t|^p, \qquad (1.8)\\ (\cot_p t)'&=-|\cot_p t|^{2-p}\bigl(1+|\cot_p t|^p\bigr).\end{aligned}$$

Hence $(\tan_p t)'>0$, $(\cot_p t)'<0$ on their definition domains and there exists the inverse functions arctan_p, arccot_p which are defined as inverse functions of $\tan_p$ and $\cot_p$ in the domains $(-\pi_p/2,\pi_p/2)$ and $(0,\pi_p)$, respectively. From (1.8) we have

$$(\arctan_p t)'=\frac{1}{1+|t|^p},$$

1.3. *Half-linear Prüfer transformation*

Using the above defined generalized trigonometric functions and their inverse functions, we can introduce the generalized Prüfer transformation as follows. Let x be a nontrivial solution of (0.1). Put

$$\rho(t)=\sqrt[p]{\bigl|x(t)\bigr|^p+r^q(t)\bigl|x'(t)\bigr|^p}$$

and let φ be a continuous function defined at all points where $x(t) \neq 0$ by the formula

$$\varphi(t) = \operatorname{arccot}_p \frac{r^{q-1}(t)x'(t)}{x(t)},$$

where q is the conjugate number of p, i.e., $\frac{1}{p} + \frac{1}{q} = 1$. Hence

$$x(t) = \rho(t)S\bigl(\varphi(t)\bigr), \qquad r^{q-1}(t)x'(t) = \rho(t)S'\bigl(\varphi(t)\bigr). \tag{1.9}$$

Differentiating the first equality in (1.9) and comparing it with the second one we get

$$r^{1-q}(t)\rho(t)S'\bigl(\varphi(t)\bigr) = \rho'(t)S\bigl(\varphi(t)\bigr) + \rho(t)S'\bigl(\varphi(t)\bigr)\varphi'(t). \tag{1.10}$$

Similarly, applying the function Φ to both sides of the second equation in (1.9), differentiating the obtained identity and substituting from (0.1) we get

$$\begin{aligned} -c(t)\rho^{p-1}(t)\Phi\bigl(S\bigl(\varphi(t)\bigr)\bigr) = (p-1)\bigl[&\rho^{p-2}(t)\rho'(t)\Phi\bigl(S'\bigl(\varphi(t)\bigr)\bigr) \\ &- \rho^{p-1}(t)\Phi\bigl(S\bigl(\varphi(t)\bigr)\bigr)\varphi'(t)\bigr]. \end{aligned} \tag{1.11}$$

Now, multiplying (1.10) by $\Phi(S')/\rho$, (1.11) by S/ρ^{p-1} and combining the obtained equations we get the first order system for φ and ρ

$$\begin{aligned} \varphi' &= \frac{c(t)}{p-1}\bigl|S(\varphi)\bigr|^p + r^{1-q}(t)\bigl|S'(\varphi)\bigr|^p, \\ \rho' &= \Phi\bigl(S\bigl(\varphi(t)\bigr)\bigr)S'\bigl(\varphi(t)\bigr)\left[r^{1-q}(t) - \frac{c(t)}{p-1}\right]\rho. \end{aligned} \tag{1.12}$$

1.4. *Existence and uniqueness*

Since the right-hand side of the last system is Lipschitzian in ρ, φ, the initial value problem for this system is uniquely solvable and its solution exists on the whole interval where r, c are continuous and $r(t) > 0$. Hence, the same holds for (0.1). This statement is summarized in the next theorem.

THEOREM 1.1. *Suppose that the functions r, c are continuous in an interval $I \subseteq \mathbb{R}$ and $r(t) > 0$ for $t \in I$. Given $t_0 \in I$ and $A, B \in \mathbb{R}$, there exists a unique solution of* (0.1) *satisfying $x(t_0) = A$, $x'(t_0) = B$ which is extensible over the whole interval I. This solution depends continuously on the initial values A, B.*

REMARK 1.1. The half-linear Prüfer transformation and the resulting existence and uniqueness theorem are presented in [85]. Another pioneering work in the theory of half-linear equations is the paper of Mirzov [176]. In that paper, the first order system

$$u_1' = a_1(t)|u_2|^{\lambda_1}\operatorname{sgn} u_2, \qquad u_2' = a_2(t)|u_1|^{\lambda_2}\operatorname{sgn} u_1 \tag{1.13}$$

is considered, where $\lambda_1, \lambda_2 > 0$ and $\lambda_1\lambda_2 = 1$. Comparing with (1.2), system (1.13) is slightly more general since the function a_1 (in contrast to the function r^{1-q} in (1.2)) can attain zero values. First, using Bellman's lemma [17, p. 46], Mirzov proves that the trivial solution $(u_1, u_2) \equiv (0,0)$ is the only solution satisfying the zero initial condition $u_1(t_0) = 0 = u_2(t_0)$, and that under the assumption of local integrability of the functions a_1, a_2 every solution of (1.13) is extensible over the whole interval where the integrability assumption is satisfied. Then, under the assumption that the functions a_1, a_2 do not change their signs, the Sturmian type theorem with respect to both components is proved using essentially the same method as described in the next section.

2. Sturmian theory

In this section we establish the basic oscillatory properties of half-linear equation (0.1). In particular, we show that the methods of the half-linear oscillation theory are similar to those of the oscillation theory of Sturm–Liouville linear equations (1.1), and that the Sturmian theory extends verbatim to (0.1).

2.1. *Riccati equation*

Let x be a solution of (0.1) such that $x(t) \neq 0$ in an interval I. Then $w(t) = \frac{r(t)\Phi(x'(t))}{\Phi(x(t))}$ is a solution of the Riccati-type differential equation

$$w' + c(t) + (p-1)r^{1-q}(t)|w|^q = 0, \tag{2.1}$$

where q is the conjugate number of p, i.e., $q = \frac{p}{p-1}$. Indeed, in view of (0.1) we have

$$\begin{aligned} w' &= \frac{(r\Phi(x'))'\Phi(x) - (p-1)r\Phi(x')|x|^{p-2}x'}{\Phi^2(x)} = -c - (p-1)\frac{r|x'|^p}{|x|^p} \\ &= -c - (p-1)r^{1-q}|w|^q. \end{aligned}$$

REMARK 2.1. (i) Using the above Riccati equation (2.1) one can derive the first equation in (1.12) as follows. From (1.9) we have

$$w = \frac{r(t)\Phi(x'(t))}{\Phi(x(t))} = \frac{\Phi(S'(\varphi(t))}{\Phi(S(\varphi(t))} =: v\big(\varphi(t)\big).$$

The function $v = \Phi(S')/\Phi(S)$ satisfies the Riccati equation corresponding to (1.4). This implies

$$\begin{aligned} \big[v\big(\varphi(t)\big)\big]' = v'\big(\varphi(t)\big)\varphi'(t) &= \left[-(p-1) - (p-1)\left|\frac{\Phi(S'(\varphi(t)))}{\Phi(S(\varphi(t)))}\right|^q\right]\varphi'(t) \\ &= -(p-1)\left[1 + \left|\frac{S'(\varphi(t))}{S(\varphi(t))}\right|^p\right]\varphi'(t), \end{aligned}$$

Substituting from (2.1)

$$w'(t) = -c(t) - (p-1)r^{1-q}(t)\bigl|w(t)\bigr|^q = -c(t) - (p-1)r^{1-q}(t)\left|\frac{S'(\varphi)}{S(\varphi)}\right|^p$$

and hence

$$c(t) + (p-1)r^{1-q}(t)\left|\frac{S'(\varphi(t))}{S(\varphi(t))}\right|^p = (p-1)\left[1 + \left|\frac{S'(\varphi(t))}{S(\varphi(t))}\right|^p\right]\varphi'(t).$$

Multiplying this equation by $|S(\varphi(t))|^p$ and using (1.5) we get really the first equation in (1.12).

(ii) Sometimes it is convenient to use a more general Riccati substitution

$$v(t) = \frac{f(t)r(t)\Phi(x'(t))}{\Phi(x(t))},$$

where f is a differentiable function. By a direct computation one can verify that v satisfies the first order Riccati-type equation

$$v' - \frac{f'(t)}{f(t)}v + f(t)c(t) + (p-1)r^{1-q}(t)f^{1-q}(t)|v|^q = 0. \tag{2.2}$$

The application of this more general Riccati substitution in the oscillation theory of (0.1) will be shown in Chapter 3C.

(iii) If we consider a slightly more general half-linear equation

$$\bigl(r(t)\Phi(x')\bigr)' + b(t)\Phi(x') + c(t)\Phi(x) = 0, \tag{2.3}$$

the Riccati substitution $w = r\Phi(x')/\Phi(x)$ leads to the equation

$$w' + c(t) + \frac{b(t)}{r(t)}w + (p-1)r^{1-q}(t)|w|^q = 0. \tag{2.4}$$

Multiplying this equation by $\exp\{\int^t \frac{b(s)}{r(s)}\,\mathrm{d}s\} =: g(t)$ and denoting $v = gw$, Equation (2.4) can be written in the same form as (2.1)

$$v' + c(t)g(t) + (p-1)r^{1-q}(t)g^{1-q}(t)|v|^q = 0.$$

The same effect we achieve if we multiply the original equation (2.3) by g since then this equation can be again written in the form (0.1).

2.2. *Picone's identity*

The original Picone's identity [190] for the linear second order equation (1.1) was established in 1910. Since that time, this identity has been extended in various directions and the half-linear version of this identity (proved in [115]) reads as follows.

THEOREM 2.1. *Consider a pair of half-linear differential operators*

$$l(x) = \big(r(t)\Phi(x')\big)' + c(t)\Phi(x), \qquad L(y) = \big(R(t)\Phi(y')\big)' + C(t)\Phi(y)$$

and let x, y *be continuously differentiable functions such that* $r\Phi(x')$, $R\Phi(y')$ *are also continuously differentiable and* $y(t) \neq 0$ *in an interval* $I \subset \mathbb{R}$. *Then in this interval*

$$\begin{aligned}\bigg\{\frac{x}{\Phi(y)}\big[\Phi(y)r\Phi(x') - \Phi(x)R\Phi(y')\big]\bigg\}'&\\ = (r-R)|x'| + (C-c)|x|^p + pR^{1-q}P\big(R^{q-1}x', R\Phi(xy'/y)\big)&\\ + \frac{x}{\Phi(y)}\big[\Phi(y)l(x) - \Phi(x)L(y)\big],&\end{aligned} \tag{2.5}$$

where

$$P(u,v) := \frac{|u|^p}{p} - uv + \frac{|v|^q}{q} \geqslant 0 \tag{2.6}$$

with equality if and only if $v = \Phi(u)$.

PROOF. The identity (2.5) can be verified by a direct computation, inequality (2.6) is the classical Young inequality, see, e.g., [105]. ☐

In the particular case when $r = R$, $C = c$, y is a nonzero solution of the equation $L(y) = 0$ and $w = R\Phi(y')/\Phi(y)$, then (2.5) reduces to the identity

$$r(t)|x'|^p - c(t)|x|^p = \big(w(t)|x|^p\big)' + pr^{1-q}(t)P\big(r^{q-1}(t)x', \Phi(x)w(t)\big). \tag{2.7}$$

This reduced Picone's identity will be used frequently in the sequel.

We will also need the following auxiliary statement which compares the function P with a certain quadratic function.

LEMMA 2.1. *The function* $P(u,v)$ *defined in* (2.6) *satisfies the following inequalities*

$$P(u,v) \gtrless \frac{1}{2}|u|^{2-p}\big(\Phi(u) - v\big)^2, \quad p \lessgtr 2,\ \Phi(u) \neq v,\ u \neq 0 \tag{2.8}$$

and

$$P(u,v) \lessgtr \frac{1}{2(p-1)}|u|^{2-p}\big(\Phi(u) - v\big)^2, \quad p \lessgtr 2,\ \big|\Phi(u)\big| > |v|,\ uv > 0. \tag{2.9}$$

More generally, for every $T > 0$ there exists a constant $K = K(T)$ such that

$$P(u,v) \gtrless K(T)|u|^{2-p}\big(\Phi(u) - v\big)^2, \quad p \gtrless 2,\ \Phi(u) \neq v,\ \left|\frac{v}{\Phi(u)}\right| < T. \tag{2.10}$$

PROOF. We present an outline of the proof only, for details we refer to [66,72]. We have

$$P(u,v) = |u|^p \left\{ \frac{1}{q}\left|\frac{v}{\Phi(u)}\right|^q - \frac{v}{\Phi(u)} + \frac{1}{p} \right\}$$

and

$$|u|^{2-p}\big(v - \Phi(u)\big)^2 = |u|^p \left(\frac{v}{\Phi(u)} - 1 \right)^2.$$

Denote $F(t) = \frac{1}{q}|t|^q - t + \frac{1}{p}$, $G(t) = \frac{1}{2}(t-1)^2$. The function $H = F - G$ satisfies $H(-1) = 0 = H(1)$, $H(0) = \frac{1}{p} - \frac{1}{2} \gtrless 0$ for $p \lessgtr 2$ and a closer investigation of the graph of this function shows that (2.8) and (2.10) really hold. □

2.3. *Energy functional*

The p-degree functional

$$\mathcal{F}(y; a, b) = \int_a^b \big[r(t)|y'|^p - c(t)|y|^p\big]\, \mathrm{d}t \tag{2.11}$$

considered over the Sobolev space $W_0^{1,p}(a,b)$ is usually called the *energy functional* of (0.1). Recall that the Sobolev space $W_0^{1,p}(a,b)$ consists of absolutely continuous functions x whose derivative is in $L^p(a,b)$ and $x(a) = 0 = x(b)$, with the norm $\|x\| = (\int_a^b [|x'|^p + |x|^p]\, \mathrm{d}t)^{1/p}$. Equation (0.1) is the Euler–Lagrange equation of the first variation of the functional $\mathcal{F}$. Moreover, if x is a solution of (0.1) satisfying $x(a) = 0 = x(b)$, then using integration by parts we have

$$\mathcal{F}(x; a, b) = \big[r(t)x(t)\Phi\big(x'(t)\big)\big]_a^b - \int_a^b x(t)\big[\big(r(t)\Phi(x')\big)' + c(t)\Phi(x)\big]\, \mathrm{d}t = 0. \tag{2.12}$$

2.4. *Roundabout theorem*

This theorem relates Riccati equation (2.1), the energy functional (2.11) and the basic oscillatory properties of solutions of (0.1). The terminology Roundabout theorem is motivated by the fact that the proof of this theorem consists of the "roundabout" proof of several equivalent statements.

Equation (0.1) is said to be *disconjugate* on the closed interval $[a,b]$ if the solution x given by the initial condition $x(a)=0$, $r(a)\Phi(x'(a))=1$ has no zero in $(a,b]$, in the opposite case (0.1) is said to be *conjugate* on $[a,b]$.

THEOREM 2.2. *The following statements are equivalent.*

(i) *Equation* (0.1) *is disconjugate on the interval* $I=[a,b]$. (0.1) *has at most one.*

(ii) *There exists a solution of* (0.1) *having no zero in* $[a,b]$.

(iii) *There exists a solution* w *of the generalized Riccati equation* (2.1) *which is defined on the whole interval* $[a,b]$.

(iv) *The energy functional* $\mathcal{F}(y;a,b)$ *is positive for every* $0\not\equiv y\in W_0^{1,p}(a,b)$.

PROOF. (i) ⇒ (ii): Consider the solution $\tilde{x}$ of (0.1) given by the initial condition $\tilde{x}(a)=\varepsilon$, $r(a)\Phi(\tilde{x}'(a))=1$, where $\varepsilon>0$ is sufficiently small. Then, according to the continuous dependence of solutions of (0.1) on the initial condition, disconjugacy of (0.1) on $[a,b]$ implies that this solution is positive on this interval.

(ii) ⇒ (iii): This implication is the immediate consequence of the Riccati substitution from Section 2.1.

(iii) ⇒ (iv): If there exists a solution w of (2.1) defined in the whole interval $[a,b]$, then by integrating the reduced Picone identity (2.7) with $x\in W_0^{1,p}(a,b)$ we get

$$\mathcal{F}(x;a,b)=p\int_a^b r^{1-q}(t)P\big(r^{q-1}(t)x',\Phi(x)w(t)\big)\,\mathrm{d}t\geqslant 0$$

with equality only if and only if $\Phi(r^{q-1}(t)x')=\Phi(x)w(t)$, i.e., $x'=\Phi^{-1}(w(t)/r(t))x$ in $[a,b]$, thus

$$x(t)=x(a)\exp\left\{\int_a^t\Phi^{-1}\left(\frac{w(s)}{r(s)}\right)\mathrm{d}s\right\}\equiv 0$$

since $x(a)=0$. This means that $\mathcal{F}(x;a,b)\geqslant 0$ over $W_0^{1,p}(a,b)$ with equality only if $x(t)\equiv 0$.

(iv) ⇒ (i): Suppose that $\mathcal{F}>0$ over nontrivial $y\in W_0^{1,p}(a,b)$ and (0.1) is not disconjugate in $[a,b]$, i.e., the solution x of (0.1) given by the initial condition $x(a)=0$, $r(a)\Phi(x'(a))=1$ has a zero $c\in[a,b]$. Define the function $y\in W_0^{1,p}(a,b)$ as follows

$$y(t)=\begin{cases}x(t), & t\in[a,c],\\ 0, & t\in[c,b].\end{cases}$$

Then by (2.12)

$$\mathcal{F}(y;a,b)=\mathcal{F}(y;a,c)=\mathcal{F}(x;a,c)=0$$

which contradicts the positivity of $\mathcal{F}$. □

REMARK 2.2. Similarly to the linear case, two points $t_1, t_2 \in \mathbb{R}$ are said to be *conjugate relative to* (0.1) if there exists a nontrivial solution x of this equation such that $x(t_1) = 0 = x(t_2)$. Due to Theorem 2.2, disconjugacy and conjugacy of (0.1) on a bounded interval $I \subset \mathbb{R}$ can be equivalently defined as follows. Equation (0.1) is said to be *disconjugate* on an interval I if this interval contains no pair of points conjugate relative to (0.1) (i.e., every nontrivial solution has at most one zero in I), in the opposite case (0.1) is said to be *conjugate* on I (i.e., there exists a nontrivial solution with at least two zeros in I). In Section 7 we will show that using the concept of the principal solution of (0.1), this equivalent definition applies also to unbounded intervals.

2.5. *Sturmian separation and comparison theorems*

The Roundabout theorem shows that the linear Sturmian theorems extend verbatim to half-linear equation (0.1). Indeed, the following theorem follows from the equivalence of (i) and (ii) in Theorem 2.2. We will prove it by using an alternative method.

THEOREM 2.3. *Let $t_1 < t_2$ be two consecutive zeros of a nontrivial solution x of* (0.1). *Then any other solution of this equation which is not proportional to x has exactly one zero on (t_1, t_2).*

PROOF. Among several possible methods of the proof we choose that one based on the Riccati substitution and the resulting equation (2.1). Let $w = r\Phi(x')/\Phi(x)$, then w is a solution of (2.1) which is defined on (t_1, t_2) and satisfies $w(t_1+) = \infty$, $w(t_2-) = -\infty$. Suppose that there exists a solution $\tilde{x}$ of (0.1), linearly independent of x, which has no zero in (t_1, t_2) and let $\tilde{w} = r\Phi(\tilde{x}'/\tilde{x})$. Since $\tilde{x}(t_1) \neq 0$, $\tilde{x}(t_2) \neq 0$ (otherwise $\tilde{x}$ would be a multiple of x), we have $\tilde{w}(t_1) < \infty$, $\tilde{w}(t_2) > -\infty$. Hence the graph of $\tilde{w}$ has to intersect the graph of w at some point in (t_1, t_2), but this contradicts the unique solvability of (2.1) (which follows from the fact that the function $c + (p-1)r^{1-q}|w|^q$ is Lipschitzian with respect to w). ☐

The Sturmian separation theorem also justifies the following definition of oscillation and nonoscillation of (0.1) which is the same as in the linear case. Equation (0.1) is said to be *nonoscillatory* (more precisely, nonoscillatory at ∞), if there exists $T_0 \in \mathbb{R}$ such that (0.1) is disconjugate on $[T_0, T_1]$ for every $T_1 > T_0$, in the opposite case (0.1) is said to be *oscillatory.*

According to Theorem 2.3, the above definition is correct, in the sense that oscillation of (0.1) means oscillation of its every nontrivial solution (i.e., the existence of a sequence of zeros of this solution tending to ∞). Note also that similarly to the linear case, if the functions r, c are continuous and $r(t) > 0$ in an interval $[T, \infty)$, then according to the unique solvability of the initial value problem associated with (0.1), the sequence of zeros of any nontrivial solution of (0.1) cannot a have a finite cluster point.

Along with (0.1) consider another equation of the same form

$$\big(R(t)\Phi(y')\big)' + C(t)\Phi(y) = 0, \tag{2.13}$$

where the functions R, C satisfy the same assumptions as r, c, respectively, in (0.1).

The next statement is an extension of well-known Sturm comparison theorem to (0.1).

THEOREM 2.4. *Let* $t_1 < t_2$ *be consecutive zeros of a nontrivial solution* x *of* (0.1) *and suppose that*

$$C(t) \geqslant c(t), \qquad r(t) \geqslant R(t) > 0 \tag{2.14}$$

for $t \in [t_1, t_2]$. *Then any solution of* (2.13) *has a zero in* (t_1, t_2) *or it is a multiple of the solution* x. *The last possibility is excluded if one of the inequalities in* (2.14) *is strict on a set of positive measure.*

PROOF. Let x be a solution of (0.1) having zeros at $t = t_1$ and $t = t_2$. Then by (2.12) we have $\mathcal{F}(x; t_1, t_2) = 0$ and according to (2.14)

$$\mathcal{F}_{RC}(x; t_1, t_2) := \int_{t_1}^{t_2} \left[R(t)|x'|^p - C(t)|x|^p\right] \mathrm{d}t \leqslant 0. \tag{2.15}$$

This implies, by Theorem 2.2, that the solution y of (2.13) given by the initial condition $y(t_1) = 0$, $y'(t_1) > 0$ has a zero in $(t_1, t_2]$ and by Theorem 2.3 any linearly independent solution of (2.13) has a zero in (t_1, t_2). Finally, suppose that the first zero of y to the right of t_1 is just at t_2, i.e., $y(t_1) = 0 = y(t_2)$. Let $v = R\Phi(y')/\Phi(y)$ be the solution of the Riccati equation associated with (2.13). Then, since $\lim_{t\to t_1+}[x(t)/y(t)] = \lim_{t\to t_1+}[x'(t)/y'(t)] = x'(t_1)/y'(t_1)$ exists finite, we have

$$\lim_{t\to t_1+} v(t)\big|x(t)\big|^p = \lim_{t\to t_1+} R(t)x(t)\Phi\big(y'(t)\big)\frac{\Phi(x(t))}{\Phi(y(t))} = 0.$$

Similarly $\lim_{t\to t_2-} v(t)\,|x(t)|^p = 0$. This implies $\mathcal{F}_{RC}(x; t_1, t_2) \geqslant 0$ (in view of (2.7) with R, C, v instead of r, c, w respectively), but this contradicts to (2.15) since $\mathcal{F}_{RC}(x; t_1, t_2) < 0$ if one of inequalities in (2.14) is strict on an interval of positive length. □

We will employ the same terminology as in the linear case. If (2.14) are satisfied in a given interval I, then (2.13) is said to be the *majorant* equation of (0.1) on I and (0.1) is said to be the *minorant* equation of (2.13) on I.

2.6. *Transformation of independent variable*

Let us introduce the new independent variable $s = \varphi(t)$ and the new function $y(s) = x(t)$, where φ is a differentiable function such that $\varphi'(t) \neq 0$ in some interval I where we consider Equation (0.1). Then $\frac{\mathrm{d}}{\mathrm{d}t} = \varphi'(t)\frac{\mathrm{d}}{\mathrm{d}s}$ and hence this transformation transforms (0.1) into the equation of the same form

$$\frac{\mathrm{d}}{\mathrm{d}s}\left[r(t)\Phi\big(\varphi'(t)\big)\Phi\left(\frac{\mathrm{d}}{\mathrm{d}s}y\right)\right] + \frac{c(t)}{\varphi'(t)}\Phi(y) = 0, \quad t = \varphi^{-1}(s), \tag{2.16}$$

where φ^{-1} is the inverse function of φ. In particular, if

$$\varphi(t)=\int_T^t r^{1-q}(\tau)\,\mathrm{d}\tau, \quad T\in I, \tag{2.17}$$

then the resulting equation (2.16) is the equation of the form (0.1) with $r(t)\equiv 1$.

Observe that if $\int^\infty r^{1-q}(t)\,\mathrm{d}t=\infty$, then (2.17) transforms an unbounded interval $[T,\infty)$ into the interval $[0,\infty)$ which is of the same form as $[T,\infty)$. If $\int^\infty r^{1-q}(t)\,\mathrm{d}t<\infty$ then the interval $[T,\infty)$ is transformed into the bounded interval $[0,b)$, $b=\int_T^\infty r^{1-q}(t)\,\mathrm{d}t$. This fact, coupled with the remark about cluster points of an oscillatory solution of (0.1) given below Theorem 2.3, shows why some (non)oscillation criteria and asymptotic formulas for solutions of (0.1) substantially depend on the divergence (convergence) of $\int^\infty r^{1-q}(t)\,\mathrm{d}t$.

2.7. *Reciprocity principle*

Suppose that the function c in (0.1) does not change its sign in an interval I and let $u=r\Phi(x')$. Then by a simple computation one can verify that u is a solution of the so-called *reciprocal equation*

$$\bigl(c^{1-q}(t)\Phi^{-1}(u')\bigr)'+r^{1-q}(t)\Phi^{-1}(u)=0, \tag{2.18}$$

where $\Phi^{-1}(s)=|s|^{q-1}\operatorname{sgn}s$, $q=\frac{p}{p-1}$, is the inverse function of Φ. The terminology reciprocal equation is motivated by the linear case $p=2$. The reciprocal equation to (2.18) is again the original equation (0.1)

If $t_1<t_2$ are consecutive zeros of a solution x of (0.1), then by the Rolle mean value theorem u has at least one zero in (t_1,t_2). Conversely, if $\tilde t_1<\tilde t_2$ are consecutive zeros of u, then $u'=-c(t)x$ and hence also x has a zero in (t_1,t_2). This means that (0.1) is oscillatory if and only if the reciprocal equation (2.18) is oscillatory. This fact we will refer to as the *reciprocity principle.*

2.8. *Leighton–Wintner oscillation criterion*

In this section we formulate a simple oscillation criterion for (0.1). Even if we will devote a special chapter to oscillation criteria for (0.1), we prefer to formulate this criterion already here. In the linear case $p=2$, this criterion was proved first by Leighton [146] under the addition assumption $c(t)\geqslant 0$. This restriction was later removed by Wintner, see, e.g., [208].

THEOREM 2.5. *Equation* (0.1) *is oscillatory provided*

$$\int^\infty r^{1-q}(t)\,\mathrm{d}t=\infty \quad \textit{and} \quad \int^\infty c(t)\,\mathrm{d}t=\lim_{b\to\infty}\int^b c(t)\,\mathrm{d}t=\infty. \tag{2.19}$$

PROOF. According to the definition of oscillation of (0.1), we need to show that this equation is not disconjugate on any interval of the form $[T, \infty)$. To illustrate typical methods of the half-linear oscillation theory, we present here two different proofs.

(i) *Variational proof.* We will find, for every $T \in \mathbb{R}$, a nontrivial function $y \in W_0^{1,p}(T, \infty)$ such that

$$\mathcal{F}(y; T, \infty) = \int_T^\infty \big[r(t)|y'|^p - c(t)|y|^p\big]\,\mathrm{d}t \leqslant 0. \tag{2.20}$$

The function which satisfies (2.20) can be constructed as follows

$$y(t) = \begin{cases} 0, & T \leqslant t \leqslant t_0, \\ \int_{t_0}^t r^{1-q}(s)\,\mathrm{d}s\big(\int_{t_0}^{t_1} r^{1-q}(s)\,\mathrm{d}s\big)^{-1}, & t_0 \leqslant t \leqslant t_1, \\ 1, & t_1 \leqslant t \leqslant t_2, \\ \int_t^{t_3} r^{1-q}(s)\,\mathrm{d}s\big(\int_{t_2}^{t_3} r^{1-q}(s)\,\mathrm{d}s\big)^{-1}, & t_2 \leqslant t \leqslant t_3, \\ 0, & t_3 \leqslant t < \infty, \end{cases}$$

where $T < t_0 < t_1 < t_2 < t_3$ will be specified later. Denote

$$K := \mathcal{F}(y; t_0, t_1) = \int_{t_0}^{t_1} \big[r(t)|y'|^p - c(t)|y|^p\big]\,\mathrm{d}t.$$

Then by a direct computation we have

$$\mathcal{F}(y; T, \infty) = K - \int_{t_1}^{t_2} c(t)\,\mathrm{d}t + \left(\int_{t_2}^{t_3} r^{1-q}(t)\,\mathrm{d}t\right)^{-1} - \int_{t_2}^{t_3} c(t)|y|^p\,\mathrm{d}t.$$

Since the function $g(t) = \int_t^{t_3} r^{1-q}(s)\,\mathrm{d}s(\int_{t_2}^{t_3} r^{1-q}(s)\,\mathrm{d}s)^{-1}$ is monotonically decreasing on $[t_2, t_3]$ with $g(t_2) = 1$ and $g(t_3) = 0$, by the second mean value theorem of integral calculus there exists $\xi \in (t_2, t_3)$ such that

$$\int_{t_2}^{t_3} c(t)g^p(t)\,\mathrm{d}t = \int_{t_2}^{\xi} c(t)\,\mathrm{d}t.$$

Using this equality, we have

$$\mathcal{F}(y; t_0, t_3) = K - \int_{t_1}^{\xi} c(t)\,\mathrm{d}t + \left(\int_{t_2}^{t_3} r^{1-q}(t)\,\mathrm{d}t\right)^{-1}.$$

Let $\varepsilon > 0$ and $t_1 > t_0$ be arbitrary. The second condition in (2.19) implies that t_2 can be chosen in such a way that $\int_{t_1}^{\xi} c(t)\,\mathrm{d}t > K + \varepsilon$ and the first condition in (2.19) implies that $t_3 > t_2$ can be taken such that $(\int_{t_2}^{t_3} r^{1-q}(t)\,\mathrm{d}t)^{-1} < \varepsilon$. Summarizing these estimates, we have

$$\mathcal{F}(y; t_0, t_3) \leqslant K - (K + \varepsilon) + \varepsilon \leqslant 0,$$

hence (0.1) is oscillatory by Theorem 2.2.

(ii) *Proof by the Riccati technique*. Suppose, by contradiction, that (2.19) holds and (0.1) is nonoscillatory. Then there exists $T \in \mathbb{R}$ and a solution w of Riccati equation (2.1) which is defined in the whole interval $[T, \infty)$. By integrating (2.1) from T to t we get

$$w(t) = w(T) - \int_T^t c(s)\,\mathrm{d}s - (p-1)\int_T^t r^{1-q}(s)\bigl|w(s)\bigr|^q\,\mathrm{d}s.$$

The second condition in (2.19) implies the existence of $T_1 > T$ such that we have $w(T) - \int_T^t c(s)\,\mathrm{d}s \leqslant 0$ for $t > T_1$ and hence

$$w(t) \leqslant -(p-1)\int_T^t r^{1-q}(s)\bigl|w(s)\bigr|^q\,\mathrm{d}s \quad \text{for } t > T_1.$$

Denote $G(t) = \int_T^t r^{1-q}(s)|w(s)|^q\,\mathrm{d}s$, then $|w| = [G' r^{q-1}]^{1/q}$ and the last inequality reads

$$\frac{G'(t)}{G^q(t)} \geqslant (p-1)^q r^{1-q}(t).$$

By integrating this inequality from T_1 to t we get

$$\frac{1}{q-1}G^{1-q}(T_1) > \frac{1}{q-1}\bigl[G^{1-q}(T_1) - G^{1-q}(t)\bigr] \geqslant (p-1)^q\int_{T_1}^t r^{1-q}(s)\,\mathrm{d}s.$$

Letting $t \to \infty$ we have a contradiction with the first condition in (2.19). ☐

3. Differences between linear and half-linear equations

The basic difference between linear and half-linear equations has already been mentioned at the beginning of this chapter, it the fact that the solution space of (0.1) is only homogeneous but not additive. In this subsection we point out some other differences (some of them are more or less consequences of this lack of the additivity of the solution space).

3.1. *Wronskian identity*

If x_1, x_2 are two solutions of the *linear* Sturm–Liouville differential equation (1.1), then by a direct differentiation one can verify the so-called *Wronskian identity*

$$r(t)\bigl[x_1(t)x_2'(t) - x_1'(t)x_2(t)\bigr] \equiv \omega, \tag{3.1}$$

where ω is a real constant. We have no half-linear version of this identity. More precisely, the above identity (3.1) can be regarded as an identity $W(x_1, x_1', x_2, x_2') := r(x_1'x_2 - x_1x_2') = \omega$ along solutions x_1, x_2 of (1.1). Elbert [86] showed that for $p \neq 2$ there exists no function of 4 variables $W(x_1, x_2, x_3, x_4)$ which is constant along solutions of (0.1), with the properties:

(i) W is continuously differentiable with respect to each variable.
(ii) W is not identically constant on $\mathbb{R}^4$.
(iii) W has antisymmetry property $W(x_3, x_4, x_1, x_2) = -W(x_1, x_2, x_3, x_4)$.

This result was proved using the generalized Prüfer transformation, we refer to [86] for details. The absence of a Wronskian-type identity implies that we have also no analogue of the linear reduction of order formula: given a solution $\tilde{x}$ of (1.1) with $\tilde{x}(t) \neq 0$ in an interval I, then

$$x(t) = \tilde{x}(t) \int^t \frac{\mathrm{d}s}{r(s)\tilde{x}^2(s)}$$

is another solution of (1.1)

3.2. *Transformation formula*

Let $h(t) \neq 0$ be a differentiable function such that rh' is also differentiable and let us introduce a new dependent variable y which is related to the original variable x by the formula $x = h(t)y$. Then we have the following (linear) identity which is the basis of the linear transformation theory (see [5,32,183])

$$h(t)\big[\big(r(t)x'\big)' + c(t)x\big] = \big(r(t)h^2(t)y'\big)' + h(t)\big[\big(r(t)h'(t)\big)' + c(t)h(t)\big]y. \tag{3.2}$$

In particular, if x a solution of (1.1) then y is a solution of

$$\big(R(t)y'\big)' + C(t)y = 0$$

with $R = rh^2$ and $C = h[(rh')' + ch]$. Since the function Φ is not additive, we have no half-linear analogue of this transformation identity. This has the following important consequence. Many oscillation results for linear equation (1.1) are based on the so-called *trigonometric transformation* which reads as follows. Let x_1, x_2 be two (linearly independent) solutions of (1.1) such that $r(x_1x_2' - x_1'x_2) = 1$ and let $h = \sqrt{x_1^2 + x_2^2}$. Then we have the identity (which can be verified by a direct computation)

$$h\big[(rh')' + ch\big] = \frac{1}{rh^2}.$$

This means that the transformation $x = h(t)y$ transforms (1.1) into the equation

$$\left(\frac{1}{q(t)}y'\right)' + q(t)y = 0, \quad q(t) = \frac{1}{r(t)h^2(t)}. \tag{3.3}$$

Equation (3.3) can be solved explicitly and $y_1 = \sin(\int^t q(s)\,\mathrm{d}s)$, $y_2 = \cos(\int^t q(s)\,\mathrm{d}s)$ are its linearly independent solutions, in particular, (3.3) and hence also (1.1) is oscillatory if and only if

$$\int^\infty \frac{\mathrm{d}t}{r(t)[x_1^2(t) + x_2^2(t)]} = \infty$$

for any pair of linearly independent solutions x_1, x_2 of (1.1). This fact is used in proofs of many oscillation results for (1.1), see [192,208] and references given therein. Since half-linear version of the transformation formula (3.2) is missing, analogous results for half-linear equation (0.1) are not known.

3.3. *Fredholm alternative*

Consider the linear Dirichlet boundary value problem associated with (1.1)

$$\begin{cases} \big(r(t)x'\big)' + c(t)x = f(t), & t \in [a,b] \\ x(a) = 0 = x(b). \end{cases} \tag{3.4}$$

It is well known that if the homogeneous problem with $f(t) \equiv 0$ has only the trivial solution, then (3.4) has a solution for any (sufficiently regular) right-hand side f (the so-called *nonresonant case*). If the homogeneous problem has a solution φ_0, problem (3.4) has a solution if and only if

$$\int_a^b f(t)\varphi_0(t)\,\mathrm{d}t = 0.$$

In particular, the problem

$$x'' + x = f(t), \qquad x(0) = 0 = x(\pi), \tag{3.5}$$

has a solution if and only if $\int_0^\pi f(t) \sin t\,\mathrm{d}t = 0$.

Now consider the half-linear version of the boundary value problem (3.5)

$$\begin{cases} \big(\Phi(x')\big)' + (p-1)\Phi(x) = f(t), & t \in [a,b], \\ x(0) = 0 = x(\pi_p), \end{cases} \tag{3.6}$$

where the generalized π_p is given by (1.7). A natural question is whether

$$\int_0^{\pi_p} f(t) \sin_p t\,\mathrm{d}t = 0 \tag{3.7}$$

is a necessary and sufficient condition for solvability of (3.6). This problem attracted considerable attention in last years, see [73] and the references given therein. It was shown that (3.7) is sufficient but no longer necessary for solvability of (3.6). We will deal with this problem in more details in the last chapter.

4. Some elementary half-linear equations

In this section we focus our attention to half-linear equations with constant coefficients and to Euler-type half-linear equation. The results presented here are essentially contained in the paper of Elbert [88]. This Elbert's paper deals with system (1.2) (with constant coefficients or coefficients corresponding to the generalized Euler equation), but results can be easily reformulated to equations of the form (0.1).

4.1. *Equations with constant coefficients*

Before passing to half-linear equations with constant coefficients, consider the equation

$$\big(r(t)\Phi(x')\big)' = 0. \tag{4.1}$$

The situation is here the same as in case of linear equations. The solution space of this equation is a two-dimensional linear space and the basis of this space is formed by the functions $x_1(t) \equiv 1$, $x_2(t) = \int^t r^{1-q}(s)\,\mathrm{d}s$, where $q = \frac{p}{p-1}$ is the conjugate number of p.

Now, consider Equation (0.1) with $r(t) \equiv r > 0$ and $c(t) \equiv c$. This equation can be written in the form

$$\big(\Phi(x')\big)' + \frac{c}{r}\Phi(x) = 0 \tag{4.2}$$

and the transformation of independent variable $t \mapsto \lambda t$ with $\lambda = (\frac{|c|}{r(p-1)})^{1/p}$ transforms (4.2) into the equation

$$\big(\Phi(x')\big)' + (p-1)\operatorname{sgn} c\Phi(x) = 0$$

If $c > 0$, this equation already appeared in Section 1.3 as Equation (1.4). Its solution given (uniquely) by the initial condition $x(0) = 0$, $x'(0) = 1$ was denoted by $S = S(t)$ (or by $\sin_p t$) and it is called the *half-linear sine function*. The function $C(t) = S'(t)$ is called the *the half-linear cosine function*. Consequently, taking into account homogeneity of the solution space of (0.1), we have the following statement concerning solvability of (1.4).

THEOREM 4.1. *For any $t_0 \in \mathbb{R}$, $x_0, x_1 \in \mathbb{R}$, the unique solution of* (1.4) *satisfying $x(t_0) = x_0$, $x'(t_0) = x_1$ is of the form $x(t) = \alpha S(t - t_1)$, where α, t_1 are real constants depending on t_0, x_0, x_1.*

Now let $c < 0$, i.e., we consider the equation

$$\big(\Phi(x')\big)' - (p-1)\Phi(x) = 0. \tag{4.3}$$

Multiplying this equation by x' and integrating the obtained equation over $[0, t]$ we have the identity

$$\left|x'(t)\right|^p - \left|x(t)\right|^p = \left|x'(0)\right|^p - \left|x(0)\right|^p = C, \tag{4.4}$$

where C is a real constant. If $C = 0$ then $x' = \pm x$ thus $x_1 = \mathrm{e}^t$ and $x_2 = \mathrm{e}^{-t}$ are solutions of (4.3).

In the remaining part of this subsection we focus our attention to the generalized half-linear hyperbolic sine and cosine functions. In the linear case $p = 2$, these functions are linear combinations of e^t and e^{-t}. However, in the half-linear case the additivity of the solution space of (0.1) is lost and one has to use a more complicated method. Let $E = E_p(t)$ be the solution of (4.3) with the initial conditions $E(0) = 0$, $E'(0) = 1$, and similarly let $F = F_p(t)$ be the solution given by the initial conditions $F(0) = 1$, $F'(0) = 0$. Let us also observe that the function E corresponds to $C = 1$ and F to $C = -1$ in (4.4), respectively. Moreover, for $p = 2$, i.e., if differential equation (4.3) is linear, we have $E_2(t) = \sinh t$ and $F_2(t) = \cosh t$.

Due to (4.4), the function E satisfies also the relation

$$E' = \sqrt[p]{1 + |E|^p},$$

hence $E' > 1$ for $t > 0$. Consequently

$$t = \int_0^t \frac{E'(s)\,\mathrm{d}s}{\sqrt[p]{1 + E^p(s)}} = \int_0^{E(t)} \frac{\mathrm{d}s}{\sqrt[p]{1 + s^p}}. \tag{4.5}$$

In order to compare (asymptotically) the function $E(t)$ with e^t, let the function $f(s)$ be defined by

$$f(s) = \begin{cases} 1 & \text{for } 0 \leqq s \leqq 1, \\ \frac{1}{s} & \text{for } s \geqq 1. \end{cases}$$

Then by (4.5) we obtain

$$t - \lg E(t) = 1 + \int_0^{E(t)} \frac{\mathrm{d}s}{\sqrt[p]{1 + s^p}} - \int_0^{E(t)} f(s)\,\mathrm{d}s \quad \text{for } E(t) > 1.$$

Hence

$$\lg \delta_p := \lim_{t \to \infty} \big[t - \lg E(t)\big] = 1 - \int_0^\infty \left[f(s) - \frac{1}{\sqrt[p]{1 + s^p}} \right] \mathrm{d}s. \tag{4.6}$$

The integral in the right-hand side of (4.6) can be interpreted as the area of the domain on the (s, y) plane given by the inequalities

$$\frac{1}{\sqrt[p]{1 + s^p}} \leqslant y \leqslant f(s) \quad \text{for } 0 \leqslant s < \infty.$$

Taking y as an independent variable, we find for the integral in (4.6)

$$\lg \delta_p = 1 - \int_0^1 \frac{1 - \sqrt[p]{1 - y^p}}{y}\,\mathrm{d}y = 1 - \frac{1}{p} \int_0^1 \frac{1 - u^{1/p}}{1 - u}\,\mathrm{d}u. \tag{4.7}$$

Since $0 < 1 - \sqrt[p]{1-y^p} < y$ for $0 < y < 1$ we have $0 < \lg\delta_p < 1$, i.e., $1 < \delta_p < \mathrm{e}$. On the other hand, the integral in (4.7) can be expressed by means of the function $\Psi(z) = d\lg\Gamma(z)/\mathrm{d}z$ as

$$\Psi(z) = -\widetilde{C} + \int_0^1 \frac{1-t^{z-1}}{1-t}\,\mathrm{d}t \quad \text{for } \operatorname{Re} z > 0,$$

where $\widetilde{C}$ is the Euler constant and $\Gamma(z) = \int_0^\infty t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t$ denotes the usual Euler gamma function. Making use of this relation we obtain

$$\lg\delta_p = 1 - \frac{1}{p}\left[\widetilde{C} + \Psi\left(\frac{p+1}{p}\right)\right]. \tag{4.8}$$

Finally, the relation (4.6) can be rewritten as

$$\lim_{t\to\infty} \frac{\mathrm{e}^t}{E_p(t)} = \delta_p \quad \text{where } 1 < \delta_p < \mathrm{e}. \tag{4.9}$$

A similar relation is expected also for the function $F_p(t)$. We will use the following auxiliary statement which we present without the proof, this proof can be found in [85].

LEMMA 4.1. *Let $I_1(R)$, $I_2(R)$ be integrals defined by*

$$I_1(R) = \int_0^R \frac{\mathrm{d}\xi}{\sqrt[p]{1+\xi^p}}, \quad R > 0, \qquad I_2(R) = \int_1^R \frac{\mathrm{d}\xi}{\sqrt[p]{\xi^p - 1}}, \quad R > 1.$$

Then

$$\lim_{R\to\infty} \left[I_1(R) - I_2(R)\right] = \frac{\pi}{p}\cot\frac{\pi}{p}.$$

Now we return to the asymptotic formula for F_p. We want to obtain a relation similar to (4.9). Since F fulfills the differential equation

$$|F'|^p - |F|^p = -1,$$

we have

$$\frac{F'}{\sqrt[p]{F^p - 1}} = 1 \quad \text{for } t > 0.$$

Integrating the last equality yields

$$\int_1^{F(t)} \frac{\mathrm{d}\xi}{\sqrt[p]{\xi^p - 1}} = t \quad \text{for } t > 0.$$

This relation implies that $\lim_{t\to\infty} F(t) = \infty$. On the other hand

$$\begin{aligned}\lim_{t\to\infty}\big[t - \lg F(t)\big] &= \lim_{t\to\infty}\int_1^{F(t)}\left[\frac{1}{\sqrt[p]{\xi^p - 1}} - f(\xi)\right]\mathrm{d}\xi \\ &= \int_1^{\infty}\left[\frac{1}{\sqrt[p]{\xi^p - 1}} - \frac{1}{\xi}\right]\mathrm{d}\xi \qquad (4.10)\end{aligned}$$

because the integral on the right-hand side is convergent. Let $\varDelta_p$ be introduced by

$$\lg \varDelta_p := \int_1^{\infty}\left[\frac{1}{\sqrt[p]{\xi^p - 1}} - \frac{1}{\xi}\right]\mathrm{d}\xi. \qquad (4.11)$$

It is clear that $\varDelta_p > 1$. The relation (4.10) can be rewritten as

$$\lim_{t\to\infty}\frac{\mathrm{e}^t}{F_p(t)} = \varDelta_p \quad \text{with } \varDelta_p > 1. \qquad (4.12)$$

Now we want to establish a connection between $\varDelta_p$ and δ_p. By (4.6), (4.10), and taking into account the definition of the function $f(\xi)$, we have

$$\begin{aligned}\lg\frac{\varDelta_p}{\delta_p} &= \lim_{R\to\infty}\left[\int_1^R\left(\frac{1}{\sqrt[p]{\xi^p - 1}} - \frac{1}{\xi}\right)\mathrm{d}\xi - 1 + \int_0^R\left(f(\xi) - \frac{1}{\sqrt[p]{1+\xi^p}}\right)\mathrm{d}\xi\right] \\ &= \lim_{R\to\infty}\big[I_2(R) - I_1(R)\big],\end{aligned}$$

where the functions $I_1(R)$, $I_2(R)$ were introduced in Lemma 4.1. Then by this lemma we get the wanted relation as

$$\lg\frac{\delta_p}{\varDelta_p} = \frac{\pi}{p}\cot\frac{\pi}{p}. \qquad (4.13)$$

We may observe here that this relation implies that $\delta_2 = \varDelta_2$ in the linear case $(p = 2)$. In fact, we have $\delta_2 = 2 = \varDelta_2$.

By (4.8) the value of δ_p can be considered to be known, consequently by the relation (4.13) the value of $\varDelta_p$ is known as well.

Finally, there are interesting functional relations between the half-linear hyperbolic sine and cosine functions $E_p(t)$, $F_p(t)$ as follows

$$\begin{aligned}E_p'(t) &= \big\{F_q\big((p-1)t\big)\big\}^{q-1} = \varPhi^{-1}\big(F_q\big((p-1)t\big)\big), \\ F_p'(t) &= \varPhi^{-1}\big(E_q\big((p-1)t\big)\big). \qquad (4.14)\end{aligned}$$

To prove these relations it is sufficient to show that the functions on both sides of the equalities satisfy the same differential equation and fulfill the same initial conditions, this is a matter of a direct computation (use, e.g., the result of Section 2.7).

The relations (4.14) provide another connections between the values of δ_p and Δ_p. Indeed, by (4.4) and (4.9) we have

$$\lim_{t\to\infty}\frac{E'_p(t)}{\mathrm{e}^t}=\lim_{t\to\infty}\frac{E_p(t)}{\mathrm{e}^t}=\frac{1}{\delta_p}. \tag{4.15}$$

On the other hand, this, (4.12) and (4.14) imply (taking into account that $(p-1)(q-1)=1$ for conjugate pair p,q)

$$\frac{1}{\delta_p^{p-1}}=\lim_{t\to\infty}\frac{(E'_p(t))^{p-1}}{\mathrm{e}^{(p-1)t}}=\lim_{t\to\infty}\frac{F_q((p-1)t)}{\mathrm{e}^{(p-1)t}}=\frac{1}{\Delta_q}$$

hence

$$\Delta_q=\delta_p^{p-1}, \tag{4.16}$$

and similarly

$$\delta_q=\Delta_p^{p-1}. \tag{4.17}$$

We remark that the last two relations are equivalent since replacing p by q we get each from the other. By relations (4.13), (4.16) (or (4.17)) it is sufficient to know one of the values of Δ_p, δ_p, Δ_q, δ_q, and then all the other values can be obtained easily.

As in the linear case where the function $\sinh t$ is odd and the function $\cosh t$ is even, the functions $E_p(t)$, $F_p(t)$ behave in a similar way:

$$E_p(-t)=-E_p(t),\qquad F_p(-t)=F_p(t). \tag{4.18}$$

To prove this statement it is sufficient to show that the functions on the both side of the equality are solutions of differential equation (4.3) and satisfy the same initial conditions at $t=0$. Then the uniqueness of the initial value problem (see Section 1.4) proves (4.18).

Now we know all the solutions of differential equations (4.3). We display them in the next theorem.

THEOREM 4.2. *The solutions of* (4.3) *are*:

$$K\mathrm{e}^t,\qquad K\mathrm{e}^{-t},\qquad KE_p(t+t_0),\qquad KF_p(t+t_0), \tag{4.19}$$

where K and t_0 are real parameters. More precisely, there are two one-parameter families of solutions $x(t)=K\mathrm{e}^t$, $x(t)=K\mathrm{e}^{-t}$ and two two-parameter families satisfying the following asymptotic formula

$$\lim_{t\to\infty}\frac{x(t)}{e^t}=L,$$

where $L=\frac{K\mathrm{e}^{t_0}}{\delta_p}$ or $L=\frac{K\mathrm{e}^{t_0}}{\Delta_p}$ with K from (4.19).

PROOF. Since Equation (4.3) is autonomous, it is sufficient to consider solutions whose initial values are prescribed at $t=0$, i.e., $x(0)=x_0$, $x'(0)=x_1$. If $x_0=0=x_1$, then according to the unique solvability $x(t)\equiv 0$. If at least one of the constants x_0, x_1 is nonzero, distinguish the cases according to the value of the constant C in (4.4). In case $C=0$, the only possibilities are $x(t)=K\mathrm{e}^t$ or $x(t)=K\mathrm{e}^{-t}$. More precisely, if $x_0x_1>0$, then $x(t)=x_0\mathrm{e}^t$, if $x_0x_1<0$, then $x(t)=x_0\mathrm{e}^{-t}$. Now, if $C>0$, then by the definition of the function E_p we have $x(t)=KE_p(t+t_0)$. Let $K=\operatorname{sgn}x_0\sqrt[p]{C}$. Since the function E_p is strictly increasing (observe that $E'(0)=1$, E' is continuous and $|E'|^p=1+|E|^p$) there exists $t_0\in\mathbb{R}$ such that $CE_p(t_0)=x_1$. Concerning the initial condition for the derivative x', we have

$$|x_1|^p=C+|x_0|^p=C+C|E|^p=C\big(1+|E|^p\big)=C|E'|^p,$$

hence $x_1=\pm KE'(t_0)$. But $E'(t_0)>0$ and $\operatorname{sgn}K=\operatorname{sgn}x_1$, the sign $+$ is the correct one. If $C<0$, let $K=\operatorname{sgn}x_0\sqrt[p]{-C}$ and t_0 be the solution of $CF'(t_0)=x_1$. Then the function $KF_p(t+t_0)$ is the solution we looked for.

Finally, concerning the asymptotic formula, any solution which is not proportional to e^t or e^{-t} satisfies by (4.9) or (4.12)

$$\lim_{t\to\infty}\frac{x(t)}{\mathrm{e}^t}=\lim_{t\to\infty}\frac{KE(t+t_0)}{\mathrm{e}^{t+t_0}}\mathrm{e}^{t_0}=\frac{K\mathrm{e}^{t_0}}{\delta_p}\quad\text{or}\quad\lim_{t\to\infty}\frac{x(t)}{\mathrm{e}^t}=\frac{K\mathrm{e}^{t_0}}{\Delta_p},$$

hence $L=\frac{K\mathrm{e}^{t_0}}{\delta_p}$ or $L=\frac{K\mathrm{e}^{t_0}}{\Delta_p}$. □

4.2. *Euler-type half-linear differential equation*

In this subsection we deal with the Euler-type differential equation

$$\big(\Phi(x')\big)'+\frac{\gamma}{t^p}\Phi(x)=0,\tag{4.20}$$

where γ is a real constant. By the analogue with the linear Euler equation we look first for solutions in the form $x(t)=t^\lambda$. Substituting into (4.20) we get the algebraic equation for λ

$$G(\lambda):=(p-1)|\lambda|^p-(p-1)\Phi(\lambda)+\gamma=0.$$

The function G is convex, hence the equation $G(\lambda)=0$ has two, one or no (real) root according to the value of γ. However, even if the first possibility happens, since the additivity of the solution space of is lost in the half-linear case, we are not able to compute other solutions explicitly. To get a more detailed information about their asymptotic behavior, we use the procedure which is also typical in the linear case, namely the transformation of (4.20) into an equation with constant coefficients.

The change of independent variable $s=\lg t$ converts (4.20) into the equation (where the dependent variable will be denoted again by x and $'=\frac{\mathrm{d}}{\mathrm{d}s}$)

$$\big(\Phi(x')\big)'-(p-1)\Phi(x')+\gamma\Phi(x)=0.\tag{4.21}$$

The Riccati equation corresponding to (4.20) and (4.21) are

$$w' = -\gamma t^{-p} - (p-1)|w|^p \tag{4.22}$$

and

$$v' = -\gamma + (p-1)v - (p-1)|v|^q =: F(v), \tag{4.23}$$

respectively. The solutions w and v are related by the formula $w(t) = t^{1-p} v(\lg t)$ and, moreover, we have $G(\Phi^{-1}(v)) = -F(v)$.

The function F is concave on $\mathbb{R}$ with the global maximum at $\tilde{v} = (\frac{p-1}{p})^{p-1}$ and the value of this maximum is $\tilde{\gamma} = (\frac{p-1}{p})^p - \gamma$. We distinguish the following 3 cases with respect to the value of the constant γ.

(I) $\gamma < \tilde{\gamma}$. Then the equation $F(v) = 0$ has two real roots $v_1 < \tilde{v} < v_2$;
(II) $\gamma = \tilde{\gamma}$. Then the equation $F(v) = 0$ has the double root $v = \tilde{v}$;
(III) $\gamma > \tilde{\gamma}$. Then $F(v) < 0$ for every $v \in \mathbb{R}$.

Case (I). The constant functions $v(s) \equiv v_1$, $v(s) \equiv v_2$ are solutions of (4.23). Clearly, if v is a solution of (4.23) such that $v(s) < v_1$, for some $s \in \mathbb{R}$, then $v'(s) < 0$, if $v(s) \in (v_1, v_2)$, then $v'(s) > 0$, and $v'(s) < 0$ for $v(s) > v_2$, a picture of the direction field of (4.23) helps to visualize the situation. Any solution of (4.23) different from $v(s) = v_{1,2}$ can be expressed (implicitly) in the form ($S \in \mathbb{R}$ being fixed)

$$\int_{v(S)}^{v(s)} \frac{\mathrm{d}v}{F(v)} = s - S. \tag{4.24}$$

Observe that the integral $\int_{s_1}^{s_2} \frac{\mathrm{d}s}{F(s)}$ is convergent whenever the integration interval does not contain zeros $v_{1,2}$ of F, in particular, for any $\varepsilon > 0$

$$\int_{-\infty}^{v_1-\varepsilon} \frac{\mathrm{d}v}{F(v)} > -\infty, \qquad \int_{v_2+\varepsilon}^{\infty} \frac{\mathrm{d}v}{F(v)} > -\infty.$$

Case (Ia). $v(S) < v_1$; then $v(s) < v(S)$ for $s > S$ and v is decreasing. If v were extensible up to ∞, we would have a contradiction with (4.24) since the right-hand side tends to ∞ while the left one is bounded. Later we will show that $v(s) = v_1$ is the so-called *eventually minimal solution* of (4.23).

Case (Ib). $v(S) \in (v_1, v_2)$; the solution v is increasing in this case, and as $s \to \infty$, we have $v(s) \to v_2$, otherwise we have the same contradiction as in the previous case.

Next we compute the asymptotic formula for the difference $v_2 - v(s)$. We have

$$F(v) = F'(v_2)(v - v_2) + \mathrm{O}\big((v - v_2)^2\big), \quad \text{as } v \to v_2,$$

hence

$$\frac{1}{F(v)} = \frac{1}{F'(v_2)(v - v_2)[1 + \mathrm{O}(v - v_2)]} = \frac{1}{F'(v_2)(v - v_2)}\big[1 + \mathrm{O}(v - v_2)\big]$$

$$= \frac{1}{F'(v_2)(v - v_2)} + \mathrm{O}(1)$$

and therefore (since $|v(s) - v(S)| < v_2 - v_1 = \mathrm{O}(1)$), substituting into (4.24)

$$\int_{v(S)}^{v(s)} \frac{\mathrm{d}v}{F(v)} = \frac{1}{F'(v_2)} \lg \frac{v_2 - v(s)}{v_2 - v(S)} + \mathrm{O}(1) = s - S,$$

i.e.,

$$v_2 - v(s) = K \exp\bigl\{F'(v_2)s\bigr\},$$

where K is a positive constant (depending on $v(S)$), and substituting $w(t) = t^{1-p}v(\lg t)$ we have

$$v_2 - t^{p-1}w(t) = Kt^{F'(v_2)} \to 0 \quad \text{as } t \to \infty \tag{4.25}$$

since $F'(v_2) < 0$. Substituting $w = \frac{\Phi(x')}{\Phi(x)}$ in (4.25) we have

$$\frac{\Phi(x'(t))}{\Phi(x(t))} = t^{1-p}\bigl(v_2 - Kt^{F'(v_2)}\bigr)$$

and using the formula $(1 + \alpha)^{p-1} = 1 + (p - 1)\alpha + \mathrm{o}(\alpha)$ as $\alpha \to 0$, we obtain (with $q = \frac{p}{p-1}$)

$$\frac{x'(t)}{x(t)} = \frac{\Phi^{-1}(v_2)}{t}\bigl(1 - \widetilde{K}t^{F'(v_2)}\bigr)^{q-1} \sim \frac{\Phi^{-1}(v_2)}{t}\bigl(1 - \widetilde{K}t^{F'(v_2)}\bigr),$$

as $t \to \infty$, here $f \sim g$ for a pair of functions f, g means $\lim_{t\to\infty} \frac{f(t)}{g(t)} = 1$, and $\widetilde{K}$ is a real constant. Thus

$$x(t) = t^{\lambda_2} \exp\bigl\{\overline{K}t^{F'(v_2)}\bigr\} \sim t^{\lambda_2} \quad \text{as } t \to \infty,$$

since $F'(v_2) < 0$, where $\overline{K}$ is another real constant and $\lambda_2 = \Phi^{-1}(v_2)$ is the larger of roots of the equation $G(\lambda) = 0$.

Case (Ic). $v(S) > v_2$; then $v'(s) < 0$ and $v(s) \in (v_2, v(S))$ for $s > S$. Using the same argument as in (Ib) we have

$$t^{p-1}w(t) - v_2 = \widetilde{K}t^{F'(v_2)} \to 0 \quad \text{as } t \to \infty,$$

$\widetilde{K}$ being a positive constant, and this implies the same asymptotic formula for the solution x of (4.20) which determines the solution w of (4.22).

Case (II). $\gamma = \tilde{\gamma} = (\frac{p-1}{p})^p$. Then the function F has the double root $\tilde{v} = (\frac{p-1}{p})^{p-1}$. Equation (4.20) has a solution $x(t) = t^{\Phi^{-1}(\tilde{v})} = t^{\frac{p-1}{p}}$. This means that (4.20) with $\gamma = \tilde{\gamma}$ is still nonoscillatory. In the linear case $p = 2$, $\tilde{\gamma} = \frac{1}{4}$ and we are able to compute a linearly independent solution using the reduction of order formula. This solution is $\tilde{x}(t) = \sqrt{t}\lg t$. The reduction of order formula is missing in the half-linear case, but using essentially the same method as in Case (I) we are able to show that all solutions which are not proportional to $t^{\frac{p-1}{p}}$ behave asymptotically as $t^{\frac{p-1}{p}} \lg^{\frac{2}{p}} t$.

To this end, we proceed as follows. Since $F(\tilde{v}) = 0 = F'(\tilde{v})$,

$$F(v) = \frac{1}{2} F''(\tilde{v})(v - \tilde{v})^2 + \mathrm{O}\big((v - \tilde{v})^3\big) \quad \text{as } v \to \tilde{v},$$

hence, taking into account that $F''(\tilde{v}) = -\frac{1}{\tilde{v}}$,

$$\begin{aligned} \frac{1}{F(v)} &= \frac{1}{\frac{1}{2}F''(\tilde{v})(v - \tilde{v})^2[1 + \mathrm{O}(v - \tilde{v})]} \\ &= -\frac{2\tilde{v}}{(v - \tilde{v})^2} + \mathrm{O}\big((v - \tilde{v})^{-1}\big) \quad \text{as } v \to \tilde{v}. \end{aligned}$$

On the other hand, using the same argument as in the previous part, we see from (4.24) that any solution v which starts with the initial value $v(S) < \tilde{v}$ fails to be extensible up to ∞ and solutions with $v(S) > \tilde{v}$ tend to $\tilde{v}$ as $t \to \infty$. Substituting for $F(v)$ in (4.24) we have

$$\frac{2\tilde{v}}{v - \tilde{v}} + \mathrm{O}\big(\lg|v - \tilde{v}|\big) = s - S,$$

hence

$$2\tilde{v} + (v - \tilde{v})\mathrm{O}\big(\lg|v - \tilde{v}|\big) = (v - \tilde{v})(s - S).$$

Since $\lim_{v \to \tilde{v}} (v - \tilde{v})\mathrm{O}(\lg|v - \tilde{v}|) = 0$, we have

$$\lim_{s \to \infty} (s - S)\big(v(s) - \tilde{v}\big) = \lim_{s \to \infty} s\big(v(s) - \tilde{v}\big) = 2\tilde{v}.$$

Consequently,

$$\mathrm{O}\big(\lg|v(s) - \tilde{v}|\big) = \mathrm{O}\big(\lg s^{-1}\big) = \mathrm{O}(\lg s) \quad \text{as } s \to \infty,$$

and thus $(v(s) - \tilde{v})^{-1} = \frac{s}{2\tilde{v}} + \mathrm{O}(\lg s)$, which means

$$v(s) - \tilde{v} = \frac{2\tilde{v}}{s} \frac{1}{1 + \mathrm{O}(\frac{\lg s}{s})} = \frac{2\tilde{v}}{s}\left(1 + \mathrm{O}\left(\frac{\lg s}{s}\right)\right) = \frac{2\tilde{v}}{s} + \mathrm{O}\left(\frac{\lg s}{s^2}\right).$$

Now, taking into account that solutions of (4.22) and (4.23) are related by $w(t) = t^{1-p} v(\lg t)$, we have

$$t^{p-1} w(t) - \tilde{v} = \frac{2\tilde{v}}{\lg t} + \mathrm{O}\left(\frac{\lg(\lg t)}{\lg^2 t}\right),$$

which means that the solution x of (4.20) which determines the solution w of (4.22) satisfies

$$\frac{x'(t)}{x(t)} \sim \frac{\Phi^{-1}(\tilde{v})}{t}\left(1+\frac{2}{\lg t}\right)^{\frac{1}{p-1}} \sim \frac{p-1}{pt}+\frac{2}{pt\lg t}$$

and thus

$$x(t) \sim t^{\frac{p-1}{p}} \lg^{\frac{2}{p}} t.$$

Case (III). The equation $F(v)=0$ has no real root and $F(v)<0$ for every $v \in \mathbb{R}$. Again

$$\int_{v(S)}^{v(s)} \frac{\mathrm{d}v}{F(v)} = s-S, \quad \text{and} \quad v(s)<v(S) \quad \text{for } s>S.$$

Since the left-hand side of the last equality is bounded for any value $v(s)$, while the right one tends to ∞ as $s \to \infty$, no solution of (4.23) and hence also of (4.22) is extensible up to ∞, which means that (4.20) is oscillatory. We will show that oscillatory solutions of (4.21) are periodic and we will determine the value of their period.

To this end, we use the Prüfer transformation mentioned in Section 1 (compare also Remark 1.1 applied to (4.21)). Any nontrivial solution of this equation can be expressed in the form

$$x(s)=\rho(s)S\big(\varphi(s)\big), \qquad x'(s)=\rho(s)S'\big(\varphi(s)\big),$$

where S is the generalized half-linear sine function. The angular and radial variables φ, ρ satisfy the first order differential system

$$\varphi' = \left|S'(\varphi)\right|^p - S(\varphi)\Phi\big(S'(\varphi)\big) + \frac{\gamma}{p-1}\left|S(\varphi)\right|^p, \tag{4.26}$$

$$\rho' = S'(\varphi)\left[\Phi\big(S'(\varphi)\big) + \left(1-\frac{\gamma}{p-1}\right)\Phi\big(S(\varphi)\big)\right]\rho. \tag{4.27}$$

Oscillation of (4.21) implies that $\lim_{s\to\infty}\varphi(s)=\infty$. Denote

$$\Psi(\varphi) := \left|S'(\varphi)\right|^p - S(\varphi)\Phi\big(S'(\varphi)\big) + \frac{\gamma}{p-1}\left|S(\varphi)\right|^p. \tag{4.28}$$

Then

$$\int_{\varphi(S)}^{\varphi(s)} \frac{\mathrm{d}\varphi}{\Psi(\varphi)} = s-S. \tag{4.29}$$

Here we have used the fact that $\Psi(\varphi) > 0$ for $\varphi \in \mathbb{R}$ since if $\varphi = 0 \pmod{\pi_p}$, then $\Psi(\varphi) = 1$ (observe that $|S(\varphi)|^p + |S'(\varphi)|^p = 1$), and if $S(\varphi) \neq 0$, then

$$\Psi(\varphi) = \left|S(\varphi)\right|^p \left[|\alpha|^p - \Phi(\alpha) + \frac{\gamma}{p-1}\right] > 0, \quad \alpha := \frac{S'(\varphi)}{S(\varphi)}.$$

Further, let

$$\tau = \int_0^{2\pi_p} \frac{\mathrm{d}\varphi}{\Psi(\varphi)} = 2\int_0^{\pi_p} \frac{\mathrm{d}\varphi}{\Psi(\varphi)}. \tag{4.30}$$

By (4.28) we have $\Psi(\varphi + \pi_p) = \Psi(\varphi)$, hence $\varphi(s+\tau) = \varphi(s) + 2\pi_p$ and the substitution $t := \frac{S(\varphi)}{S'(\varphi)}$ gives

$$\tau = 2\int_{-\infty}^{\infty} \frac{\mathrm{d}t}{\frac{\gamma}{p-1}|t|^p - t + 1}, \tag{4.31}$$

which is the quantity depending only on γ.

Finally, we will estimate the radial variable ρ. Denote

$$R(\varphi) := S'(\varphi)\left[\Phi\bigl(S'(\varphi)\bigr) + \left(1 - \frac{\gamma}{p-1}\right)\Phi\bigl(S(\varphi)\bigr)\right].$$

By (4.27) and the identity $\varphi(s+\tau) = \varphi(s) + 2\pi_p$ we have

$$\lg\frac{\rho(s+\tau)}{\rho(s)} = \int_s^{s+\tau} R\bigl(\varphi(s)\bigr)\,\mathrm{d}s = \int_{\varphi(s)}^{\varphi(s+\tau)} \frac{R(\varphi)}{\Psi(\varphi)}\,\mathrm{d}\varphi = \int_0^{2\pi_p} \frac{R(\varphi)}{\Psi(\varphi)}\,\mathrm{d}\varphi.$$

Now, using the identity $\Psi' + pR = p - 1$, we get

$$\lg\frac{\rho(s+\tau)}{\rho(s)} = \int_0^{2\pi_p} \frac{\frac{p}{p-1} - \frac{1}{p}\Psi'(\varphi)}{\Psi(\varphi)}\,\mathrm{d}\varphi = \frac{p-1}{p}\tau = \frac{\tau}{q}. \tag{4.32}$$

A consequence of (4.32) is that the function $\rho(s)\exp\{-\frac{s}{q}\}$ is periodic with the period τ because of

$$\frac{\rho(s+\tau)\exp\{-\frac{s+\tau}{q}\}}{\rho(s)\exp\{-\frac{s}{q}\}} = \frac{\rho(s+\tau)}{\rho(s)}\exp\left\{-\frac{\tau}{q}\right\} = 1.$$

The previous computations in Case (III) are summarized in the next theorem.

THEOREM 4.3. *If* $\gamma > \tilde{\gamma} = (\frac{p-1}{p})^p$ *then Equation* (4.21) *is oscillatory and* $x(s) = \rho(s)\exp\{-\frac{s}{q}\}$ *is a periodic solution of* (4.21) *with the period* τ *given by* (4.31).

REMARK 4.1. Consider the half-linear differential equation

$$\left(t^{\alpha}\Phi(x')\right)' + \frac{\gamma}{t^{p-\alpha}}\Phi(x) = 0. \tag{4.33}$$

If $\alpha \neq p-1$ and we look for a solution of this equation in the form $x(t) = t^{\lambda}$, then substituting into (4.33) we get the algebraic equation for the exponent λ

$$(p-1)|\lambda|^p - (p-1-\alpha)\Phi(\lambda) + \gamma = 0. \tag{4.34}$$

This equation has a real root if and only if $\gamma \leqslant \tilde{\gamma}_\alpha := (\frac{|p-1-\alpha|}{p})^p$ and hence (4.33) with $\alpha \neq p-1$ is nonoscillatory if $\gamma \leqslant \tilde{\gamma}_\alpha$. If $\gamma > \tilde{\gamma}_\alpha$, using the same ideas as in case $\alpha = 0$ treated in main part of this section one can see that (4.33) is oscillatory.

If $\alpha = p-1$ and $\gamma > 0$, Equation (4.34) has no real root and in this case we consider the modified Euler-type equation

$$\left(t^{p-1}\Phi(x')\right)' + \frac{\gamma}{t\lg^p t}\Phi(x) = 0. \tag{4.35}$$

The change of independent variable $t \mapsto \lg t$ transforms (4.35) into Equation (4.20) and the interval $[1,\infty)$ is transformed into the interval $[\mathrm{e},\infty)$. The situation is summarized in the next theorem.

THEOREM 4.4. *If* $\alpha \neq p-1$, *Equation* (4.33) *is nonoscillatory if and only if*

$$\gamma \leqslant \left(\frac{|p-1-\alpha|}{p}\right)^p.$$

Equation (4.35) *is nonoscillatory if and only if*

$$\gamma \leqslant \left(\frac{p-1}{p}\right)^p.$$

4.3. *Kneser-type oscillation and nonoscillation criteria*

As an immediate consequence of the Sturmian comparison theorem and the above result concerning oscillation of Euler equation (4.20), we have the following half-linear version of the classical Kneser oscillation and nonoscillation criterion.

THEOREM 4.5. *Suppose that*

$$\liminf_{t\to\infty} t^p c(t) > \tilde{\gamma} = \left(\frac{p-1}{p}\right)^p. \tag{4.36}$$

Then the equation

$$\left(\Phi(x')\right)' + c(t)\Phi(x) = 0 \tag{4.37}$$

is oscillatory. If

$$\liminf_{t\to\infty} t^p c(t) < \tilde{\gamma}, \tag{4.38}$$

then (4.37) *is nonoscillatory.*

PROOF. If (4.36) holds, then $c(t) > \frac{\tilde{\gamma}+\varepsilon}{t^p}$ for some $\varepsilon > 0$ and since the Euler equation (4.20) with $\gamma = \tilde{\gamma} + \varepsilon$ is oscillatory, (4.37) is also oscillatory by the Sturm comparison theorem (Theorem 2.4). The nonoscillatory part of theorem can be proved using the same argument. □

REMARK 4.2. Using the results of Theorem 4.4 and the Sturm comparison theorem, one can prove various extensions of the previous theorem. For example, if $\alpha \neq p-1$,

$$\liminf_{t\to\infty} t^{-\alpha} r(t) > 1, \qquad \limsup_{t\to\infty} t^{\alpha-p} c(t) < \tilde{\gamma}_\alpha$$

then (0.1) is nonoscillatory. An oscillation counterpart of this result, as well as the criteria in case $\alpha = p-1$ can be formulated in a similar way.

This Page Intentionally Left Blank

CHAPTER 3B

Nonoscillatory Equations

In this chapter we concentrate our attention to nonoscillatory half-linear differential equations while oscillatory equations is the main concern of Chapter 3C. However, we formulate some oscillation criteria already here, mainly in situations when they are natural complements of their nonoscillation counterparts. First we present nonoscillation criteria for (0.1) which are based on the Riccati technique and the variational principle. Section 6 is devoted to the investigation of asymptotic properties of nonoscillatory solutions of (0.1). Then we deal with the important concept of the principal solution of nonoscillatory equation (0.1) and the last section of this chapter presents criteria for conjugacy and disconjugacy of (0.1) in a given interval.

Recall that (0.1) is said to be *nonoscillatory* if there exists $T \in \mathbb{R}$ such that (0.1) is disconjugate on $[T, \infty)$, i.e., every nontrivial solution of this equation has at most one zero in this interval and this means that every nontrivial solution is eventually positive or negative. Equation (0.1) is said to be *oscillatory* in the opposite case, i.e., when every nontrivial solution has infinitely many zeros tending to ∞.

Note that from Theorem 2.2 or from Theorem 2.3 follows that this classification of half-linear equations is correct; all solution of (0.1) are either oscillatory or nonoscillatory.

5. Nonoscillation criteria

From the Roundabout theorem (Theorem 2.2) and also from the Sturmian comparison theorem (Theorem 2.4) it is easy to see that for nonoscillation of (0.1), the function c cannot be "too positive", comparing with the positivity of the function r. The criteria of this section characterize in a quantitative way the vague expression "not too positive".

The Roundabout theorem offers two basic methods of the investigation of oscillatory properties of (0.1). The first one, usually referred to as the *variational principle*, is based on the equivalence of disconjugacy of (0.1) and the positivity of the associated energy functional $\mathcal{F}$. The second main method—*the Riccati technique*—uses the equivalence of disconjugacy (0.1) and solvability of the generalized Riccati equation (2.1).

5.1. *Variational principle and Wirtinger's inequality*

As an immediate consequence of the equivalence of (i) and (iv) in Theorem 2.2 we have the following statement which is used in the proofs of oscillation and nonoscillation criteria based on the variational principle.

THEOREM 5.1. *Equation* (0.1) *is nonoscillatory if and only if there exists* $T \in \mathbb{R}$ *such that*

$$\mathcal{F}(y; T, \infty) := \int_T^\infty \left[r(t)|y'|^p - c(t)|y|^p\right] \mathrm{d}t > 0$$

for every nontrivial $y \in W_0^{1,p}(T, \infty)$.

A useful tool in the variational technique is the following half-linear version of the Wirtinger inequality.

LEMMA 5.1. *Let M be a positive continuously differentiable function for which* $M'(t) \neq 0$ *in* $[a, b]$ *and let* $y \in W_0^{1,p}(a, b)$. *Then*

$$\int_a^b \left|M'(t)\right||y|^p \,\mathrm{d}t \leqslant p^p \int_a^b \frac{M^p(t)}{|M'(t)|^{p-1}}|y'|^p \,\mathrm{d}t. \tag{5.1}$$

PROOF. Suppose that $M'(t) > 0$ in $[a, b]$, in the opposite case the proof is similar. Using integration by parts, the fact that y has a compact support in (a, b), and the Hölder inequality, we have

$$\begin{aligned}\int_a^b \left|M'(t)\right||y|^p \,\mathrm{d}t &\leqslant p \int_a^b M|y|^{p-1}|y'| \,\mathrm{d}t \\ &\leqslant p\left(\int_a^b |M'||y|^{(p-1)q} \,\mathrm{d}t\right)^{\frac{1}{q}} \left(\int_a^b \frac{M^p}{|M'|^{p-1}}|y'|^p \,\mathrm{d}t\right)^{\frac{1}{p}} \\ &= p\left(\int_a^b |M'||y|^p \,\mathrm{d}t\right)^{\frac{1}{q}} \left(\int_a^b \frac{M^p}{|M'|^{p-1}}|y'|^p \,\mathrm{d}t\right)^{\frac{1}{p}},\end{aligned}$$

hence (5.1) holds. ☐

5.2. *Nonoscillation criteria via Wirtinger inequality*

The Wirtinger inequality from the previous subsection is used in the next nonoscillation criterion.

THEOREM 5.2. *Denote* $c_+(t) = \max\{0, c(t)\}$. *If* $\int^\infty r^{1-q}(t) \,\mathrm{d}t = \infty$, $\int^\infty c_+(t) \,\mathrm{d}t < \infty$, *and*

$$\limsup_{t\to\infty}\left(\int^t r^{1-q}(s) \,\mathrm{d}s\right)^{p-1} \left(\int_t^\infty c_+(s) \,\mathrm{d}s\right) < \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}, \tag{5.2}$$

or $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t < \infty$ *and*

$$\limsup_{t\to\infty}\left(\int_t^{\infty} r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\left(\int^{t} c_{+}(s)\,\mathrm{d}s\right) < \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}, \tag{5.3}$$

then (0.1) *is nonoscillatory.*

PROOF. We will prove the statement in case $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty$. If this integral converges, the proof is analogous. Denote

$$\nu := \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}, \qquad M(t) := \left(\int^{t} r^{1-q}(s)\,\mathrm{d}s\right)^{1-p}$$

and let $T \in \mathbb{R}$ be such that the expression in (5.2) is less than ν for $t > T$. Using (5.2), the Hölder inequality and the Wirtinger inequality, we have for any nontrivial $y \in W_0^{1,p}(T,\infty)$

$$\begin{aligned}
\int_T^{\infty} c(t)|y|^p\,\mathrm{d}t &\leqslant \int_T^{\infty} c_{+}(t)|y|^p\,\mathrm{d}t = p\int_T^{\infty} c_{+}(t)\left(\int_T^{t} y'\Phi(y)\,\mathrm{d}s\right)\mathrm{d}t \\
&\leqslant p\int_T^{\infty} |y'|\Phi(y)M(t)\frac{\int_t^{\infty} c_{+}(s)\,\mathrm{d}s}{M(t)}\,\mathrm{d}t \\
&< p\nu\int_T^{\infty} M(t)|y'|\big|\Phi(y)\big|\,\mathrm{d}t \\
&\leqslant p\nu\left(\int_T^{\infty} \big|M'(t)\big||y|^p\,\mathrm{d}t\right)^{\frac{1}{q}}\left(\int_T^{\infty} \frac{|M(t)|^p}{|M'(t)|^{p-1}}|y'|^p\,\mathrm{d}t\right)^{\frac{1}{p}} \\
&\leqslant p^p\nu\int_T^{\infty} \frac{|M(t)|^p}{|M'(t)|^{p-1}}|y'|^p\,\mathrm{d}t = \int_T^{\infty} r(t)|y'|^p\,\mathrm{d}t
\end{aligned}$$

since one may directly verify that

$$\frac{|M(t)|^p}{|M'(t)|^{p-1}} = (p-1)^{1-p}r(t).$$

Hence we have

$$\mathcal{F}(y;T,\infty) = \int_T^{\infty}\big[r(t)|y'|^p - c(t)|y|^p\big]\,\mathrm{d}t > 0$$

for any nontrivial $y \in W_0^{1,p}(T,\infty)$. □

REMARK 5.1. (i) Later we will show that the constant $\frac{1}{p}(\frac{p-1}{p})^{p-1}$ in the previous nonoscillation criterion is sharp in the sense that if the limit in (5.2) (or in (5.3)) is greater than this constant, then (0.1) is oscillatory. Also, when the previous criterion is applied to

Euler-type differential equation (4.20) in the previous chapter, we again reveal the critical constant $\tilde{\gamma} = (\frac{p-1}{p})^p$.

(ii) In (5.2) and (5.3), the nonnegative part of the function c appeared. In the next subsection we present an improvement of the previous nonoscillation criterion, where instead of c_+ the function c directly appears.

(iii) If, in addition to assumptions of the previous theorem, we suppose that $c(t) > 0$ for large t then the second part of this theorem (the case when $\int^\infty r^{1-q}(t)\,\mathrm{d}t < \infty$) can be deduced from the first one using the reciprocity principle from Section 2.7. Indeed, if (5.3) holds, then $\int^\infty c(t)\,\mathrm{d}t = \infty$ and consider the reciprocal equation to (0.1)

$$\big(c^{1-q}(t)\Phi_q(u')\big)' + r^{1-q}(t)\Phi_q(u) = 0, \quad \Phi_q(u) = |u|^{q-1}\operatorname{sgn} u = \Phi^{-1}(u),$$

where q is again the conjugate number of p. This equation, with p replaced by q, r by c^{1-q} and c by r^{1-q} satisfies assumptions of the first part of Theorem 5.2, since (using $(p-1)(q-1) = 1$) $\int^\infty (c^{1-q}(t))^{1-p}\,\mathrm{d}t = \int^\infty c(t)\,\mathrm{d}t = \infty$ and

$$\begin{aligned}
&\limsup_{t\to\infty}\left(\int^t c(s)\,\mathrm{d}s\right)^{q-1}\left(\int_t^\infty r^{1-q}(s)\,\mathrm{d}s\right)\\
&\quad= \limsup_{t\to\infty}\left[\left(\int^t c(s)\,\mathrm{d}s\right)\left(\int_t^\infty r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\right]^{q-1}\\
&\quad< \left[\frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}\right]^{q-1} = \frac{1}{q}\left(\frac{q-1}{q}\right)^{q-1}
\end{aligned}$$

which is just the "reciprocal" counterpart of (5.2).

(iv) If $b = \infty$ and $M(t) = t$ in Lemma 5.1, we get the inequality

$$\int_a^\infty |y'|^p\,\mathrm{d}t \geqslant \tilde{\gamma}\int_a^\infty \frac{|y|^p}{t^p}\,\mathrm{d}t, \quad \tilde{\gamma} = \left(\frac{p-1}{p}\right)^p,$$

which is a Hardy-type inequality, see, e.g., [131]. This inequality has been extended in many directions and some of these extensions could be perhaps used to establish more sophisticated nonoscillation criteria than that given in Theorem 5.2. This problem is a subject of the present investigation.

5.3. *Riccati inequality*

From the Roundabout theorem (Theorem 2.2) it follows that nonoscillation of (0.1) is equivalent to solvability of the associated Riccati equation (2.1). Due to the Sturm comparison theorem, nonoscillation of (0.1) is actually equivalent to solvability of the Riccati *inequality*. This is formulated in the next statement.

THEOREM 5.3. *Equation* (0.1) *is nonoscillatory if and only if there exists a continuously differentiable function v defined on an interval* $[T,\infty)$ *and satisfying there the inequality*

$$v' + c(t) + (p-1)r^{1-q}(t)|v|^q \leqslant 0. \tag{5.4}$$

PROOF. Let v be a solution of (5.4). Denote $C(t) := -v' - (p-1)r^{1-q}(t)|v|^q$. Then v is a solution of $v' + C(t) + (p-1)r^{1-q}(t)|v|^q = 0$ which is the Riccati equation associated with a Sturmian majorant of (0.1) (since $C(t) \geqslant c(t)$). This majorant equation is nonoscillatory and hence (0.1) is nonoscillatory as well. □

The previous statement is used in the half-linear version of the classical Wintner criterion. This linear criterion claims that if $\int^\infty r^{-1}(t)\,\mathrm{d}t = \infty$ and $\int^\infty c(t)\,\mathrm{d}t$ converges, then the linear Sturm–Liouville equation (1.1) is nonoscillatory provided

$$\limsup_{t\to\infty}\left(\int^t r^{-1}(s)\,\mathrm{d}s\right)\left(\int_t^\infty c(s)\,\mathrm{d}s\right) < \frac{1}{4}$$

and

$$\liminf_{t\to\infty}\left(\int^t r^{-1}(s)\,\mathrm{d}s\right)\left(\int_t^\infty c(s)\,\mathrm{d}s\right) > -\frac{3}{4}.$$

The next half-linear extension of this linear criterion is proved in [60].

THEOREM 5.4. *Suppose that* $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$ *and* $\int^\infty c(t)\,\mathrm{d}t = \lim_{b\to\infty}\int^b c(t)\,\mathrm{d}t$ *converges. If*

$$\limsup_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\left(\int_t^\infty c(s)\,\mathrm{d}s\right) < \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}, \tag{5.5}$$

$$\liminf_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\left(\int_t^\infty c(s)\,\mathrm{d}s\right) > -\frac{2p-1}{p}\left(\frac{p-1}{p}\right)^{p-1}, \tag{5.6}$$

then (0.1) *is nonoscillatory.*

PROOF. We will find a solution of the Riccati type inequality

$$v' \leqslant -c(t) - (p-1)r^{1-q}(t)|v|^q \tag{5.7}$$

which is extensible up to ∞, i.e., it exists on some interval $[T,\infty)$. To find this solution v of (5.7), we show that there exists an extensible up to ∞ solution of the differential inequality

$$\rho' \leqslant (1-p)r^{1-q}(t)\big|\rho + C(t)\big|^q, \quad C(t) := \int_t^\infty c(s)\,\mathrm{d}s \tag{5.8}$$

related to (5.7) by the substitution $\rho = v - C$. This solution ρ is

$$\rho(t) = \beta\biggl(\int^t r^{1-q}(s)\,\mathrm{d}s\biggr)^{1-p}, \quad \beta := \biggl(\frac{p-1}{p}\biggr)^p.$$

Indeed, $\rho' = (1-p)\beta r^{1-q}(t)(\int^t r^{1-q}(s)\,\mathrm{d}s)^{-p}$ and the right-hand side of inequality (5.8) is

$$\begin{aligned}
&(1-p)r^{1-q}(t)\bigl|\rho + C(t)\bigr|^q \\
&\quad= (1-p)r^{1-q}(t)\biggl|\beta\biggl(\int^t r^{1-q}\biggr)^{1-p} + C(t)\biggr|^q \\
&\quad= (1-p)r^{1-q}(t)\biggl|\beta + \biggl(\int^t r^{1-q}\biggr)^{p-1} C(t)\biggr|^q \biggl(\int^t r^{1-q}\biggr)^{(1-p)q} \\
&\quad= (1-p)r^{1-q}(t)\biggl|\beta + \biggl(\int^t r^{1-q}\biggr)^{p-1} C(t)\biggr|^q \biggl(\int^t r^{1-q}\biggr)^{-p}.
\end{aligned}$$

Consequently, (5.8) is equivalent to the inequality

$$\beta \geqslant \biggl|\beta + \biggl(\int^t r^{1-q}\biggr)^{p-1} C(t)\biggr|^q. \tag{5.9}$$

However, since (5.5) and (5.6) hold, there exists $\varepsilon > 0$ such that

$$-\frac{2p-1}{p}\biggl(\frac{p-1}{p}\biggr)^{p-1} + \varepsilon < \biggl(\int^t r^{1-q}\biggr)^{p-1} C(t) < \frac{1}{p}\biggl(\frac{p-1}{p}\biggr)^{p-1} - \varepsilon$$

for large t and by a direct computation it is not difficult to verify that (5.9) really holds. □

If the integral $\int^\infty r^{1-q}(t)\,\mathrm{d}t$ is convergent, the previous statement can be modified as follows.

THEOREM 5.5. *Suppose that* $\int^\infty r^{1-q}(t)\,\mathrm{d}t < \infty$. *If*

$$\limsup_{t\to\infty}\biggl(\int_t^\infty r^{1-q}(s)\,\mathrm{d}s\biggr)^{p-1}\biggl(\int^t c(s)\,\mathrm{d}s\biggr) < \frac{1}{p}\biggl(\frac{p-1}{p}\biggr)^{p-1}$$

and

$$\liminf_{t\to\infty}\biggl(\int_t^\infty r^{1-q}(s)\,\mathrm{d}s\biggr)^{p-1}\biggl(\int^t c(s)\,\mathrm{d}s\biggr) > -\frac{2p-1}{p}\biggl(\frac{p-1}{p}\biggr)^{p-1},$$

then (0.1) *is nonoscillatory.*

PROOF. One can show in the same way as in the proof of Theorem 5.4 that the function

$$\rho(t) = -\left(\frac{p-1}{p}\right)^{p}\left(\int_t^{\infty} r^{1-q}(s)\,\mathrm{d}s\right)^{1-p},$$

satisfies the inequality

$$\rho' \leqslant (1-p)r^{1-q}(t)\left|\rho - \widetilde{C}(t)\right|^{q}, \quad \widetilde{C}(t) = \int^{t} c(s)\,\mathrm{d}s,$$

which implies that $v = \rho - \widetilde{C}$ satisfies the Riccati inequality (5.4). □

5.4. *Half-linear Hartman–Wintner theorem*

The next theorem is a half-linear extension of the classical Hartman–Wintner theorem [106] which relates the square integrability of the solutions of the Riccati equation

$$w' + c(t) + w^2 = 0$$

corresponding to (1.1) with $r(t) \equiv 1$ to the finiteness of a certain limit involving the function c. The half-linear extension of this theorem (also in case $r(t) \equiv 1$) is proved in [153], but the modification of this proof to (0.1) as presented here is straightforward.

THEOREM 5.6. *Suppose that*

$$\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty \tag{5.10}$$

and (0.1) *is nonoscillatory. Then the following statements are equivalent.*

(i) *It holds*

$$\int^{\infty} r^{1-q}(t)\left|w(t)\right|^{q}\,\mathrm{d}t < \infty \tag{5.11}$$

for every solution w of (2.1).

(ii) *There exists a finite limit*

$$\lim_{t\to\infty} \frac{\int^{t} r^{1-q}(s)\left(\int^{s} c(\tau)\,\mathrm{d}\tau\right)\mathrm{d}s}{\int^{t} r^{1-q}(s)\,\mathrm{d}s}. \tag{5.12}$$

(iii) *For the lower limit we have*

$$\liminf_{t\to\infty} \frac{\int^{t} r^{1-q}(s)\left(\int^{s} c(\tau)\,\mathrm{d}\tau\right)\mathrm{d}s}{\int^{t} r^{1-q}(s)\,\mathrm{d}s} > -\infty. \tag{5.13}$$

PROOF. (i) ⇒ (ii): Nonoscillation of (0.1) implies that the Riccati equation (2.1) has a solution which is defined on an interval $[T, \infty)$. Integrating this equation from T to t and using (5.11) we have

$$\begin{aligned} w(t) &= w(T) - \int_T^t c(\tau)\,\mathrm{d}\tau - (p-1)\int_T^t r^{1-q}(\tau)\big|w(\tau)\big|^q\,\mathrm{d}\tau \\ &= w(T) - \int_T^t c(\tau)\,\mathrm{d}\tau - (p-1)\int_T^\infty r^{1-q}(\tau)\big|w(\tau)\big|^q\,\mathrm{d}\tau \\ &\quad + (p-1)\int_t^\infty r^{1-q}(\tau)\big|w(\tau)\big|^q\,\mathrm{d}\tau \\ &= C - \int_T^t c(\tau)\,\mathrm{d}\tau + (p-1)\int_t^\infty r^{1-q}(\tau)\big|w(\tau)\big|^q\,\mathrm{d}\tau, \end{aligned} \tag{5.14}$$

where $C = w(T) + (p-1)\int_T^\infty r^{1-q}(\tau)|w(\tau)|^q\,\mathrm{d}\tau$. Multiplying (5.14) by r^{1-q} and integrating the resulting equation from T to t, and then dividing by $\int_T^t r^{1-q}(s)\,\mathrm{d}s$, we get

$$\begin{aligned} \frac{\int_T^t r^{1-q}(s)w(s)\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s} &= C - \frac{\int_T^t r^{1-q}(s)(\int_T^s c(\tau)\,\mathrm{d}\tau)\,\mathrm{d}s}{\int_T^t r^{1-q}(\tau)\,\mathrm{d}\tau} \\ &\quad + \frac{\int_T^t r^{1-q}(s)(\int_s^\infty r^{1-q}(\tau)|w(\tau)|^q\,\mathrm{d}\tau)\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s}. \end{aligned} \tag{5.15}$$

By the Hölder inequality we have

$$\begin{aligned} \left|\int_T^t r^{1-q}(s)w(s)\,\mathrm{d}s\right| &= \left|\int_T^t r^{\frac{1-q}{p}}(s) r^{\frac{1-q}{q}}(s)w(s)\,\mathrm{d}s\right| \\ &\leqslant \left(\int_T^t r^{1-q}(s)\,\mathrm{d}s\right)^{\frac{1}{p}}\left(\int_T^t r^{1-q}(s)\big|w(s)\big|^q\,\mathrm{d}s\right)^{\frac{1}{q}}, \end{aligned} \tag{5.16}$$

and hence, taking into account (5.11)

$$\begin{aligned} \left|\frac{\int_T^t r^{1-q}(s)w(s)\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s}\right| &\leqslant \frac{(\int_T^t r^{1-q}(s)\,\mathrm{d}s)^{\frac{1}{p}}(\int_T^t r^{1-q}(s)|w(s)|^q)^{\frac{1}{q}}}{\int_T^t r^{1-q}(s)\,\mathrm{d}s} \\ &= \left(\frac{\int_T^t r^{1-q}(s)|w(s)|^q}{\int_T^t r^{1-q}(s)\,\mathrm{d}s}\right)^{\frac{1}{q}} \to 0, \quad t \to \infty. \end{aligned}$$

Since also the last term in (5.15) tends to zero as $t \to \infty$ (again in view of (5.11)), we have that

$$\lim_{t\to\infty} \frac{\int_T^t r^{1-q}(s) \int_T^s c(\tau)\,\mathrm{d}\tau}{\int_T^t r^{1-q}(s)\,\mathrm{d}s} = C \quad \text{exists finite.}$$

(ii) $\Rightarrow$ (iii): This implication is trivial.

(iii) $\Rightarrow$ (i): Let w be any solution of (2.1) which exists on $[T, \infty)$. Then by (5.14) and using computation from the first part of this proof

$$\begin{aligned}\frac{\int_T^t r^{1-q}(s) w(s)\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s} &= w(T) - \frac{\int_T^t r^{1-q}(s) \int_T^s c(\tau)\,\mathrm{d}\tau}{\int_T^t r^{1-q}(\tau)\,\mathrm{d}\tau} \\ &\quad - (p-1)\frac{\int_T^t r^{1-q}(s)(\int_T^t r^{1-q}(\tau)|w(\tau)|^q\,\mathrm{d}\tau)\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s}.\end{aligned}$$

Taking into account (5.13) and applying again the Hölder inequality, there exists a real constant K such that

$$\begin{aligned}&-\frac{(\int_T^t r^{1-q}(s)|w(s)|^q\,\mathrm{d}s)^{\frac{1}{q}}}{(\int_T^t r^{1-q}(s)\,\mathrm{d}s)^{\frac{1}{q}}} \\ &\quad \leqslant K - (p-1)\frac{\int_T^t r^{1-q}(s)(\int_T^s r^{1-q}(\tau)|w(\tau)|^q\,\mathrm{d}\tau)\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s}.\end{aligned}$$

Suppose that (5.11) fails to holds. Then by L'Hospital's rule the last term in the previous inequality tends to ∞ and hence

$$\left(\frac{\int_T^t r^{1-q}(s)|w(s)|^q\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s}\right)^{\frac{1}{q}} \geqslant \frac{p-1}{p}\frac{\int_T^t r^{1-q}(s)(\int_T^s r^{1-q}(\tau)|w(\tau)|^q\,\mathrm{d}\tau)\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s}$$

for large t. Denote $S(t) = \int_T^t r^{1-q}(s)(\int_T^s r^{1-q}(\tau)|w(\tau)|^q\,\mathrm{d}\tau)\,\mathrm{d}s$. Then the last inequality reads

$$\left[\frac{S'(t) r^{q-1}(t)}{\int^t r^{1-q}(s)\,\mathrm{d}s}\right]^{\frac{1}{q}} \geqslant \frac{1}{q}\frac{S(t)}{\int_T^t r^{1-q}(s)\,\mathrm{d}s},$$

hence

$$\frac{S'(t)}{S^q(t)} \geqslant \left(\frac{1}{g}\right)^q \frac{r^{1-q}(t)}{(\int_T^t r^{1-q}(s)\,\mathrm{d}s)^{q-1}}. \tag{5.17}$$

If $q \leqslant 2$, we integrate (5.17) from T_1 to t, $T_1 > T$, and we get

$$\begin{aligned}\frac{1}{q-1}S^{1-q}(T_1) &> \frac{1}{q-1}\big[S^{1-q}(T_1) - S^{1-q}(t)\big]\\ &\geqslant \left(\frac{1}{q}\right)^q \begin{cases} \lg\big(\int_T^t r^{1-q}(s)\,\mathrm{d}s\big) & \text{if } q = 2,\\ \frac{1}{2-q}\big(\int_T^t r^{1-q}(s)\,\mathrm{d}s\big)^{2-q} & \text{if } q < 2.\end{cases}\end{aligned}$$

Letting $t \to \infty$ we have a contradiction with the assumption that $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$. If $q > 2$, we integrate (5.17) from t to ∞ and we obtain

$$\frac{1}{(q-1)S^{q-1}(t)} \geqslant \left(\frac{1}{q}\right)^q \frac{1}{(q-2)(\int^t r^{1-q}(s)\,\mathrm{d}s)^{q-2}},$$

hence

$$\frac{q^q(q-2)}{q-1} \geqslant \left(\frac{S(t)}{\int^t r^{1-q}(s)\,\mathrm{d}s}\right)^{q-1}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right),$$

which is again a contradiction since $S(t)(\int^t r^{1-q}(s)\,\mathrm{d}s)^{-1} \to \infty$ as $t \to \infty$. □

As a direct consequence of the previous theorem we have the following oscillation criterion.

THEOREM 5.7. *Suppose that* $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$. *Then each of the following two conditions is sufficient for oscillation of* (0.1):

$$\lim_{t\to\infty} \frac{\int^t r^{1-q}(s)(\int^s c(\tau)\,\mathrm{d}\tau)\,\mathrm{d}s}{\int^t r^{1-q}(s)\,\mathrm{d}s} = \infty, \tag{5.18}$$

$$\begin{aligned}-\infty &< \liminf_{t\to\infty} \frac{\int^t r^{1-q}(s)(\int^s c(\tau)\,\mathrm{d}\tau)\,\mathrm{d}s}{\int^t r^{1-q}(s)\,\mathrm{d}s}\\ &< \limsup_{t\to\infty} \frac{\int^t r^{1-q}(s)(\int^s c(\tau)\,\mathrm{d}\tau)\,\mathrm{d}s}{\int^t r^{1-q}(s)\,\mathrm{d}s}.\end{aligned} \tag{5.19}$$

PROOF. We will prove sufficiency of (5.18) only, the proof of sufficiency of (5.19) is similar. Suppose that (0.1) is nonoscillatory and (5.18) holds. Then (5.13) holds and by the previous theorem the integral (5.11) converges for every solution w of (2.1) and hence limit (5.12) exists as a finite number which contradicts to (5.18). □

5.5. *Riccati integral equation and Hille–Wintner comparison theorem*

The results of this subsection are taken essentially from the paper [136]. In that paper, it is supposed that $r(t) \equiv 1$ and $c(t) \geqslant 0$ for large t. However, the results as presented here can be extended to (0.1) with r satisfying the below assumption (5.20), and without the sign restriction on the function c.

LEMMA 5.2. *Suppose that*

$$\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty \tag{5.20}$$

and the integral $\int^{\infty} c(t)\,\mathrm{d}t$ *is convergent. Then* (0.1) *is nonoscillatory if and only if there exists a solution of the Riccati integral equation*

$$w(t) = \int_t^{\infty} c(s)\,\mathrm{d}s + (p-1)\int_t^{\infty} r^{1-q}(s)\big|w(s)\big|^q\,\mathrm{d}s. \tag{5.21}$$

PROOF. Suppose that (0.1) is nonoscillatory and let w be a solution of the associated Riccati equation (2.1) which is defined on some interval $[T_0, \infty)$. The convergence of the integral $\int^{\infty} c(t)\,\mathrm{d}t$ and (5.20) imply that (5.11) holds by Theorem 5.6. Integrating (2.1) from t to T, $t \geqslant T_0$ and letting $T \to \infty$ we see that $\lim_{T\to\infty} w(T)$ exists and since (5.20) holds, this limit equals zero, i.e., w satisfies also (5.21). Conversely, if w is a solution of (5.21), than it is also a solution of (2.1) and hence (0.1) is nonoscillatory. □

The previous lemma is used in the proof of the following half-linear extension of the Hille–Wintner comparison theorem. For its linear version see [208, Theorem 2.14].

THEOREM 5.8. *Together with* (0.1) *consider the equation*

$$\big(r(t)\Phi(y')\big)' + \tilde{c}(t)\Phi(y) = 0, \tag{5.22}$$

where r satisfies (5.20), $\tilde{c}$ *is continuous for large and* $\int^{\infty} \tilde{c}(t)\,\mathrm{d}t$ *converges. If*

$$0 \leqslant \int_t^{\infty} c(s)\,\mathrm{d}s \leqslant \int_t^{\infty} \tilde{c}(s)\,\mathrm{d}s \quad \textit{for large } t \tag{5.23}$$

and (5.22) *is nonoscillatory, then* (0.1) *is also nonoscillatory.*

PROOF. We construct a solution w of (5.21). Nonoscillation of (5.22) implies the existence of a solution v of the associated integral equation

$$v(t) = \int_t^{\infty} \tilde{c}(s)\,\mathrm{d}s + (p-1)\int_t^{\infty} r^{1-q}(s)\big|v(s)\big|^q\,\mathrm{d}s. \tag{5.24}$$

Let $T \in \mathbb{R}$ be such that (5.23) holds for $t \geqslant T$ and the solution v of (5.24) exists on $[T, \infty)$. Define the function set U and the mapping F by

$$U := \big\{u \in C[T,\infty):\ 0 \leqslant u(t) \leqslant v(t),\ t \geqslant T\big\}$$

and

$$(Fw)(t) = \int_t^\infty c(s)\,\mathrm{d}s + (p-1)\int_t^\infty r^{1-q}(s)\bigl|w(s)\bigr|^q\,\mathrm{d}s, \quad t \geqslant T.$$

The set U is the convex and closed subset of the Fréchet space $C[T,\infty)$ with the topology of the uniform convergence on compact subintervals of $[T,\infty)$. It can be shown without difficulty that F maps U into itself, that F is continuous and that $F(U)$ is relatively compact subset of $C[T,\infty)$. Therefore, it follows from the Schauder–Tychonov fixed point theorem that there exists $w \in U$ such that $w = Fw$ and this function is by definition of F a solution of (5.21). ☐

5.6. *Hille–Nehari criteria*

In Section 4.2 we have shown that the Euler equation (4.20) is nonoscillatory if and only if $\gamma \leqslant \tilde{\gamma} = (\frac{p-1}{p})^p$. Consider now the equation

$$\bigl(r(t)\Phi(x')\bigr)' + \frac{\gamma r^{1-q}(t)}{(\int^t r^{1-q}(s)\,\mathrm{d}s)^p}\Phi(x) = 0 \tag{5.25}$$

with r satisfying (5.20). The transformation of independent variable $t \mapsto \int^t r^{1-q}(s)\,\mathrm{d}s$ transforms this equation into the Euler equation (4.20). Hence also (5.25) is nonoscillatory if and only if $\gamma \leqslant \tilde{\gamma}$. This fact, combined with Theorem 5.6, leads to the following nonoscillation and oscillation criteria which are the half-linear extension of the Hille–Nehari (non)oscillation criteria, see [208, Chapter II].

THEOREM 5.9. *Suppose that* $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$ *and the integral* $\int^\infty c(t)\,\mathrm{d}t$ *is convergent.*

(i) *If*

$$0 \leqslant \left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\left(\int_t^\infty c(s)\,\mathrm{d}s\right) \leqslant \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}$$

for large t, then (0.1) *is nonoscillatory.*

(ii) *If*

$$\liminf_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\left(\int_t^\infty c(s)\,\mathrm{d}s\right) > \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}, \tag{5.26}$$

then (0.1) *is oscillatory.*

PROOF. First of all observe that (5.26) implies that $\int_t^\infty c(s)\,\mathrm{d}s > 0$ for large t. Now, since

$$\int_t^\infty \frac{\tilde{\gamma} r^{1-q}(s)}{(\int^s r^{1-q}(\tau)\,\mathrm{d}\tau)^p} = \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}\left(\int^t r^{1-q}(s)\right)^{1-p},$$

the statement follows from Theorem 5.8 with $\tilde{c}(t) = \tilde{\gamma} r^{1-q}(t)(\int^t r^{1-q}(s)\,\mathrm{d}s)^{-p}$. □

5.7. *Modified Hille–Nehari's criteria*

The results of this subsection are taken from [135]. It is supposed that $c(t) \geqslant 0$ for large t and that

$$\int^{\infty} r^{1-q}(t)\,\mathrm{d}t < \infty. \tag{5.27}$$

We denote

$$\varrho(t) = \int_t^{\infty} r^{1-q}(s)\,\mathrm{d}s.$$

The first auxiliary statement concerns boundedness of solutions of (0.1) and of the associated Riccati equation.

LEMMA 5.3. *Let x be a nonoscillatory solution of* (0.1) *and let $w = r\Phi(x')/\Phi(x)$ be the associated solution of* (2.1). *Then x and the function*

$$z(t) := \varrho^{p-1} w(t) \tag{5.28}$$

are bounded. Moreover,

$$\varrho^{p-1}(t) w(t) \geqslant -1 \quad \textit{for large } t \tag{5.29}$$

and

$$\limsup_{t\to\infty} \varrho^{p-1}(t) w(t) \leqslant 0. \tag{5.30}$$

PROOF. Without loss of generality we can suppose that $x(t) > 0$ for $t \in [t_0, \infty)$. The function $r(t)\Phi(x')$ is nonincreasing (since its derivative equals $-c(t)\Phi(x) \leqslant 0$), the derivative x' is eventually of constant sign. That is, $x'(t) > 0$ for $t \geqslant t_0$ or there is $t_1 > t_0$ such that $x'(t) < 0$ for $t \geqslant t_1$, and that

$$r^{q-1}(s)x'(s) \leqslant r^{q-1}(t)x'(t) \quad \text{for } s \geqslant t \geqslant t_0.$$

Dividing this inequality by $r^{q-1}(s)$ and integrating it over $[t, \tau]$ gives

$$x(\tau) \leqslant x(t) + r^{q-1}(t)x'(t) \int_t^{\tau} r^{1-q}(s)\,\mathrm{d}s. \tag{5.31}$$

If $x'(t) > 0$ for $t \geqslant t_0$, then we have from (5.31)

$$x(\tau) \leqslant x(t) + r^{q-1}(t)x'(t)\varrho(t)$$

which shows that x is bounded on $[t_0, \infty)$. If $x'(t) < 0$ for $t \geqslant t_1$, then x is clearly bounded and, letting $\tau \to \infty$ in (5.31) we have

$$0 \leqslant x(t) + r^{q-1}(t)x'(t)\varrho(t), \quad t \geqslant t_0.$$

In either case we obtain

$$\varrho(t)r^{q-1}(t)\frac{x'(t)}{x(t)} \geqslant -1,$$

which immediately implies (5.29).

The limit inequality (5.30) trivially holds if $x'(t) < 0$ for $t \geqslant t_1$, since in this case the function (5.28) itself is negative for $t \geqslant t_1$. If $x'(t) > 0$ for $t \geqslant t_0$, then there exist positive constants c_1, c_2 such that

$$x(t) \geqslant c_1 \quad \text{and} \quad r(t)\Phi(x'(t)) \leqslant c_2 \quad \text{for } t \geqslant t_0,$$

which implies

$$w(t) \leqslant \frac{c_2}{c_1^{p-1}}, \quad t \geqslant t_0.$$

Since $\varrho(t) \to 0$ as $t \to \infty$, we then conclude that

$$\lim_{t\to\infty} \varrho^{p-1}(t)w(t) = 0.$$

This completes the proof. □

Based on the previous lemma we show that nonoscillation of (0.1) is equivalent to solvability of a certain modified Riccati integral inequality.

THEOREM 5.10. *Equation* (0.1) *is nonoscillatory if and only if*

$$\int^{\infty} \varrho^p(t)c(t)\,\mathrm{d}t < \infty \tag{5.32}$$

and there exists a continuous function v such that

$$\varrho^{p-1}v(t) \quad \textit{is bounded}, \qquad \varrho^{p-1}(t)v(t) \geqslant -1, \tag{5.33}$$

and

$$\begin{aligned}\varrho^p(t)v(t) \geqslant \int_t^{\infty} \varrho^p(s)c(s)\,\mathrm{d}s + p\int_t^{\infty} r^{1-q}(s)\varrho^{p-1}(s)v(s)\,\mathrm{d}s \\ + (p-1)\int_t^{\infty} r^{1-q}(s)\varrho^p(s)\big|v(s)\big|^q\,\mathrm{d}s\end{aligned} \tag{5.34}$$

for large t.

PROOF. "⇒": Let x be a solution of (0.1) such that $x(t) \neq 0$ for $t \geqslant t_0$ and let $w = r\Phi(x')/\Phi(x)$ be the corresponding solution of the Riccati equation (2.1). Multiplying this equation by $\rho^p(t)$ and integrating over $[t, \tau]$, $\tau \geqslant t \geqslant t_0$, we get

$$\rho^p(\tau)w(\tau) - \rho^p(t)w(t) = -p\int_t^\tau r^{1-q}(s)\rho^{p-1}(s)w(s)\,\mathrm{d}s - \int_t^\tau \rho^p(s)c(s)\,\mathrm{d}s \\ - (p-1)\int_t^\tau r^{1-q}(s)\rho^p(s)\big|w(s)\big|^q\,\mathrm{d}s. \tag{5.35}$$

In view of boundedness of the function $\rho^{p-1}w(t)$ (compare the previous lemma), we see that $\rho^p(\tau)w(\tau) = \rho(\tau)\rho^{p-1}(\tau)w(\tau) \to 0$ as $\tau \to \infty$, and

$$\left|\int_t^\infty r^{1-q}(s)\rho^{p-1}(s)w(s)\,\mathrm{d}s\right| \leqslant \int_t^\infty r^{1-q}(s)\big|\rho^{p-1}(s)w(s)\big|\,\mathrm{d}s < \infty,$$

$$\int_t^\infty r^{1-q}(s)\rho^p(s)\big|w(s)\big|^q\,\mathrm{d}s < \infty$$

for $t \geqslant t_0$. Therefore, letting $\tau \to \infty$ in (5.35), we find that $\int_t^\infty \rho^p(s)c(s)\,\mathrm{d}s$ is convergent, i.e., (5.32) holds, and

$$\rho^p(t)w(t) = \int_t^\infty \rho^p(s)c(s)\,\mathrm{d}s + p\int_t^\infty r^{1-q}(s)\rho^{p-1}(s)w(s)\,\mathrm{d}s \\ + (p-1)\int_t^\infty r^{1-q}(s)\rho^p(s)\big|w(s)\big|^q\,\mathrm{d}s$$

hence (5.34) holds as equality. The inequality $\rho^p(t)w(t) \geqslant -1$ follows from the previous lemma.

"⇐": Suppose that (5.32) holds and let w be a continuous function satisfying conditions of theorem. Further, let us denote $C[t_0, \infty)$ the Fréchet space of continuous functions with the topology of uniform convergence on compact subintervals of $[t_0, \infty)$. Consider the space

$$\mathcal{V} := \big\{v \in C[t_0, \infty)\colon -1 \leqslant v(t) \leqslant \rho^{p-1}(t)w(t),\ t \geqslant t_0\big\}, \tag{5.36}$$

which is a closed convex subset of $C[t_0, \infty)$. Define the mapping $F : \mathcal{V} \to C[t_0, \infty)$ by

$$\rho(t)(Fv)(t) = \int_t^\infty \rho^p(s)c(s)\,\mathrm{d}s + p\int_t^\infty r^{1-q}(s)v(s)\,\mathrm{d}s \\ + (p-1)\int_t^\infty r^{1-q}(s)\big|v(s)\big|^q\,\mathrm{d}s. \tag{5.37}$$

If $v \in \mathcal{V}$, then from (5.36), (5.37) and the inequality stated in theorem

$$(Fv)(t) \leqslant F\big(\rho^{p-1}w\big)(t) \leqslant \rho^{p-1}(t)w(t), \quad t \geqslant t_0,$$

and

$$\rho(t)\big[(Fv)(t)+1\big] \geqslant \int_t^\infty r^{1-q}(s)\big[(p-1)|v(s)|^q + pv(s) + 1\big]\,\mathrm{d}s \geqslant 0,$$

where we have used also the property that the function $(p-1)|\xi|^q + p\xi$ is strictly increasing for $\xi \geqslant -1$, i.e.,

$$(p-1)|\xi|^q + p\xi + 1 \geqslant 0, \quad \text{for } \xi \geqslant -1.$$

This shows that F maps $\mathcal{V}$ into itself. It can be shown in a routine manner that F is continuous and $F(\mathcal{V})$ is relatively compact in the topology of $C[t_0,\infty)$. Therefore, by the Schauder–Tychonov fixed point theorem, there exists an element $v \in \mathcal{V}$ such that $v(t) = (Fv)(t)$. Define w by $w(t) = \frac{v(t)}{\rho^{p-1}(t)}$. Then, in view of (5.37), w satisfies the integral equation

$$\rho^p(t)w(t) = \int_t^\infty \rho^p(t)c(s)\,\mathrm{d}s + p\int_t^\infty r^{1-q}(s)\rho^{p-1}w(s)\,\mathrm{d}s \\ + (p-1)\int_t^\infty r^{1-q}(s)\rho^p(s)\big|w(s)\big|^q\,\mathrm{d}s.$$

Differentiating this equality and then dividing by $\rho^p(t)$ shows that w solves Riccati equation (2.1) and hence (0.1) is nonoscillatory. □

As an immediate consequence of the previous theorem we have the following oscillation criterion.

COROLLARY 5.1. *Equation* (0.1) *is oscillatory if*

$$\int^\infty c(s)\rho^p(t)\,\mathrm{d}t = \infty. \tag{5.38}$$

This oscillation criterion opens a natural question about oscillatory nature of (0.1) when the integral in (5.38) is convergent. In answering this question a useful role is played by the following modification of the Hille–Wintner comparison theorem. Recall that we assume that $c(t) \geqslant 0$ for large t throughout this subsection.

THEOREM 5.11. *Consider the pair of Equations* (0.1) *and*

$$\big(r(t)\Phi(y')\big)' + \tilde{c}(t)\Phi(y) = 0, \tag{5.39}$$

where $\tilde{c}(t) \geqslant 0$*, subject to the conditions*

$$\int^{\infty} c(t)\rho^p(t)\,\mathrm{d}t < \infty, \qquad \int^{\infty} \tilde{c}(t)\rho^p(t)\,\mathrm{d}t < \infty. \tag{5.40}$$

Suppose that

$$\int_t^{\infty} c(t)\rho^p(t)\,\mathrm{d}t \leqslant \int_t^{\infty} \tilde{c}(t)\rho^p(t)\,\mathrm{d}t. \tag{5.41}$$

Then nonoscillation of (5.39) *implies that of* (0.1)*, or equivalently, oscillation of* (0.1) *implies oscillation of* (5.39)*.*

PROOF. Assume that (5.39) is nonoscillatory. Then, by the part “only if” of Theorem 5.10 there exists a continuous function w satisfying (5.33) and

$$\begin{aligned}\rho^p(t)w(t) \geqslant{}& \int_t^{\infty} \rho^p(s)\tilde{c}(s)\,\mathrm{d}s + p\int_t^{\infty} r^{1-q}(s)\rho^p(s)w(s)\,\mathrm{d}s\\ &+ (p-1)\int_t^{\infty} r^{1-q}(s)\rho^p(s)\big|w(s)\big|^q\,\mathrm{d}s.\end{aligned}$$

Using (5.40) and (5.41) we see that w satisfies the integral inequality (5.34) and hence (0.1) is nonoscillatory by the part “if” of Theorem 5.10. □

In Remark 4.1 we have shown that the Euler-type half-linear differential equation (4.33) is nonoscillatory if and only if $\gamma < (\frac{|p-1-\alpha|}{p})^p$. The transformation of independent variable $u(s) = x(t)$, $s = s(t) = (\rho(t))^{\frac{p-1}{p-1-\alpha}}$, transforms (0.1) into the equation

$$\big(s^{\alpha}\Phi(u')\big)' + Q(s)\Phi(u) = 0, \tag{5.42}$$

where

$$Q(s) = \left(\frac{|\alpha - p + 1|}{p}\right)^p r^{1-q}\big(t(s)\big)\big[\rho\big(t(s)\big)\big]^{\frac{p-1}{\alpha-p+1}} c\big(t(s)\big),$$

$t = t(s)$ being the inverse function of $s = s(t)$. Now suppose that $\alpha > p - 1$. Then we have $\int^{\infty} s^{\alpha(1-q)}\,\mathrm{d}s < \infty$, so (5.42) satisfies assumption (5.27). Comparing (5.42) with the Euler equation (4.33) (using Theorem 5.11) we have the following result which we present without proof.

THEOREM 5.12. *Suppose that* (5.27) *holds and the integral* $\int^{\infty} \rho^p(t)c(t)\,\mathrm{d}t$ *is convergent.*

(i) *Equation* (0.1) *is oscillatory if*

$$\liminf_{t\to\infty} \rho^{-1}(t)\int_t^{\infty} c(s)\rho^p(s)\,\mathrm{d}s > \left(\frac{p-1}{p}\right)^p.$$

(ii) *Equation* (0.1) *is nonoscillatory if*

$$\rho^{-1}(t)\int_t^{\infty} c(s)\rho^p(s)\,\mathrm{d}s \leqslant \left(\frac{p-1}{p}\right)^p$$

for large t.

5.8. *Comparison theorem with respect to p*

Along with (0.1) we consider another half-linear equation with a different power function $\Phi_\alpha(x) = |x|^{\alpha-1}\operatorname{sgn} x$, $\alpha > 1$,

$$\big(r(t)\Phi_\alpha(x')\big)' + c(t)\Phi_\alpha(x) = 0, \tag{5.43}$$

we denote by β the conjugate number of α, i.e., $\beta = \frac{\beta}{\beta-1}$ (recall also that q is the conjugate number of p, i.e., $q = \frac{p}{p-1}$).

The main statement of this subsection gives a kind of comparison theorem with respect to the power of Φ. This statement is proved in [201] in a more general setting than presented here (in the scope of the so-called half-linear dynamic equations on time scales, compare Section 16.5 in the last chapter). However, we prefer here the formulation for (0.1) and (5.43).

THEOREM 5.13. *Let* $r^{1-\beta}(t)\,\mathrm{d}t = \infty$, $\int^\infty c(t)\,\mathrm{d}t$ *converges and* $\liminf_{t\to\infty} r(t) > 0$. *If* $\alpha \geqslant p$ *and Equation* (5.43) *is nonoscillatory, then* (0.1) *is also nonoscillatory.*

PROOF. Denote $S(w, r, p) = (p-1)r^{1-q}|w|^q$ (this is the third term in the Riccati equation (2.1)). Then S can be rewritten as

$$S(w, r, p) = (p-1)|w|\left(\frac{|w|^p}{r}\right)^{\frac{1}{p-1}}$$

and it is easy to compute that

$$\frac{\partial S}{\partial p}(w, r, p) = |w|\left(\frac{|w|}{r}\right)^{\frac{1}{p-1}}\left[1 - \lg\left(\frac{|w|}{r}\right)^{\frac{1}{p-1}}\right].$$

Hence, for $\frac{|w|}{r} \leqslant 1$ the function S is nondecreasing with respect to p.

Since (5.43) is nonoscillatory, by Lemma 5.2 there exists a function v satisfying the Riccati equation

$$v' + c(t) + S\big(v, r(t), \alpha\big) = 0,$$

and such that $v(t) \to 0$ as $t \to \infty$. Since $\liminf_{t\to\infty} r(t) > 0$, we have $\frac{|w(t)|}{r(t)} \leqslant 1$ for large t, and thus

$$0 = v' + c(t) + S\big(v, r(t), \alpha\big) \geqslant v'(t) + c(t) + S\big(v, r(t), p\big).$$

Hence (0.1) is nonoscillatory by Theorem 5.3. □

6. Asymptotic of nonoscillatory solutions

This section is devoted to the asymptotic properties of nonoscillatory solutions of Equation (0.1) when the function c does not change its sign. In this case it is possible associate with (0.1) its reciprocal equation

$$\big(c^{1-q}(t)\Phi^{-1}(u')\big)' + r^{1-q}(t)\Phi^{-1}(u) = 0. \tag{6.1}$$

Recall that the so-called reciprocity principle says that (6.1) is nonoscillatory if and only if (0.1) is nonoscillatory, see Section 2.7. Note also that if $c(t) \leqslant 0$ for large t then (0.1) is nonoscillatory since the equation $(r(t)\Phi(x'))' = 0$ is its nonoscillatory majorant.

6.1. *Integral conditions and classification of solutions*

If c is different from zero for large t, then all solutions of nonoscillatory equation (0.1) are eventually monotone, as the following result shows.

LEMMA 6.1. *Let $c(t) \neq 0$ for large t and x be a solution of nonoscillatory equation* (0.1) *defined on some interval* (α_x, ∞), $\alpha_x \geqslant 0$. *Then either $x(t)x'(t) > 0$ or $x(t)x'(t) < 0$ for large t.*

PROOF. The monotonicity of x follows from the reciprocity principle which ensures that the so-called *quasiderivative* $x^{[1]} := r\Phi(x')$ does not change its sign for large t. □

Then it is possible, a priori, to divide the set of solutions of (0.1) into the following two classes:

$$\mathbb{M}^+ = \big\{x \text{ solution of (0.1): } \exists t_x \geqslant 0\text{: } x(t)x'(t) > 0 \text{ for } t > t_x\big\},$$
$$\mathbb{M}^- = \big\{x \text{ solution of (0.1): } \exists t_x \geqslant 0\text{: } x(t)x'(t) < 0 \text{ for } t > t_x\big\}.$$

Clearly, solutions in $\mathbb{M}^+$ are eventually either positive increasing or negative decreasing and solutions in $\mathbb{M}^-$ are either positive decreasing or negative increasing. The existence of solutions in these classes depends on the sign of the function c, as the following results show.

PROPOSITION 6.1. *Assume $c(t) < 0$ for large t.*

(i) *Equation* (0.1) *has solutions in the class* $\mathbb{M}^-$. *More precisely, for every pair* $(t_0, a) \in [0,\infty) \times \mathbb{R}\backslash\{0\}$ *there exists a solution* x *of* (0.1) *in the class* $\mathbb{M}^-$ *such that* $x(t_0) = a$.
(ii) *Equation* (0.1) *has solutions in the class* $\mathbb{M}^+$. *More precisely, for every pair* $(a_0, a_1) \in \mathbb{R}^2$, $a_0 a_1 > 0$ *and for any* t_0 *sufficiently large, there exists a solution* x *of* (0.1) *in the class* $\mathbb{M}^+$ *such that* $x(t_0) = a_0$, $x'(t_0) = a_1$.

PROOF. Claim (i) follows, for instance, from [44, Theorem 1] and [178, Theorems 9.1, 9.2]. Concerning the claim (ii), let x be a solution of (0.1) such that $x(0)x'(0) > 0$. Since the auxiliary function

$$F_x(t) = r(t)\Phi\bigl(x'(t)\bigr)x(t) \tag{6.2}$$

is nondecreasing, we obtain $x(t)x'(t) > 0$ for $t > 0$. The assertion follows taking into account that every solution is continuable up to ∞, see Section 1. □

In the opposite case, i.e., when $c(t) > 0$ for large t, the existence in the classes $\mathbb{M}^+$, $\mathbb{M}^-$ may be characterized by means of the convergence or divergence of the following two integrals

$$J_r = \int_0^\infty r^{1-q}(t)\,\mathrm{d}t, \qquad J_c = \int_0^\infty \bigl|c(t)\bigr|\,\mathrm{d}t,$$

as the following results show.

LEMMA 6.2. *Assume* $c(t) > 0$ *for large* t.
(i) *If* $J_c = \infty$, *then* $\mathbb{M}^+ = \emptyset$.
(ii) *If* $J_r = \infty$, *then* $\mathbb{M}^- = \emptyset$.

PROOF. (i) Let x be a solution of (0.1) in the class $\mathbb{M}^+$ and, without loss of generality, suppose $x(t) > 0$, $x'(t) > 0$ for $t \geqslant T \geqslant 0$. From (0.1) we obtain for $t \geqslant T$

$$r(t)\Phi\bigl(x'(t)\bigr) \leqslant r(T)\Phi\bigl(x'(T)\bigr) - \Phi\bigl(x(T)\bigr)\int_T^t c(s)\,\mathrm{d}s$$

that gives a contradiction as $t \to \infty$. Claim (ii) follows by applying (i) to (6.1) and using the reciprocity principle. □

PROPOSITION 6.2. *Assume* $c(t) > 0$ *for large* t.
(i) *If* (0.1) *is nonoscillatory and* $J_r = \infty$, $J_c < \infty$, *then* $\mathbb{M}^+ \neq \emptyset$.
(ii) *If* (0.1) *is nonoscillatory and* $J_r < \infty$, $J_c = \infty$, *then* $\mathbb{M}^- \neq \emptyset$.
(iii) *If* $J_r < \infty$, $J_c < \infty$, *then* $\mathbb{M}^+ \neq \emptyset$, $\mathbb{M}^- \neq \emptyset$.

PROOF. Claims (i), (ii) follows from Lemma 6.2. The assertion (iii) follows, for instance, as a particular case from [218, Theorem 3.1, Theorem 3.3] and their proofs by choosing $R(t) = 1$ and observing that assumptions (3.2), (3.11) in [218] are not necessary in the half-linear case. □

In view of Lemma 6.2, if (0.1) is nonoscillatory, c is eventually positive and $J_r + J_c = \infty$, then all the solutions of (0.1) belong to the same class ($\mathbb{M}^+$ or $\mathbb{M}^-$). In addition, from the same Lemma 6.2, the well-known Leigthon-type oscillation result can be obtained: *Let c be eventually positive; if $J_r = J_c = \infty$, then* (0.1) *is oscillatory* (compare Theorem 2.5).

As in the quoted papers [38–41,45,176,181], in both cases $c > 0$ and $c < 0$ eventually, the classes $\mathbb{M}^+$, $\mathbb{M}^-$ may be divided, a priori, into the following four subclasses, which are mutually disjoint:

$$\mathbb{M}_B^- = \Bigl\{x \in \mathbb{M}^-\colon \lim_{t\to\infty} x(t) = \ell_x \neq 0\Bigr\},$$

$$\mathbb{M}_0^- = \Bigl\{x \in \mathbb{M}^-\colon \lim_{t\to\infty} x(t) = 0\Bigr\},$$

$$\mathbb{M}_B^+ = \Bigl\{x \in \mathbb{M}^+\colon \lim_{t\to\infty} x(t) = \ell_x,\ |\ell_x| < \infty\Bigr\},$$

$$\mathbb{M}_\infty^+ = \Bigl\{x \in \mathbb{M}^+\colon \lim_{t\to\infty} |x(t)| = \infty\Bigr\}.$$

In the following subsections we consider both cases $c(t) > 0$, $c(t) < 0$ and we describe the above classes in terms of certain integral conditions. Similarly to the linear case, we are going to show that the convergence or divergence of the two integrals

$$J_1 = \lim_{T\to\infty} \int_0^T r^{1-q}(t)\Phi^{-1}\left(\int_0^t \bigl|c(s)\bigr|\,\mathrm{d}s\right)\mathrm{d}t,$$

$$J_2 = \lim_{T\to\infty} \int_0^T r^{1-q}(t)\Phi^{-1}\left(\int_t^T \bigl|c(s)\bigr|\,\mathrm{d}s\right)\mathrm{d}t,$$

fully characterize the above four classes.

The following lemma describes relations between J_1, J_2, J_r, J_c.

LEMMA 6.3. *The following statements hold.*

(a) *If $J_1 < \infty$, then $J_r < \infty$.*
(b) *If $J_2 < \infty$, then $J_c < \infty$.*
(c) *If $J_2 = \infty$, then $J_r = \infty$ or $J_c = \infty$.*
(d) *If $J_1 = \infty$, then $J_r = \infty$ or $J_c = \infty$.*
(e) *$J_1 < \infty$ and $J_2 < \infty$ if and only if $J_r < \infty$ and $J_c < \infty$.*

PROOF. Claim (a). Let $t_1 \in (0, T)$. Because

$$\int_0^T r^{1-q}(s)\Phi^{-1}\left(\int_0^s \bigl|c(t)\bigr|\,\mathrm{d}t\right)\mathrm{d}s$$
$$> \int_0^{t_1} r^{1-q}(s)\Phi^{-1}\left(\int_0^s \bigl|c(t)\bigr|\,\mathrm{d}t\right)\mathrm{d}s + \Phi^{-1}\left(\int_0^{t_1} \bigl|c(s)\bigr|\,\mathrm{d}s\right)\int_{t_1}^T r^{1-q}(s)\,\mathrm{d}s$$

the assertion follows. Claim (b) follows in a similar way. Claims (c), (d) follow from the inequalities

$$\int_0^T r^{1-q}(t)\Phi^{-1}\left(\int_t^T |c(s)|\,\mathrm{d}s\right)\mathrm{d}t \leqslant \int_0^T r^{1-q}(t)\,\mathrm{d}t\ \Phi^{-1}\left(\int_0^T |c(s)|\,\mathrm{d}s\right),$$

$$\int_0^T r^{1-q}(t)\Phi^{-1}\left(\int_0^t |c(s)|\,\mathrm{d}s\right)\mathrm{d}t \leqslant \int_0^T r^{1-q}(t)\,\mathrm{d}t\ \Phi^{-1}\left(\int_0^T |c(s)|\,\mathrm{d}s\right).$$

Finally, the claim (e) immediately follows from (a)–(d). □

6.2. *The case c negative*

When the function $c(t)$ is eventually negative, the asymptotic properties of nonoscillatory solutions have been deeply studied and interesting contributions are due to the Georgian and Russian mathematical school [43,44,125,126,128,127,143,176,193]. Other recent developments can be found in [45,141,156,181]. Here we deal with some results that can be obtained, as a particular case, from recent criteria in [33,35] and, under the assumption that J_r is convergent or divergent, can be found in [180,210].

We start by noting that if $c(t) < 0$ in the whole interval $[0, \infty)$, then for any solution $x \in \mathbb{M}^-$ we have $x(t)x'(t) < 0$ on $[0, \infty)$. This property can be proved using the auxiliary function F_x given in (6.2). Since, as claimed, F is a nondecreasing function and x is not eventually constant, there are only two possibilities: (a) F_x does not have zeros; (b) there exists $t_x \geqslant \alpha_x$ such that $F_x(t) > 0$ for all $t > t_x$. Thus the assertion follows.

The following hold.

THEOREM 6.1. *Let $c(t) < 0$ for large t.*
(i) *Equation* (0.1) *has solutions in the class $\mathbb{M}_B^-$ if and only if $J_2 < \infty$.*
(ii) *Equation* (0.1) *has solutions in the class $\mathbb{M}_B^+$ if and only if $J_1 < \infty$.*

PROOF. Claim (i) “⇒”: Let $x \in \mathbb{M}_B^-$. Without loss of generality we can assume $x(t) > 0$, $x'(t) < 0$ for $t \geqslant T \geqslant 0$. Integrating (0.1) in (t, ∞), $t > T$, we obtain

$$-\lambda_x - r(t)\Phi\bigl(x'(t)\bigr) = \int_t^\infty |c(\tau)|\Phi_p\bigl(x(\tau)\bigr)\,\mathrm{d}\tau, \tag{6.3}$$

where $-\lambda_x = \lim_{t\to\infty}[r(t)\Phi(x'(t))]$. Since $x(\tau) > x(\infty) > 0$ and $\lambda_x \geqslant 0$, (6.3) implies

$$-r(t)\Phi\bigl(x'(t)\bigr) \geqslant \Phi\bigl(x(\infty)\bigr)\int_t^\infty |c(\tau)|\,\mathrm{d}\tau.$$

Hence

$$x(t) \leqslant x(T) - x(\infty)\int_T^t \Phi^{-1}\left(\frac{1}{r(s)}\int_s^\infty |c(\tau)|\,\mathrm{d}\tau\right)\mathrm{d}s.$$

As $t \to \infty$ we obtain the assertion.

Claim (i) “$\Leftarrow$”: Choose $t_0 \geqslant 0$ such that

$$\int_{t_0}^{\infty} \Phi^{-1}\left(\frac{1}{r(t)} \int_t^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right) \mathrm{d}t \leqslant \frac{1}{2}. \tag{6.4}$$

Denote by $C[t_0, \infty)$ the Fréchet space of all continuous functions on $[t_0, \infty)$ endowed with the topology of the uniform convergence on compact subintervals of $[t_0, \infty)$. Let Ω be the nonempty subset of $C[t_0, \infty)$ given by

$$\Omega = \left\{u \in C[t_0, \infty)\colon \ \frac{1}{2} \leqslant u(t) \leqslant 1\right\}. \tag{6.5}$$

Clearly Ω is bounded, closed and convex. Now consider the operator $T : \Omega \to C[t_0, \infty)$ which assigns to any $u \in \Omega$ the continuous function $T(u) = y_u$ given by

$$y_u(t) = T(u)(t) = \frac{1}{2} + \int_t^{\infty} \Phi^{-1}\left(\frac{1}{r(s)}\left(\int_s^{\infty} \left|c(\tau)\right|\Phi\big(u(\tau)\big)\,\mathrm{d}\tau\right)\right) \mathrm{d}s. \tag{6.6}$$

We have

$$\frac{1}{2} \leqslant T(u)(t) \leqslant \frac{1}{2} + \int_t^{\infty} \Phi^{-1}\left(\frac{1}{r(s)} \int_s^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right) \mathrm{d}s$$

which implies, by virtue of (6.4), $T(\Omega) \subseteq \Omega$. In order to apply the Tychonov fixed point theorem to operator T, it is sufficient to prove that T is continuous in $\Omega \subseteq C[t_0, \infty)$ and that $T(\Omega)$ is relatively compact in $C[t_0, \infty)$. Let $\{u_j\}$, $j \in \mathbb{N}$, be a sequence in Ω which is convergent to $\bar{u}$ in $C[t_0, \infty)$, $\bar{u} \in \overline{\Omega} = \Omega$. Since for $s \geqslant t_0$

$$\Phi^{-1}\left(\frac{1}{r(s)}\left(\int_s^{\infty} \left|c(\tau)\right|\Phi\big(u_j(\tau)\big)\,\mathrm{d}\tau\right)\right) \leqslant \Phi^{-1}\left(\frac{1}{r(s)}\left(\int_s^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right)\right) < \infty,$$

the Lebesgue dominated convergence theorem gives the continuity of T in Ω. It remains to prove that $T(\Omega)$ is relatively compact in $C[t_0, \infty)$, i.e., that functions in $T(\Omega)$ are equibounded and equicontinuous on every compact subinterval of $[t_0, \infty)$. The equiboundedness easily follows taking into account that Ω is a bounded subset of $C[t_0, \infty)$. In order to prove the equicontinuity, for any $u \in \Omega$ we have

$$\begin{aligned} 0 < -\big(T(u)(t)\big)' &= \Phi^{-1}\left(\frac{1}{r(t)}\left(\int_t^{\infty} \left|c(\tau)\right|\Phi\big(u(\tau)\big)\,\mathrm{d}\tau\right)\right) \\ &\leqslant \Phi^{-1}\left(\frac{1}{r(t)}\left(\int_t^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right)\right) \end{aligned} \tag{6.7}$$

which implies that functions in $T(\Omega)$ are equicontinuous on every compact subinterval of $[t_0, \infty)$. From the Tychonov fixed point theorem there exists $x \in \Omega$ such that $x = T(x)$ or, from (6.6),

$$x(t) = \frac{1}{2} + \int_t^\infty \Phi^{-1}\left(\frac{1}{r(s)}\left(\int_s^\infty \left|c(\tau)\right|\Phi\big(x(\tau)\big)\,\mathrm{d}\tau\right)\right)\mathrm{d}s.$$

It is easy to show that x is a positive solution of (0.1) in $[t_0, \infty)$ and, from (6.7), $x'(t) < 0$. Finally, clearly x satisfies the inequality $x(t)x'(t) < 0$ in its maximal interval of existence and the proof of claim (i) is complete.

Claim (ii) "$\Rightarrow$": Assume, by contradiction, $J_1 = \infty$. Without loss of generality let x be a solution of (0.1) in the class $\mathbb{M}_B^+$ such that $x(t) > 0, x'(t) > 0$ for $t \geqslant t_0 \geqslant 0$. Integrating (0.1) on (t_0, t) we obtain for $t > t_0$

$$x^{[1]}(t) = x^{[1]}(t_0) + \int_{t_0}^t \left|c(s)\right|\Phi\big(x(s)\big)\,\mathrm{d}s > \Phi_p\big(x(t_0)\big)\int_{t_0}^t \left|c(s)\right|\mathrm{d}s,$$

where $x^{[1]} = r\Phi(x')$. Hence

$$x'(t) > x(t_0)\Phi^{-1}\left(\frac{1}{r(s)}\int_{t_0}^t \left|c(s)\right|\mathrm{d}s\right).$$

Integrating again over (t_0, t) we obtain a contradiction.

Claim (ii) "$\Leftarrow$": The argument is similar to that given in Claim (i) "$\Rightarrow$". It is sufficient to consider in the same set Ω, defined in (6.5), the operator $T : \Omega \to C[t_0, \infty)$ given by

$$y_u(t) = T(u)(t) = \frac{1}{2} + \int_{t_0}^t \Phi^{-1}\left(\frac{1}{r(s)}\left(\int_{t_0}^s \left|c(\tau)\right|\Phi\big(u(\tau)\big)\,\mathrm{d}\tau\right)\right)\mathrm{d}s$$

and to apply the Tychonov fixed point theorem. □

THEOREM 6.2. *Let* $c(t) < 0$ *for large* t.

(i) *If* $J_1 = \infty$ *and* $J_2 < \infty$, *then* $\mathbb{M}_0^- = \emptyset$.

(ii) *If* $J_1 < \infty$, *then* $\mathbb{M}_\infty^+ = \emptyset$.

PROOF. Claim (i). Let x be a solution of (0.1) in the class $\mathbb{M}^-$ such that $0 < x(t) < \varepsilon$, $x'(t) < 0$ for $t \geqslant T$ and $\lim_{t\to\infty} x(t) = 0$. By Lemma 6.3, $J_r = \infty$ and thus, by [33, Lemma 1], $\lim_{t\to\infty} r(t)\Phi(x'(t)) = 0$. Taking into account this fact and integrating (0.1) over (t, ∞), $t > T$, we obtain

$$\frac{x'(t)}{x(t)} > -\Phi^{-1}\left(\frac{1}{r(t)}\int_t^\infty \left|c(\tau)\right|\mathrm{d}\tau\right).$$

Integrating over (T, t) we have

$$\lg\frac{x(t)}{x(T)} > -\int_T^t \Phi^{-1}\left(\frac{1}{r(s)}\int_s^\infty \left|c(\tau)\right|\mathrm{d}\tau\right)\mathrm{d}s,$$

from which, as $t \to \infty$, we obtain a contradiction.

Claim (ii). Let $x \in \mathbb{M}_{\infty}^{+}$ and assume $x(t) > 0, x'(t) > 0$ for $t \geqslant T \geqslant 0$. From (2.1) we have (with $w = r\Phi(x')/\Phi(x)$)

$$\begin{aligned}\frac{r(t)\Phi(x'(t))}{\Phi(x(t))} &= -(p-1)\int_T^t r^{1-q}(s)\left|w(s)\right|^q \mathrm{d}s + k + \int_T^t \left|c(s)\right| \mathrm{d}s \\ &\leqslant k + \int_T^t \left|c(s)\right| \mathrm{d}s, \qquad (6.8)\end{aligned}$$

where $k = r(T)\Phi(x'(T))/\Phi(x(T))$. If $J_c < \infty$, then there exists a positive constant k_1 such that

$$\frac{r(t)\Phi(x'(t))}{\Phi(x(t))} \leqslant k_1$$

or

$$\frac{x'(t)}{x(t)} \leqslant \Phi^{-1}(k_1)\Phi^{-1}\left(\frac{1}{r(t)}\right).$$

Integrating again over (T, t) we obtain

$$\lg\frac{x(t)}{x(T)} \leqslant \Phi^{-1}(k_1)\int_T^t r^{1-q}(s)\,\mathrm{d}s$$

which implies $x \in \mathbb{M}_B^{+}$, i.e., a contradiction. If $J_c = \infty$, choose $t_1 > T$ such that $k < \int_T^{t_1} c(s)\,\mathrm{d}s$. Then from (6.8) we obtain for $t \geqslant t_1$

$$\frac{r(t)\Phi(x'(t))}{\Phi(x(t))} \leqslant 2\int_T^t \left|c(s)\right| \mathrm{d}s,$$

or

$$\frac{x'(t)}{x(t)} \leqslant \Phi^{-1}(2)\Phi^{-1}\left(\frac{1}{r(t)}\int_T^t \left|c(s)\right| \mathrm{d}s\right).$$

Integrating over (t_1, t) we have

$$\lg\frac{x(t)}{x(t_1)} \leqslant \Phi^{-1}(2)\int_T^t \Phi^{-1}\left(\frac{1}{r(s)}\int_T^s \left|c(\tau)\right| \mathrm{d}\tau\right)\mathrm{d}s$$

which gives the assertion. □

THEOREM 6.3. *Let $c(t) < 0$ for large t. If $J_1 < \infty$ and $J_2 < \infty$, then Equation* (0.1) *has solutions in both classes $\mathbb{M}_0^{-}$ and $\mathbb{M}_B^{-}$.*

PROOF. The statement $\mathbb{M}_B^- \neq \emptyset$ follows from Theorem 6.1. The existence in the class $\mathbb{M}_0^-$ can be proved by using a similar argument as that given in the proof of Theorem 6.1. It is sufficient to consider the set

$$\Omega = \left\{ u \in C[t_0, \infty)\colon\ 0 \leqslant u(t) \leqslant \int_t^\infty r^{1-q}(s)\,\mathrm{d}s \right\}$$

and the operator $T : \Omega \to C[t_0, \infty)$ given by

$$y_u(t) = T(u)(t) = \int_t^\infty r^{1-q}(t)\Phi^{-1}\left(1 - \int_{t_0}^s \left|c(\tau)\right|\Phi\bigl(u(\tau)\bigr)\,\mathrm{d}\tau\right)\mathrm{d}s$$

and to apply the Tychonov fixed point theorem, details are omitted. □

REMARK 6.1. The behavior of quasiderivatives of solutions (i.e., of expressions $x^{[1]} = r\Phi(x')$) plays an important role in the study of principal solutions, especially in their limit characterization, see the next section. Concerning the solution $x \in \mathbb{M}_B^-$, defined as a fixed point in the proof of Theorem 6.1, we have $\lim_{t\to\infty} x^{[1]}(t) = 0$. Concerning the solution $x \in \mathbb{M}_0^-$, defined in the proof of Theorem 6.3, it is easy to show that $\lim_{t\to\infty} x^{[1]}(t) = c_x < 0$. Indeed, the limit

$$\lim_{t\to\infty} x'(t) r^{q-1}(t) \tag{6.9}$$

exists finite and it is different from zero, because

$$\begin{aligned} -x'(t) &= r^{1-q}(t)\Phi^{-1}\left(1 - \int_{t_0}^t \left|c(\tau)\right|\Phi\bigl(u(\tau)\bigr)\,\mathrm{d}\tau\right) \geqslant \Phi^{-1}\left(\frac{1}{2}\right)\Phi^{-1}\left(\frac{1}{r(t)}\right) \\ &= 2^{1-q} r^{1-q}(t). \end{aligned}$$

and the function $x'r^{q-1}$is negative increasing.

From Theorems 6.1, 6.2, 6.3, we can summarize the situation in the following way. Clearly, as regards the convergence or divergence of J_1, J_2, the possible cases are the following:

$$\begin{aligned} &(\mathrm{A}_1)\colon\quad J_1 = \infty, \quad J_2 = \infty,\\ &(\mathrm{A}_2)\colon\quad J_1 = \infty, \quad J_2 < \infty,\\ &(\mathrm{A}_3)\colon\quad J_1 < \infty, \quad J_2 = \infty,\\ &(\mathrm{A}_4)\colon\quad J_1 < \infty, \quad J_2 < \infty. \end{aligned}$$

Then the following result holds.

THEOREM 6.4. *Let $c(t) < 0$ for large t.*

(i) *Assume case* (A_1). *Then any solution of* (0.1) *in the class* $\mathbb{M}^-$ *tends to zero as* $t \to \infty$ *and any solution of* (0.1) *in the class* $\mathbb{M}^+$ *is unbounded.*

(ii) *Assume case* (A$_2$). *Then any solution of* (0.1) *in the class* $\mathbb{M}^-$ *tends to a nonzero limit as* $t \to \infty$ *and any solution of* (0.1) *in the class* $\mathbb{M}^+$ *is unbounded.*

(iii) *Assume case* (A$_3$). *Then any solution of* (0.1) *in the class* $\mathbb{M}^-$ *tends to zero as* $t \to \infty$ *and any solution of* (0.1) *in the class* $\mathbb{M}^+$ *is bounded.*

(iv) *Assume case* (A$_4$). *Then both solutions of* (0.1) *converging to zero and solutions of* (0.1) *tending to a nonzero limit* (*as* $t \to \infty$) *exist in the class* $\mathbb{M}^-$. *Further solutions of* (0.1) *in the class* $\mathbb{M}^+$ *are bounded.*

From Theorem 6.4 we obtain immediately the following result which generalizes a well-know one stated for the linear equation in [173, Theorems 3 and 4]; see also [106, Chapters VI, XI]).

COROLLARY 6.1. *Let* $c(t) < 0$ *for large* t.

(a) *Any solution* x *of* (0.1) *in the class* $\mathbb{M}^-$ *tends to zero as* $t \to \infty$ *if and only if* $J_2 = \infty$.

(b) *Any solution of* (0.1) *is bounded if and only if* $J_1 < \infty$.

Following another classification used in [180,210], we distinguish these types of eventually positive solutions x of (0.1) (clearly a similar classification holds for eventually negative solutions):

Type (1) $\lim_{t\to\infty} x(t) = 0$, $\lim_{t\to\infty} x^{[1]}(t) = 0$;
Type (2) $\lim_{t\to\infty} x(t) = 0$, $\lim_{t\to\infty} x^{[1]}(t) = c_1 < 0$;
Type (3) $\lim_{t\to\infty} x(t) = c_0 > 0$, $\lim_{t\to\infty} x^{[1]}(t) = c_1 \leqslant 0$;
Type (4) $\lim_{t\to\infty} x(t) = c_0 > 0$, $\lim_{t\to\infty} x^{[1]}(t) = c_1 > 0$;
Type (5) $\lim_{t\to\infty} x(t) = c_0 > 0$, $\lim_{t\to\infty} x^{[1]}(t) = \infty$;
Type (6) $\lim_{t\to\infty} x(t) = \infty$, $\lim_{t\to\infty} x^{[1]}(t) = c_1$;
Type (7) $\lim_{t\to\infty} x(t) = \infty$, $\lim_{t\to\infty} x^{[1]}(t) = \infty$.

Eventually positive solutions in $\mathbb{M}^-$ are of the Types (1)–(3), eventually positive solutions in $\mathbb{M}^+$ are of the Types (4)–(7). From Theorem 6.4 and the reciprocity principle (see Section 2.7), necessary and/or sufficient conditions for their existence can be obtained. To this end observe that the integral J_r $[J_c]$ for (0.1) plays the same role as J_c $[J_r]$ for the reciprocal equation (6.1). Similarly, for the reciprocal equation (6.1) the integrals J_1, J_2 becomes

$$R_1 = \lim_{T\to\infty} \int_0^T |c(t)| \Phi\left(\int_0^t r^{1-q}(s)\,\mathrm{d}s \right) \mathrm{d}t$$

$$R_2 = \lim_{T\to\infty} \int_0^T |c(t)| \Phi\left(\int_t^T r^{1-q}(s)\,\mathrm{d}s \right) \mathrm{d}t,$$

respectively. Then the following holds.

THEOREM 6.5. *Let $c(t) < 0$ for large t. Then the following statements hold*:

(a) *Every eventually positive solution in $\mathbb{M}^-$ is of Type* (1) *if and only if $J_2 = \infty$ and $R_2 = \infty$.*
(b) *Equation* (0.1) *has solutions of Type* (2) *if and only if $R_2 < \infty$.*
(c) *Equation* (0.1) *has solutions of Type* (3) *if and only if $J_2 < \infty$.*
(d) *Equation* (0.1) *has solutions of Type* (4) *if and only if $J_1 < \infty$ and $R_1 < \infty$.*
(e) *Equation* (0.1) *has solutions of Type* (5) *if and only if $J_1 < \infty$ and $R_1 = \infty$.*
(f) *Equation* (0.1) *has solutions of Type* (6) *if and only if $J_1 = \infty$ and $R_1 < \infty$.*
(g) *Every eventually positive solution in $\mathbb{M}^+$ is of Type* (7) *if and only if $J_1 = \infty$ and $R_1 = \infty$.*

6.3. *Uniqueness in $\mathbb{M}^-$.*

The uniqueness in the class $\mathbb{M}^-$ plays a crucial role in the study of the limit characterization of principal solutions (see Theorem 7.6 in the next section).

Proposition 6.1 states that, when c is eventually negative, the class $\mathbb{M}^-$ is nonempty. In the linear case, the assumption

$$\int_0^\infty \left(\frac{1}{r(t)} + \left|c(t)\right|\right) \mathrm{d}t = \infty \tag{6.10}$$

is necessary and sufficient for uniqueness in $\mathbb{M}^-$ of such a solution, when the initial value of the solution is given (see [173, Theorems 3,4]). We will show that also for (0.1) such a property is assured by a natural extension of condition (6.10).

THEOREM 6.6. *Let $c(t) < 0$ for large t. For any $(t_0, x_0) \in [0, \infty) \times \mathbb{R}\backslash\{0\}$, there exists a unique solution x of* (0.1) *in the class $\mathbb{M}^-$ such that $x(t_0) = x_0$ if and only if*

$$\int_0^\infty \left(r^{1-q}(t) + \left|c(t)\right|\right) \mathrm{d}t = \infty. \tag{6.11}$$

The following result can be easily proved and will be useful in the proof of Theorem 6.6.

LEMMA 6.4. *Let $c(t) < 0$ for large t. If $J_r = \infty$, then for every solution x of* (0.1) *in the class $\mathbb{M}^-$ we have $\lim_{t\to\infty} x^{[1]}(t) = 0$.*

PROOF OF THEOREM 6.6. *Necessity.* Assume (6.11) does not hold, i.e.,

$$\int_0^\infty \left|c(\tau)\right| \mathrm{d}\tau < \infty, \qquad \int_0^\infty r^{1-q}(t)\, \mathrm{d}t < \infty,$$

and let t_0 be large such that

$$\Phi^{-1}\left(\int_{t_0}^\infty \left|c(\tau)\right| \mathrm{d}\tau\right) \int_{t_0}^\infty \Phi^{-1}\left(\frac{1}{r(t)}\right) \mathrm{d}t < \frac{1}{\Phi^{-1}(2)}. \tag{6.12}$$

Consider the solutions x_1, x_2 of (0.1) with the initial values $x_1(t_0) = x_2(t_0) = 1$ and

$$x_1'(t_0) = -\Phi^{-1}\left(\frac{c_1}{r(t_0)} \int_{t_0}^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right),$$
$$x_2'(t_0) = -\Phi^{-1}\left(\frac{c_2}{r(t_0)} \int_{t_0}^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right), \tag{6.13}$$

where c_i are positive constants such that $c_1 \neq c_2$ and

$$1 \leqslant c_i \leqslant 2. \tag{6.14}$$

Let us show that $x_i \in \mathbb{M}^-$, $i = 1, 2$. It is easy to show that solutions x_i are positive decreasing on $[0, t_0]$. In order to prove that $x_i \in \mathbb{M}^-$, it will be sufficient to show that $x_i(t)x_i'(t) < 0$ for any $t \geqslant t_0$. Clearly solutions x_i are positive decreasing in a right neighborhood of t_0. Assume there exists $t_i > t_0$ such that $x_i(t_i)x_i'(t_i) = 0, x_i(t) > 0, x_i'(t) < 0$ for $t_0 \leqslant t < t_i$. Integrating (0.1) on (t_0, t_i) we have

$$r(t_i)\Phi\big(x_i'(t_i)\big) - r(t_0)\Phi\big(x_i'(t_0)\big) = \int_{t_0}^{t_i} \left|c(\tau)\right|\Phi\big(x_i(\tau)\big)\,\mathrm{d}\tau. \tag{6.15}$$

If $x_i'(t_i) = 0$, from (6.13) and (6.15) we obtain

$$\begin{aligned} c_i \int_{t_0}^{\infty} \left|c(\tau)\right| \mathrm{d}\tau &= \int_{t_0}^{t_i} \left|c(\tau)\right|\Phi\big(x_i(\tau)\big)\,\mathrm{d}\tau \leqslant \Phi\big(x_i(t_0)\big) \int_{t_0}^{t_i} \left|c(\tau)\right| \mathrm{d}\tau \\ &= \int_{t_0}^{t_i} \left|c(\tau)\right| \mathrm{d}\tau \end{aligned}$$

which implies

$$\int_{t_i}^{\infty} \left|c(\tau)\right| \mathrm{d}\tau \leqslant 0,$$

that is a contradiction. Now suppose $x_i(t_i) = 0$. For $t \in (t_0, t_i)$ from

$$r(t)\Phi\big(x_i'(t)\big) \geqslant r(t_0)\Phi\big(x_i'(t_0)\big) = -c_i \int_{t_0}^{\infty} \left|c(\tau)\right| \mathrm{d}t,$$

we obtain

$$x_i'(t) \geqslant -\Phi^{-1}(c_i) r^{1-q}(t) \Phi^{-1}\left(\int_{t_0}^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right)$$

or

$$x_i(t_i) - x(t_0) = -1 \geqslant -\Phi^{-1}(c_i)\Phi^{-1}\left(\int_{t_0}^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right) \int_{t_0}^{t_i} r^{1-q}(t)\,\mathrm{d}t.$$

Thus, by virtue of (6.14),

$$1 \leqslant \Phi^{-1}(2)\Phi^{-1}\left(\int_{t_0}^{\infty} \left|c(\tau)\right| \mathrm{d}\tau\right) \int_{t_0}^{\infty} r^{1-q}(t)\,\mathrm{d}t,$$

which contradicts (6.12) and the necessity of (6.11) is proved.

Sufficiency. Let us show that for any $(t_0, x_0) \in [0, \infty) \times \mathbb{R} \setminus \{0\}$, there exists at most one solution x of (0.1) in the class $\mathbb{M}^-$ such that $x(t_0) = x_0$ when $J_r = \infty$ or $J_c = \infty$. Let x, y be two solutions of (0.1) in the class $\mathbb{M}^-$ such that $x(t_0) = y(t_0)$, $x'(t_0) > y'(t_0)$. Consider the function d given by $d(t) = x(t) - y(t)$. Then $d(t_0) = 0, d'(t_0) > 0$. We claim that d does not have positive points of maximum greater than t_0, i.e.,

$$d(t) > 0, \qquad d'(t) > 0 \quad \text{for } t > t_0.$$

Assume there exists $t_1 > t_0$ such that $d(t_1) > 0$, $d'(t_1) = 0$ and $d'(t) > 0$ in a suitable left neighborhood I of t_1. Without loss of generality suppose that $d(t) > 0$ for $t \in I$. Now consider the function G given by

$$G(t) = r(t)\left[\Phi\left(x'(t)\right) - \Phi\left(y'(t)\right)\right].$$

Hence $G(t_1) = 0$. Taking into account that Φ is increasing and $d'(t) > 0$, we have $G(t) > 0, t \in I$. In addition, from

$$G'(t) = \left|c(t)\right|\left[\Phi\left(x(t)\right) - \Phi\left(y(t)\right)\right],$$

we obtain $G'(t) > 0$, $t \in I$, which gives a contradiction, because $G(t_1) = 0$. Hence the function d is increasing.

If $J_c = \infty$ then, by Lemma 6.3 we have $J_2 = \infty$. Then, in view of Corollary 6.1(a), we obtain $d(\infty) = 0$, that is a contradiction. If $J_r = \infty$, then taking into account that $d'(t) > 0$ for $t > t_0$, the function G satisfies $G(t) > 0$, $G'(t) > 0$ for $t > t_0$ and, by Lemma 6.4, $\lim_{t\to\infty} G(t) = 0$, that is a contradiction. Finally the existence of at least one solution $x \in \mathbb{M}^-$ such that $x(t_0) = x_0$ is assured by Proposition 6.1(i). □

6.4. *The case c positive*

As already stated before, when c is eventually positive, Equation (0.1) may be either oscillatory or nonoscillatory. In the nonoscillation case, the asymptotic behavior of solutions has been considered by many authors. Here we refer in particular to [37,90,99,112,121,157, 179,218] and references therein. In these papers certain asymptotic properties of nonoscillatory solutions are examined, under various assumptions on functions r, c, for equation (0.1) or, sometimes, for a more general equation (which includes (0.1)). In this section, similarly to Section 6.2, we will show how it is possible to obtain from these results, with a very simple argument, a complete description of the asymptotic behavior of solutions of (0.1) also when c is positive.

The following holds.

PROPOSITION 6.3. *Let $c(t) > 0$ for large t and $J_r = \infty$.*

(i) *If $J_2 = \infty$, then $\mathbb{M}_B^+ = \emptyset$.*

(ii) *If $J_2 < \infty$, then Equation* (0.1) *is nonoscillatory and $\mathbb{M}_B^+ \neq \emptyset$.*

PROOF. In view of Lemma 6.2(ii) any nonoscillatory solution of (0.1) is in the class $\mathbb{M}^+$. Then claims (i) and (ii) follow from [112, Theorem 4.2]. ☐

The following "uniqueness" result will be useful in the proof of the existence of unbounded solutions. It is, in some sense, the analogous one to Theorem 6.6.

THEOREM 6.7 ([112, Theorem 4.3]). *Let $c(t) > 0$ for large t. Let $\eta \neq 0$ be a given constant and assume $J_r = \infty$, $J_2 < \infty$. Then there exists a unique solution x of* (0.1)*, $x \in \mathbb{M}^+$, such that* $\lim_{t\to\infty} x(t) = \eta$.

By using such a result we obtain the following.

PROPOSITION 6.4. *Let $c(t) > 0$ for large t and assume $J_r = \infty$, $J_2 < \infty$. Then* (0.1) *has unbounded solutions, i.e., $\mathbb{M}_\infty^+ \neq \emptyset$.*

PROOF. Assume, by contradiction, that $M_\infty^+ = \emptyset$. In view of Proposition 6.3, let u be a solution of (0.1) in the class $\mathbb{M}_B^+$ and let x be another solution of (0.1) such that $x(0) = u(0), x'(0) \neq u'(0)$. Hence $x \neq u$ and from Lemma 6.2(ii) we have $x \in \mathbb{M}_B^+$. In view of Theorem 6.7 we have $u(\infty) \neq x(\infty)$. Now consider the solution w of (0.1) given by

$$w(t) = \frac{u(\infty)}{x(\infty)} x(t).$$

We have $w \in \mathbb{M}_B^+$. But $w(\infty) = u(\infty)$, that gives a contradiction. ☐

PROPOSITION 6.5. *Let $c(t) > 0$ for large t. If $J_r < \infty$, then* (0.1) *does not have unbounded nonoscillatory solutions, i.e., $\mathbb{M}_\infty^+ = \emptyset$.*

PROOF. The assertion follows, with minor changes, from [99, Lemma 2]. ☐

Concerning the existence in the class $\mathbb{M}^-$, the following hold.

PROPOSITION 6.6. *Let $c(t) > 0$ for large t.*

(i) *If $J_r < \infty$, $J_c = \infty$, $J_1 < \infty$, then* (0.1) *is nonoscillatory and $\mathbb{M}_B^- \neq \emptyset$.*

(ii) *If $J_1 = \infty$, then $\mathbb{M}_B^- = \emptyset$.*

PROOF. Claim (i) follows from [99, Theorem 4]. As for the claim (ii), let $x \in \mathbb{M}_B^-$ and, without loss of generality, assume $x(t) > 0, x'(t) < 0$ for $t \geqslant T$ and $x(\infty) = c_x > 0$. From (0.1) we have

$$x^{[1]}(t) = x^{[1]}(T) - \int_T^t c(s)\Phi\bigl(x(s)\bigr)\,\mathrm{d}s < -\Phi(c_x)\int_T^t c(s)\,\mathrm{d}s$$

or

$$x'(t) < -c_x \Phi^{-1}\left(\frac{1}{r(t)} \int_T^t c(s)\,\mathrm{d}s\right).$$

Integrating over (T, t) we obtain

$$x(t) - x(T) < -c_x \int_T^t \Phi^{-1}\left(\frac{1}{r(s)} \int_T^s c(\tau)\,\mathrm{d}\tau\right)\mathrm{d}s$$

that gives a contradiction as $t \to \infty$. $\square$

PROPOSITION 6.7. *Let* $c(t) > 0$ *for large* t. *If* (0.1) *is nonoscillatory and* $J_r < \infty$, *then* $\mathbb{M}_0^- \neq \emptyset$.

PROOF. If $J_c < \infty$, the assertion follows, as a particular case, from [179, Theorem 2.2]. When $J_c = \infty$ the assertion follows from [37, Lemma 2(ii)]. $\square$

By considering the mutual behavior of integrals J_r, J_c, J_1, J_2 it is possible to summarize the situation in a complete way. Indeed, as regards the convergence or divergence of the above integrals, in view of Lemma 6.3, we have the following six possible cases:

(C1) $J_r = J_c = J_1 = J_2 = \infty$,
(C2) $J_r = J_1 = J_2 = \infty$, $J_c < \infty$,
(C3) $J_r = J_1 = \infty$, $J_c < \infty$, $J_2 < \infty$,
(C4) $J_c = J_1 = J_2 = \infty$, $J_r < \infty$,
(C5) $J_c = J_2 = \infty$, $J_r < \infty$, $J_1 < \infty$,
(C6) $J_r < \infty$, $J_c < \infty$, $J_1 < \infty$, $J_2 < \infty$.

If (C1) holds, then, as already claimed in Section 6.1, Equation (0.1) is oscillatory. In the remaining cases, from the above results, we obtain the following theorem, which is a natural extension of the previous one stated in the linear case [39, Theorem 1].

THEOREM 6.8. *Let* $c(t) > 0$ *for large* t.
If (C2) *holds and* (0.1) *is nonoscillatory, then* $\mathbb{M}_\infty^+ \neq \emptyset$, $\mathbb{M}_B^+ = \mathbb{M}_B^- = \mathbb{M}_0^- = \emptyset$.
If (C3) *holds, then* (0.1) *is nonoscillatory and* $\mathbb{M}_\infty^+ \neq \emptyset$, $\mathbb{M}_B^+ \neq \emptyset$, $\mathbb{M}_B^- = \mathbb{M}_0^- = \emptyset$.
If (C4) *holds and* (0.1) *is nonoscillatory, then* $\mathbb{M}_\infty^+ = \mathbb{M}_B^+ = \mathbb{M}_B^- = \emptyset$, $\mathbb{M}_0^- \neq \emptyset$.
If (C5) *holds, then* (0.1) *is nonoscillatory and* $\mathbb{M}_\infty^+ = \mathbb{M}_B^+ = \emptyset$, $\mathbb{M}_B^- \neq \emptyset$, $\mathbb{M}_0^- \neq \emptyset$.
If (C6) *holds, then* (0.1) *is nonoscillatory and* $\mathbb{M}_\infty^+ = \emptyset$, $\mathbb{M}_B^+ \neq \emptyset$, $\mathbb{M}_B^- \neq \emptyset$, $\mathbb{M}_0^- \neq \emptyset$.

PROOF. The proof follows from the previous statements of this section.

(C2) From Lemma 6.2(ii) and Proposition 6.3(i) we have $\mathbb{M}_B^+ = \mathbb{M}_B^- = \mathbb{M}_0^- = \emptyset$. Since (0.1) is nonoscillatory, we obtain $\mathbb{M}^+ = \mathbb{M}_\infty^+ \neq \emptyset$.
(C3) The assertion follows from Lemma 6.2(ii), Proposition 6.3(ii), Proposition 6.4.
(C4) From Lemma 6.2(i) and Proposition 6.6(ii) we have $\mathbb{M}_\infty^+ = \mathbb{M}_B^+ = \mathbb{M}_B^- = \emptyset$. Since (0.1) is nonoscillatory, we obtain $\mathbb{M}^- = \mathbb{M}_0^- \neq \emptyset$.
(C5) The assertion follows from Lemma 6.2(i), Proposition 6.6(i), Proposition 6.7.

(C6) From Proposition 6.2(iii), Proposition 6.5, Proposition 6.7, we obtain $\mathbb{M}^+_\infty = \emptyset$, $\mathbb{M}^+_B \neq \emptyset$, $\mathbb{M}^-_0 \neq \emptyset$. Finally the existence in $\mathbb{M}^-_B$ can be proved using an argument similar to that given in the proof of Theorem 6.1 (see also [218, Theorem 3.3] and its proof). □

Taking into account that the possible cases concerning the convergence or divergence of J_r, J_c, J_1, J_2 are the cases (C1)–(C6), from Theorem 6.8 we easily obtain the following interesting result, which gives a necessary and sufficient condition for the existence of nonoscillatory solutions of (0.1) in the classes $\mathbb{M}^+_\infty, \mathbb{M}^+_B, \mathbb{M}^-_B, \mathbb{M}^-_0$.

THEOREM 6.9. *Let $c(t) > 0$ for large t.*

(i) *Assume* (0.1) *nonoscillatory. The class $\mathbb{M}^+_\infty$ is nonempty if and only if $J_r = \infty$.*
(ii) *The class $\mathbb{M}^+_B$ is nonempty if and only if $J_2 < \infty$.*
(iii) *Assume* (0.1) *nonoscillatory. The class $\mathbb{M}^-_0$ is nonempty if and only if $J_r < \infty$.*
(iv) *The class $\mathbb{M}^-_B$ is nonempty if and only if $J_1 < \infty$.*

REMARK 6.2. Interesting results on asymptotic properties of nonoscillatory solutions of the equation

$$\big(\Phi(x')\big)' + c(t)\Phi(x) = 0, \tag{6.16}$$

based on the concepts of *slowly* and *regularly* varying functions, can be found in the recent paper [117]. Since the presentation of these results of this paper requires the introduction of several auxiliary statements, we will not formulate these results here and we refer to the above mentioned paper [117].

7. Principal solution

The concept of the principal solution of the linear second order differential equation (1.1) was introduced in 1936 by Leighton and Morse [148] and plays an important role in the oscillation and asymptotic theory of (1.1). In this section we show that this concept can be introduced also for (nonoscillatory) half-linear equation (0.1).

7.1. *Principal solution of linear equations*

First we recall basic properties of the principal solution of linear equation (1.1). Suppose that this equation is nonoscillatory, i.e., any solution of this equation is eventually positive or negative. Then, using the below described method, one can distinguish among all solutions of this equation a solution $\tilde{x}$, called the *principal solution*, (determined uniquely up to a multiplicative factor) which is near ∞ less than any other solution of this equation in the sense that

$$\lim_{t\to\infty} \frac{\tilde{x}(t)}{x(t)} = 0$$

for any solution x which is linearly independent of $\tilde{x}$.

Let x, y be eventually positive linearly independent solution of (1.1), then

$$r(t)\big[x'(t)y(t) - x(t)y'(t)\big] =: \omega \neq 0.$$

This means that the function $\frac{x}{y}$ is monotonic and hence there exists (finite or infinite) limit $\lim_{t\to\infty} \frac{x(t)}{y(t)} = L$. If $L = 0$, the solution x the principal solution of (1.1), if $L = \infty$, the principal solution is y. If $0 < L < \infty$, we set $\tilde{x} = x - Ly$. Then obviously $\lim_{t\to\infty} \frac{\tilde{x}(t)}{y(t)} = 0$ and $\tilde{x}$ is the principal solution. Observe that this construction of the principal solution is based on the linearity of the solution space of (1.1).

Using the Wronskian identity, the principal solution $\tilde{x}$ of (1.1) is equivalently characterized as a solution satisfying

$$\int^{\infty} \frac{\mathrm{d}t}{r(t)\tilde{x}^2(t)} = \infty. \tag{7.1}$$

Indeed, let y be a solution linearly independent of $\tilde{x}$. Then by the previous argument $\frac{y}{\tilde{x}}$ tends monotonically to ∞ as $t \to \infty$, hence

$$\lim_{t\to\infty} \int^{t} \frac{\mathrm{d}s}{r(s)\tilde{x}^2(s)} = \lim_{t\to\infty} \frac{y(t)}{\tilde{x}(t)} = \infty. \tag{7.2}$$

Another characterization of the principal solution of (1.1) is via the eventually minimal solution of the associated Riccati equation

$$w' + c(t) + \frac{w^2}{r(t)} = 0. \tag{7.3}$$

Let $\tilde{x}$, x be linearly independent solutions of (1.1), the solution $\tilde{x}$ being principal, and let $\tilde{w} = \frac{r\tilde{x}'}{\tilde{x}}$, $w = \frac{rx'}{x}$ be the solutions of the associated Riccati equation. Without loss of generality we may suppose that $\tilde{x}$ and x are eventually positive. We have

$$w(t) - \tilde{w}(t) = \frac{r(t)y'(t)}{y(t)} - \frac{r(t)\tilde{x}'(t)}{\tilde{x}(t)} = \frac{r(t)[x'(t)\tilde{x}(t) - \tilde{x}'(t)x(t)]}{\tilde{x}(t)x(t)}.$$

The numerator of the last fraction is a constant and this constant is positive since we have $(\frac{x}{\tilde{x}})' = \frac{r[x'\tilde{x}-\tilde{x}'x]}{r\tilde{x}^2} > 0$ which follows from the fact that $\tilde{x}$ is the principal solution, i.e., $\frac{x}{\tilde{x}}$ tends monotonically to ∞. Hence, the solution $\tilde{w}$ of the Riccati equation (7.3) given by the principal solution of (1.1) is less than any other solution of (7.3) near ∞. Conversely, let $\tilde{w} = r\tilde{x}'/\tilde{x}$ be the minimal solution of (2.1) and suppose that the solution $\tilde{x}$ of (1.1) is not principal, i.e., the integral in (7.1) is convergent. Let $T \in \mathbb{R}$ be such that $\int_T^\infty r^{-1}(t)\tilde{x}^{-2}(t)\,\mathrm{d}t < 1$ and consider the solution w of (7.3) given by the initial condition $w(T) = \tilde{w}(T) - \frac{1}{2\tilde{x}^2(T)}$. Put $v = \tilde{x}^2(\tilde{w} - w)$. Then $v(T) = \frac{1}{2}$ and by a direct computation we have

$$v' = \frac{v^2}{r(t)\tilde{x}^2(t)}.$$

Hence

$$v(t) = \frac{1}{2 - \int_T^t r^{-1}(s)\tilde{x}^{-2}(s)\,\mathrm{d}s} \leqslant \frac{1}{2 - \int_T^\infty r^{-1}(s)\tilde{x}^{-2}(s)\,\mathrm{d}s} \leqslant 1.$$

This means that v is extensible up to ∞ and hence w has the same property and at the same time $w(t) < \tilde{w}(t)$. This contradiction shows that the eventual minimality of $\tilde{w}$ implies that (7.1) holds, i.e., the associated solution $\tilde{x}$ of (0.1) is principal.

The last construction of the principal solution of (1.1) which we present here requires (in addition to nonoscillation of (1.1)) the assumption that for any t_0, the solution of (1.1) given by the initial condition $x(t_0) = 0$, $x'(t_0) \neq 0$ has a zero point rights of t_0 (later we will show that this assumption can be eliminated), we denote this zero point by $\eta(t_0)$. The function η is nondecreasing according to the Sturmian theory, hence there exists $\lim_{t\to\infty} \eta(t) =: T$ and $T < \infty$ since we suppose that (1.1) is nonoscillatory. Now, the solution $\tilde{x}$ given by the initial condition $\tilde{x}(T) = 0$, $\tilde{x}'(T) \neq 0$ is the principal solution of (1.1). This construction is used in the original paper of Leighton and Morse [148]. Concerning other papers dealing with the principal solution of (1.1) and its properties we refer to [106] and the references given therein.

7.2. *Mirzov's construction of the principal solution*

This construction defines the principal solution of half-linear equation (0.1) via the minimal solution of the associated Riccati equation (2.1). Nonoscillation of (0.1) implies that there exist $T \in \mathbb{R}$ and a solution $\hat{w}$ of (2.1) which is defined in the whole interval $[T, \infty)$, i.e., such that (0.1) is disconjugate on $[T, \infty)$. Let $d \in (T, \infty)$ and let w_d be the solution of (2.1) determined by the solution x_d of (0.1) satisfying the initial condition $x(d) = 0$, $r(d)\Phi(y'(d)) = -1$. Then $w_d(d-) = -\infty$ and $w_d(t) < \hat{w}(t)$ for $t \in (T, d)$. Moreover, if $T < d_1 < d_2$ then

$$w_{d_1}(t) < w_{d_2}(t) < \hat{w}(t) \quad \text{for } t \in (T, d_1).$$

This implies that for $t \in (T, \infty)$ there exists the limit $w_\infty(t) := \lim_{d\to\infty} w_d(t)$ and monotonicity of this convergence (with respect to the “subscript” variable) implies that this convergence is uniform on every compact subinterval of $[T, \infty)$. Consequently, the limit function w_∞ solves (2.1) as well and any solution w of this equation which is extensible up to ∞ satisfies the inequality $w(t) > w_\infty(t)$ near ∞. Indeed, if a solution $\bar{w}$ would satisfy the inequality $\bar{w}(t) < w_\infty(t)$ on some interval (T_1, ∞), then for $\bar{t} \in (T_1, \infty)$ and d sufficiently large we have $\bar{w}(\bar{t}) < w_d(\bar{t}) < w_\infty(\bar{t})$. But this contradicts the fact that $w_d(d-) = -\infty$ and that graphs of solutions of (2.1) cannot intersect (because of unique solvability of this equation).

Now, having defined the *minimal solution* w_∞ of (2.1), we define the *principal solution* of (0.1) at ∞ as the (nontrivial) solution of the first order equation

$$x' = \Phi^{-1}\left(\frac{w(t)}{r(t)}\right)x, \tag{7.4}$$

i.e., the principal solution of (0.1) at ∞ is determined uniquely up to a multiplicative factor by the formula

$$x(t) = x(T)\exp\left\{\int_T^t r^{1-q}(s)\Phi^{-1}(w(s))\,\mathrm{d}s\right\}.$$

REMARK 7.1. (i) Mirzov actually used in his paper [177] a slightly different approach which can be briefly explained as follows. Suppose that (0.1) is nonoscillatory and let $\hat{w}$ be a solution of the associated Riccati equation which exists on some interval $[T, \infty)$ and let $W := \hat{w}(T)$. Denote by

$$\mathcal{W} = \big\{v \in (-\infty, W)\colon \text{ the solution } w \text{ of (2.1) given by the initial condition } w(T) = v \text{ is not extensible up to } \infty\big\},$$

i.e., $\mathcal{W}$ are initial values of solutions of (2.1) at $t = T$ which blow down to $-\infty$ at some finite time $t > T$. Note that the set $\mathcal{W}$ is nonempty what can be seen as follows. Let $T_1 > T$ be arbitrary and consider a solution x of (0.1) given by $x(T_1) = 0$, $x'(T_1) \neq 0$. Disconjugacy of (0.1) on $[T, \infty)$ implies that $x(t) \neq 0$ on $[T, T_1)$ and the value of the associated solution of the Riccati equation $w = \frac{r\Phi(x')}{\Phi(x)}$ at $t = T$ clearly belongs to $\mathcal{W}$. Now, let $\tilde{v} := \sup \mathcal{W}$ and let w_∞ be the solution of (2.1) given by the initial condition $w(T) = \tilde{v}$. Then this solution is extensible up to ∞ (supposing that this is not the case, we would get a contradiction with the definition of the number $\tilde{v}$) and the principal solution of (0.1) is defined again by (7.4) with w_∞ substituted for w.

(ii) Let b be a *regular* point of Equation (0.1) in the sense that for any $A, B \in \mathbb{R}$ the initial condition $x(b) = A$, $r(b)\Phi(x'(b)) = B$ determines uniquely a solution of (0.1). Let x_b be a solution given by $x_b(b) = 0$, $x_b'(b) \neq 0$. Replacing in the above construction the point $t = \infty$ by $t = b$, i.e., $w_b(t) := \lim_{d \to b-} w_d(t)$, it not difficult to see that $w_b = \frac{r\Phi(x_b')}{\Phi(x_b)}$. Consequently, what we call the principal solution x_b of (0.1) at a regular point $b \in \mathbb{R}$ is the nontrivial solution satisfying the condition $x_b(b) = 0$.

7.3. *Construction of Elbert and Kusano*

This construction was introduced (independently of Mirzov's approach) in the paper [90] and it is based on the half-linear Prüfer transformation.

Let (0.1) be nonoscillatory and let T be such that this equation is disconjugate on $[T, \infty)$. Take a solution x which is positive on $[T, \infty)$. By the generalized Prüfer transformation (see Section 1.3) this solution can be expressed in the form

$$x(t) = \rho(t)\sin_p \varphi(t), \qquad r^{q-1}(t)x'(t) = \rho(t)\cos_p \varphi(t), \tag{7.5}$$

where ρ is a positive function, the half-linear sine and cosine functions $\sin_p$, $\cos_p$ were defined in Section 1.3 and the function φ is a solution of the first order equation

$$\varphi' = r^{1-q}(t)\left|\cos_p \varphi(t)\right|^p + \frac{c(t)}{p-1}\left|\sin_p \varphi(t)\right|^p. \tag{7.6}$$

The fact that $x(t) > 0$ for $t \in [T, \infty)$ implies that $\varphi(t) \in (k\pi_p, (k+1)\pi_p)$ for some even $k \in \mathbb{Z}$ and without loss of generality we can suppose that $k = -1$. Now, let $\tau \in (T, \infty)$ and let φ_τ be the solution of (7.6) given by the initial condition $\varphi_\tau(\tau) = 0$. Since any solution of (7.6) satisfies $\varphi'(t) > 0$ whenever $\varphi(t) = 0 \pmod{\pi_p}$, the unique solvability of (7.6) (compare again Section 1.3) implies that

$$\varphi(t) < \varphi_{\tau_2}(t) < \varphi_{\tau_1}(t) \quad \text{for } t \geqslant T, \text{ whenever } T < \tau_1 < \tau_2$$

(drawing a picture helps to visualize the situation). Consequently, the monotonicity of φ_τ with respect to τ implies that there exists a finite limit

$$\lim_{\tau \to \infty} \varphi_\tau(T) = \varphi^*.$$

Now, the *principal solution* is the solution of (0.1) given by the initial condition

$$\tilde{x}(T) = \sin\varphi^*, \qquad \tilde{x}'(T) = r^{1-q}(T) \cos_p \varphi^*.$$

This means that we take $\rho(T) = 1$ in the definition of $\tilde{x}$, this can be done according to the homogeneity of the solution space of (0.1).

THEOREM 7.1. *A solution $\tilde{x}$ of a nonoscillatory equation* (0.1) *is principal in sense of Mirzov's construction if and only if it is principal in the sense of Elbert and Kusano.*

PROOF. Let x_τ be a nontrivial solution of (0.1) satisfying $x_\tau(\tau) = 0$. This solution can be expressed in the form

$$x_\tau(t) = \rho(t) \sin_p \varphi_\tau(t), \qquad r^{q-1}(t) x_\tau'(t) = \rho(t) \cos_p \varphi_\tau(t),$$

where φ_τ is the solution of (7.6) satisfying $\varphi_\tau(\tau) = 0$. The corresponding solution of the associated Riccati equation (2.1)

$$w_\tau(t) = \frac{r(t)\Phi(x_\tau'(t))}{\Phi(x_\tau(t))} = \Phi\big(\cot_p \varphi_\tau(t)\big)$$

satisfies $w_\tau(\tau-) = -\infty$. The minimal solution of (2.1) (which defines the principal solution of (0.1) in Mirzov's definition) is given by $\tilde{w}(t) = \lim_{\tau\to\infty} w_\tau(t)$, i.e., it is just the solution satisfying $\tilde{w}(T) = \Phi(\cot_p \varphi^*)$ and this is the solution of Riccati equation (2.1) given by the principal solution obtained by Elbert–Kusano's construction. $\square$

We finish this subsection with some examples of equations whose principal solution can be computed explicitly.

EXAMPLE 7.1. (i) Consider the one-term half-linear equation

$$\big(r(t)\Phi(x')\big)' = 0. \tag{7.7}$$

As we have mentioned in Section 4, the solution space of this equation is a two-dimensional linear space with the basis $x_1(t) \equiv 1$, $x_2(t) = \int^t r^{1-q}(s)\,\mathrm{d}s$. The Riccati equation associated with (7.7) is $w' + (p-1)r^{1-q}|w|^q = 0$ and the general solution of this equation is

$$w(t) = \frac{1}{\Phi(C + \int_T^t r^{1-q}(s)\,\mathrm{d}s)}, \qquad w(t) \equiv 0. \tag{7.8}$$

If $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$, then by an easy computation one can verify that $\tilde{w}(t) \equiv 0$ is the eventually minimal solution of this equation and hence $\tilde{x}(t) = 1$ is the principal solution of (0.1). If $\int^\infty r^{1-q}(t)\,\mathrm{d}t < \infty$, the eventually minimal solution of the Riccati equation is $w(t) = -\frac{1}{\Phi(\int_t^\infty r^{1-q}(s)\,\mathrm{d}s)}$ (we take $C = -\int_T^\infty r^{1-q}(s)\,\mathrm{d}s$ in formula (7.8)) and $\tilde{x}(t) = \int_t^\infty r^{1-q}(s)\,\mathrm{d}s$ is the principal solution of (0.1).

(ii) The nonoscillatory equation $(\Phi(x'))' - (p-1)\Phi(x) = 0$ investigated in Section 4.1 has solutions $x(t) = \mathrm{e}^{\pm t}$ and all other solutions are asymptotically equivalent to e^t. Consequently, the solution $\tilde{x}(t) = \mathrm{e}^{-t}$ is the principal solution at ∞.

(iii) The Euler-type equation

$$\bigl(\Phi(x')\bigr)' + \frac{\gamma}{t^p}\Phi(x) = 0 \tag{7.9}$$

is nonoscillatory if and only if $\gamma \leqslant \tilde{\gamma} = (\frac{p-1}{p})^p$, see Section 4.2. If $\gamma = \tilde{\gamma}$, then (7.9) has a solution $x(t) = t^{\frac{p-1}{p}}$ and all linearly independent solutions are asymptotically equivalent to $t^{\frac{p-1}{p}} \lg^{\frac{2}{p}} t$. Consequently, $\tilde{x}(t) = t^{\frac{p-1}{p}}$ is the principal solution of (7.9). If $\gamma < \tilde{\gamma}$, then $x_1(t) = \mathrm{e}^{\lambda_1 t}$, $x_2(t) = \mathrm{e}^{\lambda_2 t}$, where $\lambda_1 < \lambda_2$ are the roots of the algebraic equation $(p-1)[|\lambda|^p - \Phi(\lambda)] + \gamma = 0$, are solutions of (7.9), and all other linearly independent solutions are asymptotically equivalent to $x_2(t)$. Consequently, $\tilde{x}(t) = x_1(t) = \mathrm{e}^{\lambda_1 t}$.

7.4. *Comparison theorem for eventually minimal solutions of Riccati equations*

Similarly as in the linear case we have the following inequalities for solutions of a pair of Riccati equations corresponding to nonoscillatory half-linear equations.

THEOREM 7.2. *Consider a pair of half-linear equations* (0.1), (2.13), *and suppose that* (2.13) *is a Sturmian majorant of* (0.1) *for large t, i.e., there exists $T \in \mathbb{R}$ such that $0 < R(t) \leqslant r(t)$, $c(t) \leqslant C(t)$ for $t \in [T, \infty)$. Suppose that the majorant equation* (2.13) *is nonoscillatory and denote by $\tilde{w}$, $\tilde{v}$ eventually minimal solutions of* (2.1) *and of*

$$v' + C(t) + (p-1)R^{1-q}(t)|v|^q = 0, \tag{7.10}$$

respectively. Then $\tilde{w}(t) \leqslant \tilde{v}(t)$ for large t.

PROOF. Nonoscillation of (2.13) implies the existence of $T \in \mathbb{R}$ such that $\tilde{w}$ and $\tilde{v}$ exist on $[T, \infty)$. Suppose that there exists $t_1 \in [T, \infty)$ such that $\tilde{w}(t_1) > \tilde{v}(t_1)$. Let w be the solution of (2.1) given by the initial condition $w(t_1) = \tilde{v}(t_1)$. Then according to the standard theorem on differential inequalities (see, e.g., [145]) we have $w(t) \geqslant \tilde{v}(t)$ for $t \geqslant t_1$, i.e., w is extensible up to ∞. At the same time $w(t) < \tilde{w}(t)$ for $t \geqslant t_1$ since graphs of solutions of (2.1) cannot intersect (because of the unique solvability). But this contradicts the eventual minimality of $\tilde{w}$. ☐

In some oscillation criteria, we will need the following immediate consequence of the previous theorem.

COROLLARY 7.1. *Let* $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$, $c(t) \geqslant 0$ *for large* t *and suppose that* (0.1) *is nonoscillatory. Then the eventually minimal solution of the associated Riccati equation* (2.1) *satisfies* $\tilde{w}(t) \geqslant 0$ *for large* t.

PROOF. Under the assumptions of corollary, (0.1) is the majorant of the one-term equation $(r(t)\Phi(y'))' = 0$. Since $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$, $\tilde{y} \equiv 1$ is the principal solution of this equation (compare Example 7.1). Hence $\tilde{v}(t) = 0$ is the eventually minimal solution of the associated Riccati equation which implies the required statement. ☐

7.5. *Sturmian property of the principal solution*

In this short subsection we briefly show that the principal solution of (0.1) has a Sturmian-type property and that the largest zero point of this solution (if any) behaves like the left conjugate point of ∞, in a certain sense.

THEOREM 7.3. *Suppose that Equation* (0.1) *is nonoscillatory and its principal solution* $\tilde{x}$ *has a zero point and let* T *be the largest of them. Further suppose that Equation* (2.13) *is a Sturmian majorant of* (0.1) *on* $[T, \infty)$, *i.e.,* $0 < R(t) \leqslant r(t)$ *and* $C(t) \geqslant c(t)$ *for* $t \in [T, \infty)$. *Then any solution* y *of* (2.13) *has a zero point in* (T, ∞) *or it is a constant multiple of* $\tilde{x}$. *The latter possibility is excluded if one of the inequalities between* r, R *and* c, C, *respectively, is strict on an interval of positive length.*

PROOF. If (2.13) is oscillatory, the statement of theorem trivially holds, so suppose that (2.13) is nonoscillatory and let $\tilde{y}$ be its principal solution. Denote by $\tilde{w}$ and $\tilde{v}$ the minimal solutions of corresponding Riccati equations (2.1) and (7.10), respectively. According to the comparison theorem for minimal solutions of Riccati equations presented in the previous subsection, we have $\tilde{w}(t) \geqslant \tilde{v}(t)$ on the interval of existence of $\tilde{w}$. Since we suppose that $\tilde{x}(T) = 0$, this implies that $\tilde{w}(T+) = \infty$, so the interval of existence of $\tilde{v}$ must be a subinterval of $[T, \infty)$, say $[T_1, \infty)$, i.e., T_1 is the largest zero of the principal solution $\tilde{y}$ of (2.13). If one of the inequalities between r, R and c, C is strict it can be shown that the possibility $T = T_1$ is excluded. Now let y be any nontrivial solution of (2.13). If $y(t) \neq 0$ for $t \in [T_1, \infty)$, then the associated solution $v = R\Phi(y')/\Phi(y)$ exists on $[T_1, \infty)$ and satisfies there the inequality $v(t) < \tilde{v}(t)$ on $[T_1, \infty)$ (since $\tilde{v}(T_1+) = \infty$ and $v(T-) < \infty$) and this a contradiction with minimality of $\tilde{v}$. ☐

REMARK 7.2. (i) If T is the largest zero of the principal solution $\tilde{x}$ of (0.1), i.e., the same as in the previous theorem, and suppose that $R(t) = r(t)$ and $C(t) = c(t)$ for $t \in [T, \infty)$.

Then Theorem 7.3 shows that T plays the role of the left conjugate point of ∞ in the sense that any nonprincipal solution of (0.1), i.e., a solution linearly independent of $\tilde{x}$, has exactly one zero in (T, ∞). Note that the fact that this zero is exactly one (and not more) follows from the classical Sturm separation theorem (Theorem 2.3).

(ii) In Remark 2.2 we have pointed out that the disconjugacy of (0.1) on a bounded interval $I = [a, b]$ (which, by definition, means that the solution x given by $x(a) = 0$, $x'(a) \neq 0$ has no zero in $(a, b]$) is actually equivalent to the existence of a solution without any zero in $[a, b]$. Theorem 7.3 shows that we have the same situation with unbounded intervals or an interval whose endpoints are singular points of (0.1). For example, if $I = \mathbb{R} = (-\infty, \infty)$, then disconjugacy of (0.1) on this interval (defined as disconjugacy on $[-T, T]$ for every $T > 0$) is equivalent to the existence of a solution without any zero on $\mathbb{R}$, the solution having this property is, e.g., the principal solution (at ∞).

7.6. *Integral characterization of the principal solution*

Among all (equivalent) characterizations of the principal solution of linear equation (1.1), the most suitable seems be the integral one (7.1), since it needs to know just only one solution and according to the divergence/convergence of the characterizing integral it is possible to decide whether or not it is the principal solution. The remaining characterizations require to know other solutions since they are of comparison type. In the linear case, this is not serious disadvantage because of the reductions of order formula which enables to compute all solutions (at least locally) of the linear second order equation when one solution is already known. However, in the half-linear case we have no reduction of order formula as pointed in Section 3, so some kind of the integral characterization would be very useful. In the next theorem we present one candidate for the integral characterization of the principal solution of (0.1). The parts (i), (ii) and (iii) are proved in [66] and the part (iv) in [36].

THEOREM 7.4. *Suppose that Equation* (0.1) *is nonoscillatory and $\tilde{x}$ is its solution such that $\tilde{x}'(t) \neq 0$ for large t.*

(i) *Let $p \in (1, 2)$. If*

$$I(\tilde{x}) := \int^{\infty} \frac{\mathrm{d}t}{r(t)\tilde{x}^2(t)|\tilde{x}'(t)|^{p-2}} = \infty, \tag{7.11}$$

then $\tilde{x}$ is the principal solution.

(ii) *Let $p > 2$. If $\tilde{x}$ is the principal solution then* (7.11) *holds.*

(iii) *Suppose that $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty$, the function $\gamma(t) := \int_t^{\infty} c(s)\,\mathrm{d}s$ exists and $\gamma(t) \geqslant 0$, but $\gamma(t) \not\equiv 0$ eventually. Then $\tilde{x}(t)$ is the principal solution if and only if* (7.11) *holds.*

(iv) *Let $c(t) > 0$ for large t, $\int^{\infty} r^{1-q}(t)\,dt < \infty$, $\int^{\infty} c(t)\,\mathrm{d}t = \infty$. Then $\tilde{x}(t)$ is the principal solution if and only if* (7.11) *holds.*

PROOF. (i) Suppose, by contradiction, that a (positive) solution x of (0.1) satisfying (7.11) is not principal. Then the corresponding solution $w_x = r\Phi(x'/x)$ of the associated Riccati equation (2.1) is not eventually minimal. Hence, there exists another nonoscillatory solution y of (0.1) such that

$$w_y = r\Phi(y'/y) < w_x \quad \text{eventually.} \tag{7.12}$$

Due to the Picone identity given in Section 2.2 we have

$$r(t)|x'|^p - c(t)x^p = \left[x^p w_y\right]' + pr^{1-q}(t)x^p P\left(\Phi^{-1}(w_x), w_y\right)$$

and at the same time

$$r(t)|x'|^p - c(t)x^p = \left(x^p w_x\right)' - x\left[\left(r(t)\Phi(x')\right)' + c(t)\Phi(x)\right] = \left(x^p w_x\right)'.$$

Subtracting the last two equalities, we get

$$\left[x^p(w_x - w_y)\right]' = pr^{1-q}(t)x^p P\left(\Phi^{-1}(w_x), w_y\right).$$

Let $f(t) = x^p(w_x - w_y)$. By (7.12) there exists T sufficiently large such that $f(t) > 0$ for $t \geqslant T$. Then by Lemma 2.1 we have

$$\begin{aligned}\frac{f'}{f^2} &= \frac{p}{f^2} r^{1-q}(t)x^p P\left(\Phi_q(w_x), w_y\right)\\ &> \frac{p}{2}\,\frac{x^p r^{1-q}}{[x^p(w_x - w_y)]^2}\left|r^{q-1}(x'/x)\right|^{2-p}(w_x - w_y)^2\\ &= \frac{p}{2r(t)x^2|x'|^{p-2}}.\end{aligned}$$

Integrating the last inequality from T to T_1 $(T_1 > T)$, we have

$$\frac{1}{f(T)} > \frac{1}{f(T)} - \frac{1}{f(T_1)} \geqslant \frac{p}{2}\int_T^{T_1} \frac{\mathrm{d}t}{r(t)x^2(t)|x'(t)|^{p-2}}$$

and letting $T_1 \to \infty$ we are led to contradiction. Hence a solution satisfying (7.11) is principal.

(ii) We proceed again by contradiction. Suppose that $\tilde{x}$ is the principal solution and $I(\tilde{x}) < \infty$. Let T be chosen so large that $\tilde{x}(t) > 0$ for $t \geqslant T$ and

$$\int_T^\infty \frac{\mathrm{d}t}{r(t)\tilde{x}^2(t)|\tilde{x}'(t)|^{p-2}} < \frac{1}{p}.$$

Consider the solution $\bar{w}(t)$ of the Riccati equation (2.1) given by the initial condition

$$\bar{w}(T) = \tilde{w}(T) - \frac{1}{2\tilde{x}^p(T)},$$

where $\tilde{w} = r\Phi(\tilde{x}'/\tilde{x})$, i.e., $\bar{w}(T) < w(T)$. We want to show that $\bar{w}(t)$ is extensible up to ∞. To this end, denote $f(t) = \tilde{x}^p(t)(\tilde{w}(t) - \bar{w}(t))$. Then $f(T) = \frac{1}{2}$ and using the Picone identity, we have

$$\frac{f'(t)}{f^2(t)} = \frac{pr^{1-q}(t)}{f^2(t)} P\big(r^{q-1}\tilde{x}', \bar{w}\Phi(\tilde{x})\big),$$

hence, integrating this identity from T to t

$$\frac{1}{f(T)} - \frac{1}{f(t)} = p\int_T^t \frac{r^{1-q}(s)\tilde{x}^p(s)}{f^2(s)} P\left(r^{q-1}\frac{\tilde{x}'}{\tilde{x}}, \bar{w}\right) \mathrm{d}s. \tag{7.13}$$

By (2.8) of Lemma 2.1 we have

$$P\left(\frac{r^{q-1}\tilde{x}'}{\tilde{x}}, \bar{w}\right) \leqslant \frac{1}{2}\left|\frac{r^{q-1}\tilde{x}'}{\tilde{x}}\right|^{2-p} (w - \bar{w})^2,$$

which means, using (7.13) and taking into account that $f(T) = \frac{1}{2}$,

$$\begin{aligned} f(t) &\leqslant \left(2 - p\int_T^t \frac{r^{1-q}(s)\tilde{x}^p(s)}{f^2(s)} P\left(r^{q-1}\frac{\tilde{x}'}{\tilde{x}}, \bar{w}\right) \mathrm{d}s\right)^{-1} \\ &\leqslant \left(2 - p\int_T^\infty \frac{\mathrm{d}t}{2r(t)\tilde{x}^2(t)|\tilde{x}'(t)|^{p-2}}\right)^{-1} \leqslant 1. \end{aligned}$$

Consequently, $\frac{1}{2} \leqslant f(t) \leqslant 1$ and $f(t)$ can be continued to ∞, hence $\bar{w}(t)$ is a continuable up to infinity solution of (2.1) and $\bar{w}(t) < \tilde{w}(t)$ for $t \geqslant T$, i.e., $\tilde{w}(t)$ is not minimal. Thus, the solution $\tilde{x}(t)$ is not principal, which was to be proved.

(iii) The principal solution $\tilde{x}(t)$ of (0.1) is associated with the minimal solution $\tilde{w}(t)$ of (2.1) and hence it is also the minimal solution of the Riccati integral equation (the convergence of $\int^\infty r^{1-q}(t)|w(t)|^q \,\mathrm{d}t$ follows from Theorem 5.6)

$$\tilde{w}(t) = \gamma(t) + (p-1)\int_t^\infty r^{1-q}(s)\big|\tilde{w}(s)\big|^q \,\mathrm{d}s, \quad t \geqslant T_1,$$

and by the assumptions on $\gamma(t)$, there exists $T \in \mathbb{R}$ such that $\tilde{w}(t) > 0$ for $t \geqslant T$. Since $\tilde{w}$ is the minimal solution, for any other proper solution w of (2.1) we have $w(t) > \tilde{w}(t) > 0$ for $t \geqslant T_1 \geqslant T$, and hence the associated solutions $x(t)$ and $\tilde{x}(t)$ satisfy the inequalities $x'(t) > 0$, $\tilde{x}'(t) > 0$ for $t \geqslant T_1$.

Now the proof goes in different way according to $1 < p < 2$ or $p \geqslant 2$.

Case A: $1 < p < 2$. By the part (i) it is sufficient to show that the integral in (7.11) is really divergent. Suppose the contrary, i.e.,

$$\int^\infty \frac{\mathrm{d}t}{r(t)\tilde{x}^2(t)|\tilde{x}'(t)|^{p-2}} < \infty.$$

Let $T_2 \geqslant T_1$ be chosen so large that

$$\int_{T_2}^{\infty} \frac{\mathrm{d}t}{r(t)\tilde{x}^2(t)|\tilde{x}'(t)|^{p-2}} \leqslant \frac{2(p-1)}{p}.$$

Consider the solution $\bar{w}(t)$ of (2.1) with the initial condition

$$\bar{w}(T_2) = \tilde{w}(T_2) - \frac{1}{2\tilde{x}^2(T_2)},$$

and accordingly, the function $f(t)$ be defined by

$$f(t) = \tilde{x}^p(t)\bigl[\tilde{w}(t) - \bar{w}(t)\bigr].$$

Clearly, $f(T_2) = \frac{1}{2}$. Following the computation in the proof of the claim (i), we find

$$\frac{f'(t)}{f^2(t)} = p\frac{r^{1-q}(t)\tilde{x}^p(t)}{f^2(t)} P\bigl(\Phi^{-1}\bigl(\tilde{w}(t)\bigr), \bar{w}(t)\bigr)$$

hence by (2.9)

$$\frac{f'(t)}{f^2(t)} < \frac{p}{2(p-1)} \frac{1}{r(t)\tilde{x}^2(t)|\tilde{x}'(t)|^{p-2}}$$

and integrating this inequality over $[T_2, t]$ we find

$$\begin{aligned}\frac{1}{f(T_2)} - \frac{1}{f(t)} &< \frac{1}{2}\frac{p}{p-1}\int_{T_2}^{t} \frac{\mathrm{d}s}{r(s)\tilde{x}^2(s)|\tilde{x}'(s)|^{p-2}} \\ &< \frac{p}{2(p-1)}\int_{T_2}^{\infty} \frac{\mathrm{d}s}{r(s)\tilde{x}^2(s)|\tilde{x}'(s)|^{p-2}} \leqslant 1,\end{aligned}$$

consequently, $\frac{1}{2} \leqslant f(t) \leqslant 1$ for $t \geqslant T_2$. Thus, the function $\bar{w}(t)$ exists on $[T_2, \infty)$ and $\bar{w}(t) < \tilde{w}(t)$, i.e., $\tilde{w}(t)$ is not minimal solution of (2.1), hence $\tilde{x}(t)$ is not the principal solution, and this contradiction proves the first case.

Case B: $p \geqslant 2$. By the claim (ii) it is sufficient to show that if the solution x is not principal then the corresponding integral in (7.11) is convergent. Let $w(t) = r(t)\Phi(x'(t)/x(t))$ be the associated solution of (2.1). Then $w(t)$ is not minimal solution of (2.1) and let $\tilde{w}(t)$ be the minimal solution of this equation. Then we have $w(t) > \tilde{w}(t)$ for $t \geqslant T_2$ with T_2 sufficiently large. Consider the function $f(t)$ given again by

$$f(t) = x^p(t)\bigl[w(t) - \tilde{w}(t)\bigr] > 0 \quad \text{for } t \geqslant T_2.$$

By inequality (2.9) given in Lemma 2.1 we have again

$$\frac{f'}{f^2} = \frac{p}{f^2} r^{1-q} x^p P\bigl(\Phi^{-1}(w), \tilde{w}\bigr) > \frac{p}{2(p-1)} \frac{1}{rx^2|x'|^{p-2}}, \quad t \geqslant T_2,$$

hence

$$\frac{1}{f(T_2)} > \frac{1}{f(T_2)} - \frac{1}{f(t)} \geqslant \frac{p}{2(p-1)} \int_{T_2}^{t} \frac{\mathrm{d}s}{rx^2|x'|^{p-2}},$$

then letting $t \to \infty$ we obtain the desired result

$$\int^{\infty} \frac{\mathrm{d}t}{r(t)x^2(t)|x'(t)|^{p-2}} < \infty.$$

(iv) Concerning the proof of the part (iv), this proof is based on the statement which relates the principal solution of (0.1) to the principal solution of the reciprocal equation (6.1), we refer to [36] for details. □

REMARK 7.3. The equivalent integral characterization of the principal solution of (0.1) is stated in the parts (iii) and (iv) of the previous theorem is proved under some restriction on the functions r, c in (0.1). In order to better understand these restrictions, the concept of the *regular* half-linear equation has been introduced in [72] as follows. A nonoscillatory equation (0.1) is said to be regular if there exists a constant $K \geqslant 0$ such that

$$\limsup_{t\to\infty} \left| \frac{w_1(t)}{w_2(t)} \right| \leqslant K$$

for any pair of solutions w_1, w_2 of the associated Riccati equation (2.1) such that $w_2(t) > w_1(t)$ eventually. It was shown that for regular half-linear equation (7.11) holds if and only if the solution $\tilde{x}$ is principal and that under assumptions of (iii) and (iv) of the previous theorem equation (0.1) is regular.

7.7. *Another integral characterization*

The integral characterization (7.11) of the principal solution of (0.1) reduces to the usual integral characterization of the principal solution of linear equation (7.1) if $p = 2$. However, this characterization applies in case $p > 2$ only to solutions x for which $x'(t) \neq 0$ eventually. Moreover, in [36,37] examples of half-linear equations are given which show that if assumptions of the parts (iii) and (iv) of Theorem 7.4 are violated, (7.11) is *no longer equivalent characterization* of the principal solution of (0.1). For this reason, another integral characterization was suggested and the following statement is proved. The proof of this statement can be found in the above mentioned [36,37].

THEOREM 7.5. *Suppose that either*

(i) $c(t) < 0$ *for large* t, *or*

(ii) $c(t) > 0$ *for large* t *and both integrals* $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t$, $\int^{\infty} c(t)\,\mathrm{d}t$ *are convergent. Then a solution* $\tilde{x}$ *of* (0.1) *is principal if and only if*

$$\int^{\infty} \frac{\mathrm{d}t}{r^{q-1}(t)\tilde{x}^2(t)} = \infty. \qquad (7.14)$$

7.8. *Limit characterization of the principal solution*

The “most characteristic” property of the principal solution in the linear case is the limit characterization (7.2). It was proved in the above mentioned papers [36,37] that this limit characterization extends under the assumption that $c(t) \neq 0$ eventually also to half-linear equation (0.1). This statement has been proved in [36] in case $c(t) < 0$ and in [37] for $c(t) > 0$.

THEOREM 7.6. *Suppose that* (0.1) *is nonoscillatory and* $c(t) \neq 0$ *for large* t. *Then a solution* $\tilde{x}$ *is principal if and only if the limit characterization* (7.2) *holds for every solution* x *linearly independent of* $\tilde{x}$.

PROOF. The proof is based on the detailed asymptotic analysis of solutions of (0.1) made in the previous section. Since this analysis is rather complicated, we refer to the above mentioned papers [36,37] for details. □

8. Conjugacy and disconjugacy of half-linear equations

Recall that similarly as in case of linear equations, Equation (0.1) is said to be *disconjugate* in a given interval I if every nontrivial solution of this equation has at most one zero in I, in the opposite case, i.e., if there exists a nontrivial solution of (0.1) having at least two zeros in I, Equation (0.1) is said to be *conjugate* in I.

In this section we will present criteria for conjugacy and disconjugacy of equation (0.1). Some theoretical criteria of this kind have been already formulated in the previous sections and are essentially involved in the Theorem 2.2. For example, the existence of a solution w of Riccati equation (2.1) associated with (0.1) is a sufficient condition for disconjugacy of this equation in the interval of the existence of this solution w. Another criteria can be formulated as consequences of the Sturmian comparison theorem.

8.1. *Leighton's conjugacy criterion*

Consider a pair of half-linear differential equations (0.1) and (2.13). If (2.13) is a Sturmian minorant of (0.1) on $I = [a, b]$ and (2.13) is conjugate on this interval, then majorant equation (0.1) is conjugate on $[a, b]$ as well. In the next theorem we replace the pointwise comparison of coefficients by the integral one. In the linear case $p = 2$ this statement was proved by Leighton [147], the half-linear version of this statement given here can be found in [115].

THEOREM 8.1. *Suppose that points* a, b *are conjugate relative to* (2.13) *and let* y *be a nontrivial solution of this equation for which* $y(a) = 0 = y(b)$. *If*

$$\mathcal{J}(y; a, b) := \int_a^b \left[\big(r(t) - R(t)\big)|y'|^p - \big(c(t) - C(t)\big)|y|^p\right] \mathrm{d}t \leqslant 0, \tag{8.1}$$

then (0.1) *is also conjugate in* $[a,b]$.

PROOF. We have (with the notation introduced in Theorem 2.2)

$$\begin{aligned}
\mathcal{F}(y;a,b) &= \int_a^b \big[r(t)|y'|^p - c(t)|y|^p\big]\,\mathrm{d}t \\
&= \int_a^b \big[R(t)|y'|^p - C(t)|y|^p\big]\,\mathrm{d}t + \mathcal{J}(y;a,b) \\
&= \big[R(t)y\Phi(y')\big]_a^b - \int_a^b y\big[\big(R(t)\Phi(y')\big)' - C(t)\Phi(y)\big]\,\mathrm{d}t \\
&\quad + \mathcal{J}(y;a,b) \\
&= \mathcal{J}(y;a,b) \leqslant 0,
\end{aligned}$$

hence (0.1) is conjugate on $[a,b]$ by Theorem 2.2. □

8.2. *Singular Leighton's theorem*

In this subsection we show that if the points a,b are singular points of considered equations, in particular, $a=-\infty$, $b=\infty$ (or finite singularities, i.e., points where the unique solvability is violated), Leighton-type comparison theorem still holds if we replace the solution satisfying $y(a)=0=y(b)$ by the principal solution at a and b. We formulate the statement in a simplified form, as can be found in [59], a more general formulation is presented in [67].

THEOREM 8.2. *Suppose that* $\tilde{c}$ *is a continuous function such that the equation*

$$\big(r(t)\Phi(y')\big)' + \tilde{c}(t)\Phi(y) = 0 \tag{8.2}$$

has the property that the principal solutions at a *and* b *coincide and denote by* h *this simultaneous principal solution at these points. If*

$$\liminf_{s_1\downarrow a,\, s_2\uparrow b} \int_{s_1}^{s_2} \big(c(t)-\tilde{c}(t)\big)\big|h(t)\big|^p\,\mathrm{d}t \geqslant 0, \quad c(t)\not\equiv\tilde{c}(t) \text{ in } (a,b), \tag{8.3}$$

then (0.1) *is conjugate in* $I=(a,b)$, *i.e., there exists a nontrivial solution of this equation having at least two zeros in* I.

PROOF. Our proof is based on the relationship between nonpositivity of the energy functional $\mathcal{F}$ and conjugacy of (0.1) given in Theorem 2.2. We construct a nontrivial function piecewise of the class C^1, with a compact support in I, such that $\mathcal{F}(y;a,b)<0$.

Continuity of the functions $c,\tilde{c}$ and (8.3) imply the existence of $\bar{t}\in I$ and $d,\varrho>0$ such that $(c(t)-\tilde{c}(t))|h(t)|^p > d$ for $(\bar{t}-\varrho,\bar{t}+\varrho)$. Let Δ be any positive differentiable function with the compact support in $(\bar t-\varrho,\bar t+\varrho)$. Further, let $a<t_0<t_1<\bar t-\varrho<\bar t+\varrho<t_2<t_3<b$ and let f,g be the solutions of (8.2) satisfying the boundary conditions

$$f(t_0)=0,\qquad f(t_1)=h(t_1),\qquad g(t_2)=h(t_2),\qquad g(t_3)=0.$$

Note that such solutions exist if t_0,t_1 and t_2,t_3 are sufficiently close to a and b, respectively, due to nonoscillation of (8.2) near a and b (this is implied by the existence of principal solutions at these points) and the fact that the solution space of this equation is homogeneous. Define the function y as follows

$$y(t)=\begin{cases}0, & t\in(a,t_0],\\ f(t), & t\in[t_0,t_1],\\ h(t), & t\in[t_1,t_2]\backslash[\bar t-\varrho,\bar t+\varrho],\\ h(t)\bigl(1+\delta\Delta(t)\bigr), & t\in[\bar t-\varrho,\bar t+\varrho],\\ g(t), & t\in[t_2,t_3],\\ 0, & t\in[t_3,b),\end{cases}$$

where δ is a real parameter. Then we have

$$\begin{aligned}\mathcal{F}(y;t_0,t_3)&=\int_{t_0}^{t_3}\bigl[r(t)|y'|^p-c(t)|y|^p\bigr]\,\mathrm{d}t\\ &=\int_{t_0}^{t_3}\bigl[r(t)|y'|^p-\tilde c(t)|y|^p\bigr]\,\mathrm{d}t-\int_{t_0}^{t_3}\bigl[c(t)-\tilde c(t)\bigr]|y|^p\,\mathrm{d}t\\ &=\int_{t_0}^{t_1}\bigl[r(t)|f'|^p-\tilde c(t)|f|^p\bigr]\,\mathrm{d}t-\int_{t_0}^{t_1}\bigl[c(t)-\tilde c(t)\bigr]|f|^p\,\mathrm{d}t\\ &\quad+\int_{t_1}^{t_2}\bigl[r(t)|y'|^p-\tilde c(t)|y|^p\bigr]\,\mathrm{d}t-\int_{t_1}^{t_2}\bigl[c(t)-\tilde c(t)\bigr]|y|^p\,\mathrm{d}t\\ &\quad+\int_{t_2}^{t_3}\bigl[r(t)|g'|^p-\tilde c(t)|g|^p\bigr]\,\mathrm{d}t-\int_{t_2}^{t_3}\bigl[c(t)-\tilde c(t)\bigr]|g|^p\,\mathrm{d}t.\end{aligned}$$

Denote by w_f, w_g, w_h the solutions of the Riccati equation associated with (8.2)

$$w'+\tilde c(t)+(p-1)r^{1-q}(t)|w|^q=0 \tag{8.4}$$

generated by f, g and h, respectively, i.e.,

$$w_f=\frac{r\Phi(f')}{\Phi(f)},\qquad w_g=\frac{r\Phi(g')}{\Phi(g)},\qquad w_h=\frac{r\Phi(h')}{\Phi(h)}.$$

Then using Picone's identity (2.7)

$$\int_{t_0}^{t_1}\bigl[r(t)|f'|^p-\tilde c(t)|f|^p\bigr]\,\mathrm{d}t$$

$$\begin{aligned}&= w_f|f|^p\big|_{t_0}^{t_1} + p\int_{t_0}^{t_1} r^{1-q}(t)P\big(r^{q-1}f', \Phi(f)w_f\big)\,\mathrm{d}t\\&= w_f|f|^p\big|_{t_0}^{t_1},\end{aligned}$$

where $P(u,v) = \frac{|u|^p}{p} - uv + \frac{|v|^q}{q}$ (see Section 2.2). Similarly,

$$\int_{t_2}^{t_3}\big[r(t)|g'|^p - \tilde{c}(t)|g|^p\big]\,\mathrm{d}t = w_g|g|^p\big|_{t_2}^{t_3}.$$

Concerning the interval $[t_1, t_2]$, we have (again by identity (2.7))

$$\begin{aligned}\widetilde{\mathcal{F}}(y; t_1, t_2) &= \int_{t_1}^{t_2}\big[r(t)|y'|^p - \tilde{c}(t)|y|^p\big]\,\mathrm{d}t\\&= w_h|h|^p\big|_{t_1}^{t_2} + p\int_{t_1}^{t_2} r^{1-q}(t)P\big(r^{q-1}y', \Phi(y)w_h\big)\,\mathrm{d}t\\&= w_h|h|^p\big|_{t_1}^{t_2} + \int_{\bar{t}-\varrho}^{\bar{t}+\varrho}\Big\{r(t)\big|h' + \delta(\Delta h)'\big|^p\\&\quad - pr(t)\frac{\Phi(h')}{\Phi(h)}y'h^{p-1}(1+\delta\Delta)^{p-1}\\&\quad + (p-1)r^{1-q}(t)\left|\frac{r(t)\Phi(h')}{\Phi(h)}\right|^q h^p(1+\delta\Delta)^p\Big\}\,\mathrm{d}t\\&= w_h|h|^p\big|_{t_1}^{t_2} + \int_{\bar{t}-\varrho}^{\bar{t}+\varrho} r(t)\big\{|h'|^p + p\delta(\Delta h)'\Phi(h') + \mathrm{o}(\delta)\\&\quad - p\big(h' + \delta(\Delta h)'\big)\Phi(h')\big(1 + (p-1)\delta\Delta + \mathrm{o}(\delta)\big)\\&\quad + (p-1)|h'|^p\big(1 + p\delta\Delta + \mathrm{o}(\delta)\big)\big\}\,\mathrm{d}t\\&= w_h|h|^p\big|_{t_1}^{t_2} + \int_{\bar{t}-\varrho}^{\bar{t}+\varrho} r(t)\big\{|h'|^p + p\delta(\Delta h)'\Phi(h') + p|h'|^p\\&\quad - p\delta\Phi(h')(\Delta h)' - p(p-1)\delta\Delta|h'|^p + (p-1)|h'|^p\\&\quad + (p-1)p\delta\Delta|h'|^p + \mathrm{o}(\delta)\big\}\,\mathrm{d}t\\&= w_h|h|^p\big|_{t_1}^{t_2} + \mathrm{o}(\delta).\end{aligned}$$

Consequently,

$$\begin{aligned}\widetilde{\mathcal{F}}(y; t_0, t_3) &= \int_{t_0}^{t_3}\big[r(t)|y'|^p - \tilde{c}(t)|y|^p\big]\,\mathrm{d}t\\&= w_f|f|^p\big|_{t_0}^{t_1} + w_h|h|^p\big|_{t_1}^{t_2} + w_g|g|^p\big|_{t_2}^{t_3} + \mathrm{o}(\delta)\\&= \big|h(t_1)\big|^p\big(w_f(t_1) - w_h(t_1)\big) + \big|h(t_2)\big|^p\big(w_h(t_2) - w_g(t_2)\big) + \mathrm{o}(\delta)\end{aligned}$$

as $\delta \to 0+$.

Further, observe that the function $\frac{f}{h}$ is monotonically increasing in (t_0, t_1) since $\frac{f}{h}(t_0) = 0$, $\frac{f}{h}(t_1) = 1$ and $(\frac{f}{h})' = \frac{f'h - fh'}{h^2} \neq 0$ in (t_0, t_1). Indeed, if $f'h - fh' = 0$ at some point $\tilde{t} \in (t_0, t_1)$, i.e., $\frac{f'}{f}(\tilde{t}) = \frac{h'}{h}(\tilde{t})$ then $w_f(\tilde{t}) = w_h(\tilde{t})$ which contradicts the unique solvability of the generalized Riccati equation. By the second mean value theorem of integral calculus there exists $\xi_1 \in (t_0, t_1)$ such that

$$\begin{aligned}\int_{t_0}^{t_1} \bigl(c(t) - \tilde{c}(t)\bigr)|f|^p \,\mathrm{d}t &= \int_{t_0}^{t_1} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \frac{|f|^p}{|h|^p} \,\mathrm{d}t \\ &= \int_{\xi_1}^{t_1} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \,\mathrm{d}t.\end{aligned}$$

By the same argument the function $\frac{g}{h}$ is monotonically decreasing in (t_2, t_3) and

$$\int_{t_2}^{t_3} \bigl(c(t) - \tilde{c}(t)\bigr)|g|^p \,\mathrm{d}t = \int_{t_2}^{\xi_2} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \,\mathrm{d}t$$

for some $\xi_2 \in (t_2, t_3)$.

Concerning the interval (t_1, t_2) we have

$$\begin{aligned}&\int_{t_1}^{t_2} \bigl(c(t) - \tilde{c}(t)\bigr)|y|^p \,\mathrm{d}t \\ &\quad = \int_{t_1}^{\bar{t}-\varrho} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \\ &\qquad + \int_{\bar{t}-\varrho}^{\bar{t}+\varrho} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p (1 + \delta\Delta)^p \,\mathrm{d}t + \int_{\bar{t}+\varrho}^{t_2} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \,\mathrm{d}t \\ &\quad = \int_{t_1}^{t_2} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \,\mathrm{d}t + \delta \int_{\bar{t}-\varrho}^{\bar{t}+\varrho} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \Delta(t) \,\mathrm{d}t + \mathrm{o}(\delta) \\ &\quad \geqslant \int_{t_1}^{t_2} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \,\mathrm{d}t + \delta K + \mathrm{o}(\delta),\end{aligned}$$

where $K = d \int_{\bar{t}-\varrho}^{\bar{t}+\varrho} \Delta(t) \,\mathrm{d}t > 0$. Therefore

$$\int_{t_0}^{t_3} \bigl(c(t) - \tilde{c}(t)\bigr)|y|^p \,\mathrm{d}t \geqslant \int_{\xi_1}^{\xi_2} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \,\mathrm{d}t + K\delta + \mathrm{o}(\delta).$$

Summarizing our computations, we have

$$\begin{aligned}\mathcal{F}(y; t_0, t_3) \leqslant{} & \bigl|h(t_1)\bigr|^p \bigl(w_f(t_1) - w_h(t_1)\bigr) + \bigl|h(t_2)\bigr|^p \bigl(w_h(t_2) - w_g(t_2)\bigr) \\ & - \int_{\xi_1}^{\xi_2} \bigl(c(t) - \tilde{c}(t)\bigr)|h|^p \,\mathrm{d}t - \bigl(K\delta + \mathrm{o}(\delta)\bigr)\end{aligned}$$

with a positive constant K.

Now, let $\delta > 0$ (sufficiently small) be such that $K\delta + \mathrm{o}(\delta) =: \varepsilon > 0$. According to (8.3) the points t_1, t_2 can be chosen in such a way that

$$\int_{s_1}^{s_2} \big(c(t) - \tilde{c}(t)\big)|h|^p \,\mathrm{d}t > -\frac{\varepsilon}{4}$$

whenever $s_1 \in (a, t_1)$, $s_2 \in (t_2, b)$. Further, since w_h is generated by the solution h of (8.2) which is principal both at $t = a$ and $t = b$, according to the Mirzov construction of the principal solution, we have (for t_1, t_2 fixed for a moment)

$$\lim_{t_0 \to a+} \big[w_f(t_1) - w_h(t_1)\big] = 0, \qquad \lim_{t_3 \to b-} \big[w_g(t_2) - w_h(t_2)\big] = 0.$$

Hence

$$\big|h(t_1)\big|^p \big[w_f(t_1) - w_h(t_1)\big] < \frac{\varepsilon}{4}, \qquad \big|h(t_2)\big|^p \big[w_h(t_2) - w_g(t_2)\big] < \frac{\varepsilon}{4}$$

if $t_0 < t_1$, $t_3 > t_2$ are sufficiently close to a and b, respectively.

Consequently, for the above specified choice of $t_0 < t_1 < t_2 < t_3$ we have

$$\begin{aligned}
\mathcal{F}(y; t_0, t_3) &= \int_{t_0}^{t_3} \big[r(t)|y'|^p - \tilde{c}(t)|y|^p\big] \,\mathrm{d}t - \int_{t_0}^{t_3} \big(c(t) - \tilde{c}(t)\big)|y|^p \,\mathrm{d}t \\
&\leqslant \big|h(t_1)\big|^p \big[w_f(t_1) - w_h(t_1)\big] + \big|h(t_2)\big|^p \big[w_h(t_2) - w_g(t_2)\big] \\
&\quad - \int_{\xi_1}^{\xi_2} \big(c(t) - \tilde{c}(t)\big)|h|^p \,\mathrm{d}t - \big(K\delta + \mathrm{o}(\delta)\big) \\
&< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} - \varepsilon < 0.
\end{aligned}$$

The proof is now complete. □

8.3. *Lyapunov inequality*

The classical Lyapunov inequality (see, e.g., [106, Chapter XI]) for the linear differential equation (1.1) states that if a, b, $a < b$, are consecutive zeros of a nontrivial solution of this equation, then

$$\int_a^b c_+(t) \,\mathrm{d}t > \frac{4}{\int_a^b r^{-1}(t) \,\mathrm{d}t}, \qquad c_+(t) = \max\big\{0, c(t)\big\}.$$

This inequality has been extended in many directions and its half-linear extension reads as follows, see [85].

THEOREM 8.3. *Let a, b, $a < b$, be consecutive zeros of a nontrivial solution of* (0.1). *Then*

$$\int_a^b c_+(t)\,\mathrm{d}t > \frac{2^p}{(\int_a^b r^{1-q}(t)\,\mathrm{d}t)^{p-1}}. \tag{8.5}$$

PROOF. According to homogeneity of the solution space of (0.1), we can suppose that $x(t) > 0$ on (a, b). Let $c \in (a, b)$ be the least point of the local maximum of x in (a, b), i.e., $x'(c) = 0$ and $x'(t) > 0$ on $[a, c)$. By the Hölder inequality we have

$$\begin{aligned} x^p(c) &= \left(\int_a^c x'(t)\,\mathrm{d}t\right)^p = \left(\int_a^c r^{-\frac{1}{p}}(t) r^{\frac{1}{p}}(t) x'(t)\,\mathrm{d}t\right)^p \\ &\leqslant \left(\int_a^c r^{-\frac{q}{p}}(t)\,\mathrm{d}t\right)^{\frac{p}{q}} \left(\int_a^b r(t)\big(x'(t)\big)^p\,\mathrm{d}t\right). \end{aligned}$$

Multiplying (0.1) by $x(t)$ and integrating from a to c by parts we get

$$\begin{aligned} \int_a^c r(t)\big(x'(t)\big)^p\,\mathrm{d}t = \int_a^c c(t)x^p(t)\,\mathrm{d}t &\leqslant \int_a^c c_+(t)x^p(t)\,\mathrm{d}t \\ &\leqslant x^p(c)\int_a^c c_+(t)\,\mathrm{d}t, \end{aligned}$$

hence

$$x^p(c) \leqslant \left(\int_a^c r^{1-q}(t)\,\mathrm{d}t\right)^{p-1} \left(\int_a^c c_+(t)\,\mathrm{d}t\right) x^p(c),$$

which yields

$$\left(\int_a^c r^{1-q}(t)\,\mathrm{d}t\right)^{1-p} \leqslant \int_a^c c_+(t)\,\mathrm{d}t.$$

Similarly, if d is the greatest point of local maximum of x in (a, b), i.e., $x'(d) = 0$ and $x'(t) < 0$ on (d, b), we have

$$\left(\int_d^b r^{1-q}(t)\,\mathrm{d}t\right)^{1-p} \leqslant \int_d^b c_+(t)\,\mathrm{d}t.$$

Consequently,

$$\int_a^b c_+(t)\,\mathrm{d}t \geqslant \left(\int_a^c r^{1-q}(t)\,\mathrm{d}t\right)^{1-p} + \left(\int_d^b r^{1-q}(t)\,\mathrm{d}t\right)^{1-p}.$$

Finally, since the function $f(u) = u^{1-p}$ is convex for $u > 0$, the Jensen inequality $f(\frac{u+v}{2}) \leqslant \frac{1}{2}[f(u) + f(v)]$ with $u = \int_a^c r^{1-q}(t)\,\mathrm{d}t$, $v = \int_d^b r^{1-q}(t)\,\mathrm{d}t$ implies

$$\begin{aligned}
&\left(\int_a^c r^{1-q}(t)\,\mathrm{d}t\right)^{1-p} + \left(\int_d^b r^{1-q}(t)\,\mathrm{d}t\right)^{1-p} \\
&\quad \geqslant 2\left[\frac{1}{2}\left(\int_a^c r^{1-q}(t)\,\mathrm{d}t + \int_d^b r^{1-q}(t)\,\mathrm{d}t\right)\right]^{1-p} \\
&\quad \geqslant \frac{2^p}{(\int_a^b r^{1-q}(t)\,\mathrm{d}t)^{p-1}},
\end{aligned}$$

what completes the proof. □

REMARK 8.1. Note that the Lyapunov-type inequality was proved for the first time in [85] for (0.1) with $r(t) \equiv 1$, but the extension to general (0.1) is straightforward. Half-linear Lyapunov inequality has been rediscovered in several later papers, e.g., in [154,224].

8.4. *Vallée Poussin-type inequality*

Another important inequality concerning disconjugacy of the linear differential equation

$$x'' + a(t)x' + b(t)x = 0 \tag{8.6}$$

was introduced by Valleé Poussin [214] in 1929 and reads as follows. Suppose that $t_1 < t_2$ are consecutive zeros of a nontrivial solution x of (8.6), then

$$2\int_0^\infty \frac{dt}{t^2 + At + B} \leqslant t_2 - t_1, \quad A := \max_{t\in[t_1,t_2]} |a(t)|, \; B := \max_{t\in[a,b]} |b(t)|.$$

The half-linear version of this criterion can be found in [68]. Here we formulate this criterion in a simplified form, to underline similarity with the original criterion of Vallée Poussin. For the same reason we consider the equation

$$\left(\Phi(x')\right)' + a(t)\Phi(x') + b(t)\Phi(x) = 0 \tag{8.7}$$

instead of (0.1) (if the function r in (0.1) is differentiable, then this equation can be easily reduced to (8.7)).

THEOREM 8.4. *Suppose that $t_1 < t_2$ are consecutive zeros of a nontrivial solution x of (8.7). Then*

$$2\int_0^\infty \frac{\mathrm{d}t}{(p-1)t^q + At + B} \leqslant t_2 - t_1, \quad A = \max_{t\in[t_1,t_2]} |a(t)|, \; B = \max_{t\in[t_1,t_2]} |b(t)|. \tag{8.8}$$

PROOF. Suppose that $x(t) > 0$ in (t_1, t_2), in case $x(t) < 0$ in (t_1, t_2) the proof is analogical. Let $c, d \in (t_1, t_2)$, $c \leqslant d$, be the least and the greatest points of the local maximum of x in (a, b), respectively, i.e., $x'(t) > 0$ for $t \in (t_1, c)$, $x'(t) < 0$ for $t \in (d, t_2)$ and $x'(c) = 0 = x'(d)$. The Riccati variable $v = \frac{\Phi(x')}{\Phi(x)}$ satisfies $v(t_1+) = \infty$, $v(c) = 0$, $v(t) > 0$, $t \in (t_1, c)$ and

$$v' = -b(t) - a(t)v - (p-1)v^q \geqslant -B - Av - (p-1)v^q. \tag{8.9}$$

Hence,

$$\int_0^\infty \frac{\mathrm{d}v}{(p-1)v^q + Av + B} \leqslant c - t_1. \tag{8.10}$$

Concerning the interval (d, t_2), we set $v = -\frac{\Phi(x')}{\Phi(x)} > 0$ for $t \in (d, t_2)$ and similarly as for $t \in (t_1, c)$ we have

$$\int_0^\infty \frac{\mathrm{d}v}{(p-1)v^q + Av + B} \leqslant t_2 - d. \tag{8.11}$$

The summation of (8.10) and (8.11) gives

$$2\int_0^\infty \frac{\mathrm{d}v}{(p-1)v^q + Av + B} \leqslant c - t_1 + t_2 - d \leqslant t_2 - t_1,$$

what we needed to prove. □

REMARK 8.2. (i) Since (8.8) is a necessary condition for conjugacy of (8.7) in $[t_1, t_2]$, the opposite inequality is a disconjugacy criterion: if

$$2\int_0^\infty \frac{\mathrm{d}v}{(p-1)v^q + Av + B} > t_2 - t_1,$$

then (8.7) is disconjugate in $[t_1, t_2]$.

(ii) A more general Riccati substitution $v = \alpha(t)\frac{\Phi(x')}{\Phi(x)}$, $t \in (t_1, c]$, $v = -\beta(t)\frac{\Phi(x')}{\Phi(x)}$, $t \in [d, t_2)$, where α, β are suitable positive functions, enables to formulate the Vallée Poussin-type criterion in a more general form than presented in Theorem 8.4, we refer to [68] for details. Concerning the various extensions of the linear Vallée Poussin criterion we refer to the survey paper [14] and the references given therein.

8.5. *Focal point criteria*

Recall that a point b is said to be the *first right focal point* of $c < b$ with respect to (0.1) if there exists a nontrivial solution x of this equation such that $x'(c) = 0 = x(b)$ and $x(t) \neq 0$ for $t \in [c, b)$. The first left focal point a of c is defined similarly by $x(a) = 0 = x'(c)$,

$x(t) \neq 0$ on $(a, c]$. Equation (0.1) is said to be *right disfocal* on $[c, b)$ if there exists no right focal point of c relative to (0.1) in (c, b), the *left disfocality* on $(a, c]$ is defined in a similar way. Consequently, (0.1) is conjugate on an interval (a, b) if there exists $c \in (a, b)$ such that this equation is neither right disfocal on $[c, b)$ nor left disfocal on $(a, c]$. This idea is illustrated in the next statement for (6.16) considered on $(a, b) = (-\infty, \infty)$, see [65]. The extension of this statement to general half-linear equation (0.1) is immediate.

THEOREM 8.5. *Suppose that the function* $c(t) \not\equiv 0$ *for* $t \in (0, \infty)$ *and there exist constants* $\alpha \in (-\frac{1}{p}, p-2]$ *and* $T \geqslant 0$ *such that*

$$\int_0^t s^\alpha \left(\int_0^s c(\tau)\,\mathrm{d}\tau \right) \mathrm{d}s \geqslant 0 \quad \textit{for } t \geqslant T. \tag{8.12}$$

Then the solution x *of* (6.16) *satisfying the initial conditions* $x(0) = 1$, $x'(0) \leqslant 0$ *has a zero in* $(0, \infty)$.

PROOF. Suppose, by contradiction, that the solution x has no zero on $(0, \infty)$, i.e., $x(t) > 0$. Let $w = -\frac{\Phi(x')}{\Phi(x)}$ be the solution of the Riccati equation

$$w' = c(t) + (p-1) r^{1-q}(y) |w|^q.$$

Since $w(0) \geqslant 0$, we have

$$\begin{aligned} w(t) &= w(0) + \int_0^t c(s)\,\mathrm{d}s + (p-1) \int_0^t \left|w(s)\right|^q \mathrm{d}s \\ &\geqslant \int_0^t c(s)\,\mathrm{d}s + (p-1) \int_0^t \left|w(s)\right|^q \mathrm{d}s \end{aligned} \tag{8.13}$$

and

$$\int_0^t s^\alpha w(s)\,\mathrm{d}s \geqslant \int_0^t \left(s^\alpha \int_0^s c(\tau)\,\mathrm{d}\tau \right) \mathrm{d}s + G(t),$$

where

$$G(t) = (p-1) \int_0^t s^\alpha \left(\int_0^s \left|w(\tau)\right|^q \mathrm{d}\tau \right) \mathrm{d}s.$$

Then

$$G'(t) = (p-1) t^\alpha \int_0^t \left|w(\tau)\right|^q \mathrm{d}\tau \geqslant 0 \quad \text{for } t \geqslant 0 \tag{8.14}$$

and according to (8.12)

$$G(t) \leqslant \int_0^t s^\alpha w(s)\,\mathrm{d}s \quad \text{for } t \geqslant T. \tag{8.15}$$

By the Hölder inequality we have

$$\begin{aligned}\int_0^t s^\alpha w(s)\,\mathrm{d}s &\leqslant \left[\int_0^t s^{p\alpha}\,\mathrm{d}s\right]^{\frac{1}{p}}\left[\int_0^t \left|w(s)\right|^q\,\mathrm{d}s\right]^{\frac{1}{q}}\\ &= \left[\frac{t^{1+p\alpha}}{1+p\alpha}\right]^{\frac{1}{p}}\left[\int_0^t \left|w(s)\right|^q\,\mathrm{d}s\right]^{\frac{1}{q}},\end{aligned}$$

hence by (8.15)

$$\frac{t^{(1+p\alpha)\frac{q}{p}}}{(1+p\alpha)^{\frac{q}{p}}}\int_0^t \left|w(s)\right|^q\,\mathrm{d}s \geqslant G^q(t).$$

Here we need the relation $G(t) > 0$ for sufficiently large t. By (8.14) $G(t)$ is nondecreasing function of t and $G(0) = 0$. The equality $G(t) = 0$ for all $t \geqslant 0$ would imply that $w(t) \equiv 0$, consequently by (8.14) $x'(t) \equiv 0$ for $t \geqslant 0$. But this may happen only if $c(t) \equiv 0$, which case has been excluded. Hence we may suppose that T is already chosen so large that the inequality $G(t) > 0$ holds for $t \geqslant T$.

Denote $\beta = \alpha - (1 + p\alpha)\frac{q}{p}$ and $K = (p-1)(1+p\alpha)^{\frac{q}{p}} > 0$. Then by (8.14) the last inequality yields $G'G^{-q} \geqslant Kt^\beta$. Integrating this inequality from T to t, we get

$$\frac{1}{q-1}G^{1-q}(T) > \frac{1}{q-1}\left[G^{1-q}(T) - G^{1-q}(t)\right] \geqslant K\int_T^t s^\beta\,\mathrm{d}s,$$

where the integral on the right-hand side tends to ∞ as $t \to \infty$ because an easy computation shows that $\alpha \leqslant p-2$ implies $\beta \geqslant -1$. This contradiction proves that x must have a positive zero. ☐

REMARK 8.3. (i) Clearly, in Theorem 8.5 the starting point $t_0 = 0$ can be shifted to any other value $t_0 \in \mathbb{R}$ if the condition (2.4) is modified to

$$\int_{t_0}^t (s-t_0)^\alpha\left(\int_{t_0}^s c(\tau)\,\mathrm{d}\tau\right)\mathrm{d}s \geqslant 0 \quad \text{for } t \geqslant T \geqslant t_0.$$

A similar statement can be formulated on the interval $(-\infty, t_0)$, too.

(ii) In the previous theorem we have used the weight function s^α, $\alpha \in (-\frac{1}{p}, p-2]$. The results of Section 9.2 of the next chapter suggest to use a more general weight functions. This research is a subject of the present investigation. The same remark essentially concerns also the results of Section 9.5.

Using the just established focal point criterion we can prove the following conjugacy criterion for (6.16).

THEOREM 8.6. *Suppose that* $c(t) \not\equiv 0$ *both in* $(-\infty, 0)$ *and* $(0, \infty)$ *and there exist constants* $\alpha_1, \alpha_2 \in (-\frac{1}{p}, p-2]$ *and* $T_1, T_2 \in \mathbb{R}$, $T_1 < 0 < T_2$, *such that*

$$\int_t^0 |s|^{\alpha_1} \left(\int_s^0 c(\tau)\,\mathrm{d}\tau \right) \mathrm{d}s \geqslant 0, \quad t \leqslant T_1,$$
$$\int_0^t s^{\alpha_2} \left(\int_0^s c(\tau)\,\mathrm{d}\tau \right) \mathrm{d}s \geqslant 0, \quad t \geqslant T_2. \tag{8.16}$$

Then Equation (6.16) *is conjugate in* $\mathbb{R}$, *more precisely, there exists a solution of* (6.16) *having at least one positive and one negative zero.*

PROOF. The statement follows immediately from Theorem 8.5 since by this theorem the solution x given by $x(0) = 1$, $x'(0) = 0$ has a positive zero. Using the same argument as in Theorem 8.5 and the second condition in (8.16) we can show the existence of a negative zero. ☐

REMARK 8.4. (i) Assumptions of the previous theorem are satisfied if

$$\liminf_{s_1 \downarrow -\infty, s_2 \uparrow \infty} \int_{s_1}^{s_2} c(t)\,\mathrm{d}t > 0. \tag{8.17}$$

This conjugacy criterion for the linear Sturm–Liouville equation (1.1) with $r(t) \equiv 1$ is proved in [212] and the extension to (6.16) can be found in [188].

(ii) Several conjugacy criteria for linear equation (1.1) (in terms of its coefficients r, c) are proved using the fact that this equation is conjugate on (a, b) if and only if

$$\int_a^b \frac{dt}{r(t)[x_1^2(t) + x_2^2(t)]} > \pi \tag{8.18}$$

for any pair of solutions of (1.1) for which $r(x_1' x_2 - x_1 x_2') \equiv \pm 1$. This statement is based on the trigonometric transformation of (1.1), see [55,56] and also Section 3. However, since we have in disposal no half-linear analogue of the trigonometric transformation, conjugacy criteria of this kind for (0.1) are (till now) missing.

8.6. *Lyapunov-type focal points and conjugacy criteria*

The results of this section can be found in [111,188] and concern again equation (6.16).

THEOREM 8.7. *Let* x *be a nontrivial solution of* (6.16) *satisfying* $x'(d) = 0 = x(b)$ *and* $x(t) \neq 0$ *for* $t \in [d, b)$. *Then*

$$(b-d)^{p-1} \sup_{d \leqslant t \leqslant b} \left| \int_d^t c(s)\,\mathrm{d}s \right| > 1. \tag{8.19}$$

Moreover, if there is no extreme value of x in (d, b), then

$$(b-d)^{p-1} \sup_{d \leqslant t \leqslant b} \int_d^t c(s)\,\mathrm{d}s > 1. \tag{8.20}$$

PROOF. Suppose that $x(t) > 0$ on $[d, b)$, if $x(t) < 0$ we proceed in the same way. Let $v = -\frac{\Phi(x')}{\Phi(x)}$ and $V(t) = (p-1)\int_d^t |v(s)|^q\,\mathrm{d}s$. Then we have

$$v(t) = \int_d^t c(s)\,\mathrm{d}s + V(t). \tag{8.21}$$

Thus, $v(d) = 0 = V(d)$ and $\lim_{t\to b-} v(t) = \lim_{t\to b-} V(t) = \infty$. Set

$$C^* := \sup_{d \leqslant t \leqslant b} \left| \int_d^t c(s)\,\mathrm{d}s \right|$$

and observe that $|v(t)| \leqslant C^* + V(t)$, so that

$$V'(t) = (p-1)\bigl|v(t)\bigr|^q \leqslant (p-1)\bigl(C^* + V(t)\bigr)^q,$$

and

$$\frac{V'(t)}{(p-1)(C^* + V(t))^q} \leqslant 1.$$

Integrating this inequality from d to b and using $\lim_{t\to b-} V(t) = \infty$, we obtain

$$-\frac{1}{(C^* + V(t))^{q-1}}\bigg|_d^b \leqslant b - d,$$

which implies that $(b-d)^{p-1}C^* \leqslant 1$. We remark that the equality cannot hold, for otherwise $|C(t)| = |\int_d^t c(s)\,\mathrm{d}s| = C^*$ on $[d, b)$ which implies that $c(t) \equiv 0$, a contradiction, thus (8.19) holds.

If d is the largest extreme point of x in (a, b), then $x'(t) \leqslant 0$ and hence $v(t) \geqslant 0$ on $[d, b)$. Set $C_* = \sup_{d\leqslant t\leqslant b} \int_d^t c(s)\,\mathrm{d}s$. Then we also have $C_* > 0$ since the assumption $C_* \leqslant 0$ contradicts to $V(d) = 0$, $\lim_{t\to b-} V(t) = \infty$. Hence, by (8.21), $0 \leqslant v(t) \leqslant C_* + V(t)$. The remaining part of the proof is similar to the first one. □

REMARK 8.5. (i) Similarly as above, if $a < c$, $x(a) = 0 = x'(c)$ and $x(t) \neq 0$ on $(a, c]$, we have

$$(c-a)^{p-1} \sup_{a \leqslant t \leqslant c} \left| \int_t^c c(s)\,\mathrm{d}s \right| > 1.$$

(ii) Combining Theorem 8.7 with the previous part of the remark one can prove that if $a < b$ are consecutive zeros of a nontrivial solution x of (6.16), then there exist two disjoint subintervals $I_1, I_2 \subset [a, b]$ such that

$$(b-a)^{p-1} \int_{I_1 \cup I_2} c(s)\,\mathrm{d}s \geqslant \min\big\{4, 4^{p-1}\big\}, \qquad \int_{[a,b]\setminus(I_1\cup I_2)} c(s)\,\mathrm{d}s \leqslant 0.$$

For details and related (dis)conjugacy criteria we refer to [111].

CHAPTER 3C

Oscillatory Equations

In this chapter we start with oscillation criteria for half-linear equation (0.1). Some basic oscillation criteria have already been formulated in previous chapters, here we focus our attention to “more advanced criteria”. By this we mean half-linear extensions of linear oscillation criteria which are not given in standard books on linear oscillation, e.g., in [208]. Section 10 is devoted to results which are not exactly oscillation criteria, but are closely related to oscillation theory. Section 11 deals with the half-linear Sturm–Liouville problem and the last section of this chapter, entitled “Perturbation principle”, presents a new method of the investigation of oscillatory properties of (0.1), where this equation is viewed as a perturbation of another equation of the same form.

9. Oscillation criteria

Similarly to nonoscillation criteria, the basic tools in proofs of oscillation criteria are the variational principle and the Riccati technique. However, as we will see in this and the next section, the Riccati technique is used more frequently.

9.1. *General observations*

In this section we present criteria which complete nonoscillation criteria given in Section 5. We start with some general observations.

(i) If $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty$ in (0.1), then this equation can be transformed into the equation

$$\bigl(\Phi(x')\bigr)' + c(t)\Phi(x) = 0 \tag{9.1}$$

and this transformation transforms the interval $[T, \infty)$ into an interval of the same form. For this reason, we will formulate sometimes our results for (9.1) (mainly in situations when these results were first established for (1.1) with $r(t) \equiv 1$ in linear case), the extension to (0.1) with $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty$ is then straightforward.

(ii) If we suppose that $c(t) > 0$ for large t in (9.1), the situation is considerably simpler than in the general case when c is allowed to change its sign. In case $c(t) > 0$, Equation (9.1) is the Sturmian majorant of the equation $(\Phi(x'))' = 0$ and the minimal solution of the associated Riccati equation

$$v' + (p-1)|v|^q = 0$$

is $v(t) \equiv 0$. This means, in view of Theorem 7.2, that the minimal solution of the Riccati equation

$$w' + c(t) + (p-1)|w|^q = 0 \tag{9.2}$$

associated with (9.1) is positive and this implies that positive solutions of (9.1) are increasing. Consequently, in oscillation criteria for (9.1) proved via the Riccati technique, it is sufficient to impose such conditions on the function c that for any $T \in \mathbb{R}$ the solution given by $x(T) = 0$, $x'(T) > 0$ has eventually negative derivative $x'(t)$. Indeed, this means that the associated solution of (9.2) is eventually negative, hence less than minimal solution which is nonnegative (compare Theorem 7.2 and Corollary 7.1) and hence cannot be extensible up to ∞ which implies that x has a zero point in $[T, \infty)$.

9.2. *Coles-type criteria*

The results of this subsection are taken from the paper [153] and concern the half-linear extension of the averaging technique introduced in the linear case by Coles in [46]. The results are formulated for (9.1).

Let $\mathcal{J}$ be the class of nonnegative locally integrable functions f defined on $[0, \infty)$ and satisfying the condition

$$\limsup_{t\to\infty} \left(\int_0^t f(s)\,\mathrm{d}s \right)^{q-1-\mu} \big[F_\mu(\infty) - F_\mu(t) \big] > 0 \tag{9.3}$$

for some $\mu \in [0, q-1)$, where

$$F_\mu(t) = \int_0^t f(s) \frac{(\int_0^s f(\xi)\,\mathrm{d}\xi)^\mu}{(\int_0^s f^p(\xi)\,\mathrm{d}\xi)^{q-1}}\,\mathrm{d}s.$$

If $F_\mu(\infty) = \infty$, then $f \in \mathcal{J}$. Let $\mathcal{J}_0$ be the subclass of $\mathcal{J}$ consisting of nonnegative locally integrable functions f satisfying

$$\lim_{t\to\infty} \frac{\int_0^t f^p(s)\,\mathrm{d}s}{(\int_0^t f(s)\,\mathrm{d}s)^p} = 0. \tag{9.4}$$

Observe that if (9.3) or (9.4) holds, then

$$\int^\infty f(t)\,\mathrm{d}t = \infty. \tag{9.5}$$

On the other hand, every bounded nonnegative locally integrable function satisfying (9.5) belongs to $\mathcal{J}_0$ and $\mathcal{J}_0 \subset \mathcal{J}$. Since all nonnegative polynomials are in $\mathcal{J}_0$, this class of functions contains also unbounded functions. Elements of $\mathcal{J}$ and $\mathcal{J}_0$ will be called *weight functions*.

For $f \in \mathcal{J}$, we define

$$A_f(s,t) := \frac{\int_s^t f(\tau) \int_s^\tau c(\mu)\,\mathrm{d}\mu\,\mathrm{d}\tau}{\int_s^t f(\tau)\,\mathrm{d}\tau}.$$

The following statement reduces to the Hartman–Wintner theorem (Theorem 5.6) when the weight function f is $f(t) \equiv 1$.

THEOREM 9.1. *Suppose that* (9.1) *is nonoscillatory.*

(i) *If there exists* $f \in \mathcal{J}$ *such that for some* $T \in \mathbb{R}$

$$\liminf_{t\to\infty} A_f(T,t) > -\infty \tag{9.6}$$

then

$$\int^\infty \left|w(t)\right|^q \mathrm{d}t < \infty \tag{9.7}$$

for every solution w *of the associated Riccati equation* (9.2).

(ii) *Assume that* (9.7) *holds for some solution* w *of* (9.2). *Then for every* $f \in \mathcal{J}_0$ *and* $T \in \mathbb{R}$ *sufficiently large* $\lim_{t\to\infty} A_f(T,t)$ *exists finite.*

PROOF. The proof of this statement copies essentially the proof of Theorem 5.6.

(i) Assume, by contradiction, that

$$\int^\infty \left|w(t)\right|^q \mathrm{d}t = \infty \tag{9.8}$$

for some solution of (9.2). Integrating this equation from ξ to t, multiplying the obtained integral equation by $f(t)$ and then integrating again from ξ to t, we obtain

$$\begin{aligned}
\int_\xi^t f(s)w(s)\,\mathrm{d}s &= w(\xi)\int_\xi^t f(s)\,\mathrm{d}s - \int_\xi^t f(s)\int_\xi^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s \\
&\quad - (p-1)\int_\xi^t f(s)\int_\xi^s \left|w(\tau)\right|^q \mathrm{d}\tau\,\mathrm{d}s \\
&= w(\xi)\int_\xi^t f(s)\,\mathrm{d}s - A_f(\xi,t)\int_\xi^t f(s)\,\mathrm{d}s \\
&\quad - (p-1)\int_\xi^t f(s)\int_\xi^s \left|w(\tau)\right|^q \mathrm{d}\tau\,\mathrm{d}s \\
&= \left[w(\xi) - A_f(\xi,t)\right]\int_\xi^t f(s)\,\mathrm{d}s \\
&\quad - (p-1)\int_\xi^t f(s)\int_\xi^s \left|w(\tau)\right|^q \mathrm{d}\tau\,\mathrm{d}s,
\end{aligned}$$

where $t \geqslant \xi \geqslant T$. From (9.2) we have

$$w(\xi) = w(T) - \int_T^\xi c(s)\,\mathrm{d}s - (p-1)\int_T^\xi \left|w(s)\right|^q \mathrm{d}s.$$

Since $f \in \mathcal{J}$, (9.5) holds. This implies

$$\begin{aligned} A_f(\xi,t) &= \frac{\int_T^t f(s)\,\mathrm{d}s}{\int_\xi^t f(s)\,\mathrm{d}s} A_f(T,t) - \int_T^\xi c(s)\,\mathrm{d}s - \frac{\int_T^\xi f(s)\int_T^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s}{\int_\xi^t f(s)\,\mathrm{d}s} \\ &= \frac{\int_T^t f(s)\,\mathrm{d}s}{\int_\xi^t f(s)\,\mathrm{d}s} A_f(T,t) - \int_T^\xi c(s)\,\mathrm{d}s + \mathrm{o}(1) \quad \text{as } t \to \infty. \end{aligned}$$

Thus

$$\begin{aligned} w(\xi) - A_f(\xi,t) = w(T) &- \frac{\int_T^t f(s)\,\mathrm{d}s}{\int_\xi^t f(s)\,\mathrm{d}s} A_f(T,t) \\ &- (p-1)\int_T^t \left|w(s)\right|^q \mathrm{d}s + \mathrm{o}(1) \end{aligned} \tag{9.9}$$

as $t \to \infty$. Since $f \in \mathcal{J}$, there exists a positive number $\lambda > 0$ such that

$$\frac{\lambda^{1-q}}{p-1} < (q-1-\mu)\limsup_{t\to\infty}\left[\int^t f(s)\,\mathrm{d}s\right]^{q-1-\mu}\left[F_\mu(\infty) - F_\mu(t)\right], \tag{9.10}$$

where μ is the same as in (9.3). It follows from (9.6), (9.8) and the previous computation that there exist two numbers a and b, $b \geqslant a \geqslant T$, such that

$$w(a) - A_f(a,t) \leqslant -\lambda \tag{9.11}$$

for $t \geqslant b$. Let $z(t) := \int_a^t f(s)w(s)\,\mathrm{d}s$. Then the Hölder inequality implies

$$\int_a^t \left|w(\tau)\right|^q \mathrm{d}\tau \geqslant \frac{|z(t)|^q}{(\int_a^t f^p(\tau)\,\mathrm{d}\tau)^{q-1}}.$$

It follows from (9.9) and (9.11) that

$$\begin{aligned} z(t) &\leqslant -\lambda\int_a^t f(s)\,\mathrm{d}s - (p-1)\int_a^t f(s)\left|z(s)\right|^q\left(\int_a^s f^p(\tau)\,\mathrm{d}\tau\right)^{1-q}\mathrm{d}s \\ &=: -G(t). \end{aligned} \tag{9.12}$$

Thus

$$G'(t) = \lambda f(t) + (p-1)f(t)\left|z(t)\right|^q\left(\int_a^t f^p(s)\,\mathrm{d}s\right)^{1-q} \tag{9.13}$$

and

$$0 \leqslant \lambda \int_a^t f(s)\,\mathrm{d}s \leqslant G(t) \leqslant \bigl|z(t)\bigr|. \tag{9.14}$$

It follows from (9.12), (9.13) and (9.14) that

$$\begin{aligned} G'(t)G^{\mu-q}(t) &\geqslant G'(t)G^{\mu}(t)\bigl|z(t)\bigr|^{-q} \\ &\geqslant (p-1)\lambda^{\mu} f(t)\biggl(\int_a^t f(s)\,\mathrm{d}s\biggr)^{\mu}\biggl(\int_a^t f^p(s)\,\mathrm{d}s\biggr)^{1-q}. \end{aligned}$$

Integrating this inequality from t ($t \geqslant b$) to ∞, we get

$$\frac{1}{q-1-\mu} G^{\mu-q-1}(t) \geqslant (p-1)\lambda^{\mu}\bigl[F_{\mu}(\infty) - F_{\mu}(t)\bigr].$$

Inequality (9.14) then implies

$$\frac{\lambda^{1-q}}{p-1} \geqslant (q-1-\mu)\biggl[\int_a^t f(s)\,\mathrm{d}s\biggr]^{q-1-\mu}\bigl[F_{\mu}(\infty) - F_{\mu}(t)\bigr]$$

which contradicts (9.10).

(ii) As in the previous part of the proof, (9.9) holds. This implies that

$$\begin{aligned} A_f(\xi,t) = w(\xi) &- \frac{\int_{\xi}^t f(s)w(s)\,\mathrm{d}s}{\int_{\xi}^t f(s)\,\mathrm{d}s} \\ &- (p-1)\frac{\int_{\xi}^t f(s)\int_{\xi}^s |w(\tau)|^q\,\mathrm{d}\tau\,\mathrm{d}s}{\int_{\xi}^t f(s)\,\mathrm{d}s}. \end{aligned} \tag{9.15}$$

Since $f \in \mathcal{J}_0$, (9.5) holds. Thus,

$$\lim_{t\to\infty} \frac{\int_{\xi}^t f(s)\int_{\xi}^s |w(\tau)|^q\,\mathrm{d}\tau\,\mathrm{d}s}{\int_{\xi}^t f(s)\,\mathrm{d}s} = \int_{\xi}^{\infty} \bigl|w(s)\bigr|^q\,\mathrm{d}s < \infty.$$

By Hölder's inequality

$$0 \leqslant \lim_{t\to\infty} \frac{|\int_{\xi}^t f(s)w(s)\,\mathrm{d}s|}{\int_{\xi}^t f(s)\,\mathrm{d}s} \leqslant \lim_{t\to\infty} \frac{(\int_{\xi}^t f^p(s)\,\mathrm{d}s)^{1/p}(\int_{\xi}^t |w(s)|^q\,\mathrm{d}s)^{1/q}}{\int_{\xi}^t f(s)\,\mathrm{d}s} = 0.$$

Hence, by (9.15), $\lim_{t\to\infty} A_f(\xi,t)$ exists and

$$\lim_{t\to\infty} A_f(\xi,t) = w(\xi) - (p-1)\int_{\xi}^{\infty} \bigl|w(s)\bigr|^q\,\mathrm{d}s.$$

This completes the proof. ☐

As a consequence of the previous statement we have the following oscillation criterion which is the half-linear extension of the criterion of Coles [46], this statement can be also viewed as an extension of Theorem 5.7.

THEOREM 9.2. *The following statements hold*:

(i) *If there exists* $f \in \mathcal{J}$ *such that* (9.6) *holds, then either* (9.1) *is oscillatory, or* $\lim_{t\to\infty} A_g(\cdot, t)$ *exists finite for every* $g \in \mathcal{J}_0$.

(ii) *If there exist two nonnegative bounded functions* f, g *on an interval* $[T, \infty)$ *satisfying* $\int^\infty f(t)\,\mathrm{d}t = \infty = \int^\infty g(t)\,\mathrm{d}t$ *such that*

$$\lim_{t\to\infty} A_f(T,t) < \lim_{t\to\infty} A_g(T,t),$$

then Equation (9.1) *is oscillatory.*

PROOF. (i) Suppose that (9.1) is nonoscillatory. Then by Theorem 9.1 every solution of the associated Riccati equation (9.2) satisfies $\int^\infty |w(t)|^q\,\mathrm{d}t < \infty$ and hence $\lim_{t\to\infty} A_g(\cdot, t)$ exists finite for every $g \in \mathcal{J}_0$.

(ii) Let $\alpha, \beta \in \mathbb{R}$ be such that

$$\lim_{t\to\infty} A_f(T,t) < \alpha < \beta < \lim_{t\to\infty} A_g(T,t).$$

Let $h(t) = g(t)$ for $T \leqslant t \leqslant t_1$, where t_1 is determined such that $A_g(T, t_1) \geqslant \beta$ and $\int_T^{t_1} g(s)\,\mathrm{d}s \geqslant 1$. Let $h(t) = f(t)$ for $t_1 \leqslant t_2$ where t_2 is determined such that $A_h(T, t_2) \leqslant \alpha$ and $\int_T^{t_1} h(s)\,\mathrm{d}s \geqslant 2$. This is possible because

$$\begin{aligned} A_h(T,t_2) &= \frac{\int_T^{t_2} h(s) \int_T^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s}{\int_T^{t_2} h(s)\,\mathrm{d}s} \\ &= \frac{\int_T^{t_1} [g(s) - f(s)] \int_T^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s}{\int_T^{t_1} g(s)\,\mathrm{d}s + \int_{t_1}^{t_2} f(s)\,\mathrm{d}s} \\ &\quad + \frac{\int_T^{t_2} f(s) \int_T^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s}{\int_T^{t_2} f(s)\,\mathrm{d}s} \, \frac{\int_T^{t_2} f(s)\,\mathrm{d}s}{\int_T^{t_1} g(s)\,\mathrm{d}s + \int_{t_1}^{t_2} f(s)\,\mathrm{d}s} \\ &= A_f(T,t_2)\big[1 + o(1)\big] + o(1), \end{aligned}$$

as $t_2 \to \infty$. Continuing in this manner, we obtain a nonnegative and bounded function h defined on $[T, \infty)$ such that

$$\limsup_{t\to\infty} A_h(T,t) \geqslant \beta > \alpha \geqslant \liminf_{t\to\infty} A_h(T,t).$$

Hence, by the part (i), Equation (9.1) is oscillatory. ☐

9.3. *Generalized Hartman–Wintner's criterion*

In Section 5 we have shown that (0.1) with $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty$ is oscillatory provided

$$-\infty < \liminf_{t\to\infty} \frac{\int_T^t r^{1-q}(s) \int_T^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s} < \limsup_{t\to\infty} \frac{\int_T^t r^{1-q}(s) \int_T^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s}$$

or

$$\lim_{t\to\infty} \frac{\int_T^t r^{1-q}(s) \int_T^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s}{\int_T^t r^{1-q}(s)\,\mathrm{d}s} = \infty.$$

In this subsection we present one extension of this criterion which was established in [51]. In agreement with the remark from the introduction of this section, we formulate this criterion for (9.1). For the sake of convenience, similarly as in the original paper [51], we introduce the linear operator $A : C[0,\infty) \to C[0,\infty)$ defined by

$$(Af)(t) := \frac{1}{t} \int_0^t f(s)\,\mathrm{d}s, \quad (Af)(0) := 0. \tag{9.16}$$

By A^n we denote the nth iteration of A.

THEOREM 9.3. *Let $C(t) := \int_0^t c(s)\,\mathrm{d}s$. If there exists $n \in \mathbb{N}$ such that*

$$-\infty < \liminf_{t\to\infty} \bigl(A^n C\bigr)(t) < \limsup_{t\to\infty} \bigl(A^n C\bigr)(t), \tag{9.17}$$

or

$$\lim_{t\to\infty} \bigl(A^n C\bigr)(t) = \infty \tag{9.18}$$

then (9.1) *is oscillatory.*

PROOF. The proof is similar to that of Theorem 5.6 and of Theorem 5.7. Suppose, by contradiction, that (9.1) is nonoscillatory and let w be a solution of the associated Riccati equation (9.2). For convenience, we suppose that this solution is defined on $[0,\infty)$, this is no loss of generality since the lower integration limit 0 in the next computation can be replaced by any T sufficiently large. Integrating Equation (9.2) from 0 to t we get

$$w(t) - w(0) + C(t) + (p-1) \int_0^t \bigl|w(s)\bigr|^q\,\mathrm{d}s = 0. \tag{9.19}$$

The application of the operator A^n to the previous equation yields

$$\bigl(A^n w\bigr)(t) + \bigl(A^n C\bigr)(t) + (p-1) A^n\biggl(\int_0^t \bigl|w(s)\bigr|^q\,\mathrm{d}s\biggr) - w(0) = 0. \tag{9.20}$$

Each of the conditions (i), (ii) implies the existence of $K \geqslant 0$ such that $(A^n C)(t) \geqslant -K$ for large t. This implies that $\int^\infty |w(t)|^q \, \mathrm{d}t < \infty$. The proof of this claim goes by contradiction in the same way as in the proof of Theorem 5.6. Having proved the convergence of $\int^\infty |w(t)|^q \, \mathrm{d}t$, again in the same way as in the proof of Theorem 5.6 we prove that $\lim_{t\to\infty}(A^n C)(t)$ exists finite, a contradiction with (i) or (ii). □

The following example, taken again from [51], presents a construction of the function c for which the classical Hartman–Wintner criterion (i.e., the case $n = 1$ in the previous theorem) does not apply, while the previous theorem with $n = 2$ does.

EXAMPLE 9.1. Let $\{a_n\}_{n=1}^\infty$, $\{b_n\}_{n=1}^\infty$ be two sequences of real numbers defined by $a_n = n - 2^{-n}$, $b_n = n + 2^{-n}$, $n \in \mathbb{N}$. Let $\{g_n\}_{n=1}^\infty$ denote a sequence of functions $g_n : [0,\infty) \to \mathbb{R}$ of the class C^2 such that $g_n(t) > 0$ if $t \in (a_n, b_n)$ and $g_n(t) = 0$ otherwise. We also ask that

$$\int_0^\infty g_n(t) \, \mathrm{d}t = n. \tag{9.21}$$

Next define $g : [0,\infty) \to \mathbb{R}$ by

$$g(t) = \sum_{n=1}^\infty (-1)^n g_n(t). \tag{9.22}$$

The function g is also of the class C^2 and using this function we define

$$c(t) = \big(tg(t)\big)'', \qquad C(t) := \int_0^t c(s) \, \mathrm{d}s. \tag{9.23}$$

The reason for defining c in this form is two-fold. On one hand, from (9.23) we find that

$$(AC)(t) = \frac{1}{t} \int_0^t \int_0^s c(\tau) \, \mathrm{d}\tau = g(t), \tag{9.24}$$

and from (9.21) and the mean value theorem

$$\max_{t\in[0,\infty)} g_n(t) \geqslant n 2^{n-1}. \tag{9.25}$$

Thus, from (9.24) and (9.25) we obtain $\liminf_{t\to\infty}(AC)(t) = -\infty$, so the Hartman–Wintner theorem does not apply. Let us consider $p = 2$ for a moment (so we actually consider the linear equation) in (9.1) and we note that the Lebesgue measure of the set $\{t \mid \int_0^t c(s) \, \mathrm{d}s \neq 0\}$ is finite. This implies that

$$\lim_{t\to\infty} \text{approx} \int_0^t c(s) \, \mathrm{d}s = 0,$$

so the criterion of Olech, Opial and Wazewski [185] does not apply either. Recall that

$$\lim_{t\to\infty} \operatorname{approx} f(t) = l$$

if and only if, by definition,

$$l = \sup\{\ell\colon\ m\{t\colon\ f(t) > \ell\} = \infty\} = \inf\{\ell\colon\ m\{t\colon\ f(t) < \ell\} = \infty\}.$$

Here $m\{\cdot\}$ denoted the Lebesgue measure of the set indicated. Recall also that the oscillation criterion of [185] states that the equation $x'' + c(t)x = 0$ is oscillatory provided

$$\lim_{t\to\infty} \operatorname{approx} \int_0^t c(s)\,\mathrm{d}s = \infty.$$

Nevertheless, by Theorem 9.3, Equation (9.1) with any $p > 1$ is oscillatory. Indeed, from (9.25) we have $\limsup_{t\to\infty}(AC)(t) = \infty$. Also, for $t \in (b_n, a_{n+1}) = (n + 2^{-n}, n + 1 - 2^{-(n+1)})$, we have

$$\left(A^2C\right)(t) = \begin{cases} \frac{n}{2t}, & n \text{ even}, \\ \frac{-(n+1)}{2t}, & n \text{ odd}. \end{cases}$$

From the last equality can be easily shown that $(A^2C)(t)$ is bounded for $t \in (0, \infty)$. Hence from (9.17) Equation (9.1) is oscillatory.

9.4. *Generalized Kamenev criterion*

The classical Kamenev criterion concerns the linear equation $x'' + c(t)x = 0$ and claims that this equation is oscillatory provided there exists $\lambda > 1$ such that

$$\limsup_{t\to\infty} \frac{1}{t^\lambda} \int_0^t (t-s)^\lambda c(s)\,\mathrm{d}s = \infty. \tag{9.26}$$

The following half-linear extension concerns general half-linear equation (0.1) and it is taken from [150].

THEOREM 9.4. *Suppose that there exists* $\lambda > p - 1$ *such that*

$$\limsup_{t\to\infty} \frac{1}{t^\lambda} \int_0^t (t-s)^{\lambda-p}\left[(t-s)^p c(s) - \left(\frac{\lambda}{p}\right)^p r(s)\right] \mathrm{d}s = \infty. \tag{9.27}$$

Then (0.1) *is oscillatory.*

PROOF. Suppose that (0.1) is nonoscillatory, i.e., there exists a solution of the associated Riccati equation (2.1). Multiplying this equation by $(t-s)^{\lambda}$ and integrating it from T to t, T sufficiently large, we get

$$\begin{aligned}&-(t-T)^{\lambda}w(T)+\lambda\int_T^t (t-s)^{\lambda-1}w(s)\,\mathrm{d}s\\&\quad+(p-1)\int_T^t (t-s)^{\lambda}r^{1-q}(s)\big|w(s)\big|^q\,\mathrm{d}s+\int_T^t (t-s)^{\lambda}c(s)\,\mathrm{d}s=0.\end{aligned}\tag{9.28}$$

Using the Young inequality (2.6) with

$$u=\frac{\lambda}{p}(t-s)^{\frac{\lambda-p}{p}}r^{\frac{1}{p}},\qquad v=(t-s)^{\frac{\lambda}{q}}r^{\frac{1-q}{q}}\big|w(s)\big|,$$

we obtain

$$\begin{aligned}(t-s)^{\lambda}r^{1-q}(s)\big|w(s)\big|^q&\geqslant(q-1)\lambda(t-s)^{\lambda-1}\big|w(s)\big|\\&\quad-(q-1)\left(\frac{\lambda}{p}\right)^p(t-s)^{\lambda-p}r(s)\end{aligned}$$

for $T\leqslant s\leqslant t$. This inequality and (9.28) imply

$$\int_T^t (t-s)^{\lambda-p}\left[(t-s)^p c(s)-\left(\frac{\lambda}{p}\right)^p r(s)\right]\mathrm{d}s\leqslant(t-T)^{\lambda}w(T).$$

Thus

$$\limsup_{t\to\infty}\frac{1}{t^{\lambda}}\int_T^t (t-s)^{\lambda-p}\left[(t-s)^p c(s)-\left(\frac{\lambda}{p}\right)^p r(s)\right]\mathrm{d}s\leqslant w(T).$$

which contradicts to (9.27). □

REMARK 9.1. Clearly, if $r(t)\equiv1$, then $\lim_{t\to\infty}\frac{1}{t^{\lambda}}\int_T^t(t-s)^{\lambda-p}(\frac{\lambda}{p})^p r(s)\,\mathrm{d}s=0$ and hence Equation (9.1) is oscillatory if

$$\limsup_{t\to\infty}\frac{1}{t^{\lambda}}\int_0^t (t-s)^{\lambda-p}c(s)\,\mathrm{d}s=\infty\quad\text{for some }\lambda>p-1.$$

which is the half-linear extension of the classical Kamenev linear oscillation criterion.

9.5. *Another refinement of the Hartman–Wintner theorem*

The results formulated here are taken from [122], where the equation

$$x''+c(t)|x|^{p-1}|x'|^{2-p}\operatorname{sgn}x=0\tag{9.29}$$

with $p \in (1,2]$ is considered. Related results on Equation (9.29) can be found also in [42, 162]. Equation of the form (9.29) can be obtained from (9.1) using the identity

$$\big(\Phi(x')\big)' = (p-1)x''|x'|^{p-2},$$

so the results of this subsection apply also to (9.1) as noted in [122, Sec. 2]. We prefer here the original formulation to illustrate the variety of approaches to half-linear oscillation theory.

Denote

$$c_p(t) = \frac{(p-1)^2}{t^{p-1}} \int_1^t s^{p-2} \int_1^s c(\tau)\,\mathrm{d}\tau\,\mathrm{d}s.$$

Using essentially the same idea as in the proofs of Theorems 5.6 and 5.7 the following statement is proved.

THEOREM 9.5. *Let either* $\lim_{t\to\infty} c_p(t) = \infty$ *or*

$$-\infty < \liminf_{t\to\infty} c_p(t) < \limsup_{t\to\infty} c_p(t).$$

Then Equation (9.29) *is oscillatory.*

Consequently, in the next investigation we suppose

$$\lim_{t\to\infty} c_p(t) =: c_p(\infty) \tag{9.30}$$

exists finite. The following theorem shows that (9.1) is oscillatory if $c_p(t)$ does not tend to its limit too rapidly.

THEOREM 9.6. *Suppose that* (9.30) *holds and*

$$\limsup_{t\to\infty} \frac{t^{p-1}}{\lg t}\big(c_p(\infty) - c_p(t)\big) > \left(\frac{p-1}{p}\right)^p. \tag{9.31}$$

Then Equation (9.29) *is oscillatory.*

PROOF. Suppose, by contradiction, that (9.29) is nonoscillatory and $w = \Phi(x'/x)$ is a solution of the associated Riccati equation

$$w' + (p-1)c(t) + (p-1)|w|^q = 0.$$

Integrating this equation from t to ∞ we have a variant of the Riccati integral equation

$$w(t) = c_p(\infty) - (p-1)\int_1^t c(s)\,\mathrm{d}s + (p-1)\int_t^\infty \big|w(s)\big|^q\,\mathrm{d}s.$$

Multiplying both sides of this equation by t^{p-2} and integrating from T to t, T sufficiently large, we obtain (integrating one of the terms by parts)

$$\begin{aligned}\int_T^t s^{p-2}\left[c_p(\infty)-(p-1)\int_1^s c(\tau)\,\mathrm{d}\tau\right]\mathrm{d}s &\\ =\int_T^t \frac{s^{p-1}w(s)-s^{p-1}|w(s)|^q}{s}\,\mathrm{d}s-\ t^{p-1}\int_t^\infty \left|w(s)\right|^q\mathrm{d}s &\\ +T^{p-1}\int_T^\infty \left|w(s)\right|^q\mathrm{d}s. &\qquad (9.32)\end{aligned}$$

Since we have the inequality $|\lambda|^q-\lambda+\frac{1}{p}(\frac{p-1}{p})^{p-1}\geqslant 0$ for every $\lambda\in\mathbb{R}$, (9.32) implies

$$\begin{aligned}t^{p-1}\big(c_p(\infty)-c_p(t)\big)\leqslant\left(\frac{p-1}{p}\right)^p\lg\frac{t}{T}+T^{p-1}c_p(\infty)&\\ +(p-1)T^{p-1}\int_T^\infty\left|w(s)\right|^s\mathrm{d}s-T^{p-1}c_p(T).&\end{aligned}$$

Therefore

$$\limsup_{t\to\infty}\frac{t^{p-1}}{\lg t}\big(c_p(\infty)-c_p(t)\big)\leqslant\left(\frac{p-1}{p}\right)^p,$$

which contradicts (9.31). ☐

REMARK 9.2. Here we have presented only one of several (non)oscillation criteria proved in [122]. These criteria are formulated in term of the limit behavior (as $t\to\infty$) of the functions

$$Q_p(t):=t^{p-1}\left(c_p(\infty)-(p-1)\int_1^t c(s)\,\mathrm{d}s\right),$$

$$H_p(t):=\frac{p-1}{t}\int_1^t s^p c(s)\,\mathrm{d}s.$$

We refer to the above mentioned paper [122] for details.

9.6. *Half-linear Willet's criteria*

Results of this subsection extend the linear oscillation and nonoscillation criteria of Willet [217] and are presented in [153].

LEMMA 9.1. *Suppose that $B(s)$ and $Q(t,s)$ are nonnegative continuous functions on $[T,\infty)$ and $[T,\infty)\times[T,\infty)$, respectively.*

(i) *If*

$$\int_t^\infty Q(t,s)B^q(s)\,\mathrm{d}s \leqslant p^{-q}B(t), \quad t \geqslant T, \tag{9.33}$$

then the equation

$$v(t) = B(t) + (p-1)\int_t^\infty Q(t,s)\big|v(s)\big|^q\,\mathrm{d}s, \tag{9.34}$$

has a solution on $[T,\infty)$.

(ii) *If there exists* $\varepsilon > 0$ *such that*

$$\int_t^\infty Q(s,t)B^q(s)\,\mathrm{d}s \geqslant p^{-q}(1+\varepsilon)B(t) \not\equiv 0, \quad t \geqslant T, \tag{9.35}$$

then the inequality

$$v(t) \geqslant B(t) + (p-1)\int_t^\infty Q(t,s)\big|v(s)\big|^q\,\mathrm{d}s, \tag{9.36}$$

possesses no solution on $[T,\infty)$.

PROOF. (i) Let $v_1(t) = B(t)$ and define

$$v_{k+1}(t) = B(t) + (p-1)\int_t^\infty Q(t,s)\big|v_k(s)\big|^q\,\mathrm{d}s, \quad k \in \mathbb{N}.$$

Then by (9.33)

$$\begin{aligned} v_2(t) &= B(t) + (p-1)\int_t^\infty Q(t,s)B^q(s)\,\mathrm{d}s \\ &\leqslant B(t) + (p-1)p^{-q}B(t) \leqslant pB(t), \end{aligned}$$

and $v_1(t) \leqslant v_2(t)$. Suppose, by induction, that $v_1(t) \leqslant v_2(t) \leqslant \cdots \leqslant v_n(t) \leqslant pB(t)$. Then

$$\begin{aligned} v_{n+1}(t) &\leqslant B(t) + (p-1)\int_t^\infty Q(t,s)\big|v_n(s)\big|^q \\ &\leqslant B(t) + (p-1)p^q\int_t^\infty Q(t,s)B^q(s)\,\mathrm{d}s \\ &\leqslant B(t) + (p-1)p^q \cdot p^{-q}B(t) = pB(t). \end{aligned}$$

Thus, the sequence $\{v_n\}$ is nondecreasing and bounded above. Hence, it converges uniformly to a continuous function v which is a solution of (9.34).

(ii) Suppose, to the contrary, that v is a continuous function satisfying (9.36). Then $v(t) \geqslant B(t) \geqslant 0$ which implies $v^q(t) \geqslant B^q(t) \geqslant 0$. Thus

$$v(t) \geqslant B(t) + (p-1) \int_t^\infty Q(t,s) B^q(s)\,\mathrm{d}s \geqslant \left[1 + (p-1)(1+\varepsilon)p^{-q}\right] B(t).$$

Continuing in this way, we obtain $v(t) \geqslant a_n B(t)$, where $a_1 = 1$, $a_n < a_{n+1}$ and

$$a_{n+1} = 1 + (p-1) a_n^q p^{-q} (1+\varepsilon). \tag{9.37}$$

We claim that $\lim_{n\to\infty} a_n = \infty$. Assume, to the contrary, that $\lim_{n\to\infty} a_n = a < \infty$. Then $a \geqslant 1$ and from (9.37)

$$a = 1 + (p-1)(1+\varepsilon) a^q p^{-q},$$

but this is the contradiction since the equation $\lambda = 1 + (p-1)(1+\varepsilon)\lambda^q p^{-q}$ has no solution for which $\lambda \geqslant 0$. This contradiction proves that $\lim_{n\to\infty} a_n = \infty$ and hence $B(t) \equiv 0$. This contradiction with (9.35) proves the lemma. □

As a direct consequence of the previous lemma we have the following oscillation and nonoscillation criteria.

THEOREM 9.7. *Suppose that* $\int^\infty c(t)\,\mathrm{d}t$ *is convergent and* $C(t) := \int_t^\infty c(s)\,\mathrm{d}s \geqslant 0$ *for large* t.

(i) *If* $\int_t^\infty C^q(s)\,\mathrm{d}s \leqslant p^{-q} C(t)$ *for large* t, *then* (9.1) *is nonoscillatory.*

(ii) *If* $C(t) \not\equiv 0$ *for large* t *and there exists* $\varepsilon > 0$ *such that*

$$\int_t^\infty C^q(s)\,\mathrm{d}s \geqslant p^{-q}(1+\varepsilon) C(t)$$

for large t, *then* (9.1) *is oscillatory.*

PROOF. (i) By the part (i) of the previous lemma there exists a solution of the integral equation (5.21), hence (9.1) is nonoscillatory by Theorem 5.8.

(ii) By contradiction, suppose that (9.1) is nonoscillatory. Then (5.21) has a solution for large t, but it contradicts the part (ii) of the previous lemma. □

REMARK 9.3. The constant p^{-q} in the previous statement is the best possible as shows the Euler equation (4.20) with the critical constant $\gamma = \tilde{\gamma} = (\frac{p-1}{p})^p$.

9.7. *Equations with periodic coefficient*

The oscillation criterion presented in this subsection can be found in the paper [65].

THEOREM 9.8. *Suppose that the function $c(t)$ in* (9.1) *is a periodic function with the period ω, $c(t) \not\equiv 0$, and*

$$\int_0^\omega c(t)\,\mathrm{d}t \geqslant 0,$$

then (9.1) *is oscillatory both at $t = \infty$ and at $t = -\infty$.*

PROOF. To prove oscillation of (9.1), it is sufficient to find a solution of this equation with at least two zeros. Indeed, the periodicity of the function c implies that if x is a solution of (9.1) then $x(t \pm \omega)$ is a solution as well and hence any solution with two zeros has actually infinitely many of them, tending both to ∞ and $-\infty$.

The statement of theorem is clearly true if c is a positive constant function (since then $x(t) = \sin_p \mu t$ is a solution of this equation, where μ is a constant depending on c and p. So we need to consider the cases when $c(t)$ is not a constant only. Also, it is sufficient to deal with the case when $\int_0^\omega c(t)\,\mathrm{d}t = 0$ because otherwise we can define $c_0 = \frac{1}{\omega}\int_0^\omega c(t)\,\mathrm{d}t > 0$ and $\tilde{c}(t) = c(t) - c_0$. Clearly, we have $c(t) > \tilde{c}(t)$. If we prove (9.1) with $\tilde{c}$ instead of c to be oscillatory then by the Sturmian comparison theorem equation (9.1) is also oscillatory.

Now let

$$C(t) = \int_0^t c(s)\,\mathrm{d}s.$$

This is a continuous periodic function with the period ω. Let γ and δ be defined by

$$C(\delta) = \max_{0 \leqslant t \leqslant \omega} C(t), \qquad C(\gamma) = \min_{\delta \leqslant t \leqslant \delta+\omega} C(t).$$

Then $0 \leqslant \delta < \gamma < \delta + \omega$ and

$$\int_\gamma^t c(s)\,\mathrm{d}s \geqslant 0, \qquad \int_t^\delta c(s)\,\mathrm{d}s \geqslant 0 \quad \text{for } t \in \mathbb{R}.$$

Now, by Theorem 8.5 and the remark given below this theorem, the solution of (9.1) given by the initial condition $x(\delta) = 1$, $x'(\delta) = 0$ has a zero in $(-\infty, \delta)$. Indeed, $C(t) \not\equiv 0$ and

$$\int_t^\delta |s - \delta|^\alpha \left(\int_t^\delta c(\tau)\,\mathrm{d}\tau \right) \mathrm{d}s \geqslant 0 \quad \text{for } t \geqslant \delta,$$

with any $\alpha \in (-\frac{1}{p}, p - 2]$. Now we need to show that this solution has a zero on (δ, ∞) as well. We proceed by contradiction, suppose that $x(t) > 0$ for $t \geqslant \delta$. Consider the function $w = -\frac{\Phi(x')}{\Phi(x)}$ on $[\delta, \infty)$. This function satisfies the Riccati differential equation

$$w' = c(t) + (p - 1)|w|^q \tag{9.38}$$

and by integration we have

$$w(t+\omega) - w(t) = (p-1)\int_t^{t+\omega} |w(s)|^q \,\mathrm{d}s, \tag{9.39}$$

hence $w(t+\omega) > w(t)$. Consider now the sequence $w(\gamma)$, $w(\gamma+\omega)$, $w(\gamma+2\omega), \dots$. By Theorem 8.5 and by our indirect assumption on the solution $x(t)$, this sequence consists of negative terms:

$$w(\gamma) < w(\gamma+\omega) < w(\gamma+2\omega) < \cdots < 0.$$

Indeed, if $w(\gamma + k\omega) \geqslant 0$ for some $k \in \mathbb{N}$, then by Theorem 8.5 the solution $x(t)$ would have a zero in $(\gamma + k\omega, \infty)$. Hence $\lim_{k\to\infty} w(\gamma + k\omega) \leqslant 0$, consequently by (9.39)

$$w(\gamma) + (p-1)\int_\gamma^\infty |w(s)|^q \,\mathrm{d}s \leqslant 0,$$

i.e., the integral $\int_\gamma^\infty |w(s)|^q \,\mathrm{d}s$ is convergent. This implies by (9.38) that

$$w(t) = w(\gamma) + \int_\gamma^t c(s)\,\mathrm{d}s + (p-1)\int_\gamma^t |w(s)|^q \,\mathrm{d}s$$

and the function $w(t)$ is bounded. Again by (9.38) we find that w' is also bounded, say, $|w'(t)| < L$. Then

$$\begin{aligned} \left|\frac{|w(t_2)|^{q+1} - |w(t_1)|^{q+1}}{q+1}\right| &= \left|\int_{t_1}^{t_2} w'(s)|w(s)|^q \operatorname{sgn} w(s)\,\mathrm{d}s\right| \\ &\leqslant L\int_{t_1}^{t_2} |w(s)|^q \,\mathrm{d}s, \end{aligned}$$

$\gamma < t_1 < t_2$, hence $\lim_{t\to\infty} |w(t)|^{q+1}$ exists. Clearly, we have $\lim_{t\to\infty} w(t) = 0$.

On the other hand, $w(\delta) = 0$, and by (9.39) we have $\lim_{k\to\infty} w(\delta + k\omega) > 0$ and this contradicts the fact that $\lim_{t\to\infty} w(t) = 0$. □

9.8. *Equations with almost periodic coefficient*

Now suppose that $c : \mathbb{R} \to \mathbb{R}$ is a Besicovitch almost periodic function. Recall that this class of functions is defined as the closure of the set of finite trigonometric polynomials with the Besicovitch seminorm

$$\|c\|_B = \limsup_{t\to\infty} \frac{1}{2t}\int_{-t}^t |c(s)|\,\mathrm{d}s.$$

The mean value $M\{c\}$ of a function c is defined by

$$M\{c\} = \lim_{t\to\infty} \frac{1}{t} \int_{t_0}^{t} c(s)\,\mathrm{d}s,$$

for some $t_0 \geqslant 0$, see, e.g., [19,31] for details.

The proof of the main result of this subsection is based on the following Hartman–Wintner-type lemma. We present it without proof which is similar, in a certain sense, to the proof of Theorem 5.6. For details we refer to [151].

LEMMA 9.2. *Suppose that* $c : [t_0, \infty) \to \mathbb{R}$ *is a locally integrable function with* $M\{c\} = 0$ *and* (9.1) *is nonoscillatory. Then*

$$\lim_{t\to\infty} \frac{1}{t} \int_{t_0}^{t} \left|w(s)\right|^q \mathrm{d}s = 0$$

for every solution w *of the associated Riccati equation* (9.2).

THEOREM 9.9. *Suppose that c is a Besicovitch almost periodic function with the mean value* $M\{c\} = 0$ *and* $M\{|c|\} > 0$. *Then*

$$\big(\Phi(x')\big)' + \lambda c(t)\Phi(x) = 0 \tag{9.40}$$

is oscillatory for every $\lambda \neq 0$.

PROOF. Suppose that (9.40) is nonoscillatory for some λ and $w = -\frac{\Phi(x')}{\Phi(x)}$ be the corresponding solution of the associated Riccati equation

$$w' = \lambda c(t) + (p-1)\left|w(t)\right|^q. \tag{9.41}$$

We have used the Riccati substitution with the '$-$' sign to keep consistency with the original paper [151] and also to show that this minus sign in the Riccati substitution makes sometimes computations slightly easier (compare also the previous subsection).

Integrating (9.41) (with $\delta > 0$ and t sufficiently large) we get

$$\frac{\lambda}{\delta} \int_{t}^{t+\delta} c(s)\,\mathrm{d}s = \frac{w(t+\delta)}{\delta} - \frac{w(t)}{\delta} - \frac{p-1}{\delta} \int_{t}^{t+\delta} \left|w(s)\right|^q. \tag{9.42}$$

Applying the Besicovitch seminorm $\|\cdot\|_{B'}$ defined by

$$\|f\|_{B'} = \limsup_{t\to\infty} \frac{1}{t} \int_{t_0}^{t} \left|f(s)\right| \mathrm{d}s$$

to (9.42), we find

$$\begin{aligned}
0 &\leqslant \left\| \frac{\lambda}{\delta} \int_t^{t+\delta} c(s)\,\mathrm{d}s \right\|_{B'} \\
&\leqslant \left\| \frac{p-1}{\delta} \int_t^{t+\delta} \bigl|w(s)\bigr|^q \,\mathrm{d}s \right\|_{B'} + \left\| \frac{w(t+\delta)}{\delta} \right\|_{B'} + \left\| \frac{w(t)}{\delta} \right\|_{B'}
\end{aligned}$$

for all $\delta > 0$. From Lemma 9.2 follows that $M\{|w|^q\} = 0$, thus $\|w\|_{B'} = \|w(t+\delta)\|_{B'} = 0$ for all $\delta > 0$. Using the Fubini theorem we have for some $t_0 > 0$

$$\begin{aligned}
\frac{1}{\delta t} \int_{t_0}^{t} \int_s^{s+\delta} \bigl|w(\tau)\bigr|^q \,\mathrm{d}\tau\,\mathrm{d}s &= \frac{1}{\delta t} \int_{t_0}^{t} \int_0^{\delta} \bigl|w(\tau+s)\bigr|^q \,\mathrm{d}\tau\,\mathrm{d}s \\
&= \frac{1}{\delta t} \int_0^{\delta} \int_{t_0}^{t} \bigl|w(\tau+s)\bigr|^q \,\mathrm{d}s\,\mathrm{d}\tau \\
&\leqslant \frac{1}{\delta t} \int_0^{\delta} \int_{t_0}^{t+\delta} \bigl|w(s)\bigr|^q \,\mathrm{d}s\,\mathrm{d}\tau \\
&= \frac{1}{t} \int_{t_0}^{t+\delta} \bigl|w(s)\bigr|^q \,\mathrm{d}s
\end{aligned}$$

for any fixed $\delta > 0$. Using the last computation and Lemma 9.2 we have

$$\left\| \frac{p-1}{\delta} \int_t^{t+\delta} \bigl|w(s)\bigr|^q \,\mathrm{d}s \right\|_{B'} = 0.$$

Applying the last equality, coupled with the fact that $\|w(t)\|_{B'} = 0$ to the previous computation, we see that

$$\left\| \frac{\lambda}{\delta} \int_t^{t+\delta} c(s)\,\mathrm{d}s \right\|_{B'} = 0 \tag{9.43}$$

(for every $\delta > 0$). Since c is almost periodic, it follows from [19, p. 97]

$$\lim_{\delta \to 0+} \left\| c(t) - \frac{1}{\delta} \int_t^{t+\delta} c(s)\,\mathrm{d}s \right\|_{B'} = 0.$$

This and (9.43) imply $M\{|c|\} = \|c\|_{B'} = 0$ which is a contradiction. □

9.9. *Generalized H-function averaging technique*

Oscillation criteria of this subsection are established in [152]. Our presentation takes into account the remark of Rogovchenko [204] which shows that one of the assumptions given in the original paper [152] is redundant. In the linear case, the method used in this subsection was introduced by Philos [189].

THEOREM 9.10. *Let* $D_0 = \{(t,s)\colon\ t > s \geqslant t_0\}$ *and* $D = \{(t,s)\colon\ t \geqslant s \geqslant t_0\}$. *Assume that the function* $H \in C(D;\, R)$ *satisfies the following conditions*:

(i) $H(t,t) = 0$ *for* $t \geqslant t_0$ *and* $H(t,s) > 0$ *for* $t > s \geqslant t_0$;

(ii) H *has a continuous nonpositive partial derivative on* D_0 *with respect to the second variable.*

Suppose that $h : D_0 \to \mathbb{R}$ *is a continuous function such that*

$$-\frac{\partial H}{\partial s}(t,s) = h(t,s)\big[H(t,s)\big]^{1/q} \quad \textit{for all } (t,s) \in D_0,\ q = \frac{p}{p-1}.$$

If

$$\limsup_{t\to\infty} \frac{1}{H(t,t_0)} \int_{t_0}^{t} \left[H(t,s)c(s) - \left(\frac{1}{p}h(t,s)\right)^{p} \right] \mathrm{d}s = \infty \tag{9.44}$$

then (9.1) *is oscillatory.*

PROOF. Suppose that (9.1) is nonoscillatory and v is a solution of the associated Riccati equation (9.2) which exists on the interval $[T_0, \infty)$, $T_0 \geqslant t_0$. We have for $t \geqslant T \geqslant T_0$

$$\begin{aligned}
&\int_T^t H(t,s)c(s)\,\mathrm{d}s \\
&\quad = H(t,T)v(T) - \int_T^t \left(-\frac{\partial H}{\partial s}(t,s)\right) v(s)\,\mathrm{d}s - (p-1)\int_T^t H(t,s)\big|v(s)\big|^q\,\mathrm{d}s \\
&\quad = H(t,T)v(T) - \int_T^t \big\{h(t,s)\big[H(t,s)\big]^{1/q} v(s) + (p-1)H(t,s)\big|v(s)\big|^q\big\}\,\mathrm{d}s \\
&\quad = H(t,T)v(T) - \int_T^t \bigg\{h(t,s)\big[H(t,s)\big]^{1/q} v(s) + (p-1)H(t,s)\big|v(s)\big|^q \\
&\qquad + \left(\frac{1}{p}h(t,s)\right)^{p}\bigg\}\,\mathrm{d}s + \int_T^t \left(\frac{1}{p}h(t,s)\right)^{p}\mathrm{d}s.
\end{aligned}$$

Hence, for $t \geqslant T \geqslant T_0$, we have

$$\begin{aligned}
&\int_T^t \left\{H(t,s)c(s) - \left(\frac{1}{p}h(t,s)\right)^{p}\right\}\mathrm{d}s \\
&\quad = H(t,T)v(T) - \int_T^t \bigg\{h(t,s)\big[H(t,s)\big]^{1/q} v(s) + (p-1)H(t,s)\big|v(s)\big|^q \\
&\qquad + \left(\frac{1}{p}h(t,s)\right)^{p}\bigg\}\,\mathrm{d}s.
\end{aligned}$$

Since $q > 1$, by Young's inequality (2.6)

$$h(t,s)\big[H(t,s)\big]^{1/q}v(s) + (p-1)H(t,s)\big|v(s)\big|^q + \left(\frac{1}{p}h(t,s)\right)^p \geqslant 0$$

for $t \geqslant s \geqslant T_0$. This implies that for every $t \geqslant T_0$

$$\int_{T_0}^{t}\left\{H(t,s)c(s) - \left(\frac{1}{p}h(t,s)\right)^p\right\}\mathrm{d}s$$
$$\leqslant H(t,T_0)v(T_0) \leqslant H(t,T_0)\big|v(T_0)\big| \leqslant H(t,t_0)\big|v(T_0)\big|.$$

Therefore,

$$\begin{aligned}\int_{t_0}^{t}&\left\{H(t,s)c(s) - \left(\frac{1}{p}h(t,s)\right)^p\right\}\mathrm{d}s\\ &= \int_{t_0}^{T_0}\left\{H(t,s)c(s) - \left(\frac{1}{p}h(t,s)\right)^p\right\}\mathrm{d}s\\ &\quad + \int_{T_0}^{t}\left\{H(t,s)c(s) - \left(\frac{1}{p}h(t,s)\right)^p\right\}\mathrm{d}s\\ &\leqslant H(t,t_0)\int_{t_0}^{T_0}\big|c(s)\big|\,\mathrm{d}s + H(t,t_0)\big|v(T_0)\big|\\ &= H(t,t_0)\left\{\int_{t_0}^{T_0}\big|c(s)\big|\,\mathrm{d}s + \big|v(T_0)\big|\right\}.\end{aligned}$$

This gives

$$\begin{aligned}\limsup_{t\to\infty}\frac{1}{H(t,t_0)}&\int_{t_0}^{t}\left\{H(t,s)c(s) - \left(\frac{1}{p}h(t,s)\right)^p\right\}\mathrm{d}s\\ &\leqslant \int_{t_0}^{T_0}\big|c(s)\big|\,\mathrm{d}s + \big|v(T_0)\big|.\end{aligned}$$

This contradiction with (9.44) completes the proof. ☐

The next statement is also taken from [152]. We present it without proof. This proof, similar to the proof of the previous theorem, follows more or less the original idea of Philos [189]. For comparison with the linear case we also refer to the papers of Yan [220, 221]. Taking $H(t,s) = (t-s)^\lambda$, $\lambda > 0$, the previous statement reduces to the half-linear version of Kamenev's oscillation criterion presented in Section 9.4.

THEOREM 9.11. *Let H and h be as in the previous theorem, and let*

$$\inf_{s\geqslant t_0}\left\{\liminf_{t\to\infty}\frac{H(t,s)}{H(t,t_0)}\right\}>0. \tag{9.45}$$

Suppose that

$$\limsup_{t\to\infty}\frac{1}{H(t,t_0)}\int_{t_0}^{t}h(t,s)\,\mathrm{d}s<\infty$$

and there exists a function $A\in C[t_0,\infty)$ such that

$$\int^{\infty}A_+^q(s)\,\mathrm{d}s=\infty, \tag{9.46}$$

where $A_+(t)=\max\{A(t),0\}$. If

$$\limsup_{t\to\infty}\frac{1}{H(t,T)}\int_T^t\left\{H(t,s)c(s)-\left(\frac{1}{p}h(t,s)\right)^p\right\}\mathrm{d}s\geqslant A(T)$$

for $T\geqslant t_0$ then Equation (9.1) *is oscillatory.*

10. Various oscillation problems

In this section we collect various problems of half-linear oscillation theory. We start with an asymptotic formula for the distance between consecutive zeros of oscillatory solutions of half-linear equations. Then we turn our attention to various problems like oscillation of forced and retarded equations and to similar problems.

10.1. *Asymptotic formula for distance of zeros of oscillatory solutions*

The results of this subsection are taken from [93] and present the asymptotic formula for the distance of consecutive zeros of oscillatory solutions of the equation

$$\bigl(\Phi(x')\bigr)'+(p-1)c(t)\Phi(x)=0. \tag{10.1}$$

It is supposed that $c(t)>0$ for large t and the results are based on the generalized Prüfer transformation from Section 2. In this transformation, a nontrivial solution and its derivative are expressed via the generalized half-linear sine and cosine functions. Recall that the half-linear sine function, denoted by $S=S(t)$ or $\sin_p t$, is the solution of the equation

$$\bigl(\Phi(x')\bigr)'+(p-1)\Phi(x)=0$$

satisfying the initial condition $x(0)=0$, $x'(0)=1$ and the half-linear cosine function is defined by $\cos_p t=S'(t)$.

THEOREM 10.1. *Suppose that c is a differentiable function such that* $c(t) > 0$ *on an interval* $[T, \infty)$*, and*

$$\lim_{t\to\infty} c'(t)\big[c(t)\big]^{-\frac{p+1}{p}} = 0 \tag{10.2}$$

holds. Then (10.1) *is oscillatory. Moreover, if* $N[x; T]$ *denotes the number of zeros of a solution x of* (10.1) *in the interval* $[a, T]$*, then*

$$N[x; T] = P[x; T] + R[x; T], \tag{10.3}$$

where $P[x; T]$ *is the principal term given by*

$$P[x; T] = \frac{1}{\pi_p} \int_a^T \big[c(s)\big]^{\frac{1}{p}}\, \mathrm{d}s$$

and $R[x; T]$ *is the remainder which is of smaller order than* $P[x; T]$ *as* $T \to \infty$ *and satisfies*

$$\big|R[x; T]\big| \leqslant \frac{1}{p\pi_p} \int_a^T \frac{|c'(s)|}{c(s)}\, \mathrm{d}s + O(1).$$

PROOF. Set $C(t) := c'(t)[c(t)]^{-\frac{p+1}{p}}$ and define

$$C^*(t) = \sup\big\{|C(s)|\colon\ s \geqslant t\big\}, \quad t \geqslant a. \tag{10.4}$$

Then $C^*(t)$ is nonincreasing and satisfies $\lim_{t\to\infty} C^*(t) = 0$ by (10.2). We have

$$\big|\big[c(t+h)\big]^{-\frac{1}{p}} - \big[c(t)\big]^{-\frac{1}{p}}\big| = \frac{1}{p}\left|\int_t^{t+h} C(s)\, \mathrm{d}s\right| \leqslant \frac{|h|}{p} C^*(t),$$

which implies that

$$\limsup_{h\to\infty} \frac{[c(t+h)]^{-\frac{1}{p}}}{t+h} \leqslant \frac{C^*(t)}{p}.$$

It follows that $\lim_{t\to\infty} t^{-1}[c(t)]^{-\frac{1}{p}} = 0$, or equivalently, $\lim_{t\to\infty} t^p c(t) = \infty$. This implies, by Theorem 4.5, that (10.1) is oscillatory.

Now we turn our attention to the proof of the asymptotic formulas for numbers of zeros. By the Sturmian comparison theorem (Theorem 2.4) we have that $N[x_1; T]$ and $N[x_2; T]$ differ at most by one for any solutions x_1 and x_2 of (10.1), so we may restrict our attention to the solution x_0 of (10.1) determined by the initial conditions $x_0(a) = 0$, $x_0'(a) = 1$. This solution is oscillatory by the first part of the our theorem.

We introduce the polar coordinates $\rho(t)$, $\varphi(t)$ for $x_0(t)$ by setting

$$[c(t)]^{\frac{1}{p}} x_0(t) = \rho(t) S(\varphi(t)), \qquad x_0'(t) = \rho(t) S'(\varphi(t)). \tag{10.5}$$

It can be shown without difficulty that $\rho(t)$ and $\varphi(t)$ are continuously differentiable on $[a, \infty)$ and satisfy the differential equations

$$\begin{aligned} \frac{\rho'}{\rho} &= \frac{1}{p} \frac{c'(t)}{c(t)} \left|S(\varphi)\right|^p, \\ \varphi' &= [c(t)]^{\frac{1}{p}} + \frac{1}{p} \frac{c'(t)}{c(t)} S(\varphi) \Phi(S'(\varphi)). \end{aligned} \tag{10.6}$$

We use the notation

$$g(\varphi) = S(\varphi) \Phi(S'(\varphi)),$$

in terms of which (10.6) is written as

$$\varphi' = [c(t)]^{\frac{1}{p}} + \frac{1}{p} \frac{c'(t)}{c(t)} g(\varphi). \tag{10.7}$$

From the first equation in (10.5) we see that $x_0(t) = 0$ if and only if $\varphi(t) = j\pi_p$, $j \in \mathbb{Z}$. We may suppose that $\varphi(a) = 0$. In view of (10.2) there is no loss of generality in assuming that

$$C^*(t) < p \quad \text{for } t \geqslant a,$$

where $C^*(t)$ is defined by (10.4). Since

$$\left|g(\varphi)\right| \leqslant 1 \quad \text{for all } \varphi, \tag{10.8}$$

we have

$$[c(t)]^{\frac{1}{p}} + \frac{1}{p} \frac{c'(t)}{c(t)} g(\varphi(t)) \geqslant [c(t)]^{\frac{1}{p}} \left(1 - \frac{1}{p} C^*(t)\right) > 0,$$

which implies that $\varphi'(t) > 0$, so that $\varphi(t)$ is increasing for $t \geqslant a$.

We now integrate (10.7) over $[a, T]$, obtaining

$$\varphi(T) = \int_a^T [c(s)]^{\frac{1}{p}} \, \mathrm{d}s + \frac{1}{p} \int_a^T \frac{c'(s)}{c(s)} g(\varphi(s)) \, \mathrm{d}s = F(T) + G(T), \tag{10.9}$$

where

$$F(T) := \int_a^T [c(s)]^{\frac{1}{p}} \, \mathrm{d}s, \qquad G(T) := \frac{1}{p} \int_a^T \frac{c'(s)}{c(s)} g(\varphi(s)) \, \mathrm{d}s.$$

From (10.8) it is clear that

$$\left|G(T)\right| \leqslant \frac{1}{p}\int_a^T \frac{|c'(s)|}{c(s)}\,\mathrm{d}s. \tag{10.10}$$

Noting that the number of zeros of $x_0(t)$ in $[a, T]$ is given by

$$N[x_0; T] = \left[\frac{\varphi(T)}{\pi_p}\right] + 1,$$

where $[u]$ denotes the greatest integer not exceeding u, we see from (10.9) and (10.10) that the conclusion of the theorem holds with the choice

$$P[x_0; T] = \frac{1}{\pi_p}F(T) = \frac{1}{\pi_p}\int_a^T \left[c(s)\right]^{\frac{1}{p}}\,\mathrm{d}s.$$

That the term $R[x_0; T] = N[x_0; T] - P[x_0; T]$ is of smaller order than $P[x_0; T]$ follows from the observation that

$$\begin{aligned}\int_a^T \frac{|c'(s)|}{c(s)}\,\mathrm{d}s &= \int_a^T \left|C(s)\right|\left[c(s)\right]^{\frac{1}{p}}\,\mathrm{d}s\\ &\leqslant \int_a^T C^*(s)\left[c(s)\right]^{\frac{1}{p}}\,\mathrm{d}s = \mathrm{o}\left(\int_a^T \left[c(s)\right]^{\frac{1}{p}}\,\mathrm{d}s\right) \quad \text{as } T \to \infty.\end{aligned}$$

This completes the proof. ☐

EXAMPLE 10.1. Consider the equation

$$\left(\Phi(x')\right)' + (p-1)t^{\beta}\Phi(x) = 0, \quad t \geqslant 1, \tag{10.11}$$

where β is a constant with $p + \beta > 0$. The function $c(t) = t^{\beta}$ satisfies

$$\int_1^T \left[c(s)\right]^{\frac{1}{p}}\,\mathrm{d}s = \frac{p}{p+\beta}\left(T^{\frac{p+\beta}{p}} - 1\right),$$

$$\int_1^T \frac{|c'(s)|}{c(s)}\,\mathrm{d}s = |\beta|\log T,$$

and so we conclude from Theorem 10.1 that the quantity $P[x; T]$ can be taken to be

$$P[x; T] = \frac{p}{(p+\beta)\pi_p}T^{\frac{p+\beta}{p}}$$

and (10.3) holds with this $P[x; T]$ and $R[x; T]$ satisfying

$$R[x; T] = \frac{|\beta|}{p\pi_p}\log T + \mathrm{O}(1).$$

REMARK 10.1. (i) The results of this subsection cannot be applied to the generalized Euler equation (4.20), since the function $c(t) = \lambda(p-1)t^{-p}$ does not satisfy (10.2). A calculation of $P[x; T]$ and $R[x; T]$ for the generalized Euler equation

$$\big(\Phi(x')\big)' + \lambda(p-1)t^{-p}\Phi(x) = 0$$

shows that both of them are of the same logarithmic order as $T \to \infty$.

(ii) In [191], Piros has investigated a similar problem under a more stringent restriction on $c(t)$, namely he supposed that $c^{\nu}(t)$ is a concave function of t for some $\nu > 0$. Then he proved that the error term $R[x; T]$ in (10.3) is O(1). Exactly, the differential equation (10.11) with $\beta = \frac{1}{\nu}$ plays the exceptional role in determining the precise value of $R[x; T]$.

10.2. *Half-linear Milloux and Armellini–Tonelli–Sansone theorems*

Results of this subsection are taken from [15] and [23]. Recall that the classical Armellini–Tonelli–Sansone theorem concerns the convergence to zero of all solutions of the second order linear differential equation

$$x'' + c(t)x = 0. \tag{10.12}$$

In particular, by the theorem of Milloux [175], if the function c is continuously differentiable, nondecreasing, and

$$\lim_{t\to\infty} c(t) = \infty \tag{10.13}$$

then (10.12) has at least one solution satisfying

$$\lim_{t\to\infty} x(t) = 0. \tag{10.14}$$

The theorem of Armellini–Tonelli–Sansone deals with the situation when *all* solutions of (10.12) satisfy (10.14). This happens when c goes to infinity "regularly" (the exact definition is given below). Regular growth means, roughly speaking, that a function does not increase fast on intervals of short length.

Here we show that both theorems extend verbatim to (9.1). First we present some definitions. Let $S := \{(\alpha_k, \beta_k)\}$ be a sequence of intervals such that

$$0 \leqslant \alpha_1 < \beta_1 < \alpha_2 < \beta_2 < \cdots < \beta_k \to \infty \quad \text{as } k \to \infty. \tag{10.15}$$

Then

$$\limsup_{k\to\infty} \frac{\sum_{i=1}^{k}(\beta_i - \alpha_i)}{\beta_k} =: \delta(S) = \delta \tag{10.16}$$

is called the *density* of the sequence of intervals S. A nondecreasing positive function f tends to infinity *intermittently* (an alternative terminology is *quasi-jumping*) as $t \to \infty$, provided to every $\varepsilon > 0$ there exists a sequence of intervals S satisfying (10.15) such that $\delta(S) \leqslant \varepsilon$ and the increase of f on $\mathbb{R}_+ \setminus S$ is finite, i.e.,

$$\mathcal{S}(f; S) := \sum_{k=1}^{\infty} \big[f(\alpha_k) - f(\beta_{k-1}) \big] < \infty. \tag{10.17}$$

In the opposite case we say that $f(t) \to \infty$ *regularly* as $t \to \infty$.

THEOREM 10.2. *Suppose that c is a nondecreasing continuously differentiable function satisfying* (10.13). *Then* (10.1) *possesses at least one nontrivial solution satisfying* (10.14).

PROOF. From the variety of proofs we present that one based on the modified Prüfer transformation (compare Section 1.3). An alternative approach to the problem is presented in [107,108,123].

For any nontrivial solution x of (10.1) there exist a positive function ϱ given by the formula

$$\varrho = \left[|x|^p + \frac{1}{c} |x'|^p \right]^{\frac{1}{p}}$$

and a continuous function ϑ such that x can be expressed in the form

$$x(t) = \varrho(t) S\big(\vartheta(t)\big), \qquad x'(t) = c^{\frac{1}{p}}(t) \varrho(t) S'\big(\vartheta(t)\big),$$

where the generalized sine function S is the same as in Section 1.3. The functions ϑ and ϱ satisfy the differential system

$$\vartheta' = c^{\frac{1}{p}}(t) + \frac{c'(t)}{c(t)} f\big(\vartheta(t)\big), \qquad \frac{\varrho'}{\varrho} = -\frac{c'(t)}{c(t)} g\big(\vartheta(t)\big), \tag{10.18}$$

where

$$f(\vartheta) = \frac{1}{p} \Phi\big(S'(\vartheta)\big) S(\vartheta), \qquad g(\vartheta) = \frac{1}{p} \big|S'(\vartheta)\big|^p.$$

The right-hand side of (10.18) is Lipschitzian in ϑ hence the solution of (10.18) is uniquely determined by the initial condition. We denote by $\vartheta(t, \varphi)$, $\varrho(t, \varphi)$ the solution given by the initial condition $\vartheta(0) = \varphi$, $\varrho(0) = 1$. Then

$$\varrho(t, \varphi) = \exp\left\{ -\int_0^t \frac{c'(s)}{c(s)} g\big(\vartheta(s, \varphi)\big) \, \mathrm{d}s \right\},$$

and since $g(\vartheta) \geqslant 0$, the function $\varrho(t, \vartheta)$ is nonincreasing and tends to a nonnegative limit $\varrho(\infty, \varphi)$ as $t \to \infty$. Obviously, $\varrho(\infty, \varphi) = 0$ implies that $x(t) \to 0$ as $t \to \infty$. The converse is also true because x is oscillatory.

We have the following two possibilities.

(i) We have $\varrho(\infty, \varphi) = 0$, the corresponding solution x satisfies $x(t) \to 0$ as $t \to \infty$, and

$$\int_0^\infty \frac{c'(t)}{c(t)} g\bigl(\vartheta(t, \varphi)\bigr)\,\mathrm{d}t = \infty,$$

(ii) $\varrho(\infty, \varphi) > 0$, the solution x oscillates, its amplitude tends to a positive limit, and

$$\int_0^\infty \frac{c'(t)}{c(t)} g\bigl(\vartheta(t, \varphi)\bigr)\,\mathrm{d}t < \infty. \tag{10.19}$$

Now, the proof is based on the behavior as $t \to \infty$ of the function $\psi(t, \varphi_1, \varphi_2) = \vartheta(t, \varphi_2) - \vartheta(t, \varphi_1)$ which is described in the next two auxiliary statements. Here $\mathcal{X}$ denotes the set of φ's such that (10.19) holds, this means that the corresponding solution does not tend to zero as $t \to \infty$. The proof can be found in [15].

LEMMA 10.1. *Let $\varphi_1, \varphi_2 \in \mathcal{X}$ and $\varphi_1 < \varphi_2 < \varphi_1 + \pi_p$. Then*

$$\psi(\infty, \varphi_1, \varphi_2) := \lim_{t\to\infty} \bigl[\vartheta(t, \varphi_2) - \vartheta(t, \varphi_1)\bigr]$$

exists and equals 0 *or* π_p.

LEMMA 10.2. *Let $\varphi_0 \in \mathcal{X}$. Then for any $\varepsilon > 0$ there exists $\eta \in (0, \pi_p)$ such that if $|\varphi - \varphi_0| < \eta$, then*

$$\bigl|\vartheta(t, \varphi) - \vartheta(t, \varphi_0)\bigr| < \varepsilon \quad \textit{for } t \geqslant 0. \tag{10.20}$$

Now, returning to the proof of our theorem, suppose that $\mathcal{X} = \mathbb{R}$. Then the function φ given by $\psi(\infty, 0, \varphi) = 0$ is nondecreasing as φ increases in $[0, \pi_p]$. It must go from 0 to π_p, taking on only these two values, by Lemma 10.1. But this is impossible since by Lemma 10.2 this function is continuous, so the assumption $\mathcal{X} = \mathbb{R}$ was false and the theorem is proved. □

Now we turn our attention to the extension of the Armellini–Tonelli–Sansone theorem.

THEOREM 10.3. *If the function* $\lg c(t) \to \infty$ *regularly, then every solution of* (10.1) *satisfies* (10.14).

PROOF. For the sake of simplicity we suppose that c is continuously differentiable for large t. Consider the function

$$A(t) = \bigl|x(t)\bigr|^p + \frac{|x'(t)|^p}{c(t)},$$

where x is a nontrivial solution of (10.1). The function A is nondecreasing since

$$A'(t) = -\frac{c'(t)}{c^2(t)}\left|x'(t)\right|^p \leqslant 0. \tag{10.21}$$

Consequently, there exists a (finite or infinite) limit $A = \lim_{t\to\infty} A(t)$ and $A \geqslant 0$.

Suppose, by contradiction, that there exists a solution x of (10.1) which does not tend to zero. For this solution obviously $A > 0$. By (10.21)

$$\begin{aligned}
A(t) &= A(0) - \int_0^t \frac{c'(t)}{c^2(s)}\left|x'(s)\right|^p \mathrm{d}s \\
&= A(0) - \int_0^t \frac{c'(s)}{c(s)}\left(A(s) - \left|x(s)\right|^p\right) \mathrm{d}s \\
&= A(0) - \int_0^t \left(A(s) - \left|x(s)\right|^p\right) \frac{\mathrm{d}c(s)}{c(s)}.
\end{aligned}$$

Let $\varepsilon > 0$ be a number such that for every sequence S_ε of intervals

$$\mathcal{S} = \sum_{i=1}^{k}\left[\lg c(\alpha_{i+1}) - \lg c(\beta_i)\right] = \sum_{i=1}^{k} \lg \frac{c(\alpha_{i+1})}{c(\beta_i)} \to \infty \tag{10.22}$$

as $k \to \infty$.

In the remaining part of the proof we suppose that the following statement holds.

LEMMA 10.3. *For every* $\varepsilon_0 > 0$ *there exists* $\eta > 0$ *such that the density of the sequence S of all intervals where*

$$A(t) - \left|x(t)\right|^p \leqslant \eta \tag{10.23}$$

is less than ε_0.

Since the proof of this lemma is rather technical and follows essentially the original linear idea, we skip it and return to the proof of theorem.

Denote by (α_i, β_i) intervals, where (10.23) holds. On intervals (β_i, α_{i+1}) we have

$$A(t) - \left|x(t)\right|^p > \eta,$$

therefore

$$A(\alpha_k) \leqslant A(0) - \sum_{i=1}^{k-1} \int_{\beta_i}^{\alpha_{i+1}} \left(A(t) - \left|x(t)\right|^p\right) \frac{\mathrm{d}c(t)}{c(t)} < A(0) - \eta \sum_{i=1}^{k} \lg \frac{c(\alpha_{i+1})}{c(\beta_i)}$$

which implies by (10.22) that $A(\alpha_k)$ becomes negative for large k. This is a contradiction with $A = \lim_{t\to\infty} A(t) > 0$. □

10.3. *Strongly and conditionally oscillatory equation*

Differential equation (0.1) with a *positive* function c is said to be *conditionally oscillatory* if there exists a constant $\lambda_0 > 0$ such that (0.1) with $\lambda c(t)$ instead of $c(t)$ is oscillatory for $\lambda > \lambda_0$ and nonoscillatory if $\lambda < \lambda_0$. The value λ_0 is called the *oscillation constant* of (0.1). If equation is oscillatory (nonoscillatory) for every $\lambda > 0$, then equation is said to be *strongly oscillatory* (*strongly nonoscillatory*). The results of this subsection are presented in [133].

The examples illustrating these concepts we have already seen in the previous sections. For example, if $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$, the equation

$$\bigl(r(t)\Phi(x')\bigr)' + \frac{r^{1-q}(t)}{(\int^t r^{1-q}(s)\,\mathrm{d}s)^\mu}\Phi(x) = 0 \tag{10.24}$$

is conditionally oscillatory if $\mu = p$, strongly oscillatory if $\mu < p$ and strongly nonoscillatory if $\mu > p$. This follows from the fact that the transformation of independent variable $t \longmapsto \int^t r^{1-q}(s)\,\mathrm{d}s$ transforms (10.24) into the equation

$$\bigl(\Phi(x')\bigr)' + \frac{1}{t^\mu}\Phi(x) = 0 \tag{10.25}$$

and (10.25) is compared with the Euler equation (4.20).

THEOREM 10.4. *Suppose that* $\int^\infty c(t)\,\mathrm{d}t < \infty$ *and* $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$. *Equation* (0.1) *is strongly oscillatory if and only if*

$$\limsup_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\int_t^\infty c(s)\,\mathrm{d}s = \infty \tag{10.26}$$

and it strongly nonoscillatory if and only if

$$\lim_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\int_t^\infty c(s)\,\mathrm{d}s = 0. \tag{10.27}$$

PROOF. The proof is based on the statements which claim (upon a slight reformulation) that under the assumptions of our theorem Equation (0.1) is oscillatory provided

$$\liminf_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\int_t^\infty c(s)\,\mathrm{d}s > \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1} \tag{10.28}$$

or

$$\limsup_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\int_t^\infty c(s)\,\mathrm{d}s > 1 \tag{10.29}$$

and it is nonoscillatory provided

$$\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\int_t^\infty c(s)\,\mathrm{d}s<\frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1} \tag{10.30}$$

for large t. These statements follow essentially from comparing (0.1) with Euler-type differential equation, see, e.g., [60].

Suppose that (10.26) holds and $\lambda>0$ is arbitrary. Then clearly

$$\limsup_{t\to\infty}\left(\int^\infty r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\int_t^\infty \lambda c(s)\,\mathrm{d}s>1$$

hence (0.1) is strongly oscillatory. Conversely, suppose that (0.1) is strongly oscillatory and (10.26) fails to hold, i.e.,

$$\limsup_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\int_t^\infty c(s)\,\mathrm{d}s=L<\infty.$$

Then for $\lambda<\frac{1}{pL}(\frac{p-1}{p})^{p-1}$ we have

$$\limsup_{t\to\infty}\left(\int^t r^{1-q}(s)\,\mathrm{d}s\right)^{p-1}\int_t^\infty \lambda c(s)\,\mathrm{d}s<\frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1}$$

which means that (0.1) with λc instead of c is nonoscillatory. This contradicts strong oscillation of (0.1). The proof of the part concerning necessary and sufficient condition for strong nonoscillation of (0.1) is the same. ☐

10.4. *Oscillation of forced half-linear differential equations*

The result of this subsection is presented in [158] and concerns the forced half-linear differential equation

$$\big(r(t)\Phi(x')\big)'+c(t)\Phi(x)=f(t), \tag{10.31}$$

where f is a continuous function. It extends the oscillation criterion of Wong [215] (see also references given therein) which concerns the linear forced equation. Oscillation of a more general half-linear forced equation than (10.31) is investigated in [139], but for simplicity we here the results of [158].

THEOREM 10.5. *Suppose that for any T there exist $T\leqslant s_1<t_1\leqslant s_2<t_2$ such that $f(t)\leqslant 0$ for $t\in[s_1,t_1]$ and $f(t)\geqslant 0$ $t\in[s_2,t_2]$. Denote*

$$D(s_i,t_i)=\big\{u\in C^1[s_i,t_i]\colon\ u(t)\not\equiv 0,\ u(s_i)=0=u(t_i)\big\},\quad i=1,2.$$

If there exists a function $h \in D(s_i, t_i)$ *and a positive nondecreasing function* $\varphi \in C^1[T, \infty)$ *such that*

$$\int_{s_i}^{t_i} h^2(t)\varphi(t)c(t)\,\mathrm{d}t > \frac{1}{p^p}\int_{s_i}^{t_i} \frac{r(t)\varphi(t)}{|h(t)|^{p-2}}\left(2\left|h'(t)\right| + h(t)\frac{\varphi'(t)}{\varphi(t)}\right)^p \mathrm{d}t,$$

for $i = 1, 2$, *then every solution of* (10.31) *is oscillatory.*

Proof. Suppose that x is a nonoscillatory solution which is eventually of one sign, say $x(t) > 0$ for $t \geqslant T_0$ and let the function w be defined by the modified Riccati substitution

$$w(t) = \varphi(t)\frac{r(t)\Phi(x'(t))}{\Phi(x(t))}.$$

Then w solves the generalized Riccati equation

$$w' = -\varphi(t)c(t) + \frac{\varphi'(t)}{\varphi(t)}w - (p-1)\frac{|w|^q}{(r(t)\varphi(t))^{q-1}} + \frac{\varphi(t)f(t)}{\Phi(x(t))}. \tag{10.32}$$

By the assumptions of theorem, one can choose $s_1, t_1 \geqslant T_0$ so that $f(t) \leqslant 0$ on $I = [s_1, t_1]$ with $s_1 < t_1$. From (10.32) we have for $t \in I$

$$\varphi(t)c(t) \leqslant -w'(t) + \frac{\varphi'(t)}{\varphi(t)}w(t) - (p-1)\frac{|w(t)|^q}{(r(t)\varphi(t))^{q-1}}. \tag{10.33}$$

Multiplying (10.33) by h^2 and integrating over I we obtain

$$\begin{aligned}
&\int_{s_1}^{t_1} h^2(t)\varphi(t)c(t)\,\mathrm{d}t \\
&\quad\leqslant -\int_{s_1}^{t_1} h^2(t)w'(t)\,\mathrm{d}t + \int_{s_1}^{t_1} h^2(t)\frac{\varphi'(t)}{\varphi(t)}w(t)\,\mathrm{d}t \\
&\qquad - (p-1)\int_{s_1}^{t_1} \frac{|w(t)|^q}{(r(t)\varphi(t))^{q-1}}\,\mathrm{d}t.
\end{aligned}$$

Integrating the last inequality by parts and using the fact that $h(s_1) = 0 = h(t_1)$, we get

$$\begin{aligned}
&\int_{s_1}^{t_1} h^2(t)\varphi(t)c(t)\,\mathrm{d}t \\
&\quad\leqslant \int_{s_1}^{t_1} 2\left|h(t)\right|\left|h'(t)\right|\left|w(t)\right|\mathrm{d}t + \int_{s_1}^{t_1} h^2(t)\frac{\varphi'(t)}{\varphi(t)}w(t)\,\mathrm{d}t \\
&\qquad - (p-1)\int_{s_1}^{t_1} \frac{|w(t)|^q}{(r(t)\varphi(t))^{q-1}}\,\mathrm{d}t
\end{aligned}$$

$$\leqslant \int_{s_1}^{t_1} \Big(2|h(t)||h'(t)| + \frac{\varphi'(t)}{\varphi(t)} h^2(t)\Big)|w(t)|\,\mathrm{d}t$$
$$- (p-1)\int_{s_1}^{t_1} h^2(t)\frac{|w(t)|^q}{(r(t)\varphi(t))^{q-1}}\,\mathrm{d}t.$$

Now, the application of the Young inequality yields for $t \in [s_1, t_1]$

$$\Big(2|h(t)||h'(t)| + \frac{\varphi'(t)}{\varphi(t)} h^2(t)\Big)|w(t)| - (p-1)\frac{|w(t)|^q}{(r(t)\varphi(t))^{q-1}} h^2(t)$$
$$\leqslant \frac{1}{p^p}\frac{r(t)\varphi(t)}{|h(t)|^{p-2}}\Big(2|h'(t)| + |h(t)|\frac{\varphi'(t)}{\varphi(t)}\Big)^p,$$

thus

$$\int_{s_1}^{t_1} h^2(t)\varphi(t)c(t)\,\mathrm{d}t \leqslant \frac{1}{p^p}\int_{s_1}^{t_1} \frac{r(t)\varphi(t)}{|h(t)|^{p-2}}\Big(2|h'(t)| + |h(t)|\frac{\varphi'(t)}{\varphi(t)}\Big)^p\,\mathrm{d}t,$$

which contradicts our assumption. When x is eventually negative, we may employ the fact that $f(t) \geqslant 0$ on some interval in any neighborhood of ∞ to reach a similar contradiction. ☐

10.5. *Oscillation of retarded half-linear equations*

The results of this subsection are given [4], for related results we refer also to [2,227] and the references given therein. We consider the equation

$$\big(\Phi(x')\big)' + c(t)\Phi\big(x\big(\tau(t)\big)\big) = 0. \tag{10.34}$$

We suppose that $c(t) \geqslant 0$ for large t, $\lim_{t\to\infty}\tau(t) = \infty$ and $\tau(t) \leqslant t$. Equation (10.34) is said to be it *oscillatory* if all its solutions are oscillatory, i.e., have arbitrarily large zeros. Since the Sturmian theory generally does not extend to (10.34), oscillatory and nonoscillatory solutions (i.e., solutions which are eventually positive or negative) may coexist.

In the next theorem, Equation (10.34) is considered on the interval $[t_0, \infty)$ and it is shown that if the delay $\tau(t)$ is sufficiently close to t, in a certain sense, then some oscillation criteria for (9.1) can be extended to (10.34). Oscillation criteria presented here are half-linear extensions of some results for the linear second order retarded equations (the case $p = 2$ in (10.34)) given in [96,184].

First we present without proof a technical auxiliary statement, the proof can be found in [4].

LEMMA 10.4. *Suppose that the following conditions hold*:

(i) $x(t) \in C^2[T, \infty)$ *for some* $T > 0$,

(ii) $x(t) > 0$, $x'(t) > 0$, *and* $x''(t) \leqslant 0$ *for* $t \geqslant T$.

Then for each $k_1 \in (0,1)$ *there exists a constant* $T_{k_1} \geqslant T$ *such that*

$$x\bigl(\tau(t)\bigr) \geqslant \frac{k_1\tau(t)}{t}x(t), \quad \textit{for } t \geqslant T_{k_1},$$

and for every $k_2 \in (0,1)$ *there exists a constant* $T_{k_2} \geqslant T$ *such that*

$$x(t) \geqslant k_2 t x'(t), \quad \textit{for } t \geqslant T_{k_2}.$$

THEOREM 10.6. *Denote for* $t \geqslant t_0$

$$\gamma(t) := \sup\bigl\{s \geqslant t_0\colon\ \tau(s) \leqslant t\bigr\}.$$

Equation (10.34) *is oscillatory if either of the following holds*:

$$\limsup_{t\to\infty} t^{p-1}\int_t^\infty c(s)\left(\frac{\tau(s)}{s}\right)^{p-1} \mathrm{d}s > 1, \tag{10.35}$$

or

$$\limsup_{t\to\infty} t^{p-1}\int_{\gamma(t)}^\infty c(s)\,\mathrm{d}s = \infty. \tag{10.36}$$

PROOF. Suppose to the contrary that (10.34) has a nonoscillatory solution $x(t)$. Without loss of generality we may suppose that $x(t) > 0$ for large t, say $t \geqslant t_1$. Then also $x(\tau(t)) > 0$ on $[t_1, \infty)$ for t_1 large enough. Since $c(t) \geqslant 0$ on $[t_1, \infty)$,

$$\bigl(\Phi(x')\bigr)' = -c(t)\bigl(\Phi\bigl(x\bigl(\tau(t)\bigr)\bigr)\bigr) \leqslant 0. \tag{10.37}$$

Hence, the function $\Phi(x')$ is decreasing. Since $\sup\{c(t)\colon\ t \geqslant T\} > 0$ for any $T \geqslant 0$, we see that either

(a) $x'(t) > 0$ for all $t \geqslant t_1$, or

(b) there exists $t_2 \geqslant t_1$ such that $x'(t) < 0$ on $[t_2, \infty)$.

If (b) holds, then it follows from (10.37) that

$$0 \geqslant \bigl[\bigl|x'(t)\bigr|^{p-2}x'(t)\bigr]' = (p-1)\bigl|x'(t)\bigr|^{p-2}x''(t), \quad \text{for } t \geqslant t_2.$$

Thus, $x''(t) \leqslant 0$ for $t \in [t_2, \infty)$. This and $x'(t) < 0$ on $[t_2, \infty)$ imply that there exists $t_3 > t_2$ such that $x(t) \leqslant 0$ for $t \geqslant t_3$. This contradicts $x(t) > 0$. Thus, (a) holds.

Integrating (10.37) from $t \geqslant t_1$ to ∞, we obtain

$$\begin{aligned} -\int_t^\infty c(s)\Phi\bigl(x\bigl(\tau(s)\bigr)\bigr)\,\mathrm{d}s &= \int_t^\infty \bigl(\Phi\bigl(x'(s)\bigr)\bigr)'\,\mathrm{d}s = \int_t^\infty \bigl(\bigl[x'(s)\bigr]^{p-1}\bigr)'\,\mathrm{d}s \\ &= \lim_{s\to\infty}\bigl[x'(s)\bigr]^{p-1} - \bigl[x'(t)\bigr]^{p-1}. \end{aligned}$$

Since $x'(t) > 0$ for $t \geqslant t_1$, we find

$$\begin{aligned}\big[x'(t)\big]^{p-1} &= \lim_{s\to\infty}\big[x'(s)\big]^{p-1} + \int_t^{\infty} c(s)\Phi\big(x\big(\tau(s)\big)\big)\,\mathrm{d}s\\ &\geqslant \int_t^{\infty} c(s)\big[x\big(\tau(s)\big)\big]^{p-1}\,\mathrm{d}s.\end{aligned}\tag{10.38}$$

It follows from (ii) of Lemma 10.4 that, for each $k_2 \in (0,1)$, there exists a $T_{k_2} \geqslant t_1$ such that

$$\big[x(t)\big]^{p-1} \geqslant k_2^{p-1}t^{p-1}\big[x'(t)\big]^{p-1} \geqslant k_2^{p-1}t^{p-1}\int_t^{\infty} c(s)\big[x\big(\tau(s)\big)\big]^{p-1}\,\mathrm{d}s,\tag{10.39}$$

for $t \geqslant T_{k_2}$. By (i) of Lemma 10.4, for each $k_1 \in (0,1)$, there exists a T_{k_1}, such that

$$\big[x\big(\tau(s)\big)\big]^{p-1} \geqslant k_1^{p-1}\left(\frac{\tau(t)}{t}\right)^{p-1}\big(x(t)\big)^{p-1},\tag{10.40}$$

for $t \geqslant T_{k_1}$. Then, by (10.39) and (10.40), for $t \geqslant t_4 := \max\{T_{k_1}, T_{k_2}\}$,

$$\begin{aligned}\big[x(t)\big]^{p-1} &\geqslant k_2^{p-1}t^{p-1}\int_t^{\infty} c(s)\big[x\big(\tau(s)\big)\big]^{p-1}\,\mathrm{d}s\\ &\geqslant k_1^{p-1}k_2^{p-1}t^{p-1}\int_t^{\infty} c(s)\left(\frac{\tau(s)}{s}\right)^{p-1}\big[x(s)\big]^{p-1}\,\mathrm{d}s\\ &\geqslant k^{2(p-1)}t^{p-1}\int_t^{\infty} c(s)\left(\frac{\tau(s)}{s}\right)^{p-1}\big[x(s)\big]^{p-1}\,\mathrm{d}s,\end{aligned}\tag{10.41}$$

where $k := \min\{k_1, k_2\}$. Since $x'(t) > 0$, it follows that

$$\begin{aligned}1 &\geqslant \frac{k^{2p-2}t^{p-1}}{[x(t)]^{p-1}}\int_t^{\infty} c(s)\left(\frac{\tau(s)}{s}\right)^{p-1}\big[x(s)\big]^{p-1}\,\mathrm{d}s\\ &\geqslant k^{2p-2}t^{p-1}\int_t^{\infty} c(s)\left(\frac{\tau(s)}{s}\right)^{p-1}\mathrm{d}s,\quad \text{for } t \geqslant t_4.\end{aligned}\tag{10.42}$$

Hence,

$$\limsup_{t\to\infty} t^{p-1}\int_t^{\infty} c(s)\left(\frac{\tau(s)}{s}\right)^{p-1}\mathrm{d}s := a < \infty.$$

Suppose that (10.35) holds, then there exists a sequence $\{s_n\}$ such that $\lim_{n\to\infty} s_n = \infty$ and

$$\lim_{n\to\infty} s_n^{p-1} \int_{s_n}^{\infty} c(s)\left(\frac{\tau(s)}{s}\right)^{p-1} \mathrm{d}s = a > 1.$$

For $\varepsilon_1 := (a-1)/2 > 0$, there exists an integer $N_1 > 0$ such that

$$\frac{a+1}{2} = a - \varepsilon_1 < s_n^{p-1} \int_{s_n}^{\infty} c(s)\left(\frac{\tau(s)}{s}\right)^{p-1} \mathrm{d}s, \tag{10.43}$$

for $n > N_1$. Choose k such that

$$\left(\frac{2}{a+1}\right)^{1/2(p-1)} < k < 1.$$

By (10.42) and (10.43),

$$1 \geqslant k^{2(p-1)} s_n^{p-1} \int_{s_n}^{\infty} c(s)\left(\frac{\tau(s)}{s}\right)^{p-1} \mathrm{d}s > \left(\frac{2}{a+1}\right)\left(\frac{a+1}{2}\right) = 1,$$

for s_n large enough. This contradiction shows that (10.35) does not hold.

Now, by $\gamma(t) \geqslant t$ and (10.39), we have

$$\left[x(t)\right]^{p-1} \geqslant k_2^{p-1} t^{p-1} \int_{\gamma(t)}^{\infty} c(s)\left[x\left(\tau(s)\right)\right]^{p-1} \mathrm{d}s,$$

for $t \geqslant T_{k_2}$. Since $x(t)$ is increasing and $\tau(s) \geqslant t$ for $s \geqslant \gamma(t)$, it follows that

$$\begin{aligned}\left[x(t)\right]^{p-1} &\geqslant k_2^{p-1} t^{p-1} \int_{\gamma(t)}^{\infty} c(s)\left[x\left(\tau(s)\right)\right]^{p-1} \mathrm{d}s \\ &\geqslant k_2^{p-1} t^{p-1} \left[x(t)\right]^{p-1} \int_{\gamma(t)}^{\infty} c(s)\, \mathrm{d}s.\end{aligned}$$

Dividing $[x(t)]^{p-1}$ in both sides of the above inequality, we get

$$k_2^{p-1} t^{p-1} \int_{\gamma(t)}^{\infty} c(s)\, \mathrm{d}s \leqslant 1, \tag{10.44}$$

for $t \geqslant T_{k_2}$. Thus,

$$\limsup_{t\to\infty} t^{p-1} \int_{\gamma(t)}^{\infty} c(s)\, \mathrm{d}s := b < \infty.$$

Suppose that (10.36) holds. Then there exists a sequence $\{t_n\}$ with $\lim_{n\to\infty} t_n = \infty$ such that

$$\lim_{n\to\infty} t_n^{p-1} \int_{\gamma(t_n)}^{\infty} c(s)\,\mathrm{d}s = b > 1.$$

Thus, for $\varepsilon_2 := (b-1)/2 > 0$, there exists an integer $N_2 > 0$ such that

$$\frac{b+1}{2} = b - \varepsilon_2 < t_n^{p-1} \int_{\gamma(t_n)}^{\infty} c(s)\,\mathrm{d}s, \tag{10.45}$$

for $n > N_2$. Choose $k_2 \in (\frac{2}{(b+1)^{q-1}}, 1)$. By (10.44) and (10.45),

$$1 \geqslant k_2^{p-1} t_q^{p-1} \int_{\gamma(t_n)}^{\infty} c(s)\,\mathrm{d}s > \frac{2}{b+1}\,\frac{b+1}{2} = 1,$$

for t_n large enough. This contradiction proves that (10.36) does not hold. □

EXAMPLE 10.2. Consider the functional differential equation

$$\big[\Phi(x')\big]' + \frac{2^p(p-1)}{t^p}\Phi\big(x(t/2)\big) = 0. \tag{10.46}$$

Since

$$\begin{aligned}
&\limsup_{t\to\infty} t^{p-1} \int_t^{\infty} \frac{2^p(p-1)}{s^p}\left(\frac{(s/2)}{s}\right)^{p-1} \mathrm{d}s \\
&\quad = 2(p-1) \limsup_{t\to\infty} t^{p-1} \int_t^{\infty} \frac{1}{s^p}\,\mathrm{d}s \\
&\quad = 2(p-1) \limsup_{t\to\infty} t^{p-1}\left(\frac{1}{(p-1)t^{p-1}}\right) = 2 > 1,
\end{aligned}$$

it follows from (10.35) of Theorem 10.6 that (10.46) is oscillatory. In fact, if the coefficient $\frac{2^p(p-1)}{t^p}$ of (10.46) is replaced by $\frac{k}{t^p}$ with $k > 2^{p-1}(p-1)$, (10.46) will again be oscillatory.

In the next theorem we assume that $\tau(t) > 0$ and we denote

$$\mu(t) = \left(\frac{\tau(t)}{t}\right)^{p-1}.$$

THEOREM 10.7. *Equation* (10.34) *is oscillatory if the differential equation*

$$\big(\Phi(x')\big)' + \lambda\mu(t)c(t)\Phi(x) = 0 \tag{10.47}$$

is oscillatory for some $\lambda \in (0, 1)$.

PROOF. By contradiction, suppose that there exists eventually positive solution x of (10.34) and we may also assume that $x(\tau(t)) > 0$ on $[t_1, \infty)$ for some $t_1 \geqslant t_0$. Then $x''(t) \leqslant 0$, $x'(t) > 0$ on $[t_2, \infty)$ for some $t_2 \geqslant t_1$. Since $\lambda \in (0, 1)$, it follows from (a) of Lemma 10.4 that

$$x\bigl(\tau(t)\bigr) \geqslant \lambda^{1/(p-1)} x(t) \frac{\tau(t)}{t},$$

for t large enough. Thus,

$$\bigl[x\bigl(\tau(t)\bigr)\bigr]^{p-1} \geqslant \lambda\bigl[x(t)\bigr]^{p-1} \left(\frac{\tau(t)}{t}\right)^{p-1}, \tag{10.48}$$

for t large enough. Let

$$w(t) = \frac{\Phi(x'(t))}{\Phi(x(t))}.$$

Then, by (10.48),

$$\begin{aligned}
& w'(t) + \frac{\lambda[\tau(t)]^{p-1}}{t^{p-1}} c(t) + (p-1)\bigl|w(t)\bigr|^q \\
&\quad = \frac{[(x'(t))^{p-1}]'[x(t)]^{p-1} - [x'(t)]^{p-1}[(x(t))^{p-1}]'}{[(x(t))^{p-1}]^2} \\
&\qquad + \frac{\lambda(\tau(t))^{p-1}}{t^{p-1}} c(t) + (p-1)\left(\frac{[x'(t)]^{p-1}}{[x(t)]^{p-1}}\right)^q \\
&\quad = \frac{[(x'(t))^{p-1}]'[x(t)]^{p-1} - (p-1)[x'(t)]^{p-1}[x(t)]^{p-2}x'(t)}{[(x(t))^{p-1}]^2} \\
&\qquad + \frac{\lambda(\tau(t))^{p-1}}{t^{p-1}} c(t) + \frac{(p-1)[x'(t)]^p}{[x(t)]^p} \\
&\quad = \frac{[(x'(t))^{p-1}]'}{[x(t)]^{p-1}} - \frac{(p-1)[x'(t)]^p}{[x(t)]^p} + \frac{\lambda(\tau(t))^{p-1}}{t^{p-1}} c(t) + \frac{(p-1)[x'(t)]^p}{[x(t)]^p} \\
&\quad = \frac{-c(t)|x(\tau(t))|^{p-2}x(\tau(t))}{[x(t)]^{p-1}} + \frac{\lambda(\tau(t))^{p-1}c(t)}{t^{p-1}} \\
&\quad = \frac{c(t)}{[x(t)]^{p-1}} \left(\frac{\lambda[\tau(t)]^{p-1}}{t^{p-1}} \bigl[x(t)\bigr]^{p-1} - \bigl[x\bigl(\tau(t)\bigr)\bigr]^{p-1}\right) \\
&\quad \leqslant 0.
\end{aligned}$$

This and Theorem 5.3 imply that (10.47) is nonoscillatory, but this is a contradiction. Hence, (10.34) is oscillatory. □

REMARK 10.2. Theorem 10.7 is an extension of Theorem 2.2 of [96].

THEOREM 10.8. *If*

$$\limsup_{t\to\infty}\int^t c(s)\left(\frac{\tau(s)}{s}\right)^{p-1}\mathrm{d}s=\infty, \tag{10.49}$$

then Equation (10.34) *is oscillatory.*

PROOF. Suppose to the contrary that (10.34) has a nonoscillatory solution $x(t)$ which may be assumed to be eventually positive. As in the proof of Theorem 10.6, there exists $t_1 > t_0$ such that $x(\tau(t)) > 0$, $x'(t) > 0$, and $x''(t) < 0$ for $t > t_1$. By (i) of Lemma 10.4, there exists $t_2 \geqslant t_1$ such that

$$x\big(\tau(t)\big)\geqslant\left(\frac{1}{2}\right)^{1/(p-1)}\frac{\tau(t)}{t}x(t),$$

or

$$\big(x\big(\tau(t)\big)\big)^{p-1}\geqslant\frac{1}{2}\left(\frac{\tau(t)}{t}\right)^{p-1}\big[x(t)\big]^{p-1}$$

for $t \geqslant t_2$. Since $x'(t) > 0$ and $x(\tau(t)) > 0$ for $t \geqslant t_1$,

$$-\big(\big(x'(t)\big)^{p-1}\big)'=c(t)\big[x\big(\tau(t)\big)\big]^{p-1}\geqslant\frac{1}{2}c(t)\left(\frac{\tau(t)}{t}\right)^{p-1}\big[x(t)\big]^{p-1}$$

for $t \geqslant t_2$. Integrating the above inequality from t_2 to t and using the increasing property of $x(t)$, we get

$$\begin{aligned}\big[x'(t)\big]^{p-1}-\big[x'(t_2)\big]^{p-1}&\leqslant-\frac{1}{2}\int_{t_2}^{t}c(s)\left(\frac{\tau(s)}{s}\right)^{p-1}\big[x(s)\big]^{p-1}\,\mathrm{d}s\\&\leqslant-\frac{1}{2}\big[x(t_2)\big]^{p-1}\int_{t_2}^{t}c(s)\left(\frac{\tau(s)}{s}\right)^{p-1}\mathrm{d}s,\end{aligned}$$

or

$$\big[x'(t)\big]^{p-1}\leqslant\big[x'(t_2)\big]^{p-1}-\frac{1}{2}\big[x(t_2)\big]^{p-1}\int_{t_2}^{t}c(s)\left(\frac{\tau(t)}{t}\right)^{p-1}\mathrm{d}s$$

for $t \geqslant t_2$. This and (10.49) imply $[x'(t)]^{p-1} < 0$ for t large enough. This is a contradiction. Thus, (10.34) is oscillatory. □

EXAMPLE 10.3. Consider the functional differential equation

$$\left[\Phi\big(x'(t)\big)\right]' + \Phi\big(x(t/2)\big) = 0 \tag{10.50}$$

where $p > 1$. Clearly,

$$\limsup_{t\to\infty} \int^t \left(\frac{(s/2)}{s}\right)^{p-1} \mathrm{d}s = \frac{1}{2^{p-1}} \limsup_{t\to\infty} \int^t \mathrm{d}s = \infty.$$

Thus, it follows from Theorem 10.8 that (10.50) is oscillatory.

11. Half-linear Sturm–Liouville problem

In this section we show that the solutions of the Sturm–Liouville problem for half-linear equation (0.1) have similar properties as in the linear case. Of course, we cannot consider the problem of orthogonality of eigenfunctions since this concept has no meaning in L^p, $p \neq 2$. As far as we know, an open problem is whether the system of eigenfunctions is complete in L^p.

11.1. *Basic Sturm–Liouville problem*

We start with the problem

$$\big(\Phi(x')\big)' + \lambda c(t)\Phi(x) = 0, \quad x(a) = 0 = x(b), \tag{11.1}$$

under the assumption that $c(t) \geqslant 0$ and $c(t) \not\equiv 0$. The value λ is called the *eigenvalue* if there exists a nontrivial solution x of (11.1). The solution x is said to the *eigenfunction* corresponding to the eigenvalue λ. Clearly, according to the assumption $c(t) \geqslant 0$ and the Sturm comparison theorem, only values $\lambda > 0$ can be eigenvalues. The main statement of this subsection is taken from the classical paper of Elbert [85].

THEOREM 11.1. *The eigenvalue problem* (11.1) *has infinitely many eigenvalues* $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots$, $\lambda_n \to \infty$ *as* $n \to \infty$. *The nth eigenfunction has exactly* $n-1$ *zeros in* (a, b). *Moreover, if the function c is supposed to be positive in the whole interval* (a, b), *the eigenvalues satisfy the asymptotic relation*

$$\lim_{n\to\infty} \frac{\sqrt[p]{\lambda_n}}{n} = \frac{\pi_p}{\int_a^b \sqrt[p]{c(t)}\,\mathrm{d}t}. \tag{11.2}$$

PROOF. The proof of the first part of the theorem is a special case of the problem treated in the next subsection, so we present only its main idea. Let $x(t;\lambda)$ be a solution of (11.1) given by the initial condition $x(a;\lambda)=0$, $x'(a;\lambda)=1$ and let $\varphi(t;\lambda)$ be the continuous function given in all points where $x'(t;\lambda)\neq 0$ by the formula

$$\varphi(t;\lambda)=\arctan_p\frac{x(t;\lambda)}{x'(t;\lambda)},$$

i.e., $\varphi(t;\lambda)$ is the angular variable in the half-linear Prüfer transformation for $x(t;\lambda)$. This means that $\varphi(t;\lambda)$ satisfy the differential equation

$$\varphi'=\left|S'(\varphi)\right|^p+\frac{\lambda c(t)}{p-1}\left|S(\varphi)\right|^p,\quad \varphi(a;\lambda)=0, \tag{11.3}$$

where S is the half-linear sine function. The proof is based on the fact that $\varphi(b;\lambda)$ is a continuous function of λ and $\varphi(b;\lambda)\to\infty$ as $\lambda\to\infty$. The continuity property follows from the general theory of continuous dependence of solutions of first order differential equations on the right-hand side. The limit property of $\varphi(b;\lambda)$ is proved via the comparison of (11.1) with the "minorant" problem with a constant coefficient

$$\bigl(\Phi(y')\bigr)'+\lambda\bar{c}\Phi(y)=0,\quad y(a')=0=y(b'), \tag{11.4}$$

where $[a',b']\subset[a,b]$ is such that $c(t)>0$ on $[a',b']$ and $\bar{c}=\min_{t\in[a',b']}c(t)>0$. The eigenvalues and eigenfunctions of (11.4) can be computed explicitly and if $\theta(t;\lambda)$ is defined for this problem in the same way as $\varphi(t;\lambda)$ for (11.1), we have

$$\varphi(b;\lambda)\geqslant\theta(b;\lambda)\to\infty,\quad \text{as } \lambda\to\infty.$$

Now, the eigenvalues are those $\lambda=\lambda_n$ for which $\varphi(b;\lambda_n)=n\pi_p$ and taking into account that $\varphi(t;\lambda)$ is increasing in t (this follows from (11.3)), zero points of the associated eigenfunction $x_n(t)=x(t;\lambda_n)$ are those t_k, $k=1,\dots,n-1$, for which $\varphi(t_k,\lambda_n)=k\pi_p$, $k=1,\dots,n-1$.

Concerning the proof of the asymptotic formula (11.2), first suppose that the function $c(t)\equiv c_1>0$ is a constant function and consider the problem

$$\bigl(\Phi(z')\bigr)'+\lambda c_1\Phi(z)=0,\quad z(a)=0=z(b). \tag{11.5}$$

A nontrivial solution of the half-linear equation in this problem satisfying $z(a)=0$ is $z=S(\sqrt[p]{\lambda c_1}(t-a))$ and hence the k-th eigenvalue is given by $\sqrt[p]{\lambda_k c_1}(b-a)=k\pi_p$, thus

$$\frac{\sqrt[p]{\lambda_k}}{k}=\frac{\pi_p}{\int_a^b\sqrt[p]{c_1}\,\mathrm{d}t},$$

i.e., the asymptotic formula (11.2) is automatically satisfied in this case.

Now let us consider the original Sturm–Liouville problem (11.1) and its k-th eigenfunction $x(t;\lambda_k)$. This function has zeros at $t_0=a<t_1<t_2<\cdots<t_{k-1}<t_k=b$. Put $\lambda=\lambda_k$ in (11.1) and in (11.5) and define

$$c_{1,i}:=\min_{t_{i-1}\leqslant t\leqslant t_i}c(t),\qquad c_{2,i}:=\max_{t_{i-1}\leqslant t\leqslant t_i}c(t),\quad i=1,\dots,k.$$

Then the differential equation in (11.1) is a Sturmian majorant of the differential equation in (11.5) wit $c_1 = c_{1,i}$ on the interval $[t_{i-1}, t_i]$. Hence the solution $S(\sqrt[p]{\lambda_k c_{1,i}}(t - t_{i-1})$ of (11.5) has no zero on (t_{i-1}, t_i) so that

$$\sqrt[p]{\lambda_k c_{1,i}}(t_i - t_{i-1}) \leqslant \pi_p. \tag{11.6}$$

By a similar argument we have $\pi_p \leqslant \sqrt[p]{\lambda_k c_{2,i}}(t_i - t_{i-1})$. On the other hand,

$$\int_{t_{i-1}}^{t_i} \sqrt[p]{\lambda_k c_{1,i}}\,\mathrm{d}t \leqslant \int_{t_{i-1}}^{t_i} \sqrt[p]{\lambda_k c(t)}\,\mathrm{d}t \leqslant \int_{t_{i-1}}^{t_i} \sqrt[p]{\lambda_k c_{2,i}}\,\mathrm{d}t,$$

and consequently

$$\left|\pi_p - \int_{t_{i-1}}^{t_i} \sqrt[p]{\lambda_k c(t)}\right| \leqslant \int_{t_{i-1}}^{t_i} \sqrt[p]{\lambda_k}\left(\sqrt[p]{c_{2,i}} - \sqrt[p]{c_{1,i}}\right)\mathrm{d}t. \tag{11.7}$$

Let $\omega(f, \delta)$ be defined for any continuous function f on $[a, b]$ by

$$\omega(f, \delta) = \max\bigl\{\bigl|f(\tau_1) - f(\tau_2)\bigr|\colon\ |\tau_1 - \tau_2| \leqslant \delta,\ \tau_1, \tau_2 \in [a, b]\bigr\}.$$

Making use of this definition we deduce from (11.7) that

$$\left|\frac{\pi_p}{\sqrt[p]{\lambda_k}} - \int_{t_{i-1}}^{t_i} \sqrt[p]{c(t)}\,\mathrm{d}t\right| \leqslant \omega\left(\sqrt[p]{c}, t_i - t_{i-1}\right)|t_i - t_{i-1}|.$$

Let $c_1 = \min_{i=1,\dots,k} c_{1,i}$. Then by (11.6)

$$\left|\frac{k\pi_p}{\sqrt[p]{\lambda_k}} - \int_a^b \sqrt[p]{c(t)}\,\mathrm{d}t\right| \leqslant \omega\left(\sqrt[p]{c}, \frac{\pi_p}{\sqrt[p]{\lambda_k c_1}}\right)(b - a).$$

By the first part of the proof $\lambda_k \to \infty$ as $k \to \infty$. Therefore the continuity of the function c yields the formula which has to be proved. ☐

11.2. *Regular problem with indefinite weight*

The results presented in this subsection can be found in [133]. We consider the Sturm-Liouville problem

$$\begin{cases} \bigl(r(t)\Phi(x')\bigr)' + \lambda c(t)\Phi(x) = 0, \\ Ax(a) - A'x'(a) = 0, \quad Bx(b) + B'x'(b) = 0. \end{cases} \tag{11.8}$$

It is supposed that r, c are continuous in $[a, b]$ and $r(t) > 0$ in this interval. No sign restriction on the function c is supposed. A, A', B, B' are real numbers such that $A^2 + A'^2 > 0$, $B^2 + B'^2 > 0$, λ is a real-valued eigenvalue parameter.

THEOREM 11.2. *Suppose that* $AA' \geqslant 0$, $BB' \geqslant 0$ *and* $A^2 + B^2 > 0$. *Further suppose that the function c takes both positive and negative values in* $[a, b]$. *Then the totality of eigenvalues of* (11.8) *consists of two sequences* $\{\lambda_n^+\}_{n=0}^\infty$ *and* $\{\lambda_n^-\}_{n=0}^\infty$ *such that*

$$\cdots < \lambda_n^- < \cdots < \lambda_1^- < \lambda_0^- < 0 < \lambda_0^+ < \lambda_1^+ < \cdots < \lambda_n^+ < \cdots$$

and

$$\lim_{n\to\infty} \lambda_n^+ = \infty, \qquad \lim_{n\to\infty} \lambda_n^- = -\infty.$$

The eigenfunctions $x = x(t; \lambda_n^+)$ *and* $x = x(t; \lambda_n^-)$ *associated with* $\lambda = \lambda_n^+$ *and* λ_n^- *have exactly n zeros in* (a, b).

PROOF. The proof is again based on the half-linear Prüfer transformation. Let $\lambda \in \mathbb{R}$ and let $x(t; \lambda)$ be the solution of

$$\big(r(t)\Phi(x')\big)' + \lambda c(t)\Phi(x) = 0 \tag{11.9}$$

satisfying the initial conditions $x(a) = A'$, $x'(a) = A$. Note that this solution satisfies the boundary condition $Ax(a) - A'x'(a) = 0$. According to the continuous dependence of solutions on a perturbation of the functions r, c, in (0.1), the function $x(t; \lambda)$ depends continuously on λ. In particular, if $\lambda_i \to \lambda$ as $i \to \infty$, then $x(t; \lambda_i) \to x(t; \lambda)$ uniformly on $[a, b]$ as $i \to \infty$. If $x(t; \lambda)$ satisfies the second part of the boundary conditions $Bx(b) + B'x'(b) = 0$ for some $\lambda \in \mathbb{R}$, then λ is an eigenvalue and $x(t; \lambda)$ is the corresponding eigenfunction.

For $\lambda = 0$ we can compute $x(t; \lambda)$ explicitly,

$$x(t; 0) = A' + r^{q-1}(a)A \int_a^t r^{1-q}(s)\,\mathrm{d}s$$

and it is easy to see that this solution does not satisfy the condition at $t = b$, so $\lambda = 0$ is not an eigenvalue.

In what follows we suppose that $\lambda > 0$. For the solution $x(t; \lambda)$ we perform slightly modified Prüfer transformation, we express $x(t; \lambda)$ and its quasiderivative in the form

$$\begin{aligned} x(t; \lambda) &= \rho(t; \lambda)S\big(\varphi(t; \lambda)\big), \\ r^{q-1}(t)x'(t; \lambda) &= \lambda^{q-1}\rho(t; \lambda)C\big(\varphi(t; \lambda)\big). \end{aligned} \tag{11.10}$$

Here S, C are the generalized half-linear sine and cosine functions introduced in Section 1.1. The function $\rho(t; \lambda)$ is given by

$$\rho(t; \lambda) = \left[\big|x(t; \lambda)\big|^p + \left(\frac{r(t)}{\lambda}\right)^q \big|x'(t; \lambda)\big|^p\right]^{\frac{1}{p}}.$$

The functions ρ and φ satisfy the first order system

$$\begin{aligned}\varphi' &= \left(\frac{\lambda}{r(t)}\right)^{q-1}\left|C(\varphi)\right|^p + \frac{c(t)}{p-1}\left|S(\varphi)\right|^p,\\ \rho' &= \rho\left[\left(\frac{\lambda}{r(t)}\right)^{q-1} - \frac{c(t)}{p-1}\right]\Phi\big(S(\varphi)\big)C(\varphi)\end{aligned}\tag{11.11}$$

with the initial conditions

$$\begin{aligned}\rho(a;\lambda) &= \left[|A'|^p + \left(\frac{r(a)}{\lambda}\right)^q |A|^p\right]^{\frac{1}{p}},\\ \varphi(a;\lambda) &= \operatorname{arctan}_p\left(\left(\frac{\lambda}{r(a)}\right)^{\frac{1}{p}}\frac{A'}{A}\right),\end{aligned}\tag{11.12}$$

where $\tan_p = S/C = \sin_p/\cos_p$. Since $AA' \geqslant 0$, we may assume without loss of generality that

$$0 \leqslant \varphi(a;\lambda) < \frac{\pi_p}{2}, \quad \text{if } A \neq 0, \tag{11.13}$$

$$\varphi(a,\lambda) = \frac{\pi_p}{2}, \quad \text{if } A = 0. \tag{11.14}$$

Observe that as soon as $\varphi(t;\lambda)$ is known, $\rho = \rho(t;\lambda)$ can be computed explicitly and

$$\rho(t;\lambda) = \rho(a,\lambda)\exp\left\{\int_a^t \left[\left(\frac{\lambda}{r(s)}\right)^{q-1} - \frac{c(s)}{p-1}\right]\Phi\big(S\big(\varphi(s;\lambda)\big)\big)C\big(\varphi(s;\lambda)\big)\,\mathrm{d}s\right\}.$$

Thus, it is important to discuss the initial value problem (11.11), (11.12). We denote by $f(t,\varphi,\lambda)$ the right-hand side of (11.11). It is clear that, for each $\lambda > 0$, the function $f(t,\varphi,\lambda)$ is bounded for $t \in [a,b]$ and $\varphi \in \mathbb{R}$. In view of the Pythagorean identity (1.5) the function $f(t,\varphi,\lambda)$ can be written in the form

$$f(t,\varphi,\lambda) = \left(\frac{\lambda}{r(t)}\right)^{q-1} + \left\{-\left(\frac{\lambda}{r(t)}\right)^{q-1} + \frac{c(t)}{p-1}\right\}\left|S(\varphi)\right|^p.$$

Similarly as in the standard half-linear Prüfer transformation, the function $f(t,\varphi,\lambda)$ is Lipschitzian in φ, hence unique solvability is guaranteed and the solution $\varphi = \varphi(t;\lambda)$ depends continuously on $(t,\lambda) \in [a,b] \times (0,\infty)$.

It is easy to see that $\lambda > 0$ is an eigenvalue of (11.8) if and only if λ satisfies

$$\varphi(b;\lambda) = \operatorname{arctan}_p\left(-\left(\frac{\lambda}{r(b)}\right)^{q-1}\frac{B'}{B}\right) + (n+1)\pi_p \tag{11.15}$$

for some $n \in \mathbb{Z}$. Here, by virtue of $BB' \geqslant 0$, we assume without loss of generality that the value of the function $\arctan_p$ in (11.15) is in $(-(\pi_p/2), 0]$ if $B \neq 0$ and equals $-(\pi_p/2)$ if $B = 0$.

Observe that the function $\varphi(b;\lambda)$ is strictly increasing for $\lambda \in (0,\infty)$. Indeed, denote as before $f(t,\varphi,\lambda)$ the right-hand side of (11.11). Clearly, $f(t,\varphi,\lambda)$ is nondecreasing function of $\lambda \in (0,\infty)$, and, since $AA' \geqslant 0$, the initial value $\varphi(a;\lambda)$ given by (11.12) is also nondecreasing for $\lambda \in (0,\infty)$. Then a standard comparison theorem for the first order scalar differential equations implies that $\varphi(t;\lambda)$ is a nondecreasing function of $\lambda \in (0,\infty)$ for each fixed $t \in [a,b]$. Now, let $0 < \lambda < \mu$ be fixed. Since the function $\varphi(t;\lambda)$ is nondecreasing with respect to λ, we have $\varphi(t;\lambda) \leqslant \varphi(t;\mu)$. Assume that $\varphi(t;\lambda) \equiv \varphi(t;\mu)$ for all $t \in (a,b)$. Then $\varphi'(t;\lambda) \equiv \varphi'(t;\mu)$, and so we have $f(t,\varphi(t;\lambda),\lambda) \equiv f(t,\varphi(t;\mu),\mu)$ from which follows $C(\varphi(t;\lambda)) \equiv C(\varphi(t;\mu)) \equiv 0$. This implies that $\varphi(t;\lambda) \equiv (m+\frac{1}{2})\pi_p$ for some integer $m \in \mathbb{Z}$, and hence, by Equation (11.11), $c(t) \equiv 0$ for $t \in (a,b)$. This is a contradiction to the assumption that $c(t) > 0$ for some $t \in [a,b]$. Therefore we have $\varphi(t_0;\lambda) < \varphi(t_0;\mu)$ for some $t_0 \in (a,b)$. Then applying a standard comparison theorem again, we conclude that $\varphi(b;\lambda) < \varphi(b;\mu)$.

Now we claim that $x(t;\lambda)$ has no zeros in the interval $(a,b]$ for all sufficiently small $\lambda > 0$. As stated before, $x(t;\lambda) \to x(t;0)$ as $\lambda \to 0+$ uniformly on $[a,b]$. We note that $x(t;\lambda)$ satisfies

$$\begin{aligned} x(t;\lambda) = A' + \int_a^t &\left| \frac{r(a)}{r(s)}\Phi(A) - \frac{\lambda}{r(s)} I(s;\lambda) \right|^{q-2} \\ &\times \left\{ \frac{r(a)}{r(s)}\Phi(A) - \frac{\lambda}{r(s)} I(s;\lambda) \right\} \mathrm{d}s, \end{aligned}$$

for all $a \leqslant t \leqslant b$, where

$$I(s;\lambda) = \int_a^s c(\tau)\Phi\big(x(\tau;\lambda)\big)\,\mathrm{d}\tau, \quad a \leqslant s \leqslant b.$$

Then it is easy to find that if $A = 0$ or $AA' > 0$, then $x(t;\lambda)$ has no zero in the closed interval $[a,b]$ for all sufficiently small $\lambda > 0$, and that if $A \neq 0$ and $A' = 0$, then $x(t;\lambda)$ has no zero in the interval $(a,b]$ for all sufficiently small $\lambda > 0$. Further, since

$$r(t)\Phi\big(x'(t;\lambda)\big) = r(a)\Phi(A) - \lambda \int_a^t c(s)\Phi\big(x(s;\lambda)\big)\,\mathrm{d}s$$

for $a \leqslant t \leqslant b$, we see that if $A \neq 0$, then $x'(t;\lambda)$ has no zeros in $[a,b]$ for all sufficiently small $\lambda > 0$.

Next we claim that the number of zeros of $x(t;\lambda)$ in $[a,b]$ can be made as large as possible if $\lambda > 0$ is chosen sufficiently large. To this end, we consider the equation

$$\big(\Phi(x')\big)' + (p-1)\mu^p \Phi(x) = 0,$$

where $\mu > 0$ is a constant. Clearly, $S(\mu t)$ is a solution of this equation, and has zeros $t = j\pi_p/\mu$, $j \in \mathbb{Z}$, $S(\cdot)$ is the generalized sine function. Since c is supposed to be positive at some $t \in [a, b]$, there exists $[a', b'] \subset [a, b]$ such that $c(t) > 0$ on $[a', b']$. Let $k \in \mathbb{N}$ be any given positive integer and take $\mu > 0$ so that $S(\mu t)$ has at least $k + 1$ zeros in $[a', b']$. Let $r^* > 0$ and $\lambda_* > 0$ be numbers such that

$$r^* = \max_{t\in[a',b']} r(t), \qquad \lambda_* \min_{t\in[a',b']} c(t) = (p-1)r^*\mu^p.$$

Then, comparing the half-linear equation in (11.8) with $\lambda > \lambda_*$ and the equation

$$\big(r^*\Phi(x')\big)' + (p-1)r^*\mu^p\Phi(x) = 0, \quad a' \leqslant t \leqslant b',$$

we conclude by the Sturm comparison theorem that all solution of the equation in (11.8) with $\lambda > \lambda_*$ have at least k zeros in $[a, b]$. Since k was arbitrary, this shows that the number of zeros of $x(t; \lambda)$ in $[a, b]$ can be made as large as possible if $\lambda > 0$ is chosen sufficiently large.

Since the radial variable $\rho(t; \lambda) > 0$, it follows from (11.10) that $x(t; \lambda)$ has a zero at $t = c$ if and only if there exists $j \in \mathbb{Z}$ such that $\varphi(c; \lambda) = j\pi_p$. Moreover, if $\varphi(c; \lambda) = j\pi_p$, then by (11.11) we have $\varphi'(c; \lambda) = (\lambda/r(c))^{q-1} > 0$. Therefore we easily see that if $\varphi(c; \lambda) = j\pi_p$, then $\varphi(t; \lambda) > j\pi_p$ for $c < t \leqslant b$. Consequently, we have

(i) For all $\lambda > 0$ sufficiently small

$$0 < \varphi(b; \lambda) < \frac{\pi_p}{2}, \quad \text{if } A \neq 0,$$
$$0 < \varphi(b; \lambda) < \pi_p, \quad \text{if } A = 0.$$

(ii) $\lim_{\lambda\to\infty} \varphi(b; \lambda) = \infty$.

Now we seek $\lambda > 0$ satisfying (11.15) for some $n \in \mathbb{Z}$. The left-hand side $\varphi(b; \lambda)$ of (11.15) is a continuous function of $\lambda \in (0, \infty)$, and it is strictly increasing for $\lambda \in (0, \infty)$, moreover, it has the following properties

$$0 \leqslant \lim_{\lambda\to 0+} \varphi(b; \lambda) < \frac{\pi_p}{2}, \quad \text{if } A \neq 0,$$
$$0 \leqslant \lim_{\lambda\to 0+} \varphi(b; \lambda) < \pi_p, \quad \text{if } A = 0,$$

and

$$\lim_{\lambda\to\infty} \varphi(b; \lambda) = \infty.$$

On the other hand, by virtue of $BB' \geqslant 0$, the right-hand side of (11.15) is a nonincreasing function of $\lambda \in (0, \infty)$ for each $n \in \mathbb{Z}$. More precisely, in case $BB' > 0$, it is strictly decreasing and varies from $(n + 1)\pi_p$ to $(n + \frac{1}{2})\pi_p$ as λ varies from 0 to ∞. In the case $B' = 0$, it is the constant function $(n + \frac{1}{2})\pi_p$.

From what was observed above we find that, for each $n = 0, 1, 2, \ldots$, there exists a unique $\lambda_n^+ > 0$ such that

$$\varphi\bigl(b; \lambda_n^+\bigr) = \arctan_p\left(-\left(\frac{\lambda_n^+}{r(b)}\right)^{q-1}\frac{B'}{B}\right) + (n+1)\pi_p.$$

Then, each λ_n^+ is an eigenvalue of (11.8), and the associated eigenfunction $x(t; \lambda_n^+)$ has exactly n zeros in the open interval (a, b), where $n = 0, 1, 2, \ldots$. It is clear that

$$\lambda_0^+ < \lambda_1^+ < \cdots < \lambda_n^+ < \cdots, \qquad \lim_{n\to\infty} \lambda_n^+ = \infty.$$

The proof concerning the sequence of negative eigenvalues λ_n^- and the number of zeros of associated eigenfunctions can be proved in the same way. □

11.3. *Singular Sturm–Liouville problem*

The results of this subsection are taken from the paper [91], for related results we refer to [133,134]. We consider the equation

$$\bigl(\Phi(x')\bigr)' + \lambda c(t)\Phi(x) = 0, \quad t \in [a, \infty), \tag{11.16}$$

where $\lambda > 0$ is a real-valued parameter and c is a nonnegative piecewise continuous eventually nonvanishing function. A solution $x_0 = x_0(t; \lambda)$ of (11.16) is said to be *subdominant* if

$$\lim_{t\to\infty} x_0(t; \lambda) = k_0, \tag{11.17}$$

for some constant $k_0 \neq 0$, and a solution $x_1 = x_1(t; \lambda)$ is said to be *dominant* if

$$\lim_{t\to\infty} \bigl[x_1(t; \lambda) - k_1(t-a)\bigr] = 0 \tag{11.18}$$

for some constant $k_1 \neq 0$. We will show that the subdominant and dominant solutions are essentially unique in the sense that if $\tilde{x}_0(t; \lambda)$ and $\tilde{x}_1(t; \lambda)$ denote the solutions of (11.16) satisfying

$$\lim_{t\to\infty} \tilde{x}_0(t; \lambda) = 1 \tag{11.19}$$

and

$$\lim_{t\to\infty} \bigl[\tilde{x}_1(t; \lambda) - (t-a)\bigr] = 0 \tag{11.20}$$

then $x_0(t; \lambda) = k_0\tilde{x}_0(t; \lambda)$ and $x_1(t; \lambda) = k_1\tilde{x}_1(t; \lambda)$. According to the results presented in Section 6, any eventually positive solution (11.16) has one of the following asymptotic behavior:

(i) $\lim_{t\to\infty} \Phi(x'(t;\lambda)) = \text{const} > 0$;
(ii) $\lim_{t\to\infty} \Phi(x'(t;\lambda)) = 0$, $\lim_{t\to\infty} x(t;\lambda) = 0$;
(iii) $\lim_{t\to\infty} \Phi(x'(t;\lambda)) = 0$, $\lim_{t\to\infty} x(t;\lambda) = \text{const} > 0$.

In view of this result, the dominant and subdominant solutions investigated in this subsection correspond to cases (i) and (iii), respectively.

The proofs of three statements presented in this subsection are rather complex, so we skip them and we refer to the above mentioned paper [91]. We note only that these proofs are again based on the half-linear Prüfer transformation, this time combined with a detailed asymptotic analysis of solutions of (11.16).

THEOREM 11.3. *Suppose that*

$$\int^{\infty} \left(\int_t^{\infty} c(s)\,\mathrm{d}s \right)^{q-1} \mathrm{d}t < \infty.$$

Then for every λ *Equation* (11.16) *has a unique solution* $\tilde{x}_0(t;\lambda)$ *satisfying* (11.19) *and there exists a sequence* $\{\lambda_n^{(0)}\}_{n=0}^{\infty}$ *of positive parameters with the properties that*

(i) $0 = \lambda_0^{(0)} < \lambda_1^{(0)} < \cdots < \lambda_n^{(0)} < \cdots$, $\lim_{n\to\infty} \lambda_n^{(0)} = \infty$;
(ii) *for* $\lambda \in (\lambda_{n-1}^{(0)}, \lambda_n^{(0)})$, $n = 1, 2, \dots$, $\tilde{x}_0(t;\lambda)$ *has exactly* $n-1$ *zeros in* (a,∞) *and* $\tilde{x}_0(a;\lambda) \neq 0$;
(iii) *for* $\lambda = \lambda_n^{(0)}$, $n = 1, 2, \dots$, $\tilde{x}_0(t;\lambda)$ *has exactly* $n-1$ *zeros in* (a,∞) *and* $\tilde{x}_0(a;\lambda) = 0$.

THEOREM 11.4. *Let the sequence* $\{\lambda_n^{(0)}\}_{n=0}^{\infty}$ *be defined as in the previous theorem. Then the number of zeros of any nontrivial solution* $x(t;\lambda)$ *on* $[a,\infty)$ *can be*

(i) *exactly* n *if* $\lambda = \lambda_n^{(0)}$, $n = 1, 2, \dots$;
(ii) *either* $n-1$ *or* n *if* $\lambda_{n-1}^{(0)} < \lambda < \lambda_n^{(0)}$, *and both cases occur.*

THEOREM 11.5. *Suppose that*

$$\int^{\infty} t^p c(t)\,\mathrm{d}t < \infty.$$

Then for every $\lambda > 0$ *Equation* (11.16) *has a unique solution* $\tilde{x}_1(t;\lambda)$ *satisfying* (11.20) *and there exists a sequence* $\{\lambda_n^{(1)}\}_{n=0}^{\infty}$ *of positive parameters with the properties that*

(i) $0 = \lambda_0^{(1)} < \lambda_1^{(1)} < \cdots < \lambda_n^{(1)} < \cdots$, $\lim_{n\to\infty} \lambda_n^{(1)} = \infty$;
(ii) *for* $\lambda \in (\lambda_{n-1}^{(1)}, \lambda_n^{(1)})$, $n = 1, 2, \dots$, *the solution* $\tilde{x}_1(t;\lambda)$ *has exactly* n *zeros in* (a,∞) *and* $\tilde{x}_1(t;\lambda) \neq 0$;
(iii) *for* $\lambda = \lambda_n^{(1)}$, $n = 1, 2, \dots$, *the solution* $\tilde{x}_1(t;\lambda)$ *has exactly* n *zeros and* $\tilde{x}_1(a;\lambda) = 0$;
(iv) *the parameters* $\{\lambda_n^{(0)}\}$ *and* $\{\lambda_n^{(1)}\}$ *have the interlacing property* $0 = \lambda_0^{(1)} = \lambda_0^{(0)} < \lambda_1^{(1)} < \lambda_1^{(0)} < \cdots < \lambda_n^{(1)} < \lambda_n^{(0)} < \cdots$.

12. Perturbation principle

12.1. *General idea*

In the previous sections devoted to oscillation and nonoscillation criteria for (0.1), this equation was essentially viewed as a perturbation of the one-term equation

$$\big(r(t)\Phi(x')\big)' = 0. \tag{12.1}$$

As we have already mentioned, for oscillation (nonoscillation) of (0.1), the function c must be "sufficiently positive" ("not too positive") comparing with the function r. In this section we use a more general approach, Equation (0.1) is investigated as a perturbation of another (nonoscillatory) *two-term* half-linear equation

$$\big(r(t)\Phi(x')\big)' + \tilde{c}(t)\Phi(x) = 0 \tag{12.2}$$

with a continuous function $\tilde{c}$, i.e., (0.1) is written in the form

$$\big(r(t)\Phi(x')\big)' + \tilde{c}(t)\Phi(x) + \big(c(t) - \tilde{c}(t)\big)\Phi(x) = 0. \tag{12.3}$$

The main idea is essentially the same as before. If the difference $(c - \tilde{c})$ is sufficiently positive (not too positive), then (12.3) becomes oscillatory (remains nonoscillatory).

Note that in the linear case $p = 2$, the idea to investigate the linear Sturm–Liouville equation (1.1) as a perturbation of the nonoscillatory two-term equation

$$\big(r(t)x'\big)' + \tilde{c}(t)x = 0 \tag{12.4}$$

(and not only as a perturbation of the one-term equation $(r(t)x')' = 0$) brings essentially no new idea. Indeed, let us write (1.1) in the "perturbed" form

$$\big(r(t)x'\big)' + \tilde{c}(t)x + \big(c(t) - \tilde{c}(t)\big)x = 0. \tag{12.5}$$

Further, let h be a solution of (12.4) and consider the transformation $x = h(t)u$. This transformation transforms (12.5) into the equation

$$\big(r(t)h^2(t)u'\big)' + \big[c(t) - \tilde{c}(t)\big]h^2(t)u = 0 \tag{12.6}$$

(compare (3.2)) and this equation, whose oscillatory properties are the same as those of (1.1), can be again investigated as a perturbation of the one-term equation $(r(t)h^2(t)u')' = 0$. In the half-linear case we have in disposal no transformation which reduces nonoscillatory two-terms equation into a one-term equation, so we have to use different methods and this "perturbation principle" brings new phenomena. Note also that some ideas used in this section have already been applied in Section 8.

12.2. *Leighton–Wintner type oscillation criterion*

Recall that if $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty$ and $\int^{\infty} c(t)\,\mathrm{d}t = \infty$, then (0.1) is oscillatory. This direct extension of the classical linear Leighton–Wintner criterion has been proved in Section 2. This criterion characterizes exactly what means that for oscillation of (0.1) the function c must be sufficiently positive comparing with the function r in one-term equation (12.1). Here we extend this result to the situation when (0.1) is investigated as a perturbation of (12.2). The results of this subsection are presented in [69].

THEOREM 12.1. *Suppose that h is the principal solution of* (*nonoscillatory*) *equation* (12.2) *and*

$$\int^{\infty} \big(c(t) - \tilde{c}(t)\big)h^p(t)\,\mathrm{d}t := \lim_{b\to\infty} \int^{b} \big(c(t) - \tilde{c}(t)\big)h^p(t)\,\mathrm{d}t = \infty. \tag{12.7}$$

Then Equation (0.1) *is oscillatory.*

PROOF. According to the relationship between disconjugacy of (0.1) and positivity of the functional $\mathcal{F}$ mentioned in Section 2, to prove that (0.1) is oscillatory, it suffices to find (for any $T \in \mathbb{R}$) a function $y \in W^{1,p}(T,\infty)$, with a compact support in (T,∞), such that $\mathcal{F}(y; T,\infty) < 0$. Hence, let $T \in \mathbb{R}$ be arbitrary and $T < t_0 < t_1 < t_2 < t_3$ (these points will be specified later). Define the test function y as follows.

$$y(t) = \begin{cases} 0, & T \leqslant t \leqslant t_0, \\ f(t), & t_0 \leqslant t \leqslant t_1, \\ h(t), & t_1 \leqslant t \leqslant t_2, \\ g(t), & t_2 \leqslant t \leqslant t_3, \\ 0, & t_3 \leqslant t < \infty, \end{cases}$$

where f, g are solutions of (12.2) given by the boundary conditions $f(t_0) = 0$, $f(t_1) = h(t_1)$, $g(t_2) = h(t_2)$, $g(t_3) = 0$. Denote

$$w_f := \frac{r\Phi_p(f')}{\Phi_p(f)}, \qquad w_h := \frac{r\Phi_p(h')}{\Phi_p(h)}, \qquad w_f := \frac{r\Phi_p(g')}{\Phi_p(g)},$$

i.e., w_f, w_g, w_h are solutions of the Riccati equation associated with (12.2) generated by f, g, h respectively. Using exactly the same computations as in the proof of Theorem 1 from [58], one can show that

$$\mathcal{F}(y; T,\infty) = K - \int_{t_1}^{\xi} \big(c(t) - \tilde{c}(t)\big)h^p(t)\,\mathrm{d}t + h^p(t_2)\big[w_h(t_2) - w_g(t_2)\big], \tag{12.8}$$

where

$$K := \mathcal{F}(f; t_0, t_1) + h^p(t_1)\big[w_f(t_1) - w_h(t_1)\big]$$

and $\xi \in (t_2, t_3)$. Now, if $\varepsilon > 0$ is arbitrary and $T < t_0 < t_1$ are fixed, then, according to (12.7), t_2 can be chosen in such a way that $\int_{t_1}^{t} (c(s) - \tilde{c}(s))h^p(s)\,\mathrm{d}s > K + \varepsilon$ whenever $t > t_2$. Finally, again using the same argument as in [58] we have (observe that w_g actually depends also on t_3)

$$\lim_{t_3 \to \infty} h^p(t_2)\big[w_h(t_2) - w_g(t_2)\big] = 0,$$

hence the last summand in (12.8) is less than ε if t_3 is sufficiently large. Consequently, $\mathcal{F}(y; t_0, t_3) < 0$ if t_0, t_1, t_2, t_3 are chosen in the above specified way. □

If $r(t) \equiv 1$ in (0.1) and $\tilde{c}(t) = \frac{\tilde{\gamma}}{t^p}$, $\tilde{\gamma} = (\frac{p-1}{p})^p$, i.e., (12.2) is the generalized Euler equation with the critical coefficient (4.20), then the previous theorem reduces to the oscillation criterion given by Elbert [87].

12.3. *Hille–Nehari-type oscillation criterion*

The results of this subsection can be viewed as an extension of Theorems 5.4 and 5.9 to the situation when (0.1) (or (6.16)) is viewed as a perturbation of a one-term equation.

THEOREM 12.2. *Let* $\int^{\infty} r^{1-q}(t)\,\mathrm{d}t = \infty$ *and* $c(t) \geqslant 0$ *for large* t. *Further suppose that equation* (12.2) *is nonoscillatory and possesses a positive solution* h *satisfying*

(i) $h'(t) > 0$ *for large* t;

(ii) *it holds*

$$\int^{\infty} r(t)\big(h'(t)\big)^p\,\mathrm{d}t = \infty; \tag{12.9}$$

(iii) *there exists a finite limit*

$$\lim_{t \to \infty} r(t)h(t)\Phi\big(h'(t)\big) =: L > 0. \tag{12.10}$$

Denote by

$$G(t) = \int^{t} \frac{\mathrm{d}s}{r(s)h^2(s)(h'(s))^{p-2}} \tag{12.11}$$

and suppose that the integral

$$\int^{\infty} \big(c(t) - \tilde{c}(t)\big)h^p(t)\,\mathrm{d}t = \lim_{b \to \infty} \int^{b} \big(c(t) - \tilde{c}(t)\big)h^p(t)\,\mathrm{d}t \tag{12.12}$$

is convergent. If

$$\liminf_{t\to\infty} G(t)\int_t^\infty \bigl(c(s)-\tilde c(s)\bigr)h^p(s)\,\mathrm{d}s > \frac{1}{2q} \tag{12.13}$$

then Equation (0.1) *is oscillatory.*

PROOF. Suppose, by contradiction, that (0.1) is nonoscillatory, i.e., there exists an eventually positive principal solution x of this equation. Denote by $\rho := r(t)\frac{\Phi(x')}{\Phi(x)}$. Then ρ satisfies the Riccati equation (2.1) and using the Picone identity for half-linear equations (2.5) we have

$$\begin{aligned}&\int_T^t \bigl(r(s)|y'|^p - c(s)|y|^p\bigr)\,\mathrm{d}s\\ &\quad= \rho(s)|y|^p\big|_T^t + p\int_T^t r^{1-q}(s)P\bigl(r^{q-1}y', \rho\Phi(y)\bigr)\,\mathrm{d}s\end{aligned}$$

for any differentiable function y, where P is given by (2.6), and integration by parts yields

$$\begin{aligned}&\int_T^t \bigl[r(s)|y'|^p - c(s)|y|^p\bigr]\,\mathrm{d}s\\ &\quad= \int_T^t \bigl[r(s)|y'|^p - \tilde c(s)|y|^p\bigr]\,\mathrm{d}s - \int_T^t \bigl(c(s)-\tilde c(s)\bigr)|y|^p\,\mathrm{d}s\\ &\quad= r(s)y\Phi(y')\big|_T^t - \int_T^t y\bigl[\bigl(r(s)\Phi(y')\bigr)' - \tilde c(s)\Phi(y)\bigr]\,\mathrm{d}s\\ &\qquad - \int_T^t \bigl(c(s)-\tilde c(s)\bigr)|y|^p\,\mathrm{d}s.\end{aligned}$$

Substituting $y = h$ into the last two equalities (h being a solution of (12.2) satisfying the assumptions (i)–(iii) of theorem), we get

$$h^p(\tilde\rho - \rho)\big|_T^t = \int_T^t \bigl(c(s)-\tilde c(s)\bigr)h^p\,\mathrm{d}s + p\int_T^t r^{1-q}(s)P\bigl(r^{q-1}h', \rho\Phi(h)\bigr)\,\mathrm{d}s, \tag{12.14}$$

where $\tilde\rho = \frac{r\Phi(h')}{\Phi(h)}$. Since $\int^\infty r^{1-q}(t)\,\mathrm{d}t = \infty$, $w \equiv 0$ is the distinguished solution of the Riccati equation corresponding to the equation $(r(t)\Phi(x'))' = 0$ and since $c(t) \geqslant 0$, by Theorem 7.2 $\rho(t) \geqslant 0$ eventually. Hence, with L given by (12.10), we have

$$\begin{aligned}&L + h^p(T)\bigl(\rho(T) - \tilde\rho(T)\bigr)\\ &\quad\geqslant \int_T^t \bigl(c(s)-\tilde c(s)\bigr)h^p\,\mathrm{d}s + p\int_T^t r^{1-q}(s)P\bigl(r^{q-1}h', \rho\Phi(h)\bigr)\,\mathrm{d}s,\end{aligned}$$

and since $P(u, v) \geqslant 0$, this means that

$$\int^{\infty} r^{1-q}(t) P\big(r^{q-1}(t)h'(t), \rho(t)\Phi\big(h(t)\big)\big)\,\mathrm{d}t < \infty. \tag{12.15}$$

Now, since (12.10), (12.12), (12.15) hold, from (12.14) it follows that there exists a finite limit

$$\lim_{t\to\infty} h^p(t)\big(\rho(t) - \tilde{\rho}(t)\big) =: \beta$$

and also the limit

$$\lim_{t\to\infty} \frac{\rho(t)}{\tilde{\rho}(t)} = \lim_{t\to\infty} \frac{h^p(t)\rho(t)}{h^p(t)\tilde{\rho}(t)} = \frac{L+\beta}{L}. \tag{12.16}$$

Therefore,

$$h^p(t)\big(\rho(t) - \tilde{\rho}(t)\big) - \beta = C(t) + p\int_t^{\infty} r^{1-q}(s) P\big(r^{q-1}h', \rho\Phi(h)\big)\,\mathrm{d}s,$$

where $C(t) = \int_t^{\infty}(c(s) - \tilde{c}(s))h^p(s)\,\mathrm{d}s$.

Concerning the function $P(u, v)$, we have for $u, v > 0$

$$P(u, v) = \frac{u^p}{p} - uv + \frac{v^q}{q} = u^p\left(\frac{1}{q}\frac{v^q}{u^p} - vu^{1-p} + \frac{1}{p}\right) = u^p Q\big(vu^{1-p}\big), \tag{12.17}$$

where $Q(\lambda) = \frac{1}{q}\lambda^q - \lambda + \frac{1}{p}$ for $\lambda \geqslant 0$ with equality if and only if $\lambda = 1$ and

$$\lim_{\lambda\to 1} \frac{Q(\lambda)}{(\lambda-1)^2} = \frac{q-1}{2}. \tag{12.18}$$

Hence for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$P(u, v) \geqslant \left(\frac{q-1}{2} - \varepsilon\right) u^p \left(\frac{v}{u^{p-1}} - 1\right)^2, \tag{12.19}$$

whenever $|vu^{1-p} - 1| < \delta$. This implies that $\beta = 0$ in (12.16) since the case $\beta \neq 0$ contradicts the divergence of $\int^{\infty} r(t)(h'(t))^{p-1}\,\mathrm{d}t$. If we denote

$$f(t) := h^p(t)\big(\rho(t) - \tilde{\rho}(t)\big), \qquad H(t) := \frac{1}{r(t)h^2(t)(h'(t))^{p-2}},$$

then using

$$
\begin{aligned}
f(t) &\geqslant C(t) + \left(\frac{p(q-1)}{2} - \tilde{\varepsilon}\right) \int_t^\infty r(s)\big(h'(s)\big)^p \left(\frac{\rho(s)}{\tilde{\rho}(s)} - 1\right)^2 \mathrm{d}s \\
&= C(t) + \left(\frac{q}{2} - \tilde{\varepsilon}\right) \int_t^\infty H(s) f^2(s)\, \mathrm{d}s, \qquad (12.20)
\end{aligned}
$$

where $\tilde{\varepsilon} = p\varepsilon$. Multiplying (12.20) by $G(t)$ we get

$$
G(t)f(t) \geqslant G(t)C(t) + \left(\frac{q}{2} - \tilde{\varepsilon}\right) G(t) \int_t^\infty H(s) f^2(s)\, \mathrm{d}s. \qquad (12.21)
$$

Inequality (12.21) together with (12.13) imply that there exists a $\tilde{\delta} > 0$ such that

$$
G(t)f(t) \geqslant \frac{1}{2q} + \tilde{\delta} + \left(\frac{q}{2} - \tilde{\varepsilon}\right) G(t) \int_t^\infty \frac{H(s)}{G^2(s)} \big[G(s)f(s)\big]^2 \mathrm{d}s \qquad (12.22)
$$

for large t.

Suppose first that $\liminf_{t\to\infty} G(t)f(t) =: c < \infty$. Then for every $\bar{\varepsilon} > 0$ we have $[G(t)f(t)]^2 > (1-\bar{\varepsilon})c^2$ for large t and (12.22) implies

$$
c \geqslant \frac{1}{2q} + \tilde{\delta} + \left(\frac{q}{2} - \tilde{\varepsilon}\right)(1 - \bar{\varepsilon})c^2.
$$

Now, letting $\tilde{\varepsilon}, \bar{\varepsilon} \to 0$ we have

$$
c \geqslant \frac{1}{2q} + \tilde{\delta} + \frac{q}{2}c^2,
$$

but this is impossible since $1 - 2q(\frac{1}{2q} + \tilde{\delta}) < 0$.

Finally, if

$$
\liminf_{t\to\infty} G(t)f(t) = \infty, \qquad (12.23)
$$

denote by $m(t) = \inf_{t\leqslant s}\{G(s)f(s)\}$. Then m is nondecreasing and (12.22) implies that

$$
G(t)f(t) \geqslant K + \left(\frac{q}{2} - \tilde{\varepsilon}\right) m^2(t),
$$

where $K = \frac{1}{2q} + \tilde{\delta}$ Since m is nondecreasing, we have for $s > t$

$$
G(s)f(s) \geqslant K + \left(\frac{q}{2} - \tilde{\varepsilon}\right) m^2(s) \geqslant K + \left(\frac{q}{2} - \tilde{\varepsilon}\right) m^2(t), \quad t \leqslant s,
$$

and hence

$$m(t) \geqslant K + \left(\frac{q}{2} - \tilde{\varepsilon}\right) m^2(t)$$

which contradicts (12.23). The proof is complete. □

When (12.3) reduces to the below given Euler-type equation (12.24), some technical assumptions on the function h in the previous theorem are satisfied and Theorem 12.2 simplifies as follows.

COROLLARY 12.1. *Let* $r(t) \equiv 1$, $c(t) \geqslant 0$ *and*

$$\tilde{c}(t) = \frac{\tilde{\gamma}}{t^p}, \quad \tilde{\gamma} = \left(\frac{p-1}{p}\right)^p.$$

Then (12.3) *is the generalized Euler equation with the critical coefficient*

$$\left(\Phi(y')\right)' + \frac{\tilde{\gamma}}{t^p}\Phi(y) = 0 \tag{12.24}$$

and the solution $h(t) = t^{\frac{p-1}{p}}$ *of this equation satisfies all assumptions of Theorem* 12.2 *with*

$$G(t) = \left(\frac{p}{p-1}\right)^{p-2} \lg t.$$

Hence Equation (9.1) *is oscillatory provided*

$$\liminf_{t\to\infty} \lg t \int_t^\infty s^{p-1}\left[c(s) - \frac{\gamma_0}{s^p}\right] \mathrm{d}s > \frac{1}{2}\left(\frac{p-1}{p}\right)^{p-1}.$$

12.4. *Hille–Nehari-type nonoscillation criterion*

Now we turn our attention to a nonoscillation criterion which is proved under no sign restriction on the function c and also under no assumption concerning divergence of the integral $\int^\infty r^{1-q}(t)\,\mathrm{d}t$ (compare Theorem 12.2).

THEOREM 12.3. *Suppose that Equation* (12.2) *is nonoscillatory and possesses a solution* h *satisfying* (i), (iii) *of Theorem* 12.2. *Moreover, suppose that*

$$\int^\infty \frac{\mathrm{d}t}{r(t)h^2(t)(h'(t))^{p-2}} = \infty. \tag{12.25}$$

If $G(t)$ *is the same as in Theorem* 12.2 *and*

$$\limsup_{t\to\infty} G(t)\int_t^\infty \big(c(s)-\tilde{c}(s)\big)h^p(s)\,\mathrm{d}s < \frac{1}{2q} \tag{12.26}$$

and

$$\liminf_{t\to\infty}\ G(t)\int_t^\infty \big(c(s)-\tilde{c}(s)\big)h^p(s)\,\mathrm{d}s > -\frac{3}{2q}, \tag{12.27}$$

then (0.1) *is nonoscillatory.*

PROOF. Denote again

$$C(t)=\int_t^\infty \big(c(s)-\tilde{c}(s)\big)h^p(s)\,\mathrm{d}s.$$

To prove that (0.1) is nonoscillatory, according to Section 5.3 it suffices to find a differentiable function ρ which verifies the differential inequality (5.4) for large t. This inequality can be written in the form (with $w=h^{-p}(\rho+C)$)

$$\begin{aligned}\rho' &\leqslant -p\left[\frac{1}{q}\left|\frac{\rho+C}{h}\right|^q r^{1-q} - h'\left(\frac{\rho+C}{h}\right)+\frac{r(h')^p}{p}\right]+r(h')^p-\tilde{c}(t)h^p\\ &= -pr^{1-q}\left[\frac{1}{q}\left|\frac{\rho+C}{h}\right|^q - r^{q-1}h'\left(\frac{\rho+C}{h}\right)+\frac{1}{p}r^q(h')^p\right]+r(h')^p-\tilde{c}h^p\\ &= -pr^{1-q}P\left(r^{q-1}h',\frac{\rho+C}{h}\right)+r(h')^p-\tilde{c}h^p.\end{aligned}$$

We will show that the function

$$\rho(t)=r(t)h(t)\Phi\big(h'(t)\big)+\frac{1}{2qG(t)} \tag{12.28}$$

satisfies this inequality for large t. To this end, let $v=\frac{\rho+C}{h}$, $u=r^{q-1}h'$. The fact that the solution h of (12.2) is increasing together with (12.25), (12.26), (12.27) and the assumption (iii) of Theorem 12.2 imply that

$$\frac{v}{\Phi(u)}=\frac{\rho(t)+C(t)}{h(t)r(t)\Phi(h'(t))}=1+\frac{1+2qC(t)G(t)}{2qG(t)r(t)h(t)\Phi(h'(t))}\to 1$$

as $t\to\infty$. Hence, using (12.17) and the same argument as in the proof of Theorem 12.2, for any $\varepsilon>0$, we have (with Q satisfying (12.18))

$$pr^{1-q}\left[\frac{1}{q}\left|\frac{\rho+C}{h}\right|^q - h'r^{q-1}\left(\frac{\rho+C}{h}\right)+\frac{r^q(h')^p}{p}\right]$$

$$
\begin{aligned}
&= pr^{1-q}r^q(h')^p Q\left(\frac{\rho+C}{hr\Phi(h')}\right)\\
&\leqslant p\left(\frac{q-1}{2}+\varepsilon\right)r(h')^p\frac{(1+2qGC)^2}{4q^2r^2h^2(h')^{2p-2}G^2}\\
&= \left(\frac{q}{2}+p\varepsilon\right)\frac{1}{rh^2(h')^{p-2}}\frac{(1+2qGC)^2}{4G^2q^2}
\end{aligned}
$$

as $t \to \infty$.

Now, since (12.26), (12.27) hold, there exists $\delta > 0$ such that

$$
\frac{-3+\delta}{2q} < G(t)C(t) < \frac{1-\delta}{2q} \quad \Longleftrightarrow \quad \left|1+2qG(t)C(t)\right| < 2-\delta
$$

for large t, hence $\varepsilon > 0$ can be chosen in such a way that

$$
\left(\frac{q}{2}+\varepsilon\right)\frac{(1+2qG(t)C(t))^2}{4q^2} < \frac{1}{2q}
$$

for large t. Consequently (using the fact that h solves (12.2)), we have

$$
\begin{aligned}
&-pr^{1-q}\left[\frac{1}{q}\left|\frac{\rho+C}{h}\right|^q - r^{q-1}h'\left(\frac{\rho+C}{h}\right)+\frac{r(h')^p}{p}\right]+r(h')^p-\tilde{c}(t)h^p\\
&\geqslant -\left(\frac{q}{2}+\varepsilon\right)\frac{1}{G^2rh^2(h')^{p-2}}\frac{(1+2qGC)^2}{4q^2}+r(h')^p-\tilde{c}(t)h^p\\
&> -\frac{1}{2qG^2rh^2(h')^{p-2}}+\left[rh\Phi(h')\right]' = \left[rh\Phi(h')+\frac{1}{2qG}\right]' = \rho'.
\end{aligned}
$$

The proof is complete. □

COROLLARY 12.2. *If* (12.2) *is the generalized Euler equation* (12.24) *then by the previous theorem Equation* (9.1) *is nonoscillatory provided*

$$
\limsup_{t\to\infty} \lg t \int_t^\infty \left(c(s)-\frac{\gamma}{s^p}\right)s^{p-1}\,\mathrm{d}s < \frac{1}{2}\left(\frac{p-1}{p}\right)^{p-1}
$$

and

$$
\liminf_{t\to\infty} \lg t \int_t^\infty \left(c(s)-\frac{\gamma}{s^p}\right)s^{p-1}\,\mathrm{d}s > -\frac{3}{2}\left(\frac{p-1}{p}\right)^{p-1}.
$$

12.5. *Perturbed Euler equation*

The results of this subsection can be found in the paper [94]. If we distinguish the cases $p \in (1,2]$ and $p \geqslant 2$, the following refinement of oscillation and nonoscillation criteria from the previous subsection can be proved. We present the main results of [94] here without proofs since these proofs need several technical lemmata.

THEOREM 12.4. *Consider the half-linear equation*

$$\big(\Phi(x')\big)' + \frac{\tilde{\gamma}}{t^p}\Phi(x) + 2\left(\frac{p-1}{p}\right)^{p-1}\frac{\delta(t)}{t^p}\Phi(x) = 0, \quad \tilde{\gamma} = \left(\frac{p-1}{p}\right)^p, \tag{12.29}$$

and the linear second order equation

$$(ty')' + \frac{\delta(t)}{t}y = 0. \tag{12.30}$$

Suppose that the integral

$$\sigma(t) := \int_t^\infty \frac{\delta(s)}{s}\,\mathrm{d}s$$

is convergent and $\sigma(t) \geqslant 0$ for large t.

(i) *Let $p \geqslant 2$ and* (12.30) *is nonoscillatory. Then* (12.29) *is also nonoscillatory.*

(ii) *Let $p \in (1, 2]$ and half-linear equation* (12.29) *is nonoscillatory. Then linear equation* (12.30) *is also nonoscillatory.*

This Page Intentionally Left Blank

CHAPTER 3D

Related Equations and Problems

In this chapter we discuss various problems related to the oscillation theory of half-linear differential equations. First we deal with boundary value problems associated with (0.1), where a particular attention is focused to the half-linear Fredholm alternative. In Section 14 we briefly mention the so-called quasilinear equations, i.e., equations, where in addition to additivity, also homogeneity of the solution space of equations is lost. Section 15 is devoted to partial differential equations with p-Laplacian and the last section of this chapter presents basic facts of the oscillation theory of half-linear difference equations.

13. Half-linear boundary value problems

There is a voluminous literature dealing with boundary value problems of the form

$$\big(\Phi(x')\big)' + g(t,x) = h(t), \quad x(0) = 0 = x(\pi_p), \tag{13.1}$$

(or some other boundary conditions, e.g., periodic, Neuman, mixed,...) and even a brief survey of these results exceeds the scope of this treatment. For this reason we focus our attention only to some particular boundary value problems. We refer to [52,53,73,76,77, 80,82,164–168] as to a sample of papers dealing with half-linear boundary value problems.

13.1. *Basic boundary value problem*

Under the "basic" boundary value problem we understand the problem

$$\begin{cases} \big(\Phi(x')\big)' + \lambda\Phi(x) = 0, \\ x(0) = 0 = x(\pi_p), \end{cases} \tag{13.2}$$

where λ is the eigenvalue parameter. Here π_p is the same as in the Section 1 and its value is defined by the formula

$$\pi_p = 2\int_0^1 \frac{\mathrm{d}s}{\sqrt[p]{1-s^p}}. \tag{13.3}$$

Eigenvalue problem (13.2) is a special case of the general Sturm–Liouville problem for half-linear equations treated in Section 11, but its simple structure enables to determine completely the eigenvalues and eigenfunctions. The situation is essentially the same as in the linear case where the eigenvalues are $\lambda_n = n^2$ with the associated eigenfunctions $x_n(t) = \sin nt$.

THEOREM 13.1. *The eigenvalues of* (13.2) *are* $\lambda_n = (p-1)n^p$, $n \in \mathbb{N}$, *and the corresponding eigenfunctions are* (*up to a nonzero multiplicative factor*) $x_n(t) = \sin_p(nt)$, *where the half-linear sine function* $\sin_p$ *is defined in Section* 1.

PROOF. The proof if this statement follows immediately from the homogeneity of the solution space of half-linear equations and from the unique solvability the initial value problem for these equations. The function $x_1(t) = \sin_p t$ is a solution of (13.2) with $\lambda = (p-1)$ and satisfies $x(0) = 0 = x(\pi_p)$ (by the definition of this function in Section 1), and $x_n(t) = \sin_p(nt)$ is a solution of (13.2) with $\lambda_n = (p-1)n^p$. □

13.2. *Variational characterization of eigenvalues*

In the linear case, the Courant–Fischer minimax principle provides a variational characterization of eigenvalues of the classical Sturm–Liouville eigenvalue problem. This characterization is based on the orthogonality of the eigenfunctions corresponding to different eigenvalues. In the half-linear case the meaning of orthogonality is lost, but eigenvalues can be described using the Lusternik–Schnirelman procedure, for general facts concerning this approach we refer to [103]. The specification of this method to (13.2) presented here can be found in [80].

Let us introduce the functionals over $\mathcal{X} := W_0^{1,p}(0, \pi_p)$, endowed with the norm $\|x\| = (\int_0^{\pi_p} |x'|^p \, \mathrm{d}t)^{\frac{1}{p}}$,

$$A(x) = \frac{1}{p}\int_0^{\pi_p} |x'|^p \, \mathrm{d}t, \qquad B(x) = \frac{1}{p}\int_0^{\pi_p} |x|^p \, \mathrm{d}t.$$

Eigenfunctions and eigenvalues of (13.2) are equivalent to critical points and critical values of the functional

$$E(x) = \frac{A(x)}{B(x)}.$$

We also introduce the notation

$$\mathcal{S} = \{x \in \mathcal{X}\colon\ B(x) = 1\},$$

(hence $E(x) = A(x)$, $E'(x) = A'(x) - A(x)B'(x)$ for $x \in \mathcal{S}$, where A', B', E' are differentials of the functionals A, B, E, respectively). After some computation one can verify that $E|_{\mathcal{S}}$ satisfies the so-called *Palais–Smale condition*: if $\{x_k\} \in \mathcal{S}$ is a sequence such that $E(x_k)$ is convergent and $E'(x_k) \to 0$ in $\mathcal{X}^*$ (the dual space of $\mathcal{X}$), then $\{x_k\}$ contains a convergent subsequence (in the norm of $\mathcal{X}$).

Let us recall also the definition of the Krasnoselskii genus of a symmetric set $\mathcal{A} \subset \mathcal{X}$. Let

$$\mathcal{F} := \{\mathcal{A} \subset \mathcal{X}\colon\ \mathcal{A} \text{ is closed and } \mathcal{A} = -\mathcal{A}\}$$

and let

$$\mathcal{M} = \bigl\{m \in \mathbb{N}\colon\ \exists h \in C\bigl(\mathcal{A}; \mathbb{R}^m \setminus \{0\}\bigr) \text{ such that } h(-x) = -h(x)\bigr\}.$$

Then the Krasnoselskii genus of $\mathcal{A}$ is defined by

$$\gamma(\mathcal{A}) := \begin{cases} \inf \mathcal{M}, & \text{if } \mathcal{M} \neq \emptyset, \\ \infty, & \text{if } \mathcal{M} = \emptyset. \end{cases}$$

Using the above given concepts we can now present the formulas for variational characterization of *all* eigenvalues of (13.2).

THEOREM 13.2. *Let*

$$\mathcal{F}_k := \bigl\{\mathcal{A} \in \mathcal{F}\colon\ 0 \notin \mathcal{A},\ \gamma(\mathcal{A}) \geqslant k\bigr\},$$
$$\widetilde{\mathcal{F}}_k := \{\mathcal{A} \in \mathcal{F}_k\colon\ \mathcal{A} \subset \mathcal{S},\ \mathcal{A} \textit{ is compact}\}$$

and let

$$\beta_k = \min_{\mathcal{A} \in \widetilde{\mathcal{F}}_k} \max_{x \in \mathcal{A}} E(x). \tag{13.4}$$

Then $\beta_n = \lambda_n = (p-1)n^p$ *for* $n \in \mathbb{N}$.

PROOF. We present only a brief outline of the proof of this statement. First it is proved that the nth eigenvalue λ_n satisfies $\lambda_n \leqslant \beta_n$. Then, via the construction of a suitable set $\mathcal{A} \in \widetilde{\mathcal{F}}$, it is shown that $\lambda_n \geqslant \beta_n$. We refer to [80] for details. □

13.3. *Nonresonance problem*

In this subsection we consider the nonhomogeneous problem

$$\begin{cases} \bigl(\Phi(x')\bigr)' + \lambda \Phi(x) = f(t), \\ x(0) = 0 = x(\pi_p). \end{cases} \tag{13.5}$$

Let

$$\mathcal{J}_f^{\lambda}(x) := \frac{1}{p} \int_0^{\pi_p} \bigl|x'(t)\bigr|^p \,\mathrm{d}t - \frac{\lambda}{p} \int_0^{\pi_p} \bigl|x(t)\bigr|^p \,\mathrm{d}t - \int_0^{\pi_p} f(t)x(t) \,\mathrm{d}t$$

and suppose that λ is not an eigenvalue, i.e., $\lambda \neq \lambda_k$. For simplicity we deal with $f \in C[0, \pi_p]$ and solution of (13.5) is understood in the classical sense, i.e., it is a function x such that $\Phi(x') \in C^1[0, \pi_p]$ and equation and boundary conditions in (13.5) are satisfied. Similarly as in the previous subsection, the critical points of $\mathcal{J}_f^\lambda$ are in one to one correspondence with solutions of (13.5).

Due to the variational characterization of the least eigenvalue

$$\lambda_1 = \min \frac{\int_0^{\pi_p} |x'(t)|^p \, \mathrm{d}t}{\int_0^{\pi_p} |x(t)|^p \, \mathrm{d}t}, \tag{13.6}$$

where the minimum is taken over all nonzero elements of $W_0^{1,p}(0, \pi_p)$ and due to the monotonicity of the operators

$$A', B' : W_0^{1,p}(0, \pi_p) \to \left(W_0^{1,p}(0, \pi_p)\right)^*$$

defined by (note that these operators are differentials of operators A, B defined in the previous subsection)

$$\langle A'u, v\rangle = \int_0^{\pi_p} \Phi\bigl(u'(t)\bigr) v'(t) \, \mathrm{d}t, \qquad \langle B'u, v\rangle = \int_0^{\pi_p} \Phi\bigl(u(t)\bigr) v(t) \, \mathrm{d}t$$

(here $\langle \cdot , \cdot \rangle$ is the duality pairing between $(W_0^{1,p}(0, \pi_p))^*$ and $W_0^{1,p}(0, \pi_p)$) it is easy to prove that for $\lambda \leqslant 0$ the energy functional $\mathcal{J}_f^\lambda$ has a unique minimizer in $W_0^{1,p}(0, \pi_p)$ for arbitrary $f \in (W_0^{1,p}(0, \pi_p))^*$. In particular, it follows from here that given arbitrary $f \in C[0, \pi_p]$, the problem (13.5) has a unique solution. So, from this point of view, the situation is the same for $p = 2$ (linear case) and $p \neq 2$.

The case $\lambda > 0$ is different. It is well known that for $p = 2$ and $\lambda \neq \lambda_k$, $k = 1, 2, \ldots,$ for any $f \in C[0, \pi_p]$ the problem (13.5) has a unique solution, which follows e.g. from the Fredholm alternative. Let us consider now $p \neq 2$ and $0 < \lambda < \lambda_1$. Due to the variational characterization of λ_1 given by (13.6), the energy functional is still coercive but the monotone operators A', B' "compete" because of the positivity of λ. While in the linear case $p = 2$ this fact does not affect the uniqueness, for $p \neq 2$ the following interesting phenomenon is observed.

THEOREM 13.3. *There exists functions $f \in C[0, \pi_p]$ such that $\mathcal{J}_f^\lambda$ has at least two critical points. One of them corresponds to the global minimizer of $\mathcal{J}_f^\lambda$ on $W_0^{1,p}(0, \pi_p)$ (which does exist due to $\lambda < \lambda_1$) and the other is a critical point of saddle type.*

Examples which illustrate these facts were given in [101] for $1 < p < 2$ and in [52] for $p > 2$. The results were generalized for general $\lambda > 0$ in [82].

13.4. *Fredholm alternative for the scalar p-Laplacian*

This is perhaps the most interesting part of the qualitative theory of half-linear differential equations, since one meets there phenomena which are completely different comparing with the classical Fredholm alternative for the linear boundary value problem

$$u'' + m^2 u = h(t), \quad u(0) = 0 = u(\pi),$$

which has a solution if and only if

$$\int_0^{\pi} h(t) \sin mt \, \mathrm{d}t = 0. \tag{13.7}$$

In this subsection we discuss the extension of the Fredholm alternative to (13.5). We suppose that $\lambda = \lambda_k$ for some $k \in \mathbb{N}$, so the problem (13.2) possesses a nontrivial solution $x(t) = \sin_p(kt)$. The half-linear version of (13.7) when $\lambda = \lambda_1$ and $m = 1$ is

$$\int_0^{\pi_p} f(t) \sin_p t \, \mathrm{d}t = 0. \tag{13.8}$$

The next theorem show that (13.8) is sufficient but generally not necessary for solvability of (13.5) with $\lambda = \lambda_1$. The statements of this section are taken from [49]. We present them without proofs since these proofs are technically rather complicated.

THEOREM 13.4. *Let us assume that $f \in C^1[0, \pi_p]$, $f \not\equiv 0$ and* (13.8) *is satisfied. Then* (13.5) *with $\lambda = \lambda_1$ has at least one solution. Moreover, if $p \neq 2$, then the set of possible solutions is bounded in $C^1[0, \pi_p]$.*

Observe that this result really reveals a striking difference between the cases $p \neq 2$ and $p = 2$, since in the latter case the solution set is an unbounded continuum. It would be natural to expect the number of solutions under (13.8) be *generically finite* if $p \neq 2$.

Note also that in the proof of Theorem 13.4 it appears that the degree of the fixed point of a certain associated operator in a large ball of $C^1[0, \pi_p]$ becomes $+1$ if $p > 2$ while equals -1 if $p < 2$.

The eigenvalue problem (13.5) with $\lambda = \lambda_1$ and $f \equiv 0$ is closely related to the L^p-Poincaré inequality

$$\int_0^{\pi_p} \left|x'(t)\right|^p \mathrm{d}t \geqslant C \int_0^{\pi_p} \left|x(t)\right|^p \mathrm{d}t, \quad \text{for all } x \in W_0^{1,p}(0, \pi_p). \tag{13.9}$$

The constant $C = \lambda_1$ is precisely the largest $C > 0$ for which (13.9) holds. Then $\int_0^{\pi_p} |x'|^p - \lambda_1 \int_0^{\pi_p} |x|^p \geqslant 0$ for all $x \in W_0^{1,p}(0, \pi_p)$ while it minimizes and equals 0 exactly on the ray generated by the first eigenfunction $\sin_p t$. Now we consider the following question:

What is the sensitivity of this optimal Poincare's inequality under a linear perturbation? We consider then the functional

$$\mathcal{J}_f^{\lambda_1}(x) = \frac{1}{p}\int_0^{\pi_p} |x'|^p \,\mathrm{d}t - \frac{\lambda_1}{p}\int_0^{\pi_p} |x|^p \,\mathrm{d}t + \int_0^{\pi_p} f x \,\mathrm{d}t$$

and ask whether $\mathcal{J}_f^{\lambda_1}$ is *bounded from below*. It is easy to see that a necessary condition for this is that f satisfies the orthogonality condition (13.8) for otherwise E_1 is unbounded below along the ray generated by the first eigenfunction. If $p = 2$, an L^2-orthogonal expansion into the Fourier series yields that this condition is also sufficient for the boundedness from below. However, this approach seems to be of no use when $p \neq 2$. Under the additional assumption $f \in C^1[0, \pi_p]$, the answer answering the sufficiency is provided by the following result. Note that some of its conclusions are already implied by the previous theorem.

THEOREM 13.5. *Assume that* $f \in C^1[0, \pi_p]$, $f \not\equiv 0$, *and* (13.8) *holds.*

(i) *For* $1 < p < 2$ *the functional* $\mathcal{J}_f^{\lambda_1}$ *is bounded from below. The set of its critical points is nonempty and bounded.*

(ii) *For* $p > 2$ *the functional* $\mathcal{J}_f^{\lambda_1}$ *is bounded from below and has a global minimizer. The set of its critical points is bounded, however* $\mathcal{J}_f^{\lambda_1}$ *does not satisfy the Palais–Smale condition at the level* 0.

It is interesting to see that changing p from $p > 2$ to $p < 2$ shifts the structure of this functional $\mathcal{J}_f^{\lambda_1}$ from global minima to a saddle point geometry for its level sets. If $p = 2$ the functional is convex with a whole ray of minimizers. This result seems to be open and interesting issue concerning the geometry of L^p spaces where the absence of a good orthogonality notion makes the structure of Poincaré-type inequality fairly subtle.

On the other hand, the statement (ii) in the last theorem sets a word of warning in the use of min-max schemes based on the Palais–Smale condition in resonant problems involving the p-Laplacian. Here a very natural example arises of an equation with a priori estimates for the solutions for which the Palais–Smale condition in the associated functional fails to hold. We refer to the paper [49] for details.

Our next result states in particular another interesting difference with the linear case $p = 2$. If $p \neq 2$, then the set of functions f for which (13.5) with $\lambda = \lambda_1$ is solvable has *nonempty interior* in $L^\infty(0, \pi_p)$.

THEOREM 13.6. *Let* $p \neq 2$. *Then there exists an open cone* $\mathcal{C} \subset L^\infty(0, \pi_p)$ *such that for all* $f \in \mathcal{C}$ *problem* (13.5) *with* $\lambda = \lambda_1$ *has at least two solutions. Moreover*

$$\int_0^{\pi_p} f(t) \sin_p t \,\mathrm{d}t \neq 0 \tag{13.10}$$

for all $f \in \mathcal{C}$.

A by-product of the proof of this theorem is the following general fact. For any $f \in L^{\infty}(0, \pi_p)$ such that (13.10) holds, one has that the set of all possible solutions of (13.5) is bounded and the degree of the associated fixed point operator equals 0. Combining this and Theorem 13.4 yields in particular that for any $f \not\equiv 0$ of the class C^1 and $p \neq 2$, there are *a priori estimates* for the solution set.

The proof of the results presented in this subsection are based on the analysis of the initial value problem

$$\bigl(\Phi(x')\bigr)' + (p-1)\Phi(x) = f(t), \quad x(0) = 0,\ x'(0) = \alpha,$$

with $f \in L^{\infty}_{loc}[0, \infty)$. Here, x is a globally defined solution of this problem and for α sufficiently large (positive or negative) a first zero $t_1^{\alpha} > 0$ exists. Moreover, $t_1^{\alpha} \to \pi_p$ as $|\alpha| \to \infty$. The key matter is a detailed analysis of the relative location of t_1^{α} with respect to π_p for large $|\alpha|$. Of course, one has a solution of (13.5) with $\lambda = \lambda_1$ whenever t_1^{α} hits exactly π_p. In particular, in the asymptotic expansion of the dependence of t_1^{α} on α yields that under assumptions of Theorem 13.4, one has

$$t_a^{\alpha} < \pi_p \quad \text{if } p < 2 \quad \text{and} \quad t_1^{\alpha} > \pi_p, \quad \text{if } p > 2,$$

whenever $|\alpha|$ is sufficiently large.

13.5. *Homotopic deformation along p and Leray–Schauder degree*

The Leray–Schauder degree of a mapping associated with the investigated BVP is one of the most frequently used methods when dealing with this problem. In this subsection we briefly present the main ideas of [52] which deals with solvability of (13.1). First consider the associated problem

$$\bigl(\Phi_p(x')\bigr)' = f(t), \quad x(0) = 0 = x(\pi_p), \tag{13.11}$$

(note that $\Phi_p = \Phi$) and the energy functional corresponding to this problem

$$\Psi_p(x) = \frac{1}{p}\int_0^{\pi_p} \bigl|x'(t)\bigr|^p \,\mathrm{d}t - \int_0^{\pi_p} f(t)x(t)\,\mathrm{d}t. \tag{13.12}$$

Here we use the index p by Φ and Ψ to stress their dependence on the power p. The functional Ψ_p is coercive, continuous and convex over $W_0^{1,p}(0, \pi_p)$ and hence it possesses the unique global minimum which is the critical point and hence a solution of (13.11). This means that we have correctly defined mapping $G_p : L^q(0, \pi_p) \to C^1[0, \pi_p]$ which assigns to the right-hand side f of (13.11) the solution x of this problem. This mapping is completely continuous. Moreover, if p_n is a real sequence such that $p_n \to p$ and $f_n \in L^q(0, \pi_p)$ is such that $f_n \rightharpoonup f \in L^q(0, \pi_p)$ ($\rightharpoonup$ denotes the weak convergence), then $\lim_{n\to\infty} G_{p_n}(f_n) = G_p(f)$ as it is shown in [52].

Now, for fixed $p > 1$, we define $T_p : C[0, \pi_p] \to C[0, \pi_p]$ by $T_p(x) = x - G_p(\lambda \Phi_p(x))$ with $\lambda \in \mathbb{R}$. Obviously, the equation $T_p(x) = 0$ has a nontrivial solution if and only if $\lambda = \lambda_n(p) = (p-1)n^p$ and this solution is $x_n(t) = \alpha \sin_p(nt)$, $\alpha \in \mathbb{R}$, $\alpha \neq 0$.

The following statement concerns a homotopic deformation along the power p of the Leray–Schauder degree of the mapping T_p. Note that the classical result of the linear theory is that the Leray–Schauder degree of T_2 with respect to the ball

$$B(0,r) := \Big\{ u \in C[0, \pi_p]\colon\ \|u\|_C = \max_{t \in [0, \pi_p]} \big|u(t)\big| \leqslant r \Big\}$$

is

$$d\big(T_2, B(0,r), 0\big) = (-1)^n, \tag{13.13}$$

where n is the number of the eigenvalues of (13.2) with $p = 2$ which are less than λ.

THEOREM 13.7. *Let $p > 1$ be arbitrary, $\lambda \neq \lambda_n(p) = (p-1)n^p$, $n \in \mathbb{N}$. Then for every $r > 0$, the Leray–Schauder degree $d(T_p, B(r,0), 0)$ is well defined and satisfies*

$$d\big(T_p, B(r,0), 0\big) = (-1)^n, \tag{13.14}$$

where n is the number of eigenvalues of (13.2) *which are less than λ.*

PROOF. Suppose that $p \geqslant 2$ and $\lambda > \lambda_1 = (p-1)$, i.e., $\lambda = (p-1)(n+s)^p$ for some $s \in (0,1)$ and $n \in \mathbb{N}$. In the remaining cases the idea of the proof is the same. We will show that $d(T_p, B(r,0), 0) = (-1)^n$ for every $r > 0$.

Let $\Lambda : [p, \infty) \to \mathbb{R}$ be defined by $\Lambda(\alpha) = [(n+s)\pi_\alpha / \pi_p]^\alpha$, where π_α is given by (13.3) with α instead of p. Obviously, π_α depends continuously on α and hence Λ is continuous. Next, define the mapping

$$T(\alpha, x) = x - G_\alpha\big(\Lambda(\alpha)\Phi_\alpha(x)\big).$$

The mapping $\widetilde{G}(\alpha, x) := G_\alpha(\Lambda(\alpha)\Phi_\alpha(x))$ is completely continuous and $T(\alpha, x) \neq 0$ for all $\alpha \in [p, \infty)$ (for details see [52, Theorem 4.1]). Hence, from the invariance of the degree under homotopies and from (13.13) we obtain the required statement. ☐

THEOREM 13.8. *Suppose that there exists $n \in \mathbb{N}$ such that the nonlinearity g in* (13.1) *satisfies*

$$\lambda_n \leqslant a(t) := \liminf_{|s| \to \infty} \frac{g(t,s)}{\Phi(s)} \leqslant \limsup_{|s| \to \infty} \frac{g(t,s)}{\Phi(s)} =: b(t) \leqslant \lambda_{n+1}$$

uniformly on $[0, \pi_p]$, the first and the last inequalities being strict on a subset of positive measure in $[0, \pi_p]$. Then the BVP (13.1) *has a solution.*

PROOF. Let $\nu \in (\lambda_n, \lambda_{n+1})$. According to the previous theorem, it suffices to construct a homotopic bridge connecting (13.1) with the problem

$$\big(\Phi(x')\big)' + \nu\Phi(x) = 0, \quad x(0) = 0 = x(\pi_p).$$

The degree of the mapping associated with this problem has been computed in the previous theorem. This homotopy is defined as follows

$$H(\tau, x) = G_p\big(\tau\nu\Phi(x) + (1-\tau)g\big(t, x(t)\big)\big).$$

Using the standard method it can be proved that there exists $r > 0$ such that $x - H(\tau, x) \neq 0$ for $x \in \partial B(r, 0)$ for every $\tau \in [0, 1]$ if $r > 0$ is sufficiently large. This proof goes by contradiction. Supposing that there exists $x_n \in C[0, \pi_p]$ and with $\|x_n\|_C \to \infty$ and $\tau_n \in [0, 1]$ such that $x_n = H(\tau_n, x_n)$, functions v and c are constructed (using essentially the same construction as in the linear case) in such a way that the half-linear equation

$$\big(\Phi(v')\big)' + c(t)\Phi(v) = 0, \quad v(0) = 0 = v(\pi_p)$$

with $\lambda_n \leqslant c(t) \leqslant \lambda_{n+1}$ has a nontrivial solution. Since each of the previous inequalities is strict on the set of the positive measure, we have a contradiction with the Sturmian comparison theorem. □

13.6. *Resonance problem*

In the previous subsection, the nonlinearity was "situated" between two consecutive eigenvalues, i.e., it did nor "interacted" with the spectrum of the half-linear part. This situation is usually regarded as a *nonuniform nonresonance*. Now we deal with the situation when the nonlinearity is of the form $\lambda_n\Phi(x) + g(x)$ with a bounded function g, so the nonlinearity (perhaps, a better terminology would be "nonhalf-linearity") is "around" an eigenvalue of the half-linear part. This situation is referred to as the *resonant case*. The paper of Landesman and Lazer [145] published in 1970 is the pioneering work along this line in the linear case. Since that time, the conditions ensuring solvability of BVPs in resonance (the so-called Landesman–Lazer conditions) have been extended in many directions. The next theorem, taken from [80], establishes the Landesman–Lazer solvability conditions for the half-linear BVP

$$\big(\Phi(x')\big)' + \lambda_n\Phi(x) + g(x) = h(t), \quad x(0) = 0 = x(\pi_p). \tag{13.15}$$

It is supposed that there exist finite limits $\lim_{x\to\pm\infty} g(x) = g(\pm\infty)$. By φ_n we denote the eigenfunction corresponding to the nth eigenvalue, i.e., $\lambda_n = (p-1)n^p$, $\varphi_n(t) = \alpha_n \sin_p(nt)$, where $\alpha_n > 0$ is such that $\|\varphi_n\|_{L^p} = 1$.

THEOREM 13.9. *The boundary value problem* (13.15) *has a solution provided one of the following two conditions is satisfied*

$$\begin{aligned} g(\infty)\int_0^{\pi_p}\varphi_n^+(t)\,\mathrm{d}t + g(-\infty)\int_0^{\pi_p}\varphi_n^-(t)\,\mathrm{d}t \\ > \int_0^{\pi_p}\varphi_n(t)h(t)\,\mathrm{d}t \\ > g(\infty)\int_0^{\pi_p}\varphi_n^-(t)\,\mathrm{d}t + g(-\infty)\int_0^{\pi_p}\varphi_n^+(t)\,\mathrm{d}t, \end{aligned}$$

or

$$\begin{aligned} g(\infty)\int_0^{\pi_p}\varphi_n^+(t)\,\mathrm{d}t + g(-\infty)\int_0^{\pi_p}\varphi_n^-(t)\,\mathrm{d}t \\ < \int_0^{\pi_p}\varphi_n(t)h(t)\,\mathrm{d}t \\ < g(\infty)\int_0^{\pi_p}\varphi_n^-(t)\,\mathrm{d}t + g(-\infty)\int_0^{\pi_p}\varphi_n^+(t)\,\mathrm{d}t, \end{aligned}$$

where $\varphi_n^+ = \max\{0, \varphi_n\}$, $\varphi_n^- = \min\{0, \varphi_n\}$.

We skip the proof of this statement because of its technical complexity. This proof is based on a variant of the saddle point theorem and relies on the variational characterization of eigenvalues of (13.2).

14. Quasilinear and related differential equations

In this section we change the notation which we used throughout the whole treatment. Till now, q was the conjugate number of p, i.e., $q = \frac{p}{p-1}$. In this section q *is any real number satisfying* $q > 1$ and the conjugate number of p will be denoted by p^*. The main part of this section will be devoted to the equation

$$\bigl(r(t)\Phi_p(x')\bigr)' + c(t)\Phi_q(x) = 0, \quad \Phi_p(s) = |s|^{p-2}s,\ \Phi_q(s) := |s|^{q-2}s, \tag{14.1}$$

we will briefly treat also some more general equations. The functions r, c satisfy the same assumptions as in (0.1). Note that using the substitution $r\Phi_p(x') =: u$, Equation (14.1) can be written as the first order system of the form

$$x' = a_1(t)|u|^{\lambda_1}\operatorname{sgn} u, \qquad u' = a_2(t)|x|^{\lambda_2}\operatorname{sgn} x \tag{14.2}$$

with suitable functions a_1, a_2 and real constants λ_1, λ_2. The last system has been investigated in several papers of Mirzov and the results are summarized in his book [178]. As a sample of papers dealing with (14.1) and related equations we refer to [16,34,132,144,205] and the references given therein.

14.1. *Equation* (14.1) *with constant coefficients*

The results of this subsection are taken from [78]. First we focus our attention to the initial value problem

$$\big(\Phi_p(x')\big)' + \lambda\Phi_q(x) = 0, \quad x(0) = a,\ x'(0) = b. \tag{14.3}$$

We will modify the method used in the definition of the half-linear sine function $\sin_p$ and of other half-linear trigonometric functions.

THEOREM 14.1. *For any $\lambda \geqslant 0$, the initial value problem* (14.3) *has a unique solution defined on the whole real line* $\mathbb{R}$.

PROOF. The crucial fact used in the proof is that

$$\frac{|x'(t)|^p}{p^*} + \lambda\frac{|x(t)|^q}{q} = \frac{|b|^p}{p^*} + \lambda\frac{|a|^q}{q} \tag{14.4}$$

as can be verified by differentiation. Clearly, if $a = 0 = b$, the last identity implies that the trivial solution is the unique solution. If $a = 0$ or $b = 0$, supposing that there are two different solutions x_1, x_2 satisfying the same initial condition, we find that the absolute value of their difference $z = |x_1 - x_2|$ satisfies a Gronwall-type inequality and hence $z \equiv 0$. This idea, slightly modified, applies also to the case when both $a \neq 0$ and $b \neq 0$. □

The remaining part of this subsection will be devoted to the initial value problem

$$\big(\Phi_p(x')\big)' + \lambda\Phi_q(x) = 0, \quad x(0) = 0,\ x'(0) = \alpha > 0. \tag{14.5}$$

Denote by t_α the first positive zero of the derivative x', i.e., $x(t) > 0$, $x'(t) > 0$ for $t \in (0, t_\alpha)$. Further denote by $R := x(t_\alpha)$. Then using the same idea as above we have the identity

$$\frac{(x'(t))^p}{p^*} + \lambda\frac{x^q(t)}{q} = \lambda\frac{R^q}{q}. \tag{14.6}$$

Solving this equality for x' and integrating, we find

$$\left(\frac{q}{\lambda p^*}\right)^{\frac{1}{p}} \int_0^t \frac{x'(s)\,\mathrm{d}s}{(R^q - x^q(s))^{\frac{1}{p}}} = t, \tag{14.7}$$

which after a change of variables can be written as

$$t = \left(\frac{q}{\lambda p^*}\right)^{\frac{1}{p}} \frac{1}{R^{\frac{q-p}{p}}} \int_0^{\frac{x}{R}} \frac{\mathrm{d}s}{(1 - s^q)^{\frac{1}{p}}}. \tag{14.8}$$

For $t \in [0, \frac{q}{2}]$, let us set

$$\operatorname{arcsin}_{pq} t := \frac{q}{2} \int_0^{\frac{2t}{q}} \frac{\mathrm{d}s}{(1-s^q)^{\frac{1}{p}}}, \tag{14.9}$$

and note that this integral converges for $t \in [0, \frac{q}{2}]$. Substituting $t = \tau^{\frac{1}{q}}$ in (14.9), we obtain

$$\operatorname{arcsin}_{pq} t = \frac{1}{2} \widetilde{B}\left(\frac{1}{q}, \frac{1}{p^*}, \left(\frac{2t}{q}\right)^q\right), \tag{14.10}$$

where

$$\widetilde{B}\left(\frac{1}{q}, \frac{1}{p^*}, y\right) = \int_0^y \tau^{\frac{1}{q}-1}(1-\tau)^{-\frac{1}{p}}\, \mathrm{d}\tau$$

denotes the incomplete beta function. Next, substituting $t = \frac{q}{2}$ in (14.10), we define

$$\pi_{pq} := 2 \operatorname{arcsin}_{pq}\left(\frac{q}{2}\right) = B\left(\frac{1}{q}, \frac{1}{p^*}\right),$$

where B denoted the classical Euler beta function. When $p = q$, the definition of $\operatorname{arcsin}_{pq}$ and of π_{pq} coincides with the definition of arcsin_p and π_p in Section 1.1.

The function $\operatorname{arcsin}_{pq} : [0, \frac{q}{2}] \to [0, \frac{\pi_{pq}}{2}]$, so we can define first $\sin_{pq} : [0, \frac{\pi_{pq}}{2}] \to [0, \frac{q}{2}]$ as the inverse function of $\operatorname{arcsin}_{pq}$ and then to define this function for $t \in \mathbb{R}$ in the obvious way: $\sin_{pq} t = \sin_{pq}(\pi_{pq} - t)$ for $t \in [\frac{\pi_{pq}}{2}, \pi_{pq}]$ and then extend this function over $\mathbb{R}$ as odd and $2\pi_{pq}$ periodic function. It is a simple matter to verify that $\sin_{pq}$ is the unique (global) solution of the initial value problem

$$\left(\Phi_p(x')\right)' + \frac{2^q}{p^* q^{q-1}} \Phi_q(x) = 0, \quad x(0) = 0,\ x'(0) = 1. \tag{14.11}$$

Similarly as in case $p = q$ we denote $\cos_{pq} t = \frac{d}{dt} \sin_{pq} t$. Then from (14.4) and (14.11) we have

$$|\cos_{pq} t|^p + \left(\frac{2}{q}\right)^q |\sin_{pq} t|^q \equiv 1. \tag{14.12}$$

From (14.8) and (14.9) we find that

$$t = \frac{2}{(\lambda p^*)^{\frac{1}{p}} q^{\frac{1}{p^*}}} R^{\frac{p-q}{p}} \operatorname{arcsin}_{pq}\left(\frac{qx}{2R}\right), \tag{14.13}$$

and hence

$$x(t) = \frac{2R}{q} \sin_{pq}\left(\frac{(\lambda p^*)^{\frac{1}{p}} q^{\frac{1}{p^*}}}{2} R^{\frac{q-p}{p}} t\right), \tag{14.14}$$

for all $t \in \mathbb{R}$.

From (14.6) we can express R in term of α to obtain

$$R^{\frac{q-p}{p}} = \left(\frac{q}{\lambda p^*}\right)^{\frac{q-p}{pq}} \alpha^{\frac{q-p}{q}}.$$

Substituting this expression into (14.14), and setting

$$A_{pq}(\alpha, \lambda) := \frac{(\lambda p^*)^{\frac{1}{q}} q^{\frac{1}{q^*}}}{2} \alpha^{\frac{q-p}{q}}, \quad q^* = \frac{q}{q-1},$$

we find that the solution of (14.5) is

$$x(t) = \frac{\alpha}{A_{pq}(|\alpha|, \lambda)} \sin_{pq}\bigl(A_{pq}(|\alpha|, \lambda)t\bigr), \tag{14.15}$$

and this solution is $\tau(\alpha)$-periodic function with

$$\tau(\alpha) = \frac{2\pi_{pq}}{A_{pq}(\alpha, \lambda)} = 4t_\alpha.$$

THEOREM 14.2. *For any given $\alpha \neq 0$, the set of eigenvalues of the problem*

$$\bigl(\Phi_p(x')\bigr)' + \lambda \Phi_q(x) = 0, \quad x(0) = 0 = x(T) \tag{14.16}$$

is given by

$$\lambda(\alpha) = \left(\frac{2n\pi_{pq}}{T}\right)^q \frac{|\alpha|^{p-q}}{p^* q^{q-1}}, \quad n \in \mathbb{N}, \tag{14.17}$$

with the corresponding eigenfunctions

$$x_{n,\alpha}(t) = \frac{\alpha T}{n\pi_{pq}} \sin_{pq}\left(\frac{n\pi_{pq}}{T} t\right). \tag{14.18}$$

PROOF. For a given $\alpha \in \mathbb{R}$, by imposing that x in (14.15) satisfies the boundary conditions in (14.16), we obtain that λ is an eigenvalue of this problem if and only if

$$\frac{1}{2}(p^*)^{\frac{1}{q}} q^{\frac{1}{q^*}} \lambda^{\frac{1}{q}} |\alpha|^{\frac{q-p}{q}} T = n\pi_{pq}, \quad n \in \mathbb{N}, \tag{14.19}$$

and hence (14.17) follows. The expression for eigenfunctions follows then directly from (14.15). □

14.2. *Emden–Fowler type equation*

In this subsection we recall very briefly some results concerning asymptotic behavior of the quasilinear equation (14.1) and of the associated system (14.2). As we have mentioned before, the solution space (14.1) and of (14.2) is no longer even homogeneous, hence the investigation of these equations and systems is more complicated than in case of half-linear equations (0.1). Equation (14.1) and system (14.2) are sometimes called *Emden–Fowler* type equation (system) since if $p = 2$ in (14.1), this equation reduces to Emden–Fowler equation (1.3) mentioned in Section 1.1.

Recall that a solution x of (14.1) is called the *singular solution of the first kind* if x becomes eventually trivial, i.e., there exists $T \in \mathbb{R}$ such that $x \not\equiv 0$ for $t < T$ and $x(t) = 0$ for $t \geqslant T$, and a solution x is *singular solution of the second kind* if there exists a finite time T such $\lim_{t\to T-} |x(t)| = \infty$. The set of singular solution of the first and second kind we will denote by $\mathbb{S}_1$ and $\mathbb{S}_2$, respectively. A solution which is not singular is called *proper*. Recall also the classification of nonoscillatory solutions of (0.1) which can be extended under the assumption that $c(t) \neq 0$ for large t also to (14.1)

$$\mathbb{M}^+ = \{x \text{ solution of (14.1): } \exists t_x \geqslant 0\text{: } x(t)x'(t) > 0 \text{ for} t > t_x\},$$
$$\mathbb{M}^- = \{x \text{ solution of (14.1): } \exists t_x \geqslant 0\text{: } x(t)x'(t) < 0 \text{ for } t > t_x\}.$$

Results of [43,44,178] imply the following statement.

THEOREM 14.3. *Suppose that $r(t) > 0$, $c(t) < 0$ for large t.*
(i) *If $p = q$, i.e.,* (14.1) *reduces to* (0.1)*, then $\mathbb{S}_1 = \emptyset$, $\mathbb{S}_2 = \emptyset$, $\mathbb{M}^- \neq \emptyset$ and $\mathbb{M}^+ \neq \emptyset$.*
(ii) *If $p < q$, then $\mathbb{S}_1 = \emptyset$, $\mathbb{S}_2 \neq \emptyset$, $\mathbb{M}^- \neq \emptyset$.*
(iii) *If $p > q$, then $\mathbb{S}_1 \neq \emptyset$, $\mathbb{S}_2 = \emptyset$ and $\mathbb{M}^+ \neq \emptyset$.*

In Section 6 we have seen that certain integrals of functions r, c play an important role in the asymptotic classification of nonoscillatory solutions of (0.1). As an illustration of the extension of these results to (14.1) we give two statements. The first one is proved in [33] using the Schauder–Tychonov fixed point theorem, while the second one follows from the results of Kvinikadze [143], see also [33].

THEOREM 14.4. *Suppose that $r(t) > 0$, $c(t) < 0$ for large t, $\int^\infty r^{1-p^*}(t)\,\mathrm{d}t < \infty$ and*

$$\int^\infty \Phi_q\left(\int_t^\infty r^{1-p^*}(s)\,\mathrm{d}s\right)\mathrm{d}t < \infty,$$

where $\Phi_q(s) = |s|^{q-1}\operatorname{sgn} s$. Then there exists at least one solution x of (14.1) *in the class $\mathbb{M}^-$ such that $\lim_{t\to\infty} x(t) = 0$ and*

$$\lim_{t\to\infty} \frac{x(t)}{\int_t^\infty r^{1-p^*}(s)\,\mathrm{d}s} = \ell_x, \quad 0 < \ell_x < \infty.$$

THEOREM 14.5. *Suppose that* $r(t) > 0$, $c(t) < 0$ *for large* t, $p < q$ *and*

$$\int^{\infty} r^{1-p^*}(t)\Phi_{p^*}\left(\int^{t} \left|c(s)\right| \mathrm{d}s\right) \mathrm{d}t < \infty$$

or

$$\int^{\infty} \left|c(t)\right| \Phi_q\left(\int^{t} r^{1-p^*}(s)\, \mathrm{d}s\right) \mathrm{d}t < \infty.$$

Then $\mathbb{M}^+$ *contains a one parametric family of the so-called strongly increasing solutions, i.e., solutions satisfying*

$$\lim_{t\to\infty} x(t) = \infty, \qquad \lim_{t\to\infty} r(t)\Phi_p\bigl(x'(t)\bigr) = \infty \tag{14.20}$$

and a one parametric family of the so-called positive weakly increasing solutions, i.e., positive solutions where both limits in (14.20) *exists finitely.*

Another results concerning asymptotic properties of nonoscillatory solutions of (14.1) and of more general equations of this type can be found in [34,35,138,141,180,205,210, 218].

14.3. *More about quasilinear equations*

The results of this subsection can be found in [120] and concern the equation

$$\bigl(\Phi(x')\bigr)' + f(t, x) = 0 \tag{14.21}$$

under the assumptions that the function f satisfies $\operatorname{sgn} f(t, x) = \operatorname{sgn} x$ for $t \in [t_0, \infty)$.

THEOREM 14.6. *All proper solutions of* (14.21) *are oscillatory if one of the following three conditions is satisfied*:

(i) *for all* $\delta > 0$

$$\int^{\infty} \inf_{\delta\leqslant|y|} \left|f(t, y)\right| \mathrm{d}t = \infty,$$

(ii) *for some* $0 < \lambda < p - 1$ *and all* $\delta > 0$

$$\int^{\infty} t^{\lambda} \inf_{\delta\leqslant|y|} \frac{f(t, y)}{|y|^{p-1}}\, \mathrm{d}t = \infty,$$

(iii) *for all* δ, δ' *with* $\delta' > \delta > 0$

$$\int^{\infty} \inf_{\delta\leqslant|y|\leqslant\delta'} \left|f(t, y)\right| \mathrm{d}t = \infty,$$

and there exists a positive continuous function φ *satisfying* $\int^{\infty} \varphi(y)\,\mathrm{d}y = \infty$*, such that* $|f(t, y)| \geqslant \varphi(|y|)$ *for large* t *and large* $|y|$.

PROOF. To illustrate ideas used in the proof, we prove the part (ii), the proof of the remaining two statements is analogical. Suppose, by contradiction, that (14.21) has a proper solution x which is positive for large t (if x is negative, we proceed analogically). Then from (14.21) we have that also $x'(t) > 0$ and we put $w(t) = \frac{\Phi(x')}{\Phi(x)}$ (compare the remark (ii) in Section 9.1). Then w satisfies the Riccati-type equation

$$w' + (p-1)|w|^{p^*} + \frac{f(t, x(t))}{x^{p-1}(t)} = 0, \tag{14.22}$$

recall that p^* denotes the conjugate number of p. Multiplying (14.22) by t^{λ} and integrating over $[t_0, t]$, t_0 sufficiently large, we have

$$\begin{aligned} t^{\lambda} w(t) - \lambda \int_{t_0}^{t} s^{\lambda-1} w(s)\,\mathrm{d}s + (p-1) \int_{t_0}^{t} s^{\lambda} \big(w(s)\big)^{p^*}\,\mathrm{d}s \\ + \int_{t_0}^{t} s^{\lambda} \frac{f(s, x(s))}{x^{p-1}(s)}\,\mathrm{d}s \leqslant c, \end{aligned} \tag{14.23}$$

where $c > 0$ is a real constant.

Suppose first that $\int^{\infty} s^{\lambda-1} w(s)\,\mathrm{d}s < \infty$. Then it follows from (14.23) that

$$\int_{t_0}^{t} s^{\lambda} \frac{f(s, x(s))}{x^{p-1}(s)}\,\mathrm{d}s \leqslant c + \lambda \int_{t_0}^{t} s^{\lambda-1} w(s)\,\mathrm{d}s,$$

and taking the limit as $t \to \infty$, we get

$$\int_{t_0}^{\infty} s^{\lambda} \frac{f(s, x(s))}{x^{p-1}(s)}\,\mathrm{d}s < \infty.$$

However, this is impossible since assumptions of our theorem imply that for t_0 sufficiently large

$$\int_{t_0}^{\infty} s^{\lambda} \frac{f(s, x(s))}{x^{p-1}(s)}\,\mathrm{d}s \geqslant \int_{t_0}^{\infty} s^{\lambda} \inf_{\delta \leqslant x} \frac{f(s, x)}{x^{p-1}}\,\mathrm{d}s = \infty, \tag{14.24}$$

where $\delta = x(t_0) > 0$.

Suppose next that

$$\int^{\infty} s^{\lambda-1} w(s)\,\mathrm{d}s = \infty. \tag{14.25}$$

Then, by (14.23),

$$\int_{t_0}^{t} s^{\lambda}\frac{f(s,x(s))}{x^{p-1}(s)}\,\mathrm{d}s \leqslant c+\lambda\int_{t_0}^{t} s^{\lambda-1}w(s)\,\mathrm{d}s-(p-1)\int_{t_0}^{t} s^{\lambda}\big|w(s)\big|^{p^*}\,\mathrm{d}s. \tag{14.26}$$

Note that the second integral in Equation (14.26) is estimated by means of the Hölder inequality as follows

$$\begin{aligned}\int_{t_0}^{t} s^{\lambda-1}w(s)\,\mathrm{d}s &= \int_{t_0}^{t} s^{(\lambda-p)/p}s^{\lambda(p-1)/p}w(s)\,\mathrm{d}s\\ &\leqslant \left(\int_{t_0}^{t} s^{\lambda-p}\,\mathrm{d}s\right)^{\frac{1}{p}}\left(\int_{t_0}^{t} s^{\lambda}w^{p*}(s)\,\mathrm{d}s\right)^{\frac{1}{p^*}}\\ &\leqslant \left(\frac{t_0^{\lambda-p+1}}{p-1-\lambda}\right)^{\frac{1}{p}}\left(\int_{t_0}^{t} s^{\lambda}w^{p^*}(s)\,\mathrm{d}s\right)^{\frac{1}{p^*}}\\ &= \frac{(t_0^{\lambda-p+1}/(p-1-\lambda))^{\frac{1}{p}}}{(\int_{t_0}^{t} s^{\lambda}w^{p^*}(s)\,\mathrm{d}s)^{\frac{1}{p}}}\int_{t_0}^{t} s^{\lambda}w^{p^*}(s)\,\mathrm{d}s. \end{aligned} \tag{14.27}$$

Since (14.25) implies that

$$\int_{t_0}^{t} s^{\lambda}w^{p^*}(s)\,\mathrm{d}s \to \infty, \quad \text{as } t\to\infty,$$

we see from (14.27) that there exists $t_1 \geqslant t_0$ such that

$$\int_{t_0}^{t} s^{\lambda-1}w(s)\,\mathrm{d}s \leqslant \frac{p-1}{\lambda}\int_{t_0}^{t} s^{\lambda}w^{p^*}(s)\,\mathrm{d}s, \quad t\geqslant t_1.$$

Using this inequality in (14.26) we conclude that

$$\int_{t_0}^{\infty} s^{\lambda}\frac{f(s,x(s))}{x^{p-1}(s)}\,\mathrm{d}s \leqslant c$$

in contradiction to (14.24) which holds also in this case. □

15. Partial differential equations with p-Laplacian

Similarly to the boundary value problems for half-linear ordinary differential equations, also partial differential equations with p-Laplacian are treated in many papers. Recall that the p-Laplacian is the partial differential operator

$$\Delta_p u(x) := \operatorname{div}\big(\|\nabla u(x)\|^{p-2}\nabla u(x)\big), \quad x=(x_1,\dots,x_N)\in\mathbb{R}^N, \tag{15.1}$$

where $\operatorname{div} := \sum_{k=1}^{N} \frac{\partial}{\partial x_k}$ is the usual divergence operator and $\nabla u(x) = (\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_N})$ is the Hamilton nabla operator.

15.1. *Dirichlet BVP with p-Laplacian*

In this subsection we deal with the properties of the first eigenvalue and the associated eigenfunction of the Dirichlet boundary value problem

$$\begin{cases} \Delta_p u + \lambda \Phi(u) = 0, & x \in \Omega \subset \mathbb{R}^n, \\ u(x) = 0, & x \in \partial\Omega, \end{cases} \tag{15.2}$$

where Ω is a bounded domain in $\mathbb{R}^n$.

The solution of problem (15.2) is understood in the weak sense; we say that λ is an *eigenvalue* if there exists a function $u \in W_0^{1,p}(\Omega)$, $u \not\equiv 0$, such that

$$\int_\Omega \|\nabla u\|^{p-2} \langle \nabla u, \nabla \eta \rangle \, \mathrm{d}x = \lambda \int_\Omega \Phi(u) \eta \, \mathrm{d}x, \tag{15.3}$$

for every $\eta \in W_0^{1,p}(\Omega)$, where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^N$. The function u is called the *eigenfunction*.

The first eigenvalue $\lambda_1 = \lambda_1(\Omega)$ is obtained as the minimum of the Rayleigh quotient

$$\lambda_1 = \inf_v \frac{\int_\Omega \|\nabla v\|^p \, \mathrm{d}x}{\int_\Omega |v|^p \, \mathrm{d}x}, \tag{15.4}$$

where the infimum is taken over all $v \in W_0^{1,p}\Omega$, $v \not\equiv 0$. If u realizes the infimum in (15.4), so does also $|u|$, this leads immediately to the following statement.

THEOREM 15.1. *The eigenfunction u associated with the first eigenvalue λ_1 does not change its sign in Ω. Moreover, if $u \geqslant 0$ then actually $u > 0$ in the interior of Ω.*

PROOF. The statement concerning the positivity of u follows from the Harnack inequality [213, p. 724]. □

In the proof of the main result of this subsection we will need the following inequalities, for the proof see [160].

LEMMA 15.1. *Let $w_1, w_2 \in \mathbb{R}^N$.*

(i) *If $p \geqslant 2$, then*

$$\|w_2\|^p \geqslant p\|w_1\|^p \langle w_1, (w_2 - w_1) \rangle + \frac{\|w_2 - w_1\|^p}{2^{p-1} - 1}. \tag{15.5}$$

(ii) *If* $1 < p < 2$, *then*

$$\|w_2\|^p \geqslant p\|w_1\|^p\big\langle w_1, (w_2 - w_1)\big\rangle + C(p)\frac{\|w_2 - w_1\|^p}{(\|w_1\| + \|w_2\|)^{2-p}}, \tag{15.6}$$

where $C(p)$ *is a positive constant depending only on* p.

The main statement of this subsection reads as follows.

THEOREM 15.2. *The first eigenvalue of* (15.2) *is simple and isolated for any bounded domain* $\Omega \subset \mathbb{R}^N$.

PROOF. Here we follow Lindqvist's [160] modification of the original proof of Anane [11] where it is supposed that the boundary $\partial\Omega$ is of the Hölder class $C^{2,\alpha}$. This assumption on the boundary of Ω is removed in Lindqvist's proof by introducing the functions $u + \varepsilon$, $v + \varepsilon$ instead of u, v, respectively (used by Anane).

Suppose that u, v are eigenfunctions of (15.3) with $\lambda = \lambda_1$. Let $\varepsilon > 0$ and denote $v_\varepsilon = v + \varepsilon$, $u_\varepsilon = u + \varepsilon$. Further, let $\eta = u_\varepsilon - v_\varepsilon^p u_\varepsilon^{1-p}$, $\tilde\eta = v_\varepsilon - u_\varepsilon^p v_\varepsilon^{1-p}$. Then $\eta, \tilde\eta \in W_0^{1,p}(\Omega)$ and

$$\nabla\eta = \left\{1 + (p-1)\left(\frac{v_\varepsilon}{u_\varepsilon}\right)^p\right\}\nabla u - p\left(\frac{v_\varepsilon}{u_\varepsilon}\right)^{p-1}\nabla v.$$

A similar formula we have for $\nabla\tilde\eta$. Inserting the test functions η and $\tilde\eta$ into (15.3) and adding both equations, we get

$$\begin{aligned}
&\lambda_1 \int_\Omega \left[\frac{u^{p-1}}{u_\varepsilon^{p-1}} - \frac{v^{p-1}}{v_\varepsilon^{p-1}}\right]\big(u_\varepsilon^p - v_\varepsilon^p\big)\,\mathrm{d}x \\
&\quad= \int_\Omega \left[\left\{1 + (p-1)\left(\frac{v_\varepsilon}{u_\varepsilon}\right)^p\right\}\|\nabla u_\varepsilon\|^p + \left\{1 + (p-1)\left(\frac{u_\varepsilon}{v_\varepsilon}\right)^p\right\}\|\nabla v_\varepsilon\|^p\right]\mathrm{d}x \\
&\qquad - \int_\Omega \left[p\left(\frac{v_\varepsilon}{u_\varepsilon}\right)^{p-1}\|\nabla u_\varepsilon\|^{p-2}\langle\nabla u_\varepsilon, \nabla v_\varepsilon\rangle\right. \\
&\qquad \left. + p\left(\frac{u_\varepsilon}{v_\varepsilon}\right)^{p-1}\|\nabla v_\varepsilon\|^{p-2}\langle\nabla v_\varepsilon, \nabla u_\varepsilon\rangle\right]\mathrm{d}x \\
&\quad= \int_\Omega \big(u_\varepsilon^p - v_\varepsilon^p\big)\big(\|\nabla \lg u_\varepsilon\|^p - \|\nabla \lg v_\varepsilon\|^p\big)\,\mathrm{d}x \\
&\qquad - \int_\Omega p v_\varepsilon^p \|\nabla \lg u_\varepsilon\|^{p-2}\big\langle\nabla \lg u_\varepsilon, (\nabla \lg v_\varepsilon - \nabla \lg u_\varepsilon)\big\rangle\,\mathrm{d}x \\
&\qquad - \int_\Omega p u_\varepsilon^p \|\nabla \lg v_\varepsilon\|^{p-2}\big\langle\nabla \lg v_\varepsilon, (\nabla \lg u_\varepsilon - \nabla \lg v_\varepsilon)\big\rangle\,\mathrm{d}x
\end{aligned} \tag{15.7}$$

and the last term is nonpositive by the inequality given in Lemma 15.1.

It is obvious that

$$\lim_{\varepsilon\to 0+}\int_\Omega\left[\frac{u^{p-1}}{u_\varepsilon^{p-1}}-\frac{v^{p-1}}{v_\varepsilon^{p-1}}\right]\bigl(u_\varepsilon^p-v_\varepsilon^p\bigr)\,\mathrm{d}x=0. \tag{15.8}$$

Let us first consider the case $p\geqslant 2$. According to inequality (15.5) we have

$$\begin{aligned}0&\leqslant\frac{1}{2^{p-1}-1}\int_\Omega\left(\frac{1}{v_\varepsilon^p}+\frac{1}{u_\varepsilon^p}\right)\|v_\varepsilon\nabla u_\varepsilon-u_\varepsilon\nabla v_\varepsilon\|^p\,\mathrm{d}x\\&\leqslant-\lambda_1\int_\Omega\left[\left(\frac{u}{u_\varepsilon}\right)^{p-1}-\left(\frac{v}{v_\varepsilon}\right)^{p-1}\right]\bigl(u_\varepsilon^p-v_\varepsilon^p\bigr)\,\mathrm{d}x\end{aligned}$$

for every $\varepsilon>0$ (here we have used inequality (15.5) with $w_1=\nabla\lg u_\varepsilon$, $w_2=\nabla\lg v_\varepsilon$ and vice versa). In view of (15.8), taking a sequence $\varepsilon_k\to 0+$ as $k\to\infty$ and using Fatou's lemma in the previous computations we finally arrive to the conclusion that $v\nabla u=u\nabla v$ a.e. in Ω. Hence there is a constant κ such that $u=\kappa v$ a.e. in Ω and by continuity this equality holds everywhere in Ω.

Now we turn the attention to the case $1<p<2$ where the situation is similar as in the previous case. Applying the inequality (15.6) in (15.7) we obtain

$$\begin{aligned}0&\leqslant C(p)\int_\Omega(u_\varepsilon v_\varepsilon)^p\bigl(u_\varepsilon^p+v_\varepsilon^p\bigr)\frac{\|v_\varepsilon\nabla u_\varepsilon-u_\varepsilon\nabla v_\varepsilon\|^2}{(v_\varepsilon\|\nabla u_\varepsilon\|+u_\varepsilon\|\nabla v_\varepsilon\|)^{2-p}}\,\mathrm{d}x\\&\leqslant-\lambda_1\int_\Omega\left[\left(\frac{u}{u_\varepsilon}\right)^{p-1}-\left(\frac{v}{v_\varepsilon}\right)^{p-1}\right]\bigl(u_\varepsilon^p-v_\varepsilon^p\bigr)\,\mathrm{d}x\end{aligned}$$

for every $\varepsilon>0$. Using (15.8), we again arrive at the desired dependence $u=\kappa v$ for some constant κ.

As for the isolation of the first eigenvalue λ_1, we proceed as follows. Since λ_1 is defined as the minimum of the quotient (15.4), it is isolated from the left. If v is an eigenfunction associated with an eigenvalue $\lambda>\lambda_1$ then v changes its sign in Ω. In fact, suppose that v does not change its sign in Ω. Then using the same method as in the previous part of the proof we get (for details we refer to [11])

$$0\leqslant\int_\Omega(\lambda_1-\lambda)\bigl(u^p-v^p\bigr)\,\mathrm{d}x=(\lambda_1-\lambda)\left(\frac{1}{\lambda_1}-\frac{1}{\lambda}\right)$$

what is a contradiction.

Now, suppose, by contradiction, that there exists a sequence of eigenvalues $\lambda_n\to\lambda_1+$ and let u_n be the sequence of associated eigenfunctions such that $\|u_n\|=1$. This sequence contains a weakly convergent subsequence in $W_0^{1,p}(\Omega)$, denoted again u_n, and hence strongly convergent in $L^p(\Omega)$. Since $u_n=-\Delta_p^{-1}(\Phi(u_n))$ (this is a usual argument in the theory of partial equations with p-Laplacian, we refer, e.g., to the monograph [103]), the sequence u_n converges strongly in $W_0^{1,p}(\Omega)$ to a function of the $W_0^{1,p}$ norm equal to

1 associated with λ_1. However, by the Jegorov theorem, the sequence u_n converges uniformly to a function u except for a set of arbitrarily small Lebesgue measure. However, this is a contradiction with the fact that the eigenfunction associated with the first eigenvalue does not change its sign in Ω. ☐

15.2. *Picone's identity for equations with p-Laplacian*

Picone's identity as presented in this subsection was proved in [118]. However, this identity can be found in various modifications (sometimes implicitly) also in other papers, e.g. in [9,10,83].

Consider a pair of partial differential operators with p-Laplacian

$$l[u] := \operatorname{div}\bigl(r(x)\|\nabla u\|^{p-2}\nabla u\bigr) + c(x)\Phi(u) = 0$$

and

$$L[u] := \operatorname{div}\bigl(R(x)\|\nabla u\|^{p-2}\nabla u\bigr) + C(x)\Phi(u) = 0.$$

It is assumed that r, c, R, C are defined in some bounded domain $G \subset \mathbb{R}^N$ with piecewise smooth boundary ∂G and that $r, R \in C^1(\overline{G})$ are positive functions in $\overline{G}$, and $c, C \in C(\overline{G})$. The domain $\mathcal{D}_l(G)$ of l is defined to be the set of all functions of the class $C^1(\overline{G})$ with the property that $r\|\nabla u\|^{p-2}\nabla u \in C^1(G) \cap C(\overline{G})$. The domain $\mathcal{D}_L(G)$ of L is defined similarly.

The proof of the below given N-dimensional extension of Picone's identity is similar to that given in Section 1.

THEOREM 15.3. *Let $u \in \mathcal{D}_l(G)$, $v \in \mathcal{D}_L(G)$ and $v(x) \neq 0$ for $x \in G$. Then*

$$\begin{aligned}
&\operatorname{div}\left(\frac{u}{\Phi(v)}\bigl[\Phi(v)r(x)\|\nabla u\|^{p-2}\nabla u - \Phi(u)R(x)\|\nabla v\|^{p-2}v\bigr]\right)\\
&\quad = \bigl[r(x) - R(x)\bigr]\|\nabla u\|^p + \bigl[C(x) - c(x)\bigr]|u|^p\\
&\qquad + R(x)\left[\|\nabla u\|^p + (p-1)\left\|\frac{u}{v}\nabla v\right\|^p - p\left\|\frac{u}{v}\nabla v\right\|^{p-2}(\nabla u)\left(\frac{u}{v}\nabla v\right)\right]\\
&\qquad + \frac{u}{\Phi(v)}\bigl[\Phi(v)l[u] - \Phi(u)L[v]\bigr].
\end{aligned}$$

Taking $r = R$, $c = C$ in the previous theorem, and using the fact that if v is a solution of $l[v] = 0$ for which $v(x) \neq 0$ in G, then the function $w = \frac{r(x)\|\nabla v\|^{p-2}\nabla v}{\Phi(v)}$ is a solution of the Riccati-type partial differential equation

$$\operatorname{div} w + c(x) + (p-1)r^{1-q}(x)\|w\|^q = 0, \quad q = \frac{p}{p-1},$$

we have Picone's identity in the special form

$$\begin{aligned}&r(x)\|\nabla u\|^p - c(x)|u|^p\\&\quad = \operatorname{div}\bigl(w(x)|u|^p\bigr) + pr^{1-q}(x)\widetilde{P}\bigl(r^{q-1}(x)\nabla u, w(x)\Phi(u)\bigr),\end{aligned}$$

where

$$\widetilde{P}(x,y) = \frac{\|x\|^p}{p} - \langle x,y\rangle + \frac{\|y\|^q}{q}.$$

As a consequence of Theorem 15.3 we have the following extension of the Leighton comparison theorem. The proof of this statement is again similar to the "ordinary" case, compare Section 8.1.

THEOREM 15.4. *Suppose that the boundary ∂G is of the class C^1. If there exists a nontrivial solution $u \in \mathcal{D}_l(G)$ of $l[u]=0$ such that $u=0$ on ∂G and*

$$\int_G \bigl\{\bigl[R(x)-r(x)\bigr]\|\nabla u\|^p - \bigl[C(x)-c(x)\bigr]|u|^p\bigr\}\,\mathrm{d}x \leqslant 0,$$

then every solution $v \in \mathcal{D}_L(G)$ of $L[v]=0$ must vanish at some point of G, unless v is a constant multiple of u.

Another consequence of Picone's identity is the following Sturmian separation theorem.

THEOREM 15.5. *Suppose that G is the same as in the previous theorem and there exists a nontrivial solution $u \in \mathcal{D}_l(G)$ of $l[u]=0$ with $u=0$ on ∂G. Then every solution v of $l[u]=0$ must vanish at some point of G, unless v is a constant multiple of u.*

REMARK 15.1. (i) There exist numerous papers dealing with various oscillation and spectral properties of PDEs with p-Laplacian. We recall here at least the papers [12,18,47,48, 70,75,80], but this is only a very limited sample of papers where equations with the p-Laplacian are treated.

(ii) If we study properties of solutions of PDEs with p-Laplacian

$$\operatorname{div}\bigl(\|\nabla u\|^{p-2}\nabla u\bigr) + c(x)\Phi(u) = 0 \tag{15.9}$$

in a radially symmetric domain $G = B_R = \{x \in \mathbb{R}^N\colon\ \|x\| \leqslant R\}$ with a radially symmetric potential c, i.e., $c(x) = b(\|x\|)$ for some $b:\mathbb{R}_+ \to \mathbb{R}$, then one can look for solutions in the radial form $u(x) = v(r) = v(\|x\|)$ and then v solved the ODE of the form (0.1)

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[r^{N-1}\Phi\left(\frac{\mathrm{d}}{\mathrm{d}r}v\right)\right] + r^{N-1}b(r) = 0.$$

This method of the investigation of oscillatory properties of (15.9) has been used, e.g., in [70,136], see also the references given therein.

15.3. *Second eigenvalue of p-Laplacian*

In this subsection we briefly mention the results of the paper [13] which deals with the variational description of the second eigenvalue of p-Laplacian and with a nodal domain property of the associated eigenfunction. We again suppose that Ω is a bounded domain in $\mathbb{R}^n$.

We consider the eigenvalue problem (15.2) and we introduce the functionals

$$A(u) = \frac{1}{p}\int_\Omega \|\nabla u\|^p \,\mathrm{d}x, \qquad B(u) = \frac{1}{p}\int_\Omega |u|^p \,\mathrm{d}x,$$
$$F(u) = A^2(u) - B(u).$$

It is clear that the critical point u of F associated to a critical value c (i.e., $F(u) = c$ and $F'(u) = 0$) is an eigenfunction associated to the eigenvalue

$$\lambda = \frac{1}{2\sqrt{-c}}.$$

Conversely, if $u \neq 0$ is an eigenfunction associated to a positive eigenvalue λ, $v = (2\lambda A(u))^{-\frac{1}{p}} u$ will be also an eigenfunction associated to $\lambda = \frac{1}{2A(v)}$ and v is a critical point of F associated to the critical value $c = -\frac{1}{4\lambda^2}$.

Let us consider the sequence $\{c_n\}_{n\in\mathbb{N}}$ defined by

$$c_n = \inf_{K\in\mathcal{A}_n} \sup_{v\in K} F(v), \tag{15.10}$$

where

$$\mathcal{A}_n = \bigl\{K \in W_0^{1,p}(\Omega)\colon\ K \text{ symmetrical compact and } \gamma(K) \geqslant n\bigr\}$$

and $\gamma(K)$ denotes the Krasnoselskii genus of K, i.e., the minimal integer n such that there exists a continuous odd mapping of $K \to \mathbb{R}^n \setminus \{0\}$. It can be proved (using the Palais–Smale condition for F) that the sequence c_n consists of the critical values of F and $c_n \to 0-$. The sequence of eigenvalues λ_n defined by

$$\lambda_n = \frac{1}{2\sqrt{-c_n}} \tag{15.11}$$

is positive, nondecreasing and tends to ∞. Note that it is an open problem whether (15.11) describes *all* eigenvalues of (15.2) (in contrast to the scalar case $N = 1$, compare Section 13.2).

We denote by $Z(u) = \{x \in \Omega\colon\ u(x) = 0\}$ the so-called *nodal contour* of the function u and let $N(u)$ denote the number of components (the so-called *nodal domains*) of $\Omega \setminus Z(u)$. For each eigenfunction u associated to λ we define

$$N(\lambda) = \max\bigl\{N(u)\colon\ u \text{ is a solution of (15.2)}\bigr\}.$$

Now, at the end of this subsection, we present without proof the main result of [13]. In contrast to the scalar case, it is not known whether (15.10) and (15.11) describe *all* eigenvalues of (15.2). The next statement shows, among others, that the second eigenvalue λ_2 can be characterized by (15.11).

THEOREM 15.6. *For each eigenvalue* λ *of* (15.2) $\lambda_{N(\lambda)} \leqslant \lambda$. *Moreover, the value* λ_2 *given by* (15.11) *satisfies*

$$\lambda_2 = \inf\bigl\{\lambda\colon\ \lambda \text{ positive eigenvalue of } (15.2),\ \lambda > \lambda_1\bigr\}.$$

15.4. *Equations involving pseudolaplacian*

Another partial differential equation which reduces to half-linear equation (9.1) in the "ordinary" case is the partial differential equation with the so-called *pseudolaplacian*

$$\tilde{\Delta}_p u := \sum_{i=1}^{N} \frac{\partial}{\partial x_i} \Phi\left(\frac{\partial u}{\partial x_i}\right).$$

We consider the partial differential equation

$$\tilde{\Delta}_p u + c(x)\Phi(u) = 0 \tag{15.12}$$

and the associated energy functional

$$\begin{aligned}\mathcal{F}_p(u; \Omega) :&= \int_\Omega \left\{\sum_{i=1}^{N} \left|\frac{\partial u}{\partial x_i}\right|^p - c(x)|u|^p\right\} \mathrm{d}x \\ &= \int_\Omega \left\{\|\nabla u\|_p^p - c(x)|u|^p\right\} \mathrm{d}x,\end{aligned}$$

where $\|x\|_p = (\sum_{i=1}^N |x_i|^p)^{1/p}$ denotes the p-norm in $\mathbb{R}^N$. Another important object associated with (15.12) is a Riccati-type equation which we obtain as follows. Let u be a solution of (15.12) which is nonzero in Ω and denote

$$v := \left(\Phi\left(\frac{\partial u}{\partial x_1}\right), \dots, \Phi\left(\frac{\partial u}{\partial x_n}\right)\right), \qquad w := \frac{v}{\Phi(u)}.$$

Then, using the fact that (15.12) can be written in the form $\operatorname{div} v = -c(x)\Phi(u)$, we have

$$\begin{aligned}\operatorname{div} w &= \frac{1}{\Phi^2(u)}\bigl\{\Phi(u)\operatorname{div} v - \Phi'(u)\langle\nabla u, v\rangle\bigr\} \\ &= -c(x) - (p-1)\frac{|u|^{p-2}}{|u|^{2p-2}} \sum_{i=1}^{N} \left|\frac{\partial u}{\partial x_i}\right|^p\end{aligned}$$

$$
\begin{aligned}
&= -c(x) - (p-1)\sum_{i=1}^{N}\left|\Phi\left(\frac{\partial u/\partial x_i}{u}\right)\right|^q \\
&= -c(x) - (p-1)\|w\|_q^q,
\end{aligned}
$$

where $\langle\cdot,\cdot\rangle$ denotes the usual scalar product in $\mathbb{R}^N$, $q := \frac{p}{p-1}$ is the conjugate exponent of p and $\|x\|_q = (\sum_{i=1}^N |x_i|^q)^{1/q}$ denotes the q-norm in $\mathbb{R}^N$. Consequently, the vector variable w satisfies the Riccati-type equation

$$
\operatorname{div} w + c(x) + (p-1)\|w\|_q^q = 0. \tag{15.13}
$$

For Equation (15.12) we can establish oscillation theory and theory for eigenvalue problems similar to that for classical p-Laplacian. An important role in this theory is played by the following Picone-type identity. For its proof and other results concerning PDEs with pseudolaplacian we refer to [25,26,28,62] and the reference given therein.

THEOREM 15.7. *Let w be a solution of* (15.13) *which is defined in $\bar{\Omega}$ and $u \in W^{1,p}(\Omega)$. Then*

$$
\begin{aligned}
\mathcal{F}_p(u;\Omega) = &\int_{\partial\Omega} \left|u(x)\right|^p w(x)\,\mathrm{d}S \\
&+ p\int_{\Omega}\left\{\frac{\|\nabla u(x)\|_p^p}{p} - \left\langle \nabla u(x), \Phi\big(u(x)\big)w(x)\right\rangle \right. \\
&\left. + \frac{\|w(x)\|_q^q |\Phi(u(x))|^q}{q}\right\}\mathrm{d}x.
\end{aligned}
$$

Moreover, the last integral in this formula is always nonnegative, it equals zero only if $u \neq 0$ in $\bar{\Omega}$ and

$$
w = \frac{1}{\Phi(u)}\left(\Phi\left(\frac{\partial u}{\partial x_1}\right), \ldots, \Phi\left(\frac{\partial u}{\partial x_n}\right)\right).
$$

16. Half-linear difference equations

In the last two decades, a considerable attention has been devoted to the oscillation theory of the Sturm–Liouville difference equation

$$
\Delta(r_k \Delta x_k) + c_k x_{k+1} = 0, \tag{16.1}
$$

where $\Delta x_k = x_{k+1} - x_k$ is the usual forward difference, r, c are real-valued sequences and $r_k \neq 0$. Oscillation theory parallel to that for the Sturm–Liouville differential equation (1.1) has been established and almost all oscillation and nonoscillation criteria have now their discrete counterparts for (16.1). We refer to monographs [1,6,124] for general background. Basic tools of the linear discrete oscillation theory are the discrete quadratic functional

$$\mathcal{F}_d(x;0,N)=\sum_{k=0}^{N}\left[r_k(\Delta x_k)^2-c_kx_{k+1}^2\right],$$

the Riccati difference equation (related to (16.1) by the substitution $w=\frac{r\Delta x}{x}$)

$$\Delta w_k+c_k+\frac{w_k^2}{r_k+w_k}=0 \tag{16.2}$$

and the link between them, the (reduced) discrete Picone identity

$$\mathcal{F}_d(x;0,N)=w_ky_k^2\big|_{k=0}^{N+1}+\sum_{k=0}^{N}\frac{1}{r_k+w_k}(r_k\Delta x_k-w_kx_k)^2,$$

w being a solution of the Riccati equation, which is defined for $k=0,\dots,N+1$.

A natural idea, suggested by similarity of oscillation theories for linear equation (1.1) and half-linear equation (0.1), is to look for half-linear extension of these results and to establish a *discrete* half-linear oscillation theory parallel to that for (0.1). Therefore, the subject of this section is the half-linear *difference* equation

$$\Delta\big(r_k\Phi(\Delta x_k)\big)+c_k\Phi(x_{k+1})=0, \tag{16.3}$$

where r,c are real-valued sequences and $r_k\neq 0$. We will see that the results for (16.3) are similar to those for (0.1), but the proofs are sometimes more difficult. The reason is that the calculus of finite differences and sums is sometimes more cumbersome than the differential and integral calculus. For example, we have no discrete analogue of the chain rule for the differentiation of the composite function. On the other hand, there are some points where the discrete calculus is "easier", for example, if an infinite series $\sum^{\infty}a_k$ is convergent, we have $\lim_{n\to\infty}a_n=0$, while the convergence of the integral $\int^{\infty}f(t)\,\mathrm{d}t$ gives generally no information about $\lim_{t\to\infty}f(t)$. Most of the results of this section are taken from the papers of Řehák [194–199].

16.1. *Roundabout theorem for half-linear difference equations*

The basic results of the discrete half-linear oscillation theory are established in the series of papers [194–199]. Here we present principal results of this theory.

First of all, let us note that in contrast to the continuous case, there is no problem with the existence and uniqueness for solutions of (16.3). Expanding the forward differences, this equation can be written as

$$r_{k+1}\Phi(x_{k+2}-x_{k+1})-r_k\Phi(x_{k+1}-x_k)+c_k\Phi(x_{k+1})=0$$

and hence

$$x_{k+2} = x_{k+1} + \Phi^{-1}\left(\frac{1}{r_{k+1}}\big[r_k\Phi(x_{k+1} - x_k) - c_k\Phi(x_{k+1})\big]\right).$$

This means that given the initial conditions $x_0 = A, x_1 = B$, we can compute explicitly all other x_k. Moreover, given any $N \in \mathbb{N}$, the values $x_2, \dots, x_N$ depend continuously (in the norm of $\mathbb{R}^{N-1}$) on x_0, x_1. Let us also emphasize that general discrete oscillation theory can be established under the mere assumption $r_k \neq 0$, while we have to suppose that $r(t) > 0$ in the continuous case.

Oscillatory properties of (16.3) are defined using the concept of generalized zero points. We say that an interval $(m, m+1]$ contains a *generalized zero* of a solution x of (16.3) if $x_m \neq 0$ and $x_m x_{m+1} r_m \leqslant 0$. If $r_m > 0$, a generalized zero of x is just the zero of x at $m+1$ or the sign change $x_m x_{m+1} < 0$.

THEOREM 16.1. *The following statements are equivalent*:

(i) *Equation* (16.3) *is disconjugate on* $[0, N]$, *i.e., the solution* $\tilde{x}$ *given by the initial conditions* $\tilde{x}_0 = 0$, $r_0\Phi(\tilde{x}_1) = 1$ *has no generalized zero in* $(0, N+1]$.

(ii) *There exists a solution of* (16.3) *having no generalized zero in* $[0, N+1]$.

(iii) *There exists a solution* w *of the Riccati-type difference equation* (*related to* (16.3) *by the substitution* $w_k = r_k\Phi(\Delta x_k/x_k)$)

$$\Delta w_k + c_k + w_k\left(1 - \frac{r_k}{\Phi(\Phi^{-1}(r_k) + \Phi^{-1}(w_k))}\right) = 0 \tag{16.4}$$

which is defined for every $k \in [0, N+1]$ *and satisfies* $r_k + w_k > 0$ *for* $k \in [0, N]$.

(iv) *The discrete* p*-degree functional*

$$\mathcal{F}_d(y; 0, N) = \sum_{k=0}^{N}\big[r_k|\Delta y_k|^p - c_k|y_{k+1}|^2\big]$$

is positive for every nontrivial $y = \{y_k\}_{k=0}^{N+1}$ *satisfying* $y_0 = 0 = y_{N+1}$.

PROOF. (i) ⇒ (ii): Consider the solution x of (16.3) given by the initial condition $x_0 = \varepsilon$, $x_1 = \Phi^{-1}(1/r_0)$, where $\varepsilon > 0$ is sufficiently small. Then according to the above mentioned continuous dependence of $x_2, \dots, x_{N+1}$ on x_0, x_1 we still have $r_k x_k x_{k+1} > 0$, $k = 1, \dots, N$, when $\varepsilon > 0$ is sufficiently small, and $r_0 x_0 x_1 > 0$ as well, i.e., the solution x has no generalized zero in $[0, N+1]$.

(ii) ⇒ (iii): Let x be a solution of (16.3) having no generalized zeros in $[0, N+1]$, and let $w_k = \frac{r_k\Phi(\Delta x_k)}{\Phi(x_k)}$. Then

$$\begin{aligned}\Delta w_k &= \frac{\Delta(r_k\Phi(\Delta x_k))\Phi(x_k) - r_k\Phi(\Delta x_k)(\Phi(x_{k+1}) - \Phi(x_k))}{\Phi(x_{k+1})\Phi(x_k)} \\ &= -c_k - w_k + \frac{r_k\Phi(\Delta x_k)}{\Phi(x_k + \Delta x_k)} = -c_k - w_k + \frac{r_k\Phi(\Delta x_k)}{\Phi(x_k)\Phi(1 + \frac{\Delta x_k}{x_k})}\end{aligned}$$

$$= -c_k - w_k\left(1 - \frac{1}{\Phi(1+\Phi^{-1}(\frac{w_k}{r_k}))}\right)$$
$$= -c_k - w_k\left(1 - \frac{r_k}{\Phi(\Phi^{-1}(r_k)+\Phi^{-1}(w_k))}\right).$$

Moreover, $r_k x_k x_{k+1} > 0 \Leftrightarrow r_k\Phi(x_k)\Phi(x_{k+1}) > 0$ and

$$\begin{aligned} r_k\Phi(x_k)\Phi(x_{k+1}) &= r_k\Phi(x_k)\Phi(x_k + \Delta x_k) \\ &= \Phi^2(x_k)\Phi\left(\Phi^{-1}(r_k) + \Phi^{-1}(r_k)\frac{\Delta x_k}{x_k}\right) \\ &= \Phi^2(x_k)\Phi\big(\Phi^{-1}(r_k) + \Phi^{-1}(w_k)\big), \end{aligned}$$

hence $r_k x_k x_{k+1} > 0$ if and only if $\Phi^{-1}(r_k) + \Phi^{-1}(w_k) > 0$, i.e., if and only if $r_k + w_k > 0$.

(iii) $\Rightarrow$ (iv): Let w be a solution of (16.4) such that $r_k + w_k > 0$, $k = 0, \dots, N$. Then

$$w_{k+1} = -c_k + \frac{r_k w_k}{\Phi(\Phi^{-1}(r_k) + \Phi^{-1}(w_k))}$$

and for any sequence $y = \{y_k\}_{k=0}^{N+1}$ we have

$$\begin{aligned} \Delta\big(w_k|y_k|^p\big) &= w_{k+1}|y_{k+1}|^p - w_k|y_k|^p \\ &= -c_k|y_{k+1}|^p + \frac{r_k w_k|y_k + \Delta y_k|^p}{\Phi(\Phi^{-1}(w_k) + \Phi^{-1}(r_k))} \\ &\quad - w_k|y_k|^p + r_k|\Delta y_k|^p - r_k|\Delta y_k|^p. \end{aligned}$$

Using the fact that $y_0 = 0 = y_{N+1}$, the summation of the last equality from $k = 0$ to $k = N$ gives

$$\mathcal{F}_d(y; 0, N) = r_k|\Delta y_k|^p - \frac{r_k w_k|y_k + \Delta y_k|^p}{\Phi(\Phi^{-1}(w_k) + \Phi^{-1}(r_k))} - w_k|y_k|^p. \tag{16.5}$$

The right-hand side of (16.5) is always nonnegative and it is zero if and only if $w_k = r_k\Phi(\Delta y_k/y_k)$ (see [195]), but this means that $y \equiv 0$ since $y_0 = 0$. Hence $\mathcal{F}_d(y; 0, N) > 0$ for every nontrivial y with $y_0 = 0 - y_{N+1}$.

(iv) $\Rightarrow$ (i): Suppose that $\mathcal{F}_d > 0$ and (16.3) is not disconjugate in $[0, N+1]$, i.e., the solution x given by the initial condition $x_0 = 0$, $x_1 = \Phi^{-1}(1/r_0)$ has a generalized zero in the interval $[0, N+1]$, i.e., $r_m x_m x_{m+1} < 0$ or $x_{m+1} = 0$ for some $m \in \{1, \dots, N\}$. Define $y = \{y_k\}_{k=0}^{N+1}$ as follows

$$y_k = \begin{cases} x_k, & k = 0, \dots, m, \\ 0, & k = m+1, \dots, N+1. \end{cases}$$

Then we have (using summation by parts applied to $\mathcal{F}_d(x; 0, m-1)$)

$$\begin{aligned}\mathcal{F}_d(y; 0, N) &= \mathcal{F}_d(x; 0, m-1) + \left[r_m |\Delta y_m|^p\right] = r_k \Phi(\Delta x_k) x_k \Big|_0^m + r_m |x_m|^p \\ &= |x_m|^p \left[r_m \frac{\Phi(\Delta x_m)}{\Phi(x_m)} + r_m \right] = |x_m|^p [w_m + r_m] \leqslant 0\end{aligned}$$

since $w_m + r_m \leqslant 0$ if and only if $r_m x_m x_{m+1} \leqslant 0$ as we have shown in the previous part of this proof. □

REMARK 16.1. (i) The previous theorem shows that (16.3) can be classified as oscillatory or nonoscillatory in the same way as in the continuous case. Equation (16.3) is said to be *nonoscillatory* if there exists $N \in \mathbb{N}$ such that (16.3) is disconjugate on $[N, M]$ for every $M > N$, in the opposite case (16.3) is said to be *oscillatory*.

(ii) Theorem 16.1 also shows that not only Sturmian separation, but also Sturmian comparison theorem extends verbatim to (16.3). In particular, if $0 \neq R_k \leqslant r_k$ and $C_k \geqslant c_k$ for large k and the equations

$$\Delta\big(R_k \Phi(\Delta y_k)\big) + C_k \Phi(y_{k+1}) = 0$$

is nonoscillatory, then (16.3) is also nonoscilatory. The argument in the proof of this statement is the same as that for (0.1).

16.2. *Discrete Leighton–Wintner criterion*

In this criterion, similarly as in the continuous case, Equation (16.3) is viewed as a perturbation of the one-term equation

$$\Delta\big(r_k \Phi(\Delta x_k)\big) = 0. \tag{16.6}$$

In accordance with the continuous case, we need (16.6) to be nonoscillatory in this approach, so we suppose that $r_k > 0$ for large k, otherwise this equation is oscillatory—each sign change of r_k is a generalized zero of the constant solution $x_k \equiv 1$.

THEOREM 16.2. *Suppose that $r_k > 0$ for large k,*

$$\sum^{\infty} r_k^{1-q} = \infty \quad \textit{and} \quad \sum^{\infty} c_k = \infty. \tag{16.7}$$

Then (16.3) *is oscillatory.*

PROOF. We present here the complete proof in order to show that its idea is exactly the same as in the continuous case. Let $N \in \mathbb{N}$ be arbitrary. Define the class of sequences

$$\mathcal{D}(N) := \big\{ y = \{y_k\}_{k=N}^{\infty},\ y_N = 0,\ \exists M > N\colon\ y_k = 0 \text{ for } k \geqslant M \big\} \tag{16.8}$$

and for $N < n < m < M$ (which will be determined later) define a sequence $y \in \mathcal{D}(N)$ as follows

$$y_k = \begin{cases} 0, & k = N, \\ \big(\sum_{j=N}^{k-1} r_j^{1-q}\big)\big(\sum_{j=N}^{n-1} r_j^{1-q}\big)^{-1}, & N+1 \leqslant k \leqslant n, \\ 1, & n+1 \leqslant k \leqslant m-1, \\ \big(\sum_{j=k}^{M-1} r_j^{1-q}\big)\big(\sum_{j=m}^{M-1} r_j^{1-q}\big)^{-1}, & m \leqslant k \leqslant M-1, \\ 0, & k \geqslant M. \end{cases}$$

Then we have

$$\begin{aligned} \mathcal{F}_d(y; N, \infty) &= \sum_{k=N}^{\infty} \big[r_k |\Delta y_k|^p - c_k |y_{k+1}|^p\big] = \sum_{k=N}^{M-1} \big[r_k |\Delta y_k|^p - c_k |y_{k+1}|^p\big] \\ &= \left(\sum_{k=N}^{n-1} + \sum_{k=n}^{m-1} + \sum_{k=m}^{M-1}\right) \big[r_k |\Delta y_k|^p - c_k |y_{k+1}|^p\big] \\ &= \left(\sum_{k=N}^{n-1} r_k^{1-q}\right)^{-1} - \sum_{k=N}^{n-1} c_k |y_{k+1}|^p - \sum_{k=n}^{m-1} c_k - \sum_{k=m}^{M-1} c_k |y_{k+1}|^p \\ &\quad + \left(\sum_{k=m}^{M-1} r_k^{1-q}\right)^{-1}. \end{aligned}$$

Now, using the discrete version of the second mean value theorem of the summation calculus (see, e.g., [57]), there exists $\tilde{m} \in [m-1, M-1]$ such that

$$\sum_{k=m}^{M-1} c_k |y_{k+1}|^p \geqslant \sum_{k=m}^{\tilde{m}} c_k.$$

Let $n > N$ be fixed. Since (16.7) holds, for every $\varepsilon > 0$ there exist $M > m > n$ such that

$$\sum_{k=n}^{\tilde{m}} c_k > \mathcal{F}_d(y; N, n-1) + \varepsilon \quad \text{whenever } \tilde{m} > m \text{ and } \left(\sum_{k=m}^{M-1} r_k^{1-q}\right)^{-1} < \varepsilon.$$

Consequently, we have

$$\mathcal{F}_d(y; N, \infty) \leqslant \mathcal{F}_d(y; N, n-1) - \sum_{k=n}^{\tilde{m}} c_k + \left(\sum_{k=m}^{M-1} r_k^{1-q}\right)^{-1} < 0$$

what we needed to prove. □

In Section 9 we have presented an alternative proof of the continuous Leighton–Wintner criterion—based on the Riccati technique. Next we show difficulties in the attempt to follow this idea also in the discrete case. The "Riccati" proof goes by contradiction. Suppose that (16.7) holds and (16.3) is nonoscillatory, i.e., there exists a solution of (16.4) satisfying $r_k + w_k > 0$ for large k. The summation of (16.4) from N to $k-1$, where N, k are sufficiently large, we have

$$\begin{aligned} w_k &= w_N - \sum_{j=N}^{k-1} c_j - \sum_{j=N}^{k-1} w_j\left(1 - \frac{r_j}{\Phi(\Phi^{-1}(r_j) + \Phi^{-1}(w_j))}\right) \\ &\leqslant -\sum_{j=N}^{k-1} w_j\left(1 - \frac{r_j}{\Phi(\Phi^{-1}(r_j) + \Phi^{-1}(w_j))}\right) =: G_k. \end{aligned}$$

In the continuous case we obtained the analogous inequality

$$w(t) \leqslant -(p-1)\int_T^t r^{1-q}(s)\left|w(s)\right|^q \mathrm{d}s =: G(t)$$

which leads to the inequality

$$\frac{G'(t)}{G^q(t)} \leqslant \frac{r^{1-q}}{\int_T^t r^{1-q}(s)\,\mathrm{d}s} \tag{16.9}$$

and integrating this inequality we got $\int^\infty r^{1-q}(t)\,\mathrm{d}t < \infty$, a contradiction.

The discrete analogue of (16.9) is the inequality $w_k \leqslant G_k$ and to get a contradiction from this inequality is a difficult problem even in the linear case $p = 2$.

16.3. *Riccati inequality*

The equivalence of disconjugacy of (16.3) and solvability of (16.4) (satisfying $r_k + w_k > 0$), coupled with the Sturmian comparison theorem for (16.3) mentioned in Remark 16.1, lead to the following refinement of the Riccati equivalence.

THEOREM 16.3. *Equation* (16.3) *is nonoscillatory if and only if there exists a sequence* w_k, *with* $r_k + w_k > 0$ *for large* k, *such that*

$$R[w_k] := \Delta w_k + c_k + w_k\left(1 - \frac{r_k}{\Phi(\Phi^{-1}(w_k) + \Phi^{-1}(r_k))}\right) \leqslant 0. \tag{16.10}$$

PROOF. The part "only if" follows immediately from Theorem 16.1. For the part "if", let us denote $l[u_k] := \Delta(r_k\Phi(\Delta u_k)) + c_k\Phi(u_{k+1})$. We will show that if there exists u_k such that

$$r_k u_k u_{k+1} > 0 \quad \text{and} \quad u_{k+1} l[u_k] \leqslant 0 \tag{16.11}$$

for $k \in [N, \infty)$, $N \in \mathbb{N}$, then (16.3) is disconjugate on $[N, \infty)$ and thus nonoscillatory. Therefore, suppose that a sequence u_k satisfying (16.11) on $[N, \infty)$ exists. Then $S_k := -u_{k+1}l[u_k]$ is a nonnegative sequence on this discrete interval. Further, set $\tilde{r}_k = r_k$ and $\tilde{c}_k = c_k - S_k/|u_{k+1}|^p$. Hence $\tilde{c}_k \geqslant c_k$ and

$$\Delta\big(\tilde{r}_k\Phi(\Delta u_k)\big) + \tilde{c}_k\Phi(u_{k+1}) = \Delta\big(r_k\Phi(\Delta u_k)\big) + \left(c_k - \frac{S_k}{|u_{k+1}|^p}\right)\Phi(u_{k+1}) = 0.$$

Thus equation $\Delta(\tilde{r}_k\Phi(\Delta u_k)) + \tilde{c}_k\Phi(u_{k+1}) = 0$ is disconjugate on $[N, \infty)$ and therefore (16.3) is also disconjugate on $[N, \infty)$ by Sturm comparison theorem.

To finish the proof, it remains to find a sequence u_k satisfying (16.11). Let w_k satisfy (16.10) with $r_k + w_k > 0$ on $[N, \infty)$ and let

$$u_k = \prod_{j=N}^{k-1} \big(1 + \Phi^{-1}(w_j/r_j)\big), \quad k > N,$$

be a solution of the first order difference equation

$$\Delta u_k = \Phi^{-1}(w_k/r_k)u_k, \quad u_N = 1,$$

Then $u_k \neq 0$ since

$$1 + \Phi^{-1}(w_k/r_k) = \frac{1}{\Phi^{-1}(r_k)}\big[\Phi^{-1}(r_k) + \Phi^{-1}(w_k)\big] \neq 0,$$

recall that $\Phi^{-1}(r_k) + \Phi^{-1}(w_k) > 0$ if and only if $w_k + r_k > 0$. Further,

$$\begin{aligned} u_{k+1}l[u_k] &= u_{k+1}\big[\Delta\big(r_k\Phi(\Delta u_k)\big) + c_k\Phi(u_{k+1})\big] - \frac{|u_{k+1}|^p r_k\Phi(\Delta u_k)\Delta\Phi(u_k)}{\Phi(u_k)\Phi(u_{k+1})} \\ &\quad + \frac{|u_{k+1}|^p r_k\Phi(\Delta u_k)\Delta\Phi(u_k)}{\Phi(u_k)\Phi(u_{k+1})} \\ &= u_{k+1}\Phi(u_{k+1})\frac{\Delta(r_k\Phi(\Delta u_k))\Phi(u_k) - r_k\Phi(\Delta u_k)\Delta\Phi(u_k)}{\Phi(u_k)\Phi(u_{k+1})} \\ &\quad + |u_{k+1}|^p c_k + |u_{k+1}|^p\frac{r_k\Phi(\Delta u_k)}{\Phi(u_k)}\left(1 - \frac{\Phi(u_k)}{\Phi(u_{k+1})}\right) \\ &= |u_{k+1}|^p R[w_k] \leqslant 0, \end{aligned}$$

for $k \in [N, \infty)$, since $w_k = \frac{r_k\Phi(\Delta u_k)}{\Phi(u_k)}$ and

$$\frac{\Phi(u_k)}{\Phi(u_{k+1})} = \frac{1}{\Phi(1 + \frac{\Delta u_k}{u_k})} = \frac{r_k}{\Phi(\Phi^{-1}(r_k) + \Phi^{-1}(w_k))}.$$

This completes the proof. ☐

16.4. *Hille–Nehari nonoscillation criterion*

The following theorem is the discrete version of Theorem 5.4. This statement is one of the main results of [71].

THEOREM 16.4. *Suppose that* $r_k > 0$ *for large* k, $\sum^{\infty} c_k = \lim_{k\to\infty} \sum^{k} c_j$ *is convergent and*

$$\lim_{k\to\infty} \frac{r_k^{1-q}}{\sum^{k-1} r_j^{1-q}} = 0. \tag{16.12}$$

If

$$\limsup_{k\to\infty} \left(\sum^{k-1} r_j^{1-q}\right)^{p-1} \left(\sum_{j=k}^{\infty} c_j\right) < \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1} \tag{16.13}$$

and

$$\liminf_{k\to\infty} \left(\sum^{k-1} r_j^{1-q}\right)^{p-1} \left(\sum_{j=k}^{\infty} c_j\right) > -\frac{2p-1}{p}\left(\frac{p-1}{p}\right)^{p-1} \tag{16.14}$$

then (16.3) *is nonoscillatory.*

PROOF. It is sufficient to show that the generalized Riccati inequality (16.10) has a solution w with $r_k + w_k > 0$ in a neighborhood of infinity. We recommend the reader to compare this proof with the proof of Theorem 5.4 to see difference between discrete and continuous case.

Set

$$w_k = C\left(\sum^{k-1} r_j^{1-q}\right)^{1-p} + \sum_{j=k}^{\infty} c_j, \tag{16.15}$$

where C is a suitable constant, which will be specified later. The following equalities hold by the Lagrange mean value theorem

$$\Delta\left(\sum^{k-1} r_j^{1-q}\right)^{1-p} = (1-p) r_k^{1-q} \eta_k^{-p},$$

where $\sum^{k-1} r_j^{1-q} \leqslant \eta_k \leqslant \sum^{k} r_j^{1-q}$. Similarly

$$1 - \frac{r_k}{\Phi(\Phi^{-1}(r_k) + \Phi^{-1}(w_k))}$$

$$= \frac{1}{\Phi(\Phi^{-1}(r_k)+\Phi^{-1}(w_k))}\left\{\Phi\big(\Phi^{-1}(r_k)+\Phi^{-1}(w_k)\big)-\Phi\big(\Phi^{-1}(r_k)\big)\right\}$$

$$= \frac{p-1}{\Phi(\Phi^{-1}(r_k)+\Phi^{-1}(w_k))}|\xi|^{p-2}\Phi^{-1}(w_k),$$

where ξ_k is between $\Phi^{-1}(r_k)$ and $\Phi^{-1}(r_k)+\Phi^{-1}(w_k)$. Hence

$$\Phi^{-1}(r_k)-\left|\Phi^{-1}(w_k)\right| \leqslant \xi_k \leqslant \Phi^{-1}(r_k)+\left|\Phi^{-1}(w_k)\right|$$

and

$$\frac{|w_k|}{r_k} = \left(\frac{r_k^{1-q}}{\sum^{k-1} r_j^{1-q}}\right)^{p-1}\left|C+\left(\sum^{k-1} r_j^{1-q}\right)^{p-1}\left(\sum_{j=k}^{\infty} c_j\right)\right|.$$

Hence $w_k/r_k \to 0$ for $k\to\infty$ according to (16.12), (16.13) and (16.14).

Further, we have

$$\begin{aligned}
&\Delta w_k + c_k - w_k\left(1-\frac{r_k}{\Phi(\Phi^{-1}(r_k)+\Phi^{-1}(w_k))}\right)\\
&= (1-p)Cr_k^{1-q}\eta_k^{-p} + \frac{(p-1)|w_k|^q\xi_k^{p-2}}{\Phi(\Phi^{-1}(r_k)+\Phi^{-1}(w_k))}\\
&= (p-1)r_k^{1-q}\left\{\frac{|w_k|^q|\xi_k|^{p-2}r_k^{q-1}}{\Phi(\Phi^{-1}(r_k)+\Phi^{-1}(w_k))}-\frac{C}{\eta_k^p}\right\}\\
&= (p-1)\frac{r_k^{1-q}}{(\sum^{k-1} r_j^{1-q})^p}\left\{\frac{|C+(\sum^{k-1} r_j^{1-q})^{p-1}\sum_{j=k}^{\infty} c_j|^q|\xi_k|^{p-2}r_k^{q-1}}{\Phi(\Phi^{-1}(r_k)+\Phi^{-1}(w_k))}\right.\\
&\qquad\left.-\frac{C(\sum^{k-1} r_j^{1-q})^p}{\eta_k^p}\right\}\\
&\leqslant \frac{(p-1)r_k^{1-q}}{(\sum^{k-1} r_j^{1-q})^p}\left\{\left|C+\left(\sum^{k-1} r_j^{1-q}\right)^{p-1}\left(\sum_{j=k}^{\infty} c_j\right)\right|^q\gamma_k - \frac{C(\sum^{k-1} r_j^{1-q})^p}{(\sum^k r_k^{1-q})^p}\right\},
\end{aligned}$$

where

$$\gamma_k = \frac{|\xi_k|^{p-2}r_k^{1-q}}{\Phi(\Phi^{-1}(r_k)+\Phi^{-1}(w_k))}.$$

Concerning the asymptotic behavior of this sequence and of $(\sum^{k-1} r_j^{1-q})(\sum^{k} r_k^{1-q})^{-1}$, we have

$$\lim_{k\to\infty} \frac{\sum^k r_j^{1-q}}{\sum^{k-1} r_j^{1-q}} = \lim_{k\to\infty} \frac{r_k^{1-q} + \sum^{k-1} r_j^{1-q}}{\sum^{k-1} r_j^{1-q}} = 1$$

since (16.12) holds. Further,

$$\frac{|\xi_k|^{p-2} r_k^{1-q}}{\Phi(\Phi^{-1}(r_k) + \Phi^{-1}(w_k))} \leqslant \frac{r_k^{1-q}(\Phi^{-1}(r_k) + \Phi^{-1}(|w_k|))^{p-2}}{\Phi(\Phi^{-1}(r_k) - \Phi^{-1}(|w_k|))} \to 1$$

as $k \to \infty$ since $\frac{|w_k|}{r_k} \to 0$ as $k \to \infty$. Consequently,

$$\limsup_{k\to\infty} \gamma_k \leqslant 1. \tag{16.16}$$

Now, inequalities (16.13), (16.14) imply the existence of $\varepsilon > 0$ such that

$$\begin{aligned} &-\frac{2p-1}{p}\left(\frac{p-1}{p}\right)^{p-1} + \varepsilon \\ &\quad < \left(\sum^{k-1} r_j^{1-q}\right)^{p-1} \left(\sum_{j=k}^{\infty} c_j\right) < \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1} - \varepsilon \end{aligned} \tag{16.17}$$

for k sufficiently large. Let $\tilde{\gamma}_k = \gamma_k^{\frac{1}{q}}$, $\tilde{\varepsilon} = \varepsilon(\frac{p}{p-1})^{p-1}$ and let $C = (\frac{p-1}{p})^p$ in (16.15). According to (16.16) $\tilde{\gamma}_k < \frac{1}{1-\tilde{\varepsilon}}$ for large k and

$$\begin{aligned} \tilde{\gamma}_k < \frac{1}{1-\tilde{\varepsilon}} &\iff 1 > \left[1 - \left(\frac{p}{p-1}\right)^{p-1} \varepsilon\right]\tilde{\gamma}_k \\ &\iff \left(\frac{p-1}{p}\right)^{p-1} > \left(\frac{p-1}{p}\right)^{p-1} \left[\frac{p-1}{p} + \frac{1}{p} - \left(\frac{p}{p-1}\right)^{p-1} \varepsilon\right]\tilde{\gamma}_k \\ &\iff \left(\frac{p-1}{p}\right)^{\frac{p}{q}} > \left[\left(\frac{p-1}{p}\right)^{p} + \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1} - \varepsilon\right]\tilde{\gamma}_k \\ &\iff \frac{C^{\frac{1}{q}} - C\tilde{\gamma}_k}{\tilde{\gamma}_k} > \frac{1}{p}\left(\frac{p-1}{p}\right)^{p-1} - \varepsilon. \end{aligned}$$

Therefore, the second inequality in (16.17) implies

$$C^{\frac{1}{q}} > \left[C + \left(\sum^{k-1} r_j^{1-q}\right)^{p-1} \left(\sum_{j=k}^{\infty} c_j\right)\right]\tilde{\gamma}_k.$$

By a similar computation (using the first inequality in (16.17)) we get

$$-C^{\frac{1}{q}} < \left[C + \left(\sum^{k-1} r_j^{1-q}\right)^{p-1} \left(\sum_{j=k}^{\infty} c_j\right)\right]\tilde{\gamma}_k.$$

Consequently,

$$\left|C + \left(\sum^{k-1} r_j^{1-q}\right)^{p-1} \left(\sum_{j=k}^{\infty} c_j\right)\right|^q < C$$

for large k and hence $R[w_k] \leqslant 0$. Finally, since $r_k > 0$ and $w_k/r_k \to 0$ as $k \to \infty$, we have $r_k + w_k > 0$ for large k and the proof is complete. □

REMARK 16.2. If we compare the previous statement with Theorem 5.4 (which is a continuous counterpart of this theorem), we see that assumption (16.12) has no continuous analogue. This is a consequence of the fact that we have no equivalence of the differentiation chain rule in the discrete case, and its partial discrete substitution—the Lagrange mean value theorem—need additional assumptions.

16.5. *Half-linear dynamic equations on time scales*

By a *time scale* $\mathbb{T}$ (an alternative terminology is *measure chain*) we understand any closed subset of the real numbers $\mathbb{R}$ with the usual topology inherited from $\mathbb{R}$. Typical examples of time scales are $\mathbb{T} = \mathbb{R}$ and $\mathbb{T} = \mathbb{Z}$ – the set of integers. The operators $\rho, \sigma : \mathbb{T} \to \mathbb{T}$ are defined by

$$\sigma(t) = \inf\{s \in \mathbb{T}:\ s > t\}, \qquad \rho(t) = \sup\{s \in \mathbb{T}:\ s < t\}$$

and are called the *right jump operator* and *left jump operator*, respectively. The quantity $\mu(t) = \sigma(t) - t$ is called the *graininess* of $\mathbb{T}$. If $f : \mathbb{T} \to \mathbb{R}$, the generalized Δ-derivative is defined by

$$f^{\Delta}(t) = \lim_{s \to t,\ \sigma(s) \neq t} \frac{f(\sigma(s)) - f(t)}{\sigma(s) - t}$$

and $f^{\Delta}(t) = f'(t)$ if $\mathbb{T} = \mathbb{R}$ and $f^{\Delta}(t) = \Delta f(t) = f(t+1) - f(t)$.

Now consider the linear dynamic equation on a time scale $\mathbb{T}$

$$\big(r(t)x^{\Delta}\big)^{\Delta} + c(t)x^{\sigma} = 0, \tag{16.18}$$

where $x^{\sigma} = x \circ \sigma$ and $r, c : \mathbb{T} \to \mathbb{R}$, and its half-linear extension

$$\big(r(t)\Phi\big(x^{\Delta}\big)\big)^{\Delta} + c(t)\Phi\big(x^{\sigma}\big) = 0. \tag{16.19}$$

Obviously, (16.19) reduces to (0.1) if $\mathbb{T}=\mathbb{R}$ and to (16.3) if $\mathbb{T}=\mathbb{Z}$, respectively. The basic facts of the time scale calculus can be found in [109] and the general theory of dynamic equations on time scales is presented in [29,30].

Concerning (16.18), oscillation theory of this dynamic equation is established in [97]. The half-linear extension of this theory can be found in the recent paper [200], where it is shown that under the assumption $r(t)\neq 0$, the times scale Riccati equation

$$w^{\Delta}+c(t)+S[w,r](t)=0,$$

where

$$S[w,r](t)=\lim_{\lambda\to\mu(t)}\frac{w(t)}{\lambda}\left(1-\frac{r(t)}{\Phi(\Phi^{-1}(r(t)))+\lambda\Phi^{-1}(w(t))}\right),$$

and the p-degree functional (involving the time scale integral)

$$\mathcal{F}(y;a,b)=\int_a^b\left[r(t)|y^{\Delta}|^p-c(t)|y^{\sigma}|^p\right]\Delta t$$

play the same role as their continuous and discrete counterparts. We refer to the above mentioned paper [200] for details.

References

[1] R.P. Agarwal, *Difference Equations and Inequalities, Theory, Methods and Applications*, 2nd ed., Dekker, New York–Basel (2000).

[2] R.P. Agarwal, S.R. Grace, *Oscillation criteria for second order half-linear differential equations with deviating arguments*, Dyn. Contin. Discrete Impuls. Syst., Ser. A, Math. Anal. **9** (2002), 217–224.

[3] R.P. Agarwal, S.R. Grace, D. O'Regan, *Oscillation Theory of Second Order Linear, Half-Linear, Superlinear and Sublinear Dynamic Equations*, Kluwer Academic, Dordrecht/Boston/London (2002).

[4] R.P. Agarwal, S.L. Shieh, C.C. Yeh, *Oscillation criteria for second order retarded differential equations*, Math. Comput. Modelling **26** (1997), 1–11.

[5] C.D. Ahlbrandt, D.B. Hinton, R.T. Lewis, *The effect of variable change on oscillation and disconjugacy criteria with applications to spectral theory and asymptotic theory*, J. Math. Anal. Appl. **81** (1981), 234–277.

[6] C.D. Ahlbrandt, A.C. Peterson, *Discrete Hamiltonian Systems: Difference Equations, Continued Fractions, and Riccati Equations*, Kluwer Academic, Boston (1996).

[7] W. Allegretto, *Sturmian type theorems for solutions of elliptic nonlinear problems*, Nonlinear Differential Equations Appl. **7** (2000), 309–321.

[8] W. Allegretto, *Sturmian theorems for degenerate elliptic equations*, Proc. Amer. Math. Soc. **129** (2001), 3031–3035.

[9] W. Allegretto, Y.X. Huang, *A Picone's identity for the p-Laplacian and applications*, Nonlinear Anal. **32** (1998), 819–830.

[10] W. Allegretto, Y.X. Huang, *Principal eigenvalues and Sturm comparison via Picone's identity*, J. Differential Equations **156** (1999), 427–438.

[11] A. Anane, *Simplicite et isolation de la premiére valuer propre du p-laplacien avec poids*, C. R. Acad. Sci. Paris, Ser. I **305** (1987), 725–728.

[12] A. Anane, O. Charkone, J.-P. Gossez, *Spectre d'ordre supérieur et probléms de non-résonance*, C. R. Acad. Sci. Paris, Ser. I **325** (1997), 33–36.

[13] A. Anane, N. Tsouli, *On the second eigenvalue of the p-Laplacian*, Nonlinear Notes in Mathematics, A. Benkirane, J.-P. Gosez, eds., Pitman Research Notes in Mathematics, Vol. 343, Longman, Harlow (1996), 1–9.

[14] J. Andres, *On the criterion of de la Vallée Poussin*, Publ. Math. (Debrecen) **45** (1994), 145–152.

[15] F.V. Atkinson, Á. Elbert, *An extension of Milloux theorem to half-linear differential equations*, EJQTDE, Proc. 6th Coll. QTDE **2000** (8) (2000), 1–10.

[16] B. Aynlar, A. Tiryaki, *Oscillation theorems for nonlinear second order differential equations*, Comput. Math. Appl. **44** (2002), 529–538.

[17] R. Bellman, *Stability Theory of Differential Equations*, McGraw–Hill, New York–Toronto–London (1969).

[18] Ch. Bennewitz, Y. Saitö, *An embedding norm and the Lindqvist trigonometric functions*, Electron J. Differential Equations **2002** (86) (2002), 1–6.

[19] A. Besicovitch, *Almost Periodic Functions*, Dover, New York (1954).

[20] I. Bihari, *Ausdehung der Sturmschen Oszillations and Vergleichungsärte auf die Lösungen gewisser nichtlinearen Differenzialgleichungen zweiter Ordnung*, Publ. Math. Inst. Hungar. Acad. Sci. **2** (1957), 159–173.

[21] I. Bihari, *An oscillation theorem concerning the half-linear differential equation of the second order*. Publ. Math. Inst. Hungar. Acad. Sci. Ser. A **8** (1963), 275–279.

[22] I. Bihari, *On the second order half-linear differential equation*, Studia Sci. Math. Hungar. **3** (1968), 411–437.

[23] I. Bihari, *Asymptotic results concerning the equation* $x''|x'|^{n-1} + a(t)x^{\overset{*}{n}} = 0$, Studia Sci. Math. Hungar. **19** (1984), 151–157.

[24] P.A. Binding, P. Drábek, Y.X. Huang, *The range of the p-Laplacian*, Appl. Math. Lett. **10** (1997), 77–82.

[25] G. Bognár, *On the solution of some nonlinear boundary value problem*, Publ. Univ. Miskolc, Ser. D, Nat. Sci. Math. **36** (1995), 13–22.

[26] G. Bognár, *Existence theorems for eigenvalues of a nonlinear eigenvalue problem*, Commun. Appl. Nonlinear Anal. **4** (1997), 93–102.

[27] G. Bognár, *On the roots of the solution of generalized Bessel equation*, Proceedings of CDDE 2002, M. Bartušek, O. Došlý, eds., Masaryk University Press, Brno (2003), 33–43.

[28] G. Bognár, O. Došlý, *The application of Picone's identity for some nonlinear elliptic differential equations*, Acta Math. Univ. Comenianae **72** (2003), 45–57.

[29] M. Bohner, A. Peterson, *Dynamic Equations on Time Scales: In Introduction with Applications*, Birkhäuser, Boston (2001).

[30] M. Bohner, A. Peterson, eds., *Advanced in Dynamic Equations on Time Scales*, Birkhäuser, Boston (2003).

[31] H. Bohr, *Almost-Periodic Functions*, Chelsa, New York (1951).

[32] O. Borůvka, *Lineare Differentialtransformationen 2. Ordnung*, Deutscher Verlag der Wissenschaften, Berlin (1961).

[33] M. Cecchi, Z. Došlá, M. Marini, *On the dynamics of the generalized Emden–Fowler equations*, Georgian Math. J. **7** (2000), 269–282.

[34] M. Cecchi, Z. Došlá, M. Marini, *On decaying solutions for functional differential equations with p-Laplacian*, Nonlinear Anal. **47** (2001), 4387–4398.

[35] M. Cecchi, Z. Došlá, M. Marini, *On the nonoscillatory solutions of differential equations with p-Laplacian*, Adv. Math. Sci. Appl. **11** (2001), 419–436.

[36] M. Cecchi, Z. Došlá, M. Marini, *Principal solutions and minimal sets of quasilinear differential equations*, Dynam. Systems Appl., to appear.

[37] M. Cecchi, Z. Došlá, M. Marini, *Half-linear equations and characteristic properties of principal solutions*, (2003), submitted.

[38] M. Cecchi, M. Marini, G. Villari, *On the monotonicity property for a certain class of second order differential equations*, J. Differential Equations **82** (1) (1989), 15–27.

[39] M. Cecchi, M. Marini, G. Villari, *Integral criteria for a classification of solutions of linear differential equations*, J. Differential Equations **9** (2) (1992), 381–397.

[40] M. Cecchi, M. Marini, G. Villari, *On a nonlinear second order differential equation with oscillating coefficient*, Differential Integral Equations **6** (1) (1993), 47–61.

[41] M. Cecchi, M. Marini, G. Villari, *On some classes of continuable solutions of a nonlinear differential equation*, J. Differential Equations **118** (2) (1995), 403–419.

[42] T. Chantladze, N. Kandelaki, A. Lomtatidze, *On zeros of solutions of a second order singular half-linear equation*. Mem. Differential Equations Math. Phys. **17** (1999), 127–154.

[43] T.A. Chanturia, *On singular solutions of nonlinear systems of ordinary differential equations*, Colloq. Math. Soc. J. Bolyai **15** (1975), 107–119.

[44] T.A. Chanturia *On monotonic solutions of systems of nonlinear differential equations*, Ann. Polon. Math. **37** (1980), 59–70 (in Russian).

[45] Sh. Chen, Q. Huang, L.H. Erbe, *Bounded and zero-convergent solutions of a class of Stieltjes integro-differential equations*, Proc. Amer. Math. Soc. **113** (1991), 999–1008.

[46] W.J. Coles, *An oscillation criterion for second order differential equations*, Proc. Amer. Math. Soc. **19** (1969), 775–779.

[47] M. Cuesta, D. de Figueiredo, J.-P. Gossez, *The beginning of the Fučík spectrum for the p-Laplacian*, J. Differential Equations **159** (1999), 212–238.

[48] M. Cuesta, D. de Figueiredo, J.-P. Gossez, *A nodal domain property for the p-Laplacian*, C. R. Acad. Sci. Paris, Sér. I **330** (2000), 669–673.

[49] M. Del Pino, P. Drábek, R. Manásevich, *The Fredholm alternative at the first eigenvalue for the one dimensional p-Laplacian*, J. Differential Equations **151** (1999), 386–419.

[50] M. Del Pino, R. Manásevich, *Oscillation and nonoscillation for $(|u'|^{p-2}u')' + a(t)|u|^{p-2}u = 0,\ p > 1$*, Houston J. Math. **14** (1988), 173–177.

[51] M. Del Pino, M. Elgueta, R. Manásevich, *Generalizing Hartman's oscillation result for $(|x'|^{p-2}x')' + c(t)|x|^{p-2}x = 0,\ p > 1$*, Houston J. Math. **17** (1991), 63–70.

[52] M. Del Pino, M. Elgueta, R. Manásevich, *A homotopic deformation along p of a Leray–Schauder degree result and existence for $(|u'|^{p-2}u')' + f(t,u) = 0,\ u(0) = 0 = u(T),\ p > 1$*, J. Differential Equations **80** (1989), 1–13.

[53] M. Del Pino, R. Manásevich, *Multiple solutions for the p-Laplacian under global nonresonance*, Proc. Amer. Math. Soc. **112** (1991), 131–138.

[54] J.I. Díaz, *Nonlinear Partial Differential Equations and Free Boundaries*, Elliptic Equations, Vol. I, Pitman, London (1985).

[55] O. Došlý, *The multiplicity criteria for zero points of second order differential equations*, Math. Slovaca **42** (1992), 183–192.

[56] O. Došlý, *Conjugacy criteria for second order differential equations*, Rocky Mountain J. Math. **23** (1993), 849–861.

[57] O. Došlý, *Oscillation criteria for higher order Sturm–Liouville difference equations*, J. Difference Equations Appl. **4** (1998), 425–450.

[58] O. Došlý, *Oscillation criteria for half-linear second order differential equations*, Hiroshima Math. J. **28** (1998), 507–521.

[59] O. Došlý, *A remark on conjugacy of half-linear second order differential equations*, Math. Slovaca **50** (2000), 67–79.

[60] O. Došlý, *Methods of oscillation theory of half-linear second order differential equations*, Czechoslovak Math. J. **50** (125) (2000), 657–671.

[61] O. Došlý, *Half-linear oscillation theory*, Stud. Univ. Žilina Math.-Phys. Ser. **13** (2001), 65–73.

[62] O. Došlý, *The Picone identity for a class of partial differential equations*, Math. Bohem. **127** (2002), 581–589.

[63] O. Došlý, *Qualitative theory of half-linear second order differential equations*, Math. Bohem. **127** (2002), 181–195.

[64] O. Došlý, *Half-linear second order differential equations: oscillation theory and principal solution*, Stud. Univ. Žilina Math.-Phys. Ser. **17** (2003), 69-78.

[65] O. Došlý, Á. Elbert, *Conjugacy of half-linear second-order differential equations*, Proc. Roy. Soc. Edinburgh A **130** (2000), 517–525.

[66] O. Došlý, Á. Elbert, *Integral characterization of principal solution of half-linear differential equations*, Studia Sci. Math. Hungar. **36** (2000), 455–469.

[67] O. Došlý, J. Jaroš, *Singular version of Leighton's comparison theorem for forced quasilinear second order differential equations*, Arch. Math. (Brno) **39** (2003), 335–349.

[68] O. Došlý, A. Lomtatidze, *Disconjugacy and disfocality criteria for singular half-linear second order differential equations*, Ann. Polon. Math. **72** (1999), 273–284.

[69] O. Došlý, A. Lomtatidze, *Oscillation and nonoscillation criteria for half-linear second order differential equations*, submitted.

[70] O. Došlý, R. Mařík, *Nonexistence of positive solutions of PDE's with p-Laplacian*, Acta Math. Hungar. **90** (2001), 89–107,

[71] O. Došlý, P. Řehák, *Nonoscillation criteria for half-linear second-order difference equations*, Comput. Math. Appl. **42** (2001), 453–464.

[72] O. Došlý, J. Řezníčková, *Regular half-linear second order differential equations*, Arch. Math. **39** (2003), 233–245.

[73] P. Drábek, *Fredholm alternative for the p-Laplacian: yes or no?*, Function Spaces, Differential Operators and Nonlinear Analysis, Proceedings of the Conference FSDONA-99 (Syöte, Finland, June 10–16, 1999), Mustonen, Vesa, et al., eds., Mathematical Institute of the Academy of Sciences of the Czech Republic, Prague (2000), 57–64.

[74] P. Drábek, *Solvability and Bifurcations of Nonlinear Equations*, Research Notes in Mathematics, Vol. 264, Longman, Harlow (1992).

[75] P. Drábek, *Geometry of energy functional and the Fredholm alternative for the p-Laplacian*, Electron. J. Differential Equations **2002** (08) (2002) 103–120.

[76] P. Drábek, P. Girg, R. Manásevich, *Generic Fredholm results for the one-dimensional p-Laplacian*, Nonlinear Differential Equations Appl. **8** (2001), 285–298.

[77] P. Drábek, P. Girg, P. Takáč, M. Ulm, *The Fredholm alternative for p-Laplacian: bifurcation from infinity, existence and multiplicity*, submitted.

[78] P. Drábek, R. Manásevich, *On the closed solution to some nonhomogeneous eigenvalue problems with p-Laplacian*, Differential Integral Equations **12** (1999), 773–788.

[79] P. Drábek, M. Otani, *Global bifurcation result for the p-biharmonic operator*, Electron. J. Differential Equations **2001** (48) (2001), 1–19.

[80] P. Drábek, B. Robinson, *Resonance problems for the one-dimensional p-Laplacian*, Proc. Amer. Math. Soc. **128** (1999), 755–765.

[81] P. Drábek, B. Robinson, *On the generalization of the Courant nodal domain theorem*, J. Differential Equations **181** (2002), 58–71.

[82] P. Drábek, P. Takáč, *A counterexample to the Fredholm alternative for the p-Laplacian*, Proc. Amer. Math. Soc. **127** (1999), 1079–1087.

[83] D.R. Dunninger, *A Sturm comparison theorem for some degenerate quasilinear elliptic operators*, Boll. UMIA **9** (1995), 117–121.

[84] W. Eberhard, Á. Elbert, *On the eigenvalues of half-linear boundary value problems*, Math. Nachr. **183** (1997), 55–72.

[85] Á. Elbert, *A half-linear second order differential equation*, Colloq. Math. Soc. János Bolyai **30** (1979), 158–180.

[86] Á. Elbert, *The Wronskian and the half-linear differential equations*, Studia Sci. Math. Hungar. **15** (1980), 101–105.

[87] Á. Elbert, *Oscillation and nonoscillation theorems for some non-linear ordinary differential equations*, Lecture Notes in Math., Vol. 964 (1982), 187–212.

[88] Á. Elbert, *Asymptotic behaviour of autonomous half-linear differential systems on the plane*, Studia Sci. Math. Hungar. **19** (1984), 447–464.

[89] Á. Elbert, *On the half-linear second order differential equations*, Acta Math. Hungar. **49** (1987), 487–508.

[90] Á. Elbert, T. Kusano, *Principal solutions of nonoscillatory half-linear differential equations*, Adv. Math. Sci. Appl. **18** (1998), 745–759.

[91] Á. Elbert, T. Kusano, M. Naito, *On the number of zeros of nonoscillatory solutions to second order half-linear differential equations*, Ann. Univ. Sci. Budapest **42** (1999), 101–131.

[92] Á. Elbert, T. Kusano, M. Naito, *Counting the number of zeros of nonoscillatory solutions of second order half-linear differential equations*, Proceedings of the International Scientific Conference of Mathematics (Žilina, Slovakia, June 30–July 3, 1998), P. Marušiak, et al., eds., Žilina University Publisher (1999), 133–140.

[93] Á. Elbert, T. Kusano, T. Tanigawa, *An oscillatory half-linear differential equation*, Arch. Math. **33** (1997), 355–361.

[94] Á. Elbert, A. Schneider, *Perturbations of the half-linear Euler differential equation*, Results Math. **37** (2000), 56–83.

[95] R. Emden, *Gaskugeln, Anwendungen der mechanischen Warmentheorie auf Kosmologie und metheorologische Probleme*, Leipzig (1907).

[96] L. Erbe, *Oscillation criteria for second order nonlinear delay equations*, Canad. Math. Bull. **16** (1973), 49–56.

[97] L. Erbe, S. Hilger, *Sturmian theory on measure chains*, Differential Equations Dynam. Systems **1** (1993), 223-244.

[98] V.M. Evtuchov, *On conditions for nonoscillation of a nonlinear second order differential equation*, Mat. Zametki **67** (2000), 201–210.

[99] X. Fan, W.-T. Li, C. Zhong, *A Classification scheme for positive solutions of second order iterative differential equations*, Electron. J. Differential Equations **2002** (25) (2000), 1–14.

[100] J. Fleckinger, J.-P. Gossez, P. Takáč, *Existence, nonnexistence et principe de l'antimaximum pour le p-Laplacian*, C. R. Acad. Sci. Paris, Sér. I **321**, 731–734.

[101] J. Fleckinger, J. Hernández, P. Takáč, F. de Thélin, *Uniqueness and positivity for solutions of equations with p-Laplacian*, Proc. Conference on Reaction-Diffusion Equations (Trieste, Italy, October 1995), Marcel Dekker, New York–Basel (1997).

[102] R.H. Fowler, *The solutions of Emden's and similar differential equations*, Month. Not. Roy. Astron. Soc. **91** (1930), 63–91.

[103] S. Fučík, J. Nečas, J. Souček, V. Souček, *Spectral Analysis of Nonlinear Operators*, Lecture Notes in Math., Vol. 346, Springer-Verlag, New York–Berlin–Heidelberg (1973).

[104] M.R. Grossinho, P. Omari, *A multiplicity result for a class of quasilinear elliptic and parabolic problems*, Electron. J. Differential Equations **1997** (08) (1997), 1–16.

[105] G.H. Hardy, J.E. Littlewood, G. Pólya, *Inequalities*, Cambridge (1934).

[106] P. Hartman, *Ordinary Differential Equations*, Wiley, New York–London–Sydney (1964).

[107] L. Hatvani, *On stability properties of solutions of second order differential equations*, Proc. Colloq. Qual. Theory Differ. Equ. Szeged, Electron. J. Qual. Theory Differ. Equ. **2000**, Supl. 6 p.

[108] L. Hatvani, *On the existence of small solutions to linear second order differential equations with step function coefficients*, Dynam. Contin., Discrete Impuls. Systems **4** (1998), 321–330.

[109] S. Hilger, *Analysis on measure chains—a unified approach to continuous and discrete calculus*, Results Math. **18** (1990), 18–56.

[110] E. Hille, *Lectures on Ordinary Differential Eqautions*, Addison–Wesley, Reading (1969).

[111] H.L. Hong, W.Ch. Lian, C.C. Yeh, *The oscillation of half-linear differential equations with oscillatory coefficient*, Math. Comput. Modelling **24** (1996), 77–86.

[112] H. Hoshino, R. Imabayashi, T. Kusano, T. Tanigawa, *On second order half-linear oscillations*, Adv. Math. Sci. Appl. **8** (1998), 199–216.

[113] Ch. Huang, *Oscillation and nonoscillation for second order linear differential equations*, J. Math. Anal. Appl. **210** (1997), 712–713.

[114] Y.X. Huang, *Eigenvalues of the p-Laplacian with indefinite weight*, Comment. Math. Univ. Carolinae **36** (1995), 519–527.

[115] J. Jaroš, T. Kusano, *A Picone type identity for half-linear differential equations*, Acta Math. Univ. Comenianea **68** (1999), 137–151.

[116] J. Jaroš, T. Kusano, *Generalized Picone's formula and forced oscillations in quasilinear differential equations of the second order*, Arch. Math. (Brno) **38** (2002), 53–59.

[117] J. Jaroš, T. Kusano, T. Tanigawa, *Nonoscillation theory for second order half-linear differential equations in the framework of regular variation*, Results Math. **43** (2003), 129–149.

[118] J. Jaroš, T. Kusano, N. Yoshida, *A Picone-type identity and Sturmian comparison and oscillation theorems for a class of half-linear partial differential equations of the second order*, Nonlinear Anal. TMA **40** (2000), 381–395.

[119] J. Jianchu, *Oscillation and nonoscillation for the second order quasilinear differential equations*, Math. Sci. Res. Hot-Line **4** (2000), 39–47.

[120] W. Jingfa, *Oscillation and nonoscillation theorems for a class of second order quasilinear functional differential equations*, Hiroshima Math. J. **27** (1997), 449–466.

[121] W. Jingfa, *On second order quasilinear oscillations*, Funkcial. Ekvac. **41** (1998), 25–54.

[122] N. Kandelaki, A. Lomtatidze, D. Ugulava, *On the oscillation and nonoscillation of a second order half-linear differential equation*, Georgian Math. J. **7** (2000), 329–346.

[123] J. Karsai, *On the existence of a solution tending to zero of nonlinear differential equations (a nonlinear extension of a theorem of Milloux, Prodi, Trevisian and Hartman)*, Dynamic Systems Appl. **6** (1997), 429–440.

[124] W.G. Kelley, A. Peterson, *Difference Equations: An Introduction with Applications*, Academic Press, San Diego (1991).

[125] I. Kiguradze, *Asymptotic properties of solutions of a nonlinear differential equation of Emden–Fowler type*, Izvestiya Akad. Nauk SSSR. Ser. Mat. **29** (1965), 965–986 (in Russian).

[126] I. Kiguradze, *Some Singular Boundary Value Problems for Ordinary Differential Equations*, Tbilisi University Press, Tbilisi (1975) (in Russian).

[127] I.T. Kiguradze, T.A. Chanturia, *Asymptotic Properties of Solutions of Nonautonomous Ordinary Differential Equations*, Kluwer Academic, Dordrecht (1993).

[128] I. Kiguradze, G. Kvinikadze, *On strongly increasing solutions of nonlinear ordinary differential equations*, Ann. Mat. Pura Appl. **130** (1982), 67–87.

[129] J. Kiomura, T. Kusano, M. Naito, *Positive solutions of second order quasilinear differential equations with general nonlinearities*, Studia Sci. Math. Hungar. **35** (1999), 39–51.

[130] Q. Kong, *Oscillation criteria for second order half-linear differential equations*, Differential Equations with Applications in Biology. Proceedings of the International Conference (Halifax, Canada, July 25–29, 1997), Ruan, Shigui, et al., eds., Amer. Math. Soc., Providence, RI, Fields Inst. Commun. **21** (1999), 317–323.

[131] A. Kufner, L.-E. Persson, *Weighted Inequalities of Hardy-Type*, World Scientific, Singapore (2003).

[132] T. Kusano, M. Kitano, *On a class of second order quasilinear ordinary differential equations*, Hiroshima Math. J. **25** (1995), 321–355.

[133] T. Kusano, M. Naito, *Sturm–Liouville eigenvalue problems for half-linear ordinary differential equations*, Rocky Mountain J. Math. **31** (2001), 1039–1054.

[134] T. Kusano, M. Naito, *On the number of zeros of nonoscillatory solutions of half-linear ordinary differential equations involving a parameter*, Trans. Amer. Math. Soc. **324** (2002), 4751–4767.

[135] T. Kusano, Y. Naito, *Oscillation and nonoscillation criteria for second order quasilinear differential equations*, Acta Math. Hungar. **76** (1997), 81–99.

[136] T. Kusano, Y. Naito, A. Ogata, *Strong oscillation and nonoscillation of quasilinear differential equations of second order*, Differential Equations Dynamic Systems **2** (1994), 1–10.

[137] T. Kusano, M. Naito, T. Tanigawa, *Second order half-linear eigenvalue problem*, Fukuoka Univ. Sci. Rep. **27** (1999), 1–7.

[138] T. Kusano, A. Otaga, *Existence and asymptotic behavior of positive solutions of second order quasilinear differential equations*, Funkcial. Ekvac. **37** (1994), 345–361.

[139] T. Kusano, A. Otaga, *On oscillation of functional differential equations with forcing term*, Fukuoka Univ. Sci. Rep. **29** (1999), 225–229.

[140] T. Kusano, A. Otaga, H. Usami, *Oscillation theory for a class of second order quasilinear ordinary differential equations with application to partial differential equations*, Jpn. J. Math. **19** (1993), 131–147.

[141] T. Kusano, N. Yoshida, *Nonoscillation theorems for a class of quasilinear differential equations of second order*, J. Math. Anal. Appl. **189** (1995), 115–127.

[142] T. Kusano, J. Wang, *Oscillation properties of half-linear differential equations of the second order*, Hiroshima Math. J. **25** (1995), 371–385.

[143] G. Kvinikadze, *On strongly increasing solutions of system of nonlinear ordinary differential equations*, Trudy Inst. Prikl. Mat. I. Vekua **37** (1990), 41–49.

[144] J. Luo, L. Debnath, *Oscillation criteria for second order functional differential equations*, Comput. Math. Appl. **44** (2002), 731–739.

[145] E.M. Landesman, A.C. Lazer, *Nonlinear perturbations of linear elliptic boundary value problems at resonance*, J. Math. Mech. **19** (1970), 609–623.

[146] W. Leighton, *The detection of oscillation of solutions of a second order linear differential equations*, Duke J. Math. **17** (1950), 57–62.

[147] W. Leighton, *Comparison theorems for linear differential equations of second order*, Proc. Amer. Math. Soc. **13** (1962) 603–610.

[148] W. Leighton, M. Morse, *Singular quadratic functionals*, Trans. Amer. Math. Soc. **40** (1936), 252–286.

[149] H.J. Li, *Oscillation criteria for second order linear differential equations*, J. Math. Anal. Appl. **194** (1995), 217–234.

[150] H.J. Li, C.C. Yeh, An integral criterion for oscillation of nonlinear differential equations, Math. Japon. **41** (1995), 185–188.

[151] H.J. Li, C.C. Yeh, *An oscillation criterion of almost-periodic Sturm–Liouville equations*, Rocky Mountain J. Math. **25** (1995), 1417–1429.

[152] H.J. Li, C.C. Yeh, *Oscillations criteria for half-linear second order differential equations*, Hiroshima Math. J. **25** (1995), 571–583.

[153] H.J. Li, C.C. Yeh, *Oscillations of half-linear second order differential equations*, Hiroshima Math. J. **25** (1995), 585–594.

[154] H.J. Li, C.C. Yeh, *Sturmian comparison theorem for half-linear second order differential equations*, Proc. Roy. Soc. Edinburgh A **125** (1995), 1193–1204.

[155] H.J. Li, C.C. Yeh, *Nonoscillation of half-linear differential equations*, Publ. Math. (Debrecen) **49** (1996), 327–334.

[156] W.-T. Li, *Classification schemes for positive solutions of nonlinear differential systems*, Math. Comput. Modelling **36** (2002) 411–418.

[157] W.-T. Li, S.S. Cheng, *Limiting behavior of nonoscillatory solutions of a pair of coupled nonlinear differential equations*, Proc. Edinburgh Math. Soc. **43** (2000), 457–473.

[158] W.-T. Li, S.S. Cheng, *An oscillation criterion for nonhomogeneous half-linear differential equations*, Appl. Math. Lett. **15** (2002), 259–263.

[159] P. Lindqvist, *Note on a nonlinear eigenvalue problem*, Rocky Mountain J. Math. **23** (1993), 281–288.

[160] P. Lindqvist, *On the equation* $\operatorname{div}(|\nabla u|^{p-2}\nabla u) + \lambda |u|^{p-2}u = 0$, Proc. Amer. Math. Soc. **109** (1990), 157–164.

[161] P. Lindqvist, *Some remarkable sine and cosine functions*, Ricerche Math. **14** (1995), 269–280.

[162] A. Lomtatidze, *Oscillation and nonoscillation of Emden–Fowler type equation of second order*, Arch. Math. **32** (1996), 181–193.

[163] J.V. Manojlović, *Oscillation criteria for second-order half-linear differential equations*, Math. Comput. Modelling **30** (1999), 109–119.

[164] R. Manásevich, J. Mawhin, *The spectrum of p-Laplacian systems with various boundary conditions*, Adv. Differential Equations **5** (2000), 1289–1318.

[165] R. Manásevich, F.I. Njoku, F. Zanolin, *Positive solutions for the one-dimensional p-Laplacian*, Differential Integral Equations **8** (1995), 213–222.

[166] R. Manásevich, S. Sedziwy, *p-Laplacian and Lienard-type equation*, Rocky Mountain J. Math. **27** (1997), 611–617.

[167] R. Manásevich, P. Takáč, *On the Fredholm alternative for the p-Laplacian*, Proc. London Math. Soc., Ser. III **84** (2002), 324–342.

[168] R. Manásevich, F. Zanolin, *Time-mappings and multiplicity of solutions for one-dimensional p-Laplacian*, Nonlinear Anal. TMA **21** (1993), 269–291.

[169] R. Mařík, *Nonnegativity of functionals corresponding to the second order half-linear differential equation*, Arch. Math. (Brno) **35** (1999), 155–164.

[170] R. Mařík, *Focal points of half-linear differential equations*, Differential Equations Dynamic Systems **8** (2000), 111–124.

[171] R. Mařík, *A half-linear differential equation corresponding to variational problem*, Nonlinear Anal. A **45** (2001), 203–211.

[172] M. Marini, *Monotone solutions of a class of second order nonlinear differential equation*, Nonlinear Anal. TMA **8** (1984), 261–271.

[173] M. Marini, *On nonoscillatory solutions of a second order nonlinear differential equation*, Boll. UMI, Ser. VI **3C** (1) (1984), 189–202.

[174] M. Marini, P. Zezza, *On the asymptotic behavior of the solutions of second order linear differential equations*, J. Differential Equations **28** (1978), 1–17.

[175] H. Milloux, *Sur l'équation différentielle* $x'' + A(t)x = 0$, Prace Mat. **41** (1934), 39–53.

[176] J.D. Mirzov, *On some analogs of Sturm's and Kneser's theorems for nonlinear systems*, J. Math. Anal. Appl. **53** (1976), 418–425.

[177] J.D. Mirzov, *Principal and nonprincipal solutions of a nonoscillatory system*, Tbiliss. Gos. Univ. Inst. Prikl. Mat. Trudy **31** (1988), 100–117.

[178] J.D. Mirzov, *Asymptotic Properties of Nonlinear Systems of Nonautonomous Ordinary Differential Equations*, Majkop (1993) (in Russian).

[179] S. Matucci, *On asymptotic decaying solutions for a class of second order differential equations,* Arch. Math. (Brno) **35** (1999), 275–284.

[180] M. Mizukami, M. Naito, H. Usami, *Asymptotic behavior of solutions of a class of second order quasilinear ordinary differential equations*, Hiroshima Math. J. **32** (2002), 51–78.

[181] Y. Nagabuchi, M. Yamamoto, *On the monotonicity of solutions for a class of nonlinear differential equations of second order*, Math. Japonica **37** (1992), 665–673.

[182] M. Naito, Y. Naito, H. Usami, *Oscillation theory for semilinear elliptic equations with arbitrary nonlinearities*, Funkcial. Ekvac. **40** (1997), 41-55.

[183] F. Neuman, *Global Properties of Linear Differential Equations*, Kluwer Academic, Dordrecht (1991).

[184] J. Ohriska, *Oscillation of second order delay and ordinary differential equation*, Czech. Math. J. **34** (1984), 107–112.

[185] C. Olech, Z. Opial, T. Ważewski, *Sur le probleme d'oscillation des integrales de l'equation* $y'' + g(t)y = 0$, Bull. Acad. Polon. Sci. **5** (1957), 621–626.

[186] M. Otani, *A remark on certain nonlinear elliptic equations*, Proc. Fac. Sci. Tokai Univ. **19** (1984), 23–28.

[187] M. Otani, *On certain second order ordinary differential equations associated with Sobolev–Poincare-type inequalities*, Nonlinear Anal. Appl. **8** (1984), 1255–1270.

[188] S. Peña, *Conjugacy criteria for half-linear differential equations*, Arch. Math. (Brno) **35** (1999), 1–11.

[189] Ch. Philos, *Oscillation theorems for linear differential equations of second order*, Arch. Math. (Basel) **53** (1989), 493–492.

[190] M. Picone, *Sui valori eccezionalle di un parametro da cui dipende un' equatione differenzialle lineare del secondo ordine*, Ann. Scuola Norm. Sup. Pisa **11** (1910), 1–141.

[191] M. Piros, *On the solutions of a half-linear differential equation*, Studia Sci. Math. Hungar. **19** (1984), 193–211.

[192] M. Ráb, *Kriterien für oscillation der Lösungen der Differentialgleichung* $[p(x)y']' + q(x)y = 0$, Čas. Pěst. Mat. **84** (1957), 335–370.

[193] V.A. Rabtsevich, *On strongly increasing solutions of arbitrary order Emden–Fowler system*, Differential'nye Uravneniya **34** (1998), 222–227 (in Russian).

[194] P. Řehák, *Hartman–Wintner lemma, oscillation and conjugacy criteria for half-linear difference equations*, J. Math. Anal. Appl. **252** (2000), 813–827.

[195] P. Řehák, *Oscillatory properties of second order half-linear difference equations,* Czechoslovak. Math. J. **51** (126) (2001), 303–321.

[196] P. Řehák, *Generalized discrete Riccati equation and oscillation of half-linear difference equations*, Math. Comput. Modelling **34** (2001), 257–269.

[197] P. Řehák, *Oscillation and nonoscillation criteria for second order linear difference equations*, Fasc. Math. **31** (2001), 71–89.

[198] P. Řehák, *Oscillation criteria for second order half-linear difference equations*, J. Difference Equations Appl. **7** (2001), 483–505.

[199] P. Řehák, *Comparison theorems and strong oscillation in the half-linear discrete oscillation theory*, Rocky Mountain J. Math. **33** (2003), 333–352.

[200] P. Řehák, *Half-linear dynamic equations on time scales: IVP and oscillatory properties*, J. Nonlinear Funct. Anal. Appl. **7** (2002), 361–403.

[201] P. Řehák, *On certain comparison theorems for half-linear dynamic equations on time scales*, Abstracts Appl. Anal., to appear.

[202] W. Reichel, W. Walter, *Radial solutions of equations and inequalities involving the p-Laplacian*, J. Inequal. Appl. **1** (1997), 47–91.

[203] R.T. Rockafeller, *Convex Analysis*, Princeton (1970).

[204] Y.V. Rogovchenko, *Note on: "Oscillation criteria for second order linear differential equations"*, J. Math. Anal. Appl. **203** (1996), 560–563.

[205] S.P. Rogovchenko, Y.V. Rogovchenko, *Asymptotic behavior of certain second order differential equations*, Dynamic Systems Appl. **10** (2001), 185–200.

[206] G. Sansone, *Equazioni diferenziali nel campo reale I, II*, Zanichelli, Bologna (1949).

[207] M. Struwe, *Variational Methods; Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*, Springer-Verlag, New York (1990).

[208] C.A. Swanson, *Comparison and Oscillation Theory of Linear Differential Equations*, Academic Press, New York–London (1968).

[209] G. Talenti, *Best constant in Sobolev inequality*, Ann. Mat. Pura Appl. **110** (1976), 353–272.

[210] T. Tanigawa, *Existence and asymptotic behavior of positive solutions of second order quasilinear differential equations*, Adv. Math. Sci. Appl. **9** (1999), 907–938.

[211] L.H. Thomas, *The calculation of atomic fields*, Proc. Cambridge Philos. Soc. **23** (1927), 542–548.

[212] F.J. Tipler, *General relativity and conjugate differential equations*, J. Differential Equations **30** (1978), 165–174.

[213] N. Trudinger, *On Harnack type inequalities and their application to quasilinear elliptic equations*, Comm. Pure Appl. Math. **20** (1967), 721–747.

[214] Ch. de Valleé Poussin, *Sur l'équation différentielle linéqire du second order. Détermination d'une intégrale par deux valuers assignés. Extension aux équasions d'ordre n*, J. Math. Pures Appl. **8** (1929), 125–144.

[215] J.S. Wong, *Oscillation criteria for forced second order linear differential equation*, J. Math. Anal. Appl. **231** (1999), 235–240.

[216] Q.-R. Wang, *Oscillation and asymptotics for second-order half-linear differential equations*, Appl. Math. Comput. **122** (2001), 253–266.

[217] D. Willet, *On the oscillatory behavior of the solutions of second order differential equations*, Ann. Polon. Math. **21** (1969), 175–194.

[218] P.J.Y. Wong, R.P. Agarwal, *The oscillation and asymptotically monotone solutions of second-order quasilinear differential equations*, Appl. Math. Comput. **79** (1996), 207–237.

[219] P.J.Y. Wong, R.P. Agarwal, *Oscillation criteria for half-linear differential equations*, Adv. Math. Sci. Appl. **9** (1999), 649–663.

[220] J. Yan, *A note on an oscillation criterion for an equation with damping term*, Proc. Amer. Math. Soc. **90** (1984), 277–280.

[221] J. Yan, *Oscillation theorems for second order linear differential equations with damping*, Proc. Amer. Math. Soc. **98** (1986), 276–282.

[222] X. Yang, *Oscillation results for second order half-linear differential equations*, Math. Comput. Modelling **36** (2002), 503–507.

[223] X. Yang, *Positive solutions for the one-dimensional p-Laplacian*, Math. Comput. Modelling **35** (2002), 129–135.

[224] X. Yang, *On inequalities of Lyapunov type*, Appl. Math. Comput. **134** (2003), 293–300.

[225] X. Yang, *Nonoscillation criteria for second-order nonlinear differential equations*, Appl. Math. Comput. **131** (2002), 125–131.

[226] C.C. Yeh, *Levin's comparison theorems for second order nonlinear differential equations and inequalities*, Math. Jpn. **36** (1991), 703–710.

[227] G. Zhang, S.S. Cheng, *On connected half-linear differential equations*, Demonstratio Math. **32** (1999), 345–354.

This Page Intentionally Left Blank

CHAPTER 4

Radial Solutions of Quasilinear Elliptic Differential Equations

Jon Jacobsen

Department of Mathematics, Harvey Mudd College, Claremont, CA 91711, USA
E-mail: jacobsen@math.hmc.edu

Klaus Schmitt

Department of Mathematics, 155 South 1400 East, University of Utah, Salt Lake City, UT 84112-0090, USA
E-mail: schmitt@math.utah.edu

Contents

1. Introduction . . . 361
2. Boundary value problems on a ball . . . 363
    2.1. Introduction . . . 363
    2.2. The case when $f(0) = 0$ . . . 364
    2.3. The case when $f(0) < 0$ . . . 387
    2.4. The case when $f(0) > 0$ . . . 389
3. Problems on annular domains . . . 392
    3.1. Fixed point formulation . . . 393
    3.2. Existence results . . . 396
    3.3. Positone problems . . . 400
4. The Liouville–Gelfand equation — A case study . . . 405
    4.1. Introduction . . . 405
    4.2. Solution set for the semilinear case . . . 406
    4.3. Solution set for the quasilinear case . . . 409
5. Rellich–Pohozaev identities . . . 412
    5.1. More general equations . . . 414
    5.2. Critical dimensions . . . 418
6. Problems linear at infinity . . . 419
    6.1. Landesman–Lazer results . . . 421
    6.2. Nonlinear terms which oscillate . . . 422
7. Symmetry breaking . . . 424
8. Whole space problems . . . 427
References . . . 431

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 1
Edited by A. Cañada, P. Drábek and A. Fonda
© 2004 Elsevier B.V. All rights reserved

Abstract

This paper constitutes a short survey of the subject of radial solutions for quasilinear elliptic partial differential equations where the underlying domain is either a ball, an annular region, the exterior of a ball, or the whole space. In case the dependence of the equation on the independent variable is only in the radial direction, special solutions of such equations may be sought which depend only on the radial variable and as such are solutions of a boundary value problem for an associated nonlinear ordinary differential equation.

1. Introduction

In this paper we provide a survey of several results concerning radial solutions of quasilinear partial differential equations where the independent variable is a spatial variable varying over a domain with radial symmetry, such as a ball centered at the origin, an annular domain determined by concentric spheres centered at the origin, an exterior domain exterior to a ball, or the whole space. If the equation at hand also has the property that the dependence upon the independent variable is radial, then special radial solutions of the problem at hand may be sought and it is often the case that certain solutions having special properties, in fact, must be radial solutions.

The situation is well illustrated by the very classical problem of finding the radial eigenvalues and eigenfunctions of the Laplace operator subject to zero Dirichlet boundary conditions on the unit disk in the plane. Another illustration is the following classical Liouville–Gelfand problem which is concerned with the existence of positive solutions of the equation

$$\begin{cases} \Delta u + \lambda \mathrm{e}^u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{1.1}$$

where $\lambda > 0$ and Ω is a bounded domain in $\mathbb{R}^N$. If it is the case that $\Omega = \{x \in \mathbb{R}^N\colon |x| < 1\} := B_1(0)$, then it is reasonable to ask whether Equation (1.1) has solutions which only depend upon the radial variable. It follows from the maximum principle for elliptic equations that solutions of (1.1) can only assume positive values in the interior of the domain and then it follows by the classical result of Gidas, Ni, and Nirenberg [50] that all solutions of (1.1) are radially symmetric and (1.1) is equivalent to the ordinary differential equation's boundary value problem

$$\begin{cases} u'' + \frac{N-1}{r} u' + \lambda \mathrm{e}^u = 0, & r \in (0,1), \\ u'(0) = u(1) = 0, \end{cases} \tag{1.2}$$

for the profile $u(r) = u(|x|)$. Note that the originally discrete parameter N is now allowed to vary continuously. The results of [50] are valid for much more general situations and it follows that if $f : \mathbb{R} \to \mathbb{R}$ is a suitably smooth function (e.g., Lipschitz continuous), then any positive solution of

$$\begin{cases} \Delta u + f(u) = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{1.3}$$

with Ω a ball, must be radially symmetric about the center of the ball and similar results hold for the case that Ω is the whole space or a suitable exterior domain. It, on the other hand fails to hold for the case that Ω is an annular domain, in which case it often may happen that radial solutions undergo symmetry breaking bifurcations (some such results will be discussed in this paper).

If it is the case that

$$\Omega = \big\{x \in \mathbb{R}^N\colon\ 0 < a < |x| < b\big\},$$

then radial solutions of (1.3) are solutions of the boundary value problem

$$\begin{cases} u'' + \frac{N-1}{r}u' + f(u) = 0, & r \in (a,b), \\ u(a) = u(b) = 0. \end{cases} \tag{1.4}$$

For $N = 1$ these problems are amenable to reduction of order methods, and hence may be explicitly solved. For other values of N, this is, of course, no longer the case in general and other methods must be employed to study the solution structure of a given equation. We shall give a detailed account of problems related to (1.1) and related equations, a subject that dates back to Liouville in 1853 [71]. In 1914 Bratu [15] found an explicit solution to (1.2) when $N = 2$. Numerical progress for (1.2) when $N = 3$ was made by Frank-Kamenetskii (see [40]) in his study of thermal ignition problems. Further progress for $N = 3$ was made by Chandrasekhar [20, IV: §22–27], where (1.2) appears as a model for the temperature distribution of an isothermal gas sphere in gravitational equilibrium. Gelfand [49] built upon Frank-Kamenetskii's work when $N = 3$ and used Emden's transformation to prove the existence of a value of λ for which (1.2) has infinitely many nontrivial solutions.

In 1973 Joseph and Lundgren [61] completely characterized the solution structure of (1.2) for all N and hence, because of [50] also of the corresponding problem (1.1) in the case the domain is a ball. Other related examples arise from a larger class of partial differential operators, for example the work of Clément, de Figueiredo, and Mitidieri [22], Azorero and Alonso [44], Jacobsen [57], and Jacobsen and Schmitt [59], who consider existence and multiplicity results for the model equations

$$\begin{cases} \Delta_p u + \lambda \mathrm{e}^u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{1.5}$$

where $\Delta_p = \operatorname{div}(|\nabla u|^{p-2}\nabla u)$ is the p-Laplace operator [56,74] and

$$\begin{cases} S_k\big(D^2u\big) + \lambda \mathrm{e}^u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{1.6}$$

where $S_k(D^2u)$ is the k-Hessian operator [108], defined as the sum of all principal $k \times k$ minors of the Hessian matrix D^2u. For instance $S_1(D^2u) = \Delta u$ and $S_N(D^2u) = \det D^2u$, the Monge–Ampère operator.

Note that both equations are extensions of (1.1). In particular, the results of Joseph and Lundgren explain the radial case of (1.5) for $p = 2$ and of (1.6) when $k = 1$. In [22], the authors consider (among other topics) the radial case of both (1.5) for $p = N$ and (1.6) for $N = 2k$.

All of the above problems are simply special cases of the more general family of problems

$$\begin{cases} r^{-\gamma}\big(r^\alpha |u'|^\beta u'\big)' + f(\lambda, u) = 0, & r \in (0,1), \\ u > 0, & r \in (0,1), \\ u'(0) = u(1) = 0, \end{cases} \tag{1.7}$$

or

$$\begin{cases} r^{-\gamma}\left(r^{\alpha}|u'|^{\beta}u'\right)' + f(\lambda, u) = 0, & r \in (a,b), \\ u > 0, & r \in (a,b), \\ u(a) = u(b) = 0, \end{cases} \tag{1.8}$$

where certain inequalities are to be imposed on the parameters involved in the equation. Here $'$ denotes differentiation with respect to r. For instance, if $\Omega = B_1(0)$ is the unit ball, then Equation (1.7) with $f(\lambda, u) = \lambda e^u$ arises from (1.5) and (1.6) as a consequence of a priori symmetry results (see [35] for (1.6) and [7] for (1.5)). Similar problems may be posed also for exterior domain and whole space problems.

Much work has also been devoted to boundary value problems and other qualitative studies for more general differential operators of the form

$$r^{-\gamma}\left(r^{\alpha}\phi(u')\right)' + f(\lambda, u) = 0, \tag{1.9}$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is an increasing homeomorphism of $\mathbb{R}$, with $\phi(0) = 0$. Such problems arise in a very natural way in diffusion problems where diffusion is governed by rapidly growing terms. We shall survey some such problems below.

In most of the discussion to follow the parameter α is taken to equal the parameter γ and is denoted by $N - 1$, to indicate the partial differential equation origin of the problem, where N denotes the dimension of the underlying domain. In the discussion, however, $N - 1$ may simply denote a nonnegative parameter. The equations stated above also may depend on other parameters, denoted by λ, which dependence may be in a linear or nonlinear fashion, thus this parameter may occur as a multiplicative factor or simply as a variable in the function evaluation. Should this parameter play no role in the result at hand, we simply shall suppress the dependence.

The paper is organized as follows: We first discuss boundary value problems on a ball related to the differential operator (1.9) and rely mainly on the recent work in [45–47,54, 53]. We then proceed to discuss problems on annular domains based on some work in [9,10, 25,30,29,74]. Next, we present a detailed discussion of Gelfand type problems. Following the Gelfand case study we return to general theory and present a range of related topics including some classical oscillation and nonoscillation theorems, problems for which radial solutions can undergo symmetry breaking bifurcations (relying on work in [67,68,80–82]), and problems concerning radial ground states of problems defined in all of space.

We shall denote by $\|\cdot\|$ the supremum norm in $C[a,b]$ for any interval $[a,b] \subset \mathbb{R}$.

2. Boundary value problems on a ball

2.1. *Introduction*

In this section, we consider the existence of positive solutions for the boundary value problems

$$\begin{cases} \left(r^{N-1}\phi(u')\right)' + \lambda r^{N-1} f(u) = 0, & 0 < r < R, \\ u'(0) = u(R) = 0, \end{cases} \tag{2.1}$$

where ϕ is an odd increasing homeomorphism on $\mathbb{R}$ and λ is a positive parameter (i.e., we consider the case that $\alpha = \gamma = N - 1$).

As pointed out before, Equation (2.1) arises in the study of radial solutions for quasilinear elliptic boundary value problems of the form

$$\begin{cases} \operatorname{div}\big(A(\nabla u)\big) + \lambda f(u) = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{2.2}$$

where Ω is a ball of radius R in $\mathbb{R}^N$.

We first discuss existence results for (2.1) when $f(0) = 0$. These results are of bifurcation type and are based on the work [45,46]. We then discuss the cases when $f(0) < 0$ or $f(0) > 0$, and f is ϕ superlinear at ∞. In the former case there exists a positive, decreasing solution to (2.1) for λ small. The asymptotic behavior of the solution (as $\lambda \to 0$) will also discussed. This generalizes the results in [19,2] (see also [100]) in which $\phi(x) = x$, and complements those in [31,74] where similar problems were considered on an annulus. In the latter case, there exists a positive number λ^* such that (2.1) has at least two positive solutions for $\lambda < \lambda^*$, at least one for $\lambda = \lambda^*$, and none for $\lambda > \lambda^*$. This result complements corresponding results in [3,29] on annular domains. We refer to [100] for the literature on problem (2.2) with $A(x) = x$ on bounded domains. The approach used here depends on degree theory and also uses results about upper and lower solutions. Such results can be found in, for example, Lloyd [72], Berger [13], Deimling [32], and Schmitt [99].

2.2. *The case when* $f(0) = 0$

The prototypical example in this case is the partial differential equation

$$\begin{cases} \Delta u + u^p = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{2.3}$$

where $p > 1$. Note that $u = 0$ is a solution to (2.3). One approach to finding nontrivial solutions is to consider it as a perturbation of an eigenvalue equation

$$\Delta u + \lambda u + u^p = 0, \tag{2.4}$$

and study solution continua, i.e., curves in the solution set

$$\mathcal{S} = \big\{(\lambda, u)\colon\ (\lambda, u) \text{ satisfies (2.4)}\big\}.$$

For instance, if $\mathcal{S}$ contains a point $(0, u)$ with $u \neq 0$, then we obtain a nontrivial solution to (2.3). Viewing (2.1) with this motivation, we now consider the equation

$$\begin{cases} -r^{1-N}\big(r^{N-1}\phi(u')\big)' = \lambda\psi(u) + g(\lambda, u), & 0 < r < R, \\ u'(0) = u(R) = 0, \end{cases} \tag{2.5}$$

where $\phi:\mathbb{R}\to\mathbb{R}$ and $\psi:\mathbb{R}\to\mathbb{R}$ are odd increasing homeomorphisms of $\mathbb{R}$, which both vanish at the origin. We further assume that $g(\lambda,0)=0$ for all $\lambda\in\mathbb{R}$. A solution to Equation (2.5) is a function $u\in C^1[0,R]$ with $\phi(u')\in C^1[0,R]$ which satisfies (2.5).

Furthermore we will require that ϕ, ψ satisfy the asymptotic homogeneity conditions:

$$\lim_{s\to 0}\frac{\phi(\sigma s)}{\psi(s)}=\sigma^{p-1},\quad \text{for all } \sigma\in\mathbb{R}_+,\ \text{for some } p>1, \tag{2.6}$$

and

$$\lim_{s\to\pm\infty}\frac{\phi(\sigma s)}{\psi(s)}=\sigma^{q-1},\quad \text{for all } \sigma\in\mathbb{R}_+,\ \text{for some } q>1. \tag{2.7}$$

We note that if the pair ϕ,ψ satisfies the asymptotic homogeneity conditions (2.6) and (2.7), then the function ϕ satisfies both of these conditions with ψ replaced by ϕ and also ψ satisfies both of these conditions with ϕ replaced by ψ. Such conditions appear in a variety of contexts (see, e.g., [95]).

Let $u(r)$ be a solution of (2.5). By integrating the equation in (2.5) we see that $u(r)$ satisfies the integral equation

$$u(r)=\int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\bigl(\lambda\psi\bigl(u(\xi)\bigr)+g\bigl(\lambda,u(\xi)\bigr)\bigr)\,\mathrm{d}\xi\right]\mathrm{d}s. \tag{2.8}$$

This equation will be fundamental for much of the subsequent analysis.

2.2.1. *Index calculations.* Let us consider (2.5) with $g=0$:

$$\begin{cases} -r^{1-N}\bigl(r^{N-1}\phi(u')\bigr)'=\lambda\psi(u), & 0<r<R,\\ u'(0)=u(R)=0. \end{cases} \tag{2.9}$$

A value λ such that (2.9) has a nontrivial solution is called an eigenvalue of (2.9). From (2.8) it follows that u is a solution if and only if u is a fixed point of the completely continuous operator $T^\lambda: C[0,R]\to C[0,R]$ defined by

$$T^\lambda(u)(r)=\int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\lambda\psi\bigl(u(\xi)\bigr)\,\mathrm{d}\xi\right]\mathrm{d}s. \tag{2.10}$$

If $\phi(t)=\psi(t)=|t|^{p-2}t$, then it is known that the eigenvalue problem (2.9) has a sequence of eigenvalues $\{\lambda_m=\lambda_m(p),\ m=1,2,\ldots\}$, with $\lambda_m(p)\to\infty$, as $\lambda\to\infty$, and this set has been completely described (see, e.g., [34,36] and the earlier papers [4,69]). We need the following lemma from [46]:

LEMMA 2.1. *Consider the problem* (2.9).

- *If* (2.6) *holds, then the Leray–Schauder degree of* $I - T^\lambda$ *is defined for* $B(0, \varepsilon)$, (*the open ball centered at zero with radius* ε *in* $C[0, R]$), *for all sufficiently small* ε. *Furthermore we have*

$$\deg_{\mathrm{LS}}\bigl(I - T^\lambda, B(0, \varepsilon), 0\bigr) = \begin{cases} 1, & \lambda < \lambda_1(p), \\ (-1)^m, & \lambda \in \bigl(\lambda_m(p), \lambda_{m+1}(p)\bigr). \end{cases} \tag{2.11}$$

- *If* (2.7) *holds, then the Leray–Schauder degree for* $I - T^\lambda$ *is defined for* $B(0, M)$, *for all sufficiently large* M, *and a similar formula to* (2.11) *holds, namely*

$$\deg_{\mathrm{LS}}\bigl(I - T^\lambda, B(0, M), 0\bigr) = \begin{cases} 1, & \lambda < \lambda_1(q), \\ (-1)^l, & \lambda \in \bigl(\lambda_l(q), \lambda_{l+1}(q)\bigr), \end{cases} \tag{2.12}$$

where $\{\lambda_l(q),\ l = 1, 2, \ldots\}$ *is the set of eigenvalues of* (2.9) *with* $\phi(t) = \psi(t) = |t|^{q-2}t$.

Lemma 2.1 may be applied to obtain an existence result for nontrivial solutions.

THEOREM 2.2. *Consider problem* (2.9) *and suppose that* ϕ, ψ *are odd increasing homeomorphisms of* $\mathbb{R}$ *with* $\phi(0) = 0 = \psi(0)$, *which satisfy* (2.6) *and* (2.7) *with* $p \neq q$. *Assume that for some* $j \in \mathbb{N}$, $\lambda_j(p) \neq \lambda_j(q)$ *and that* $\lambda \in (A, B)$, *where* $A = \min\{\lambda_j(p), \lambda_j(q)\}$ *and* $B = \max\{\lambda_j(p), \lambda_j(q)\}$. *Assume furthermore that* (A, B) *does not contain any other eigenvalue from* $\{\lambda_m(p)\}$ *or* $\{\lambda_m(q)\}$. *Then problem* (2.9) *has a nontrivial solution.*

PROOF. Assuming for example that $\lambda_j(q) > \lambda_j(p)$, it follows from Lemma 2.1 that

$$\deg_{\mathrm{LS}}\bigl(I - T^\lambda, B(0, \varepsilon), 0\bigr) = (-1)^j, \tag{2.13}$$

for $\varepsilon > 0$ small, and that

$$\deg_{\mathrm{LS}}\bigl(I - T^\lambda, B(0, M), 0\bigr) = (-1)^{j-1}, \tag{2.14}$$

for M large. Thus combining (2.13) and (2.14) with the excision property of the Leray–Schauder degree, we obtain that

$$\deg_{\mathrm{LS}}\bigl(I - T^\lambda, B(0, M) \setminus \overline{B(0, \varepsilon)}, 0\bigr) \neq 0, \tag{2.15}$$

yielding the existence of a nontrivial solution with $\varepsilon < \|u\| < M$. □

This theorem suggests the existence of a branch of solutions to (2.5) connecting $(\lambda_j(p), 0)$ with $(\lambda_j(q), \infty)$, for each $j \in \mathbb{N}$, generalizing the well-known property for the homogeneous case. We will return to the structure of the eigenvalue set in Section 2.2.4.

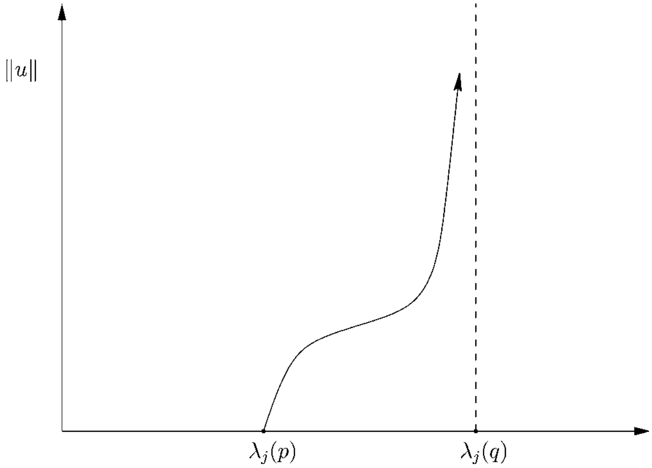


Fig. 1. Possible continua of Theorem 2.2.

2.2.2. *On initial value problems.* In this section we discuss some results for the initial value problem associated with (2.5), i.e.,

$$\begin{cases} -r^{1-N}\big(r^{N-1}\phi(u')\big)' = \lambda\psi(u) + g(\lambda, u), & 0 < r < R, \\ u(0) = d, \\ u'(0) = 0, \end{cases} \tag{2.16}$$

which will be needed in the next section. Throughout we shall assume that

$$ug(\lambda, u) \geqslant 0, \quad \text{for all } \lambda, u \in \mathbb{R}. \tag{2.17}$$

The following proposition follows from an application of the contraction mapping principle applied to (2.8):

PROPOSITION 2.3. *Suppose that $g(\lambda, u) = \mathrm{O}(|\psi(u)|)$ near zero, uniformly for r and λ in bounded intervals. Then the only solution to the problem*

$$\begin{cases} -r^{1-N}\big(r^{N-1}\phi(u')\big)' = \lambda\psi(u) + g(\lambda, u), & 0 < r < R, \\ u(r_0) = 0, \\ u'(r_0) = 0, \end{cases} \tag{2.18}$$

with $r_0 \in [0, R]$ is the trivial solution $u = 0$.

The next result concerns the oscillation of nontrivial solutions.

PROPOSITION 2.4. *Suppose that $g(\lambda, u) = \mathrm{O}(|\psi(u)|)$ near zero, uniformly for λ in bounded intervals, then nontrivial solutions of the initial value problem* (2.16) *are oscillatory, i.e., solutions are defined on all of* $[0, \infty)$ *and have infinitely many zeros.*

PROOF. Let u be a nontrivial solution of (2.16) and assume that $d > 0$. If u does not vanish in $(0, \infty)$, then u is decreasing on $(0, \infty)$. Integrating the equation from 0 to $s \in (0, r)$ we find

$$-u'(s) = \phi^{-1}\left(\frac{1}{s^{N-1}} \int_0^s \xi^{N-1}\big(\lambda\psi\big(u(\xi)\big) + g\big(\lambda, u(\xi)\big)\big)\,\mathrm{d}\xi\right). \tag{2.19}$$

Integrating (2.19) from $r/2$ to r yields

$$u(r/2) \geqslant \int_{r/2}^{r} \phi^{-1}\left(\frac{\lambda}{s^{N-1}} \int_0^s \xi^{N-1}\psi\big(u(\xi)\big)\,\mathrm{d}\xi\right)\mathrm{d}s, \tag{2.20}$$

$$\geqslant \int_{r/2}^{r} \phi^{-1}\left(\frac{\lambda}{r^{N-1}} \int_0^{r/2} \xi^{N-1}\psi\big(u(r/2)\big)\,\mathrm{d}\xi\right)\mathrm{d}s. \tag{2.21}$$

Hence

$$\frac{\phi(\frac{2}{r}u(r/2))}{\phi(u(r/2))} \geqslant \frac{\lambda}{N}\frac{r}{2^N}\frac{\psi(u(r/2))}{\phi(u(r/2))}, \tag{2.22}$$

for any $r > 0$. But for $r > 2$ the left-hand side of (2.22) is less than 1 while the right-hand side can be arbitrarily large. Hence we conclude that u must have a zero which, by Proposition 2.3, is simple. □

We remark here that by the same argument any nontrivial solution of the initial value problem

$$\begin{cases} -\big(r^{N-1}\phi(v')\big)' = r^{N-1}\lambda\psi(v) + r^{N-1}g(\lambda, v), & 0 < r < R, \\ v(r_0) = d, \\ v'(r_0) = 0, \end{cases} \tag{2.23}$$

must have a first isolated zero to the right of r_0 and such solutions may be continued to ∞ and are oscillatory on the whole real line.

For each $d \neq 0$, let ρ_d denote the first zero of a solution u of (2.5) such that $u(0) = d$. The next two propositions show that solutions to (2.5) have Sturm type separation properties as one has for linear equations (see, e.g., [55]):

PROPOSITION 2.5. *Suppose that* $g(\lambda, u) = \mathrm{O}(|\psi(u)|)$ *near zero, uniformly for* λ *in bounded intervals. Then, for each* $R > 0$ *there exists* $\Lambda(R)$ *such that for all* $\lambda > \Lambda(R)$ *and all* $d \neq 0$ *we have that* $\rho_d \leqslant R$.

PROPOSITION 2.6. *Let* ρ_{d_j} *denote the* j*th zero of a solution* u *of* (2.5) *such that* $u(0) = d$ *and suppose that* $g(\lambda, u) = \mathrm{O}(|\psi(u)|)$ *near zero, uniformly for* λ *in bounded intervals. Then, for all* $L > 0$ *and* $j \in \mathbb{N}$ *there exists* $\Lambda_j(L) > 0$ *such that for all* $\lambda > \Lambda_j(L)$ *and all* $d \neq 0$ *we have that* $\rho_{dj} \leqslant L$.

2.2.3. *Bifurcation of solutions.* In this section we consider bifurcation problems at zero and at infinity. See [21,32,94,64] for definitions.

Let u be a solution of problem (2.5). Then u satisfies

$$u = \mathcal{F}(\lambda, u) \tag{2.24}$$

where

$$\mathcal{F}(\lambda, u)(r) = \int_r^R \phi^{-1}\left(\frac{1}{s^{N-1}} \int_0^s \xi^{N-1}\big(\lambda\psi\big(u(\xi)\big) + g\big(\lambda, u(\xi)\big)\big)\, \mathrm{d}\xi\right) \mathrm{d}s. \tag{2.25}$$

It is clear that $\mathcal{F} : \mathbb{R} \times C[0, R] \to C[0, R]$ and it is a completely continuous operator. Concerning (2.24), the following theorem is proved by a standard argument which considers a sequence of solutions approaching a bifurcation point, appropriately renormalized to take advantage of the condition on g:

THEOREM 2.7. (i) *Suppose that* $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *is continuous and satisfies* $g(\lambda, s) = o(|\phi(s)|)$ *near* $s = 0$ *uniformly for* λ *in bounded intervals, and that* ϕ *and* ψ *satisfy* (2.6). *If* $(\bar{\lambda}, 0)$ *is a bifurcation point for* (2.24), *then* $\bar{\lambda} = \lambda_m(p)$, *for some* $m \in \mathbb{N}$.

(ii) *Suppose that* $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *is continuous and satisfies* $g(\lambda, s) = o(|\phi(s)|)$ *near infinity, uniformly for* λ *in bounded intervals, and that* ϕ *and* ψ *satisfy* (2.7). *If* $(\bar{\lambda}, \infty)$ *is a bifurcation point for* (2.24), *then* $\bar{\lambda} = \lambda_m(q)$, *for some* $m \in \mathbb{N}$.

The next theorem is the main result on bifurcation of solutions to (2.24):

THEOREM 2.8. *Suppose that* $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *is continuous and satisfies* $g(\lambda, s) = o(|\phi(s)|)$ *near* $s = 0$, *uniformly for* λ *in bounded intervals.*

(i) *If* ϕ *and* ψ *satisfy* (2.6), *then for each* $k \in \mathbb{N}$ *there is a connected component* $\mathcal{S}_k \subset \mathbb{R} \times C[0, 1]$ *of the set of nontrivial solutions of* (2.5) *whose closure* $\overline{\mathcal{S}_k}$ *contains* $(\lambda_k(p), 0)$. *Moreover,* $\mathcal{S}_k$ *is unbounded in* $\mathbb{R} \times C[0, 1]$ *and if* $(\lambda, u) \in \mathcal{S}_k$, *then* u *has exactly* $k - 1$ *simple zeros in* $(0, R)$.

(ii) *If* $ug(\lambda, u) \geqslant 0$, *then there exists* $M_k \in (0, \infty)$ *such that if* $(\lambda, u) \in \mathcal{S}_k$, *then* $\lambda \leqslant M_k$.

PROOF. The proof of the existence of the connected component $\mathcal{S}_k$ such that $(\lambda_k(p), 0)$ belongs to $\overline{\mathcal{S}_k}$ and that $\mathcal{S}_k$ is unbounded or contains another bifurcation point $(\lambda_j(p), 0)$, $j \neq k$, is entirely similar to that of [94]. The fact that $(\lambda_k(p), 0) \in \overline{\mathcal{S}_k}$ and that $(\lambda, u) \in \mathcal{S}_k$ implies u has exactly $k - 1$ simple zeros in $(0, R)$ is in turn similar to that of Theorem 4.1 of [34] and uses the results of Section 2.2.2. The existence of M_k such that $\lambda \leqslant M_k$ follows directly from Proposition 2.6. □

The following will serve as examples to illustrate the above results.

THEOREM 2.9. *Suppose that* ϕ, ψ *are odd increasing homeomorphisms of* $\mathbb{R}$ *with* $\phi(0) = \psi(0) = 0$, *which satisfy* (2.6) *and* (2.7) *with* $p \neq q$ *and* $ug(\lambda, u) \geqslant 0$. *Then for any* $j \in \mathbb{N}$, *there exists a connected component* S_j *of the set on nontrivial solutions of* (2.5) *connecting* $(\lambda_j(p), 0)$ *to* $(\lambda_j(q), \infty)$ *such that* $(\lambda, u) \in \mathcal{S}_j$ *implies* u *has exactly* $j - 1$ *simple zeros in* $(0, R)$.

PROOF. It follows from Theorem 2.8 that for any $j \in \mathbb{N}$, there exists an unbounded connected component S_j of the set on nontrivial solutions emanating from $(\lambda_j(p), 0)$ such that $(\lambda, u) \in \mathcal{S}_j$ implies u has exactly $j - 1$ simple zeros in $(0, R)$. Since there is an M_j such that $(\lambda, u) \in \mathcal{S}_k$ implies that $\lambda \leqslant M_k$ and since there are no nontrivial solutions of (2.5) for $\lambda = 0$, it follows that for any $M > 0$, there is $(\lambda, u) \in \mathcal{S}_k$ such that $\|u\| > M$. Hence (ii) of Theorem 2.7 implies that as $\|u\| \to \infty$ with $(\lambda, u) \in \mathcal{S}_k$, it must be that $\lambda \to \lambda_j(q)$. □

The second application is motivated by a result of [34].

THEOREM 2.10. *Consider the problem*

$$\begin{cases} -\left(r^{N-1}\phi(u')\right)' = r^{N-1}g(u), & 0 < r < R, \\ u'(0) = u(R) = 0. \end{cases} \tag{2.26}$$

Assume ϕ *is an increasing homeomorphisms of* $\mathbb{R}$ *with* $\phi(0) = 0$ *and which satisfies* (2.6) *and* (2.7) *with* $\phi = \psi$. *Further suppose that* $g : \mathbb{R} \to \mathbb{R}$ *is continuous with* $ug(u) \geqslant 0$ *and that there exist positive integers k and n, with* $k \leqslant n$, *such that*

$$\mu := \lim_{|s| \to 0} \frac{g(s)}{\phi(s)} < \lambda_k(p) \leqslant \lambda_n(q) < \nu := \lim_{|s| \to \infty} \frac{g(s)}{\phi(s)}.$$

Then for each integer $j \in (k, n)$ *Equation* (2.26) *has a solution with exactly* $j - 1$ *simple zeros in* $(0, R)$. *Thus* (2.26) *possesses at least* $n - k + 1$ *nontrivial solutions.*

2.2.4. *On principal eigenvalues.* We next present results concerning the existence of positive solutions to the problem

$$\begin{cases} -r^{1-N}\left(r^{N-1}\phi(u')\right)' = \lambda\psi(u), & 0 < r < R, \\ u'(0) = u(R) = 0, \end{cases} \tag{2.27}$$

where ϕ is an increasing homeomorphism of $\mathbb{R}$ and ψ is nondecreasing with $\phi(0) = \psi(0) = 0$. A constant λ such that (2.27) has a positive (or negative) solution is called a principal eigenvalue. Note that Theorem 2.2 established the existence of solution branches to (2.27) which may or may not correspond to positive solutions.

In the case that N is a positive integer and $\phi(u) = |u|^{p-2}u = \psi(u)$ the above problem is the problem of the existence of the principal eigenvalue of the p-Laplacian on a ball of radius R in $\mathbb{R}^N$, subject to zero Dirichlet boundary data. As such it is well understood (see, e.g., [4,14,41,43,103]).

The tools that have been used for establishing the existence of such (and higher) eigenvalues come from variational methods and are usually critical point theorems for smooth functionals defined in an appropriate Sobolev space; these methods consequently also yield theorems for the case the underlying ball domain is replaced by an arbitrary bounded domain.

Here we discuss the general case (2.27) and rely on fixed point and continuation techniques based on some work in [29,46]. For the special case $N = 1$, very detailed information is available in [34,48].

We will assume further that $N \geqslant 1$ (not necessarily an integer) and ϕ and ψ satisfy for all $x \neq 0$ and $\sigma > 0$,

$$A(\sigma) \leqslant \frac{\phi(\sigma|s|)}{\psi(|s|)} \leqslant B(\sigma), \tag{2.28}$$

where $A(\sigma)$ and $B(\sigma)$ are positive constants depending on σ only.

We introduce some notation to discuss the structure of the eigenvalue set. First, let E denote the set of all principal eigenvalues. Next, for each $d > 0$, let $\Gamma(d, R)$ be defined by

$$\Gamma(d, R) = \{\lambda > 0 \mid (2.27) \text{ has a positive solution with } u(0) = d\} \tag{2.29}$$

and set

$$\Gamma^-(d, R) = \inf \Gamma(d, R). \tag{2.30}$$

Further define

$$\Gamma_1^-(R) := \liminf_{d \to \infty} \Gamma^-(d, R), \tag{2.31}$$

$$\gamma_1^-(R) := \liminf_{d \to 0} \Gamma^-(d, R), \tag{2.32}$$

$$\lambda_1^-(R) := \inf_{d > 0} \Gamma^-(d, R). \tag{2.33}$$

The main result in this section is as follows:

THEOREM 2.11. *The set $E \neq \emptyset$ and there exists a smallest $\lambda_0 > 0$ such that for $\lambda < \lambda_0$ the eigenvalue problem* (2.27) *has no nontrivial solutions. For every $d > 0$, there exists $\lambda \in E$ and a positive solution u of* (2.27) *such that $u(0) = d$. Furthermore, $\Gamma^-(d, R) > 0$ and $\Gamma_1^-(R), \gamma_1^-(R), \lambda_1^-(R)$ are all nonincreasing functions of R, and*

$$\lim_{R \to 0^+} \Gamma_1^-(R) = \lim_{R \to 0^+} \gamma_1^-(R) = \lim_{R \to 0^+} \lambda_1^-(R) = \infty,$$

$$\lim_{R \to \infty} \Gamma_1^-(R) = \lim_{R \to \infty} \gamma_1^-(R) = \lim_{R \to \infty} \lambda_1^-(R) = 0.$$

There is a result dual to Theorem 2.11 for principal eigenvalues with associated negative solutions, the set of such eigenvalues may, of course, be different from the set whose existence is asserted in Theorem 2.11.

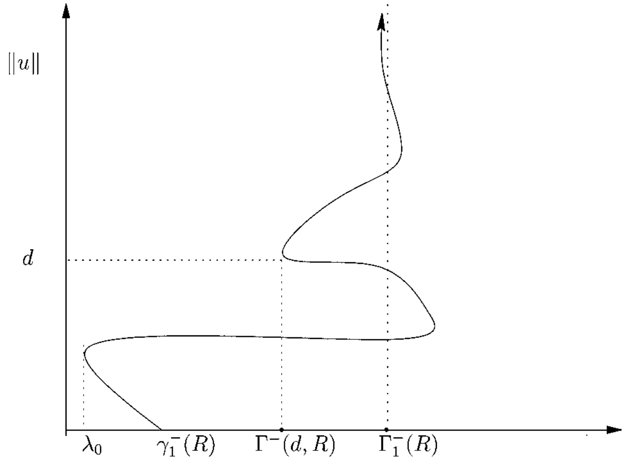


Fig. 2. Possible continua of Theorem 2.11.

We break the proof of Theorem 2.11 into several steps. First, a quick calculation shows that finding positive solutions to problem (2.27) is equivalent to finding nontrivial solutions to the problem

$$\begin{cases} \left(r^{N-1}\phi(u')\right)' + \lambda r^{N-1}\psi\left(|u|\right) = 0, & r \in (0,R), \\ u'(0) = u(R) = 0. \end{cases} \tag{2.34}$$

Let C_R denote the closed subspace of $C[0,R]$ defined by

$$C_R = \left\{u \in C[0,R] \mid u(R) = 0\right\}. \tag{2.35}$$

Then C_R is a Banach space with the inherited norm $\|\cdot\|$ from $C[0,R]$.

Similar to (2.8), we see that a solution u to (2.34) is equivalent to a fixed point of the operator

$$\mathcal{F}(\lambda,u)(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\lambda\psi\left(\left|u(\xi)\right|\right)\mathrm{d}\xi\right]\mathrm{d}s. \tag{2.36}$$

Clearly $\mathcal{F}:[0,\infty)\times C_R \to C_R$ is a well-defined operator. Define now the operator $\mathcal{F}_\epsilon:[0,\infty)\times C_R \to C_R$, by

$$\mathcal{F}_\epsilon(\lambda,u)(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\lambda\left(\psi\left(\left|u(\xi)\right|\right)+\epsilon\right)\mathrm{d}\xi\right]\mathrm{d}s, \tag{2.37}$$

where $\epsilon > 0$ is a constant. We have that $\mathcal{F}_\epsilon$ sends bounded sets of $[0,\infty)\times C_R$ into bounded sets of C_R. Moreover, $\mathcal{F}_\epsilon$ is a completely continuous operator and $\mathcal{F}_\epsilon(0,\cdot) = 0$.

Since

$$\deg_{\mathrm{LS}}\left(I - \mathcal{F}_\epsilon(0,\cdot), B(0,R), 0\right) = 1,$$

there exists a solution continuum $\mathcal{C}_\epsilon^+ \subset [0,\infty) \times C_R$ of solutions of

$$u = \mathcal{F}_\epsilon(\lambda, u) \tag{2.38}$$

with $\mathcal{C}_\epsilon^+$ unbounded in $[0,\infty) \times C_R$. In fact, this solution continuum $\mathcal{C}_\epsilon^+$ is bounded in the λ direction:

LEMMA 2.12. *There exists* $\bar{\lambda} > 0$ *such that if* (λ, u) *solve* (2.27)*, then* $\lambda \leqslant \bar{\lambda}$.

PROOF. Let (λ, u) be a solution pair to (2.27). Thus

$$r^{N-1}\phi\big(u'(r)\big) = -\lambda \int_0^r \xi^{N-1}\big(\psi\big(|u(\xi)|\big) + \epsilon\big)\,\mathrm{d}\xi \leqslant 0$$

and

$$u(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1}\lambda\big(\psi\big(|u(\xi)|\big) + \epsilon\big)\,\mathrm{d}\xi\right]\mathrm{d}s \geqslant 0.$$

Hence, $u'(r) \leqslant 0$ and $u(r) \geqslant 0$ for all $r \in [0, R]$. Also

$$u(r) \geqslant \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1}\lambda\psi\big(|u(\xi)|\big)\,\mathrm{d}\xi\right]\mathrm{d}s.$$

Thus, for all $r \in [\frac{R}{4}, \frac{3R}{4}]$, we have that

$$u(r) \geqslant \frac{R}{4}\phi^{-1}\left[\frac{\lambda R}{N4^N}\psi\big(u(r)\big)\right]$$

or equivalently,

$$\frac{\phi(\frac{4}{R}u(r))}{\psi(u(r))} \geqslant \lambda\frac{R}{N4^N}. \tag{2.39}$$

□

Note that it follows from (2.39) and (2.28) that for all $\epsilon > 0$, sufficiently small, λ is bounded independent of ϵ. Thus, there exists ϵ_0 such that for each $d > 0$ and each $0 < \epsilon \leqslant \epsilon_0$ there exists $(\lambda_\epsilon, u_\epsilon) \in \mathcal{C}_\epsilon^+$, with $\|u_\epsilon\| = d > 0$ and $0 < \lambda_\epsilon \leqslant \bar{\lambda}$. We may now let $\epsilon \to 0$ and obtain a nontrivial solution of (2.27), with $\|u\| = d$, for some $\lambda^* \in (0, \bar{\lambda}]$.

Similar to (2.30) define

$$\Gamma^+(d, R) = \sup \Gamma(d, R).$$

It follows from the above calculations that $\Gamma^+(d, R) < \infty$. We now prove $\Gamma^-(d, R) > 0$ from Theorem 2.11:

PROPOSITION 2.13. *For each* $d > 0$, $\Gamma^-(d, R) > 0$.

PROOF. If not, then there exist sequences $\{u_n\}$, $\{\lambda_n\}$, with $\|u_n\| = d$ and $\lambda_n \to 0$ such that $u_n = \mathcal{F}(\lambda_n, u_n)$. The complete continuity of $\mathcal{F}$ implies $\{u_n\}$ has a convergent subsequence, with $u_{n_j} \to u^*$, and thus $u^* = \mathcal{F}(0, u^*)$. But it follows from (2.36) that $\mathcal{F}(0, u) = 0$ for all u, a contradiction. □

Paralleling our notation above, let us set

$$\Gamma_1^+(R) = \limsup_{d\to\infty} \Gamma^+(d, R), \tag{2.40}$$

$$\gamma_1^+(R) = \limsup_{d\to 0} \Gamma^+(d, R), \tag{2.41}$$

$$\lambda_1^+(R) = \sup_{d>0} \Gamma^+(d, R). \tag{2.42}$$

Then, we have:

PROPOSITION 2.14. *Under the above hypotheses on* ϕ *and* ψ,

$$\begin{aligned} &\Gamma_1^-(R) > 0, \qquad &&\Gamma_1^+(R) < \infty, \\ &\gamma_1^-(R) > 0, &&\gamma_1^+(R) < \infty, \\ &\lambda_1^-(R) > 0, &&\lambda_1^+(R) < \infty. \end{aligned}$$

PROOF. That the numbers in the second column are finite has already been discussed. To show that the numbers in the first column are positive we argue as follows. Assume (λ, u) is a solution of (2.34) with $u(0) = d$. Then

$$\|u\| = d,$$

and

$$u(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1} \lambda \psi\big(|u(\xi)|\big)\, \mathrm{d}\xi\right] \mathrm{d}s. \tag{2.43}$$

Hence

$$d = u(0) \leqslant \int_0^R \phi^{-1}\left[\lambda \psi(d) \frac{s}{N}\right] \mathrm{d}s,$$

which implies

$$\frac{\phi(d/R)}{\psi(d)} \leqslant \lambda \frac{R}{N}. \tag{2.44}$$

Using condition (2.28) we obtain a lower bound for λ. □

The above proposition has the following corollary:

COROLLARY 2.15. *Let* (λ, u) *be a solution of* (2.34) *with* $u(0) = d$ *and let* $\theta \in (0, 1)$ *be fixed. Let* $r_0 \in (0, R)$ *be such that* $u(r_0) = \theta d$. *Then*

$$r_0 \geqslant \frac{N}{\lambda} A\left(\frac{1-\theta}{R}\right). \tag{2.45}$$

PROOF. Using Equation (2.43) we obtain the following

$$\theta d = \int_{r_0}^{R} \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1} \lambda \psi\bigl(\bigl|u(\xi)\bigr|\bigr)\,\mathrm{d}\xi\right] \mathrm{d}s, \tag{2.46}$$

and hence

$$(1-\theta)d = \int_0^{r_0} \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1} \lambda \psi\bigl(\bigl|u(\xi)\bigr|\bigr)\,\mathrm{d}\xi\right] \mathrm{d}s,$$

from which follows

$$(1-\theta)d \leqslant \int_0^{r_0} \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1} \lambda \psi(d)\,\mathrm{d}\xi\right] \mathrm{d}s,$$

and

$$(1-\theta)d \leqslant R\phi^{-1}\left[\frac{\lambda \psi(d) r_0}{N}\right].$$

The result now follows from (2.28). □

The result just proved has the following consequence concerning solutions of large norm:

COROLLARY 2.16. *Let* $\{(\lambda_n, u_n)\}$ *be a sequence of solutions of* (2.34) *with* $u_n(0) = d_n$. *If* $d_n \to \infty$ *as* $n \to \infty$, *then* $u_n(r) \to \infty$ *uniformly with respect to* r *in compact subintervals of* $[0, R)$.

PROOF. Since u_n is given by

$$u_n(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1} \lambda_n \psi\bigl(\bigl|u_n(\xi)\bigr|\bigr)\,\mathrm{d}\xi\right] \mathrm{d}s,$$

for $\theta \in (0, 1)$ we obtain for $r \geqslant r_0$ (viz. Corollary 2.15) that

$$u_n(r) \geqslant \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^{r_0} \xi^{N-1} \lambda_n \psi(\theta d_n)\,\mathrm{d}\xi\right] \mathrm{d}s.$$

The conclusion follows from this inequality. □

Let us consider the initial value problem

$$\begin{cases} -\left(r^{N-1}\phi(u')\right)' = \lambda r^{N-1}\psi(u), & r > 0, \\ u(0) = d > 0, \\ u'(0) = 0. \end{cases} \tag{2.47}$$

To obtain the (local) existence of solutions of (2.47) we obtain solutions of an equivalent integral equation whose solutions are fixed points of the operator S defined by

$$S(u)(r) = d - \int_0^r \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1}\lambda\psi\big(u(\xi)\big)\,\mathrm{d}\xi\right] \mathrm{d}s.$$

Since u is decreasing and ϕ and ψ are increasing we find that for given $\epsilon > 0$, there exists R_0 such that

$$S : \big\{u \in C[0, R_0]:\ \|u - d\| \leqslant \epsilon\big\} \to \big\{u \in C[0, R_0]:\ \|u - d\| \leqslant \epsilon\big\}$$

and that S is completely continuous. The result thus follows from Schauder's fixed point theorem. We next show that solutions of the initial value problem (2.47) exist globally and are oscillatory:

PROPOSITION 2.17. *For each* $d > 0$ *and each* $\lambda > 0$ *solutions of* (2.47) *exist globally on* $[0, \infty)$*, have only simple zeros and are oscillatory, i.e., the set of zeros is unbounded.*

PROOF. It follows from the above existence argument that a solution u of (2.47), as long as it is positive, is decreasing. If $u(r) \geqslant 0$ for $r \in [0, \infty)$, then using calculations as before, we find that for each $r > 0$

$$u(r) \geqslant r\phi^{-1}\left(\frac{\lambda r\psi(u(r))}{N2^{N-1}}\right),$$

and hence (since we may assume $r \geqslant 1$)

$$\frac{\phi(u(r))}{\psi(u(r))} \geqslant \frac{\lambda r}{N2^{N-1}},$$

which implies, using condition (2.28) that r cannot be unbounded. Therefore u must have a first zero. Easy arguments show that the zeros of u must be simple and that u is oscillatory. □

We next shall show that $\lambda_1^-(R)$, $\gamma_1^-(R)$ and $\Gamma_1^-(R)$ are nonincreasing functions of R. To this end we shall need the following elementary properties of the operator $\mathcal{F}$ defined by Equation (2.36). We note that the space C_R is a partially ordered Banach space with respect to the partial order $\leqslant$ defined by $u \leqslant v$ whenever $u(r) \leqslant v(r)$ for all $r \in [0, R]$. Further if $[u, v] = \{w \in C_R\colon\ u \leqslant w \leqslant v\}$ is an order interval in C_R, then it is a bounded closed set in C_R.

PROPOSITION 2.18. *The operator* $\mathcal{F}$ *defined by* (2.36) *is monotone with respect to the above partial order in* $C[0, R]$ *and hence in* C_R *and also monotone with respect to* λ.

From this proposition and the complete continuity of $\mathcal{F}$ immediately follows the following fixed point result:

PROPOSITION 2.19. *Assume there exists* $[\alpha, \beta] \subset C[0, R]$ *such that*

$$\mathcal{F}(\lambda, \cdot)\colon [\alpha, \beta] \to [\alpha, \beta].$$

Then $\mathcal{F}(\lambda, \cdot)$ *has a fixed point* $u \in C_R \cap [\alpha, \beta]$.

We note that the hypotheses of Proposition 2.19 will hold, whenever we can find a pair $\{\alpha, \beta\} \subset C[0, R]$ such that

$$\alpha \leqslant \beta$$

and

$$\alpha \leqslant \mathcal{F}(\lambda, \alpha), \mathcal{F}(\lambda, \beta) \leqslant \beta.$$

Using these facts we can now establish the following result:

THEOREM 2.20. λ_1^-, γ_1^-, Γ_1^- *are nonincreasing functions of* R.

PROOF. Assume there exist constants R_1 and R_2, with $R_1 < R_2$, such that $\lambda_1^-(R_1) < \lambda_1^-(R_2)$. Then there exists $\mu \in (\lambda_1^-(R_1), \lambda_1^-(R_2))$, such that (2.27) has a nontrivial solution $\tilde{\alpha}$ for $R = R_1$ and $\lambda = \mu$. Furthermore, there exists $\nu \geqslant \lambda_1^-(R_2)$ and a nontrivial solution β of (2.27) for $R = R_2$ and $\lambda = \nu$, with $\beta(0) = d$ as large as desired. It follows from Corollary 2.16 that for d sufficiently large $\beta(r) > \tilde{\alpha}(r)$, for each $r \in [0, R_1]$. Define

$$\alpha = \begin{cases} \tilde{\alpha}, & 0 \leqslant r \leqslant R_1, \\ 0, & R_1 \leqslant r \leqslant R_2. \end{cases}$$

Then the operator $\mathcal{F}(\mu, \cdot)$ for $R = R_2$ satisfies in the space $C[0, R_2]$

$$\alpha \leqslant \mathcal{F}(\mu, \alpha), \mathcal{F}(\mu, \beta) \leqslant \beta,$$

as may easily be verified. Thus by Proposition 2.19 this operator will have a fixed point in $[\alpha, \beta]$, contradicting that $\mu < \lambda_1^-(R_2)$. The monotonicity of the other functions is proved using virtually similar arguments. $\square$

REMARK 2.21.

1. Theorem 2.20 implies that problem (2.27) has no nontrivial solutions for $\lambda < \lambda_1^-(R)$.
2. Solutions of (2.27) are a priori bounded for λ in compact subintervals of $(-\infty, \Gamma_1^-(R))$.
3. The calculations and results also imply that for $\delta > 0$, sufficiently small, the degrees

$$\deg_{\mathrm{LS}}\bigl(I - \mathcal{F}(a, \cdot), B(0, \delta), 0\bigr),$$
$$\deg_{\mathrm{LS}}\bigl(I - \mathcal{F}(b, \cdot), B(0, \delta), 0\bigr)$$

are defined for $a < \gamma_1^-(R)$ and $b > \gamma_1^+(R)$ and equal 1 and 0, respectively. Hence it follows from global bifurcation theory that an unbounded continuum of positive solutions of (2.27) will bifurcate from $[\gamma_1^-(R), \gamma_1^+(R)]$.

We conclude this section with a few remarks concerning the nonhomogeneous equation

$$\begin{cases} \bigl(r^{N-1}\phi(u')\bigr)' + \lambda r^{N-1}\psi(u) = r^{N-1}h(r), & r \in (0, R), \\ u'(0) = u(R) = 0, \end{cases} \tag{2.48}$$

where $h \in L^\infty(0, R)$ is a given function. In this case, one has the following theorem:

THEOREM 2.22. *There exists* $\lambda_0 > 0$ *such that for every* $\lambda < \lambda_0$ *and every function* $h \in L^\infty(0, R)$ *the nonhomogeneous problem* (2.48) *has a solution.*

As is to be expected, λ_0 in this theorem is to the left of the set of principal eigenvalues associated with both positive and negative eigenfunctions.

PROOF. The existence of solutions to the nonhomogeneous problem is equivalent to the existence of fixed points of the operator $T(\lambda, u, h)$ defined by

$$T(\lambda, u, h)(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1}\bigl(\lambda\psi\bigl(u(\xi)\bigr) - h(r)\bigr)\,\mathrm{d}\xi\right]\mathrm{d}s. \tag{2.49}$$

One establishes the existence of a fixed point of T in the space C_R using Proposition 2.19. To this end, we define

$$\lambda_0 = \inf\bigl\{\lambda\colon\ \lambda \text{ is a principal eigenvalue of (2.27)}\bigr\}$$

and consider problem (2.48) for values of $\lambda < \lambda_0$. Let λ be so chosen and choose μ such that $\lambda < \mu < \lambda_0$ and consider the boundary value problem:

$$\begin{cases} \bigl(r^{N-1}\phi(u')\bigr)' + \mu r^{N-1}\psi(u) = 0, & 0 < r < R, \\ u'(0) = u(R) = 0. \end{cases} \tag{2.50}$$

It then follows from the results considered earlier that there exists a solution u with $u(0) = d$ as large as we like. Then u satisfies also

$$u(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\mu\psi\big(u(\xi)\big)\,\mathrm{d}\xi\right]\mathrm{d}s.$$

Further $u(r)$ may be made arbitrarily large uniformly on compact subintervals of $[0, R]$ by choosing d sufficiently large. Hence

$$\begin{aligned} u(r) &= \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\mu\psi\big(u(\xi)\big)\,\mathrm{d}\xi\right]\mathrm{d}s \\ &= \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\big(\lambda\psi\big(u(\xi)\big) - h\big)\,\mathrm{d}\xi\right. \\ &\quad \left. + \frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\big((\mu-\lambda)\psi\big(u(\xi)\big) + h\big)\,\mathrm{d}\xi\right]\mathrm{d}s \\ &\geqslant T(\lambda, u, h)(r) \end{aligned}$$

provided

$$\frac{1}{s^{N-1}}\int_0^s \xi^{N-1}\big((\mu-\lambda)\psi\big(u(\xi)\big) + h\big)\,\mathrm{d}\xi \geqslant 0, \quad \text{for } s \in [0, R],$$

which may be accomplished for a given $h \in L^\infty(0, R)$, by choosing d sufficiently large. In a similar manner we may construct a negative α such that $\alpha \leqslant T(\lambda, \alpha, h)$. We now apply Proposition 2.19 to complete the proof. □

2.2.5. *Existence and nonexistence.* We next consider existence and nonexistence questions for the following class of problems:

$$\begin{cases} \big(r^{N-1}\phi(u')\big)' + r^{N-1}g(u) = 0, & r \in (0, R), \\ u'(0) = u(R) = 0, \end{cases} \tag{2.51}$$

where ϕ is an odd increasing homeomorphism of $\mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ is a nondecreasing continuous function such that $g(0) = 0$.

Motivated by the case of the p-Laplace operator and related work (see, e.g., [29,48]), we will assume that ϕ satisfies:

$$\limsup_{s\to\infty} \frac{\phi(\sigma s)}{\phi(s)} < \infty, \quad \text{for all } \sigma > 1 \tag{2.52}$$

and

$$\limsup_{s\to 0} \frac{\phi(\sigma s)}{\phi(s)} < \infty, \quad \text{for all } \sigma > 1. \tag{2.53}$$

Let

$$\Phi(s) = \int_0^s \phi(t)\,\mathrm{d}t, \qquad G(s) = \int_0^s g(t)\,\mathrm{d}t,$$

and denote by

$$\Gamma := \limsup_{s\to\infty} \frac{\Phi(s)}{s\phi(s)},$$

and for $\theta \in (0, 1)$,

$$\delta_\theta := \liminf_{s\to\infty} \frac{G(\theta s)}{sg(s)}.$$

The main existence result is the following:

THEOREM 2.23. *Suppose that* $\Gamma < 1$ *and*

$$\lim_{s\to 0} \frac{\phi(s)}{g(s)} = \infty, \qquad \lim_{s\to\infty} \frac{g(s)}{\phi(s)} = \infty. \tag{2.54}$$

Further let there exist $\theta \in (0, 1)$ *such that* $\delta_\theta > 0$ *and*

$$N\delta_\theta > N\Gamma - 1. \tag{2.55}$$

Then (2.51) *has at least one positive solution.*

Condition (2.55) has been used in [48] to indicate subcritical growth of g with respect to ϕ. In the case of powers, i.e., when $\phi(s) = |s|^{p-2}s$ and $g(s) = |s|^{\delta-1}s$, with $1 < p < N$ and $\delta > 0$, this condition reads

$$\delta < \frac{N(p-1)+p}{N-p}, \tag{2.56}$$

which is known to be optimal in the sense that there are no positive solutions to (2.51) if

$$\delta \geqslant \frac{N(p-1)+p}{N-p}. \tag{2.57}$$

On the other hand, if we let $\phi(s) = \ln(1+|s|)|s|^{p-2}s$ then one has a solution for $p < \delta \leqslant \frac{N(p-1)+p}{N-p}$. Thus if we allow slightly faster growth we can include equality in (2.56) [23].

In the case of general ϕ and g, the main nonexistence result is the following:

THEOREM 2.24. *If*

$$\sup_{s\geqslant 0} \frac{NG(s)}{sg(s)} \leqslant \inf_{s\geqslant 0} \frac{N\Phi(s)}{s\phi(s)} - 1, \tag{2.58}$$

then (2.51) *does not have a positive solution.*

Condition (2.55) is related to a condition that appears to have been first used by Castro and Kurepa in [18,17] for the linear operator, i.e., when $\phi(s) = s$. De Thélin and El Hachimi [37] extended the results in [17] to the p-Laplacian case by assuming that $NG(s) - \frac{N-p}{p} sg(s)$ is bounded from below for all $s \in \mathbb{R}$ and for some $\theta \in (0, 1)$,

$$NG(\theta d) - \frac{N-p}{p} dg(d) \geqslant 0, \quad \text{for } d \text{ large.} \tag{2.59}$$

We note here that although condition (2.59) seems to relax the conditions in [18], it can be proved to be equivalent when $p = 2$ and g is increasing (see [47]).

PROOF OF THEOREM 2.23. By radial symmetry, finding positive solutions to problem (2.51) is equivalent to finding nontrivial solutions to the problem

$$\begin{cases} \left(r^{N-1}\phi(u')\right)' + r^{N-1} g\left(|u|\right) = 0, & r \in (0, R), \\ u'(0) = u(R) = 0. \end{cases} \tag{2.60}$$

Recall from (2.35) that C_R is the closed subspace of $C[0, R]$ defined by those u that vanish at R. Let u be a solution of (2.60). If $T_0 : C_R \to C_R$ is defined by

$$T_0(u)(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1} g\left(\left|u(\xi)\right|\right) \mathrm{d}\xi\right] \mathrm{d}s, \tag{2.61}$$

then T_0 is well defined and fixed points of T_0 will provide solutions of (2.60). Define now the operator $\mathcal{G} : [0, R] \times C_R \to C_R$, by

$$\mathcal{G}(\lambda, u)(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1}\left(g\left(\left|u(\xi)\right|\right) + \lambda h\right) \mathrm{d}\xi\right] \mathrm{d}s, \tag{2.62}$$

where $h > 0$ is a constant to be fixed later. We find that $\mathcal{G}$ sends bounded sets of $[0, R] \times C_R$ into bounded sets of C_R and $\mathcal{G}(0, u) = T_0(u)$. Moreover, $\mathcal{G}$ is a completely continuous operator.

Also, let us define the operator $\mathcal{S} : [0, R] \times C_R \to C_R$,

$$\mathcal{S}(\lambda, u) = \int_r^R \phi^{-1}\left[\frac{\lambda}{s^{N-1}} \int_0^s \xi^{N-1} g\left(\left|u(\xi)\right|\right) \mathrm{d}\xi\right] \mathrm{d}s. \tag{2.63}$$

Again, we see that $\mathcal{S}$ is completely continuous and note that $\mathcal{S}(1, \cdot) = T_0$.

One proves the existence of a fixed point of T_0, and hence of a positive solution to (2.51), by using suitable a priori estimates and degree theory. Indeed, one may show that there exist $R_1 > 0$ and $\epsilon_0 > 0$ such that

$$\deg_{\mathrm{LS}}\left(I - T_0, B(0, R_1), 0\right) = 0$$

and

$$\deg_{\mathrm{LS}}\big(I - T_0, B(0, \epsilon_0), 0\big) = 1$$

and thus, using the excision property of the Leray–Schauder degree it will follow that there must exist a solution of the equation

$$u = T_0(u)$$

in $B(0, R_1) \setminus \overline{B(0, \epsilon_0)}$.

To establish these facts, we need to prove that solutions $(\lambda, u) \in [0, R] \times C_R$ of the equation

$$u = \mathcal{G}(\lambda, u) \tag{2.64}$$

are a priori bounded. To do this one considers the auxiliary problem

$$\begin{cases} \big(r^{N-1}\phi(u')\big)' + r^{N-1}\big(g\big(|u|\big) + h\big) = 0, & r \in (0, R), \\ u'(0) = u(R) = 0. \end{cases} \tag{2.65}$$

and proves the following lemmas:

LEMMA 2.25. *There exists* $h_0 > 0$ *such that problem* (2.65) *has no solutions for* $h \geqslant h_0$.

PROOF. We argue by contradiction and thus assume that there exists a sequence $\{h_n\}_{n\in\mathbb{N}}$, with $h_n \to \infty$ as $n \to \infty$ such that (2.65) has a solution u_n. Then, u_n satisfies

$$\phi\big(u_n'(r)\big) = -r^{N-1} \int_0^r \xi^{N-1}\big(g\big(\big|u_n(\xi)\big|\big) + h_n\big)\,\mathrm{d}\xi \leqslant 0$$

and

$$u_n(r) = \int_r^R \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1}\big(g\big(\big|u_n(\xi)\big|\big) + h_n\big)\,\mathrm{d}\xi\right]\mathrm{d}s \geqslant 0.$$

Hence, $u_n'(r) \leqslant 0$ and $u_n(r) \geqslant 0$ for all $r \in [0, R]$. Also,

$$u_n(r) \geqslant (R - r)\phi^{-1}\left(\frac{rh_n}{N}\right), \quad \text{for all } r \in [0, R]$$

and thus, if $r \in [\frac{R}{4}, \frac{3R}{4}]$,

$$u_n(r) \geqslant \frac{R}{4}\phi^{-1}\left(\frac{Rh_n}{4N}\right). \tag{2.66}$$

Since we also have that for $r \in [\frac{R}{4}, \frac{3R}{4}]$

$$u_n(r) \geqslant \int_r^{\frac{3R}{4}} \phi^{-1}\left[\frac{1}{s^{N-1}} \int_0^s \xi^{N-1} g\big(|u_n(\xi)|\big)\, \mathrm{d}\xi\right] \mathrm{d}s,$$

by (2.66), the fact that u_n is decreasing, and the second assumption in (2.54), we have, that given arbitrary $A > 0$, there exists $n_1 \in \mathbb{N}$ such that for $n \geqslant n_1$

$$\begin{aligned} u_n(r) &\geqslant \int_r^{\frac{3R}{4}} \phi^{-1}\left[\frac{A}{s^{N-1}} \int_0^s \xi^{N-1} \phi\big(u_n(\xi)\big)\, \mathrm{d}\xi\right] \mathrm{d}s \\ &\geqslant \int_r^{\frac{3R}{4}} \phi^{-1}\left[\frac{A}{s^{N-1}} \int_0^r \xi^{N-1} \phi\big(u_n(\xi)\big)\, \mathrm{d}\xi\right] \mathrm{d}s, \end{aligned}$$

and, since $\phi(u_n(\cdot))$ is decreasing,

$$u_n(r) \geqslant \int_r^{\frac{3R}{4}} \phi^{-1}\left[\frac{Ar^N}{Ns^{N-1}} \phi\big(u_n(r)\big)\right] \mathrm{d}s.$$

Using now that $\frac{R}{4} \leqslant r \leqslant s \leqslant \frac{3R}{4}$, we obtain

$$\frac{Ar^N}{Ns^{N-1}} \geqslant dA,$$

where d is a constant. Thus, by the monotonicity of ϕ^{-1}, for all $r \in [\frac{R}{4}, \frac{R}{2}]$, we have that

$$u_n(r) \geqslant \frac{R}{4} \phi^{-1}\big(dA\phi\big(u_n(r)\big)\big), \quad \text{for all } n \geqslant n_1,$$

or equivalently,

$$\frac{\phi(\frac{4}{R} u_n(r))}{\phi(u_n(r))} \geqslant dA, \tag{2.67}$$

for all $n \geqslant n_1$ and for all $r \in [\frac{R}{4}, \frac{R}{2}]$.

It is clear that (2.67) cannot hold for A large if $R \geqslant 4$. If $R < 4$, by (2.66), condition (2.52), and the fact that A is arbitrary, we obtain a contradiction for $n \geqslant n_0$ for some $n_0 \geqslant n_1$. □

Let us fix $h \geqslant h_0$, for h_0 given in Lemma 2.25, and consider the family of problems

$$\begin{cases} \big(r^{N-1}\phi(u')\big)' + r^{N-1}\big(g\big(|u|\big) + \lambda h\big) = 0, & r \in (0, R), \\ u'(0) = u(R) = 0, \end{cases} \tag{2.68}$$

for $\lambda \in [0, 1]$.

The next lemma has a lengthy proof, see [47] for the details.

LEMMA 2.26. *Solutions to* (2.68) *are a priori bounded.*

Let S be as defined by (2.63). In order to complete the proof of the main result of this section we need one more lemma.

LEMMA 2.27. *There exists an* $\varepsilon_0 > 0$ *such that the equation*

$$u = \mathcal{S}(\lambda, u) \tag{2.69}$$

has no solution (λ, u) *with* $u \in \partial B(0, \varepsilon_0)$ *and* $\lambda \in [0, 1]$.

PROOF. We argue by contradiction and thus we assume that there are sequences $\{u_n\}$ and $\{\lambda_n\}$ with $\|u_n\| = \varepsilon_n \to 0$ as $n \to \infty$ and $\lambda_n \in [0, 1]$ such that (λ_n, u_n) satisfies (2.69) for each $n \in \mathbb{N}$. We have that (λ_n, u_n) satisfies

$$u_n(r) = \int_r^R \phi^{-1}\left[\frac{\lambda_n}{s^{N-1}} \int_0^s \xi^{N-1} g\left(\left|u_n(\xi)\right|\right) \mathrm{d}\xi\right] \mathrm{d}s$$

which implies, by the first assumption in (2.54), that for sufficiently large n

$$\varepsilon_n \leqslant \phi^{-1}\left(\phi(\varepsilon_n)\frac{\mu R}{N}\right) R,$$

where μ is a positive arbitrarily small number. Thus

$$\phi\left(\frac{\varepsilon_n}{R}\right) \leqslant \frac{\mu R}{N}\phi(\varepsilon_n). \tag{2.70}$$

If $R \leqslant 1$, we immediately reach a contradiction. If now $R > 1$, let us set $\tilde{\varepsilon}_n = \varepsilon_n / R$. Then

$$\frac{\phi(R\tilde{\varepsilon}_n)}{\phi(\tilde{\varepsilon}_n)} \geqslant \frac{N}{\mu R},$$

and we reach a contradiction by condition (2.53) and the fact that μ is arbitrary. □

It now follows from Lemmas 2.25 and 2.26, that if u is a solution to the equation

$$u = \mathcal{G}(\lambda, u), \quad \lambda \in [0, 1],$$

then $\|u\| \leqslant C$, where C is a positive constant. Thus if $B(0, R_1)$ is the ball centered at 0 in C_R with radius $R_1 > C$, then we have that the Leray–Schauder degree of the operator

$$I - \mathcal{G}(\lambda, \cdot) : \overline{B(0, R_1)} \to C_R$$

is well defined for every $\lambda \in [0, 1]$. Then, by the properties of the Leray–Schauder degree, we have that

$$\deg_{\mathrm{LS}}\bigl(I - T_0, B(0, R_1), 0\bigr) = \deg_{\mathrm{LS}}\bigl(I - \mathcal{G}(1, \cdot), B(0, R_1), 0\bigr) = 0, \tag{2.71}$$

since (2.64) does not have solutions on $\{1\} \times \overline{B(0, R_1)}$. Also, by Lemma 2.27 and the properties of the Leray–Schauder degree, it follows that for $\varepsilon_0 > 0$ small enough,

$$\deg_{\mathrm{LS}}\bigl(I - \mathcal{S}(\lambda, \cdot), B(0, \varepsilon_0), 0\bigr) \quad \text{is constant for all } \lambda \in [0, 1].$$

Hence

$$\deg_{\mathrm{LS}}\bigl(I - T_0, B(0, \varepsilon_0), 0\bigr) = \deg_{\mathrm{LS}}\bigl(I, B(0, \varepsilon_0), 0\bigr) = 1. \tag{2.72}$$

Thus, using the excision property of the Leray–Schauder degree we conclude from (2.71) and (2.72) that there must be a solution of the equation

$$u = T_0(u)$$

with $u \in B(0, R_1) \setminus \overline{B(0, \varepsilon_0)}$. This fixed point corresponds to a nontrivial solution to (2.51), thus completing the proof of Theorem 2.23. □

We conclude this section with a proof of the nonexistence result (Theorem 2.24) and some applications:

PROOF OF THEOREM 2.24. Assume condition (2.58) and the existence of a positive solution u of (2.51). Let

$$\Phi_*^{-1}(t) = \int_0^t \frac{\Phi^{-1}(\tau)}{\tau^{(N-1)/N}}\, \mathrm{d}\tau$$

denote the Orlicz–Sobolev conjugate of Φ. Then, multiplying the equation by $ru'(r) + bu(r)$, where

$$b = \sup_{s \geqslant 0} \frac{NG(s)}{sg(s)},$$

one obtains (see [47])

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}r}\bigl[r^N\bigl(\Phi_*\bigl(\phi\bigl(u'(r)\bigr)\bigr) + G\bigl(u(r)\bigr)\bigr) + br^{N-1}u(r)\phi\bigl(u'(r)\bigr)\bigr] \\
&\quad = r^{N-1}\bigl[(b+1)\phi\bigl(u'(r)\bigr)u'(r) - N\Phi\bigl(u'(r)\bigr)\bigr] \\
&\qquad + r^{N-1}\bigl[NG\bigl(u(r)\bigr) - bg\bigl(u(r)\bigr)u(r)\bigr] \\
&\quad \leqslant 0,
\end{aligned}$$

and thus, one concludes that

$$E(r) := r^N\big(\Phi_*\big(\phi\big(u'(r)\big)\big) + G\big(u(r)\big)\big) + br^{N-1}u(r)\phi\big(u'(r)\big)$$

is nonincreasing for $r \in [0, R]$. However, $E(0) = 0$ and $E(R) = \Phi_*(\phi(u'(R)) > 0$, which is a contradiction. □

The following two examples will serve to illustrate the above existence and nonexistence results. Consider the boundary value problem

$$\begin{cases} \operatorname{div}\big(\frac{|\nabla u|^{2p-2}\nabla u}{\sqrt{1+|\nabla u|^{2p}}}\big) + u^q = 0, & x \in B_R(0), \\ u = 0, & x \in \partial B_R(0), \end{cases} \tag{2.73}$$

for $R > 0$. In this case the function ϕ is given by

$$\phi(s) = \frac{|s|^{2p-2}s}{\sqrt{1+|s|^{2p}}}$$

and one obtains nonexistence (see (2.58)) whenever

$$q \geqslant \frac{N(2p-1)+2p}{N-2p}$$

and existence whenever

$$p - 1 < q < \frac{N(p-1)+p}{N-p}$$

(see Theorem 2.23).

For the boundary value problem

$$\begin{cases} \operatorname{div}\big((1+|\nabla u|^2)^{\frac{m-1}{2}}\nabla u\big) + u^q = 0, & x \in B_R(0), \\ u = 0, & x \in \partial B_R(0), \end{cases}$$

the above conditions imply the existence of positive solutions when

$$m < q < \frac{Nm+m+1}{N-m-1}$$

and nonexistence when

$$q \geqslant \begin{cases} \frac{Nm+m+1}{N-m-1}, & m > 1, \\ \frac{N+2}{N-2}, & m \leqslant 1. \end{cases}$$

2.3. *The case when $f(0) < 0$*

We now consider (2.1), i.e.,

$$\begin{cases} \left(r^{N-1}\phi(u')\right)' + \lambda r^{N-1} f(u) = 0, & 0 < r < R, \\ u'(0) = u(R) = 0, \end{cases} \tag{2.74}$$

and allow for $f(0) < 0$. In this case the problem has no trivial solution, hence bifurcation results (relative to a trivial branch of solutions) will not apply and other methods must be used.

We make the following assumptions:

(A.1) ϕ is an odd, increasing homeomorphism on $\mathbb{R}$, and for each $\sigma > 0$ there exists a positive number $M_\sigma > 0$ such that

$$\phi(\sigma s) \leqslant M_\sigma \phi(s) \tag{2.75}$$

for every $s \in \mathbb{R}$.

(A.2) $f : \mathbb{R}^+ \to \mathbb{R}$ is continuous, nondecreasing on $[B, \infty)$ for some $B > 0$, and

$$\lim_{s\to\infty} \frac{f(s)}{\phi(s)} = \infty. \tag{2.76}$$

(A.3) There exists $\theta \in (0, 1)$ such that

$$N \liminf_{s\to\infty} \frac{F(\theta s)}{s f(s)} > \max\left\{ N \limsup_{s\to\infty} \frac{\Phi(s)}{s\phi(s)} - 1, 0 \right\}, \tag{2.77}$$

where, as usual, $F(s) = \int_0^s f(t)\,\mathrm{d}t$ and $\Phi(s) = \int_0^s \phi(t)\,\mathrm{d}t$.

The main result in this section is:

THEOREM 2.28. *Let* (A.1)–(A.3) *hold and assume that* ϕ^{-1} *is concave on* $\mathbb{R}^+$. *Then there exists a positive number* $\bar{\lambda}$ *such that for* $0 < \lambda < \bar{\lambda}$, *Equation* (2.74) *has a positive, decreasing solution* u_λ *with*

$$\lim_{\lambda\to 0} u_\lambda(r) = \infty \tag{2.78}$$

uniformly for r *in compact subsets of* $(0, R)$.

REMARK 2.29.

1. Condition (A.3) was employed in [47] to establish the existence of positive solutions to (2.74) when $f(0) = 0$.
2. In the case when $\phi(s) = |s|^{p-2}s$, (A.3) becomes

$$\liminf_{s\to\infty} \frac{F(\theta s)}{s f(s)} > \max\left(\frac{N-p}{Np}, 0 \right). \tag{2.79}$$

In particular, if $f(s) = s^q$, (2.76) and (2.77) require that

$$p - 1 < q < \frac{Np}{\max(N - p, 0)} - 1.$$

3. In the case when $\phi(s) = s$ and $N \geqslant 2$, the existence part of Theorem 2.28 was established in [19] under the assumptions that f is nondecreasing on $\mathbb{R}$ and

$$\lim_{s\to\infty}\left(NF(\theta s) - \frac{N-2}{2}sf(s)\right)\left(\frac{s}{f(s)}\right)^{N/2} = \infty, \tag{2.80}$$

for some $\theta \in (0, 1)$. Note that (2.80) can be derived from (2.79) with $p = 2$ when f is nondecreasing.

After several technical computations one proceeds to establish properties for the associated operator equations which are defined as follows. For each $v \in C[0, R]$, and $\mu \geqslant 0$, let $u = A_\mu v$ be the solution of

$$\begin{cases} \left(r^{N-1}\phi(u')\right)' = -\lambda r^{N-1}\left(f(v) + \mu\right), & 0 < r < R, \\ u'(0) = u(R) = 0. \end{cases} \tag{2.81}$$

Then $A_\mu : C[0, R] \to C[0, R]$ is a completely continuous operator and its fixed points u such that $u \geqslant w$ on $[0, R]$ will yield solutions of (2.81) and hence of (2.74). To establish the existence of such fixed points we shall list some auxiliary result, whose proofs may be found in [53].

LEMMA 2.30. *For each* $\lambda > 0$, *small enough, there exists a positive number* R_λ *with* $\lim_{\lambda\to 0} R_\lambda = \infty$ *such that all solutions of*

$$u = \theta A_0 u, \quad 0 \leqslant \theta \leqslant 1, \tag{2.82}$$

satisfy $\|u\| \neq R_\lambda$.

LEMMA 2.31. *There exists a positive number* μ_0 (*depending on* λ) *such that*: *If* A_μ *has a fixed point, then* $\mu \leqslant \mu_0$.

LEMMA 2.32. *There exists a function* $H : \mathbb{R}^+ \to \mathbb{R}$ *with* $\lim_{d\to\infty} H(d) = \infty$ *such that all solutions u of* $u = A_\mu(u)$ *satisfy*

$$u(r) \geqslant H\left(\|u\|\right)(1 - r), \quad r \in [0, 1]. \tag{2.83}$$

LEMMA 2.33. *For each* $\nu > 0$, *there exists a positive number* C_ν *such that all solutions u of*

$$u = A_\mu(u) \tag{2.84}$$

with $\lambda \geqslant \nu$ *satisfy* $\|u\| < C_\nu$.

PROOF OF THEOREM 2.28. Using the above lemmas we obtain

$$\deg_{\mathrm{LS}}\bigl(I - A_0, B(0, C_\lambda), 0\bigr) = \deg_{\mathrm{LS}}\bigl(I - A_\mu, B(0, C_\lambda), 0\bigr) = 0, \tag{2.85}$$

where C_λ is given by Lemma 2.33. On the other hand, Lemma 2.30 implies that

$$\deg_{\mathrm{LS}}\bigl(I - A_0, B(0, R_\lambda), 0\bigr) = 1. \tag{2.86}$$

Hence it follows from the excision property of the Leray–Schauder degree that there exists a solution u_λ such that

$$R_\lambda \leqslant \|u_\lambda\| \leqslant C_\lambda.$$

Since $R_\lambda \to \infty$, as $\lambda \to 0$, it follows from Lemma 2.32 that $u_\lambda \geqslant 0$, for λ small. The remainder of the proof is straightforward. □

2.4. *The case when* $f(0) > 0$

In this section we again consider (2.74):

$$\begin{cases} \bigl(r^{N-1}\phi(u')\bigr)' + \lambda r^{N-1} f(u) = 0, & 0 < r < R, \\ u'(0) = u(R) = 0, \end{cases} \tag{2.87}$$

and allow for $f(0) > 0$. We continue to assume the conditions (A.1)–(A.3) imposed in the previous section hold. The main theorem follows:

THEOREM 2.34. *Let* (A.1)–(A.3) *hold and suppose that* $f(s) > 0$ *for* $s \geqslant 0$. *Then there exists a positive number* λ^* *such that* (2.74) *has at least two solutions for* $\lambda < \lambda^*$, *at least one for* $\lambda = \lambda^*$, *and none for* $\lambda > \lambda^*$.

Since we are interested in positive solutions, we define $f(s) = f(0)$ if $s \leqslant 0$. Before proving Theorem 2.34 we outline several lemmas:

LEMMA 2.35. *There exists a positive number* $\tilde{\lambda}$ *such that* (2.74) *has no solution for* $\lambda > \tilde{\lambda}$.

PROOF. Let u be a solution of (2.74) and $r \in (0, R)$. We have

$$u(r) \geqslant \phi^{-1}\left[\lambda \int_0^r \tau^{N-1} f(u)\,\mathrm{d}\tau\right](R - r) \tag{2.88}$$

$$\geqslant \phi^{-1}\left[\frac{\lambda r^N}{N} \bar{f}\bigl(u(r)\bigr)\right](R - r), \tag{2.89}$$

where $\bar{f}(u) = \inf_{s \geqslant u} f(s)$.

Hence

$$\frac{\phi(\frac{u(r)}{R-r})}{\bar{f}(u(r))} \geqslant \frac{\lambda r^N}{N}. \tag{2.90}$$

On the other hand, since there exists a positive number ε such that $\bar{f}(x) \geqslant \varepsilon$ for every x, it follows that

$$u(r) \geqslant \phi^{-1}\left[\frac{\lambda r^N \varepsilon}{N}\right](R-r). \tag{2.91}$$

Combining the above, we reach a contradiction if λ is sufficiently large, since $\lim_{s\to\infty} \frac{\phi(s)}{\bar{f}(s)} = 0$. ☐

Lemma 2.36. *For each $\mu > 0$, there exists a positive constant C_μ such that all solutions u of* (2.74) *with $\lambda > \mu$ satisfy $\|u\| < C_\mu$.*

Lemma 2.37. *If $\Lambda = \{\lambda > 0$: (2.74) has a solution$\}$ and $\lambda^* = \sup \Lambda$, then $0 < \lambda^* < \infty$ and $\lambda^* \in \Lambda$.*

Proof. Using the Leray–Schauder continuation theorem, it follows that there is a solution for λ small, and so $\lambda^* > 0$. Also, $\lambda^* < \infty$. We verify that $\lambda^* \in \Lambda$. Let $\{\lambda_n\} \subset \Lambda$ be such that $\lambda_n \to \lambda^*$, and let $\{u_n\}$ be the corresponding solutions of (2.74) with $\lambda = \lambda_n$. From Lemma 2.36, we deduce that $\{u_n\}$ is bounded in $C^1[0, R]$. Hence $\{u_n\}$ has a subsequence which converges to $u \in C[0, R]$. By a standard limiting process, it follows that u is a solution for (2.74) with $\lambda = \lambda^*$. ☐

Lemma 2.38. *Let $0 < \lambda < \lambda^*$ and let u_{λ^*} be a solution of* (2.74). *Then there exists a positive number ε such that $u_{\lambda^*} + \varepsilon$ is an upper solution of* (2.74).

For $v \in C[0, R]$, let $A_\lambda v = u$ where u satisfies

$$\begin{cases} \left(r^{N-1}\phi(u')\right)' = -\lambda r^{N-1} f(v), & 0 < r < R, \\ u'(0) = u(R) = 0. \end{cases} \tag{2.92}$$

Define

$$\tilde{f}(v) = \begin{cases} f(u_{\lambda^*} + \varepsilon), & v \geqslant u_{\lambda^*} + \varepsilon, \\ f(v), & -\varepsilon \leqslant v \leqslant u_{\lambda^*} + \varepsilon, \\ f(0), & v \leqslant 0, \end{cases} \tag{2.93}$$

and let $\tilde{A}_\lambda$ be the operator analogous to A_λ defined by $\tilde{f}$.

Lemma 2.39. *Let u be a solution of $u = \tilde{A}_\lambda u$. Then $0 \leqslant u \leqslant u_{\lambda^*} + \varepsilon$.*

PROOF. Let u satisfy $u = \tilde{A}_\lambda u$. Then

$$\begin{cases} \left(r^{N-1}\phi(u')\right)' = -\lambda r^{N-1} f(u), & 0 < r < R, \\ u'(0) = u(R) = 0. \end{cases} \tag{2.94}$$

By the maximum principle, $u \geqslant 0$. We claim that $u \leqslant u_{\lambda^*} + \varepsilon$. Suppose that there exists $\bar{r} \in (0, R)$ such that $u(\bar{r}) > u_{\lambda^*}(\bar{r}) + \varepsilon$. Then there exists numbers r_0 and r_1 with $0 \leqslant r_0 < r_1 < R$ such that $u'(r_0) = u'_{\lambda^*}(r_0)$, $u(r_1) = u_{\lambda^*}(r_1) + \varepsilon$, and $u > u_{\lambda^*} + \varepsilon$ on (r_0, r_1).

Hence

$$\left(r^{N-1}\phi(u')\right)' = -\lambda r^{N-1} f(u_{\lambda^*} + \varepsilon), \quad \text{for } r \in (r_0, r_1]. \tag{2.95}$$

Since

$$\left(r^{N-1}\phi(u'_{\lambda^*})\right)' \leqslant -\lambda r^{N-1} f(u_{\lambda^*} + \varepsilon), \tag{2.96}$$

we deduce that

$$\left[r^{N-1}\left(\phi(u') - \phi(u'_{\lambda^*})\right)\right]' \geqslant 0, \quad \text{for } r \in (r_0, r_1], \tag{2.97}$$

and so $u - (u_{\lambda^*} + \varepsilon)$ is nondecreasing on $(r_0, r_1]$.

Consequently,

$$u(r) < u_{\lambda^*}(r) + \varepsilon, \quad r \in (r_0, r_1], \tag{2.98}$$

a contradiction. □

We now summarize a proof of the main theorem:

PROOF OF THEOREM 2.34. Let $0 < \lambda < \lambda^*$. Since 0 is a lower solution and u_{λ^*} is an upper solution (see [31,66]), there exists a solution u_λ with $0 \leqslant u_\lambda \leqslant u_{\lambda^*}$. We next establish the existence of a second solution. Define

$$B = \left\{u \in C[0, R]\colon -\varepsilon \leqslant u \leqslant u_{\lambda^*} + \varepsilon\right\}. \tag{2.99}$$

Then B is open and $u_\lambda \in B$. Since $\tilde{A}_\lambda$ is bounded for λ in compact intervals,

$$\deg_{\mathrm{LS}}\left(I - \tilde{A}_\lambda, B(u_\lambda, M), 0\right) = 1 \tag{2.100}$$

if M is sufficiently large. Here $B(u_\lambda, M)$ denotes the open ball center at $u_\lambda$with radius M in $C[0, R]$.

If there exists $u \in \partial B$ such that $u = \tilde{A}_\lambda u$ then u is a second solution. Suppose that $u \neq \tilde{A}_\lambda u$ for $u \in \partial B$. Then $\deg_{\mathrm{LS}}(I - \tilde{A}_\lambda, B, 0)$ is well defined and since $\tilde{A}_\lambda$ has no fixed point in $B(u_\lambda, M) \setminus B$, we have

$$\deg_{\mathrm{LS}}(I - A_\lambda, B, 0) = \deg(I - \tilde{A}_\lambda, B, 0) = 1. \tag{2.101}$$

On the other hand, it follows that

$$\deg_{\mathrm{LS}}\bigl(I - A_\lambda, B(0, M), 0\bigr) = 0 \tag{2.102}$$

for some large number M. Hence

$$\deg_{\mathrm{LS}}\bigl(I - A_\lambda, B(0, M) \setminus B, 0\bigr) = -1, \tag{2.103}$$

and the existence of a second solution follows. □

3. Problems on annular domains

We now consider boundary value problems for nonlinear ordinary differential equations of the form

$$\begin{cases} \bigl(\phi(u')\bigr)' + \frac{N-1}{r}\phi(u') + f(u) = 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0, \end{cases} \tag{3.1}$$

where $R_1 > 0$. The parameter λ plays no role in the considerations to follow, thus it is suppressed in Equation (3.1). We again assume $\phi : \mathbb{R} \to \mathbb{R}$ is an odd increasing homeomorphism such that for all $c > 0$, there exists $A_c > 0$, such that

$$A_c\phi(s) \leqslant \phi(cs), \quad \text{for } s \in \mathbb{R}^+, \tag{3.2}$$

and

$$\lim_{c\to\infty} A_c = \infty. \tag{3.3}$$

This assumption implies that there exists $B_c > 0$ such that

$$\phi(cs) \leqslant B_c\phi(s), \quad \text{for } s \in \mathbb{R}^+,$$

and

$$\lim_{c\to 0} B_c = 0.$$

Since the only property of the function $\frac{N-1}{r}$ we shall use here is its continuity, we will consider the more general problem

$$\begin{cases} \bigl(\phi(u')\bigr)' + b(r)\phi(u') + f(u) = 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0, \end{cases} \tag{3.4}$$

where $b : [R_1, R_2] \to \mathbb{R}$ is a continuous function. The nonlinear term f will be assumed to be continuous and satisfy

$$\limsup_{s \to 0} \frac{f(s)}{\phi(s)} \leqslant 0 \tag{3.5}$$

and to grow superlinearly (with respect to ϕ) near infinity, i.e.,

$$\lim_{s \to \infty} \frac{f(s)}{\phi(s)} = \infty. \tag{3.6}$$

For such problems we shall establish that (3.4) always has a positive solution defined for any interval $[R_1, R_2]$. The results discussed are motivated by problems studied in [30,51, 62,73]. We remark that the results given are equally valid in case f depends upon r also, provided the assumptions made are assumed uniform with respect to r.

3.1. *Fixed point formulation*

Using the integrating factor

$$p(r) = \mathrm{e}^{\int_{R_1}^{r} b(s)\,\mathrm{d}s},$$

we may rewrite problem (3.4) equivalently as

$$\begin{cases} \big(p\phi(u')\big)' + pf(u) = 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0. \end{cases} \tag{3.7}$$

The existence of solutions of (3.7) (hence (3.4)) is established by proving the existence of fixed points of a completely continuous operator $F : E \to E$, where $E = C[R_1, R_2]$, endowed with the norm $\|u\| = \max_{r \in [R_1, R_2]} |u(r)|$. The operator F is defined by the following lemma:

LEMMA 3.1. *Let ϕ be as above and let c be a given nonnegative constant. Then for each $v \in E$ the problem*

$$\begin{cases} \big(p\phi(u')\big)' - cp\phi(u) = pv, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0, \end{cases} \tag{3.8}$$

has a unique solution $u = T(v)$, and the operator $T : E \to E$ is completely continuous.

PROOF. For each $w \in E$ let $u = B(w)$ be the unique solution of

$$\begin{cases} \big(p\phi(u')\big)' - cp\phi(w) = pv, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0, \end{cases}$$

given by

$$u(r) = \int_{R_1}^{r} \phi^{-1}\left(\frac{1}{p}\left\{q - \int_{R_1}^{s} p\big(c\phi(w) + v\big)\, \mathrm{d}\xi\right\}\right) \mathrm{d}s,$$

where q is the unique number such that $u(R_2) = 0$. The existence of a fixed point u of B, hence a solution u to Equation (3.8), follows from the continuation theorem of Leray–Schauder. One may then define the operator T by

$$T(v) = u,$$

where u is the solution of (3.8).

To accomplish this, assume $u \in E$ and $\lambda \in (0, 1)$ such that

$$u = \lambda B(u).$$

Then

$$\begin{cases} \big(p\phi\big(\frac{u'}{\lambda}\big)\big)' - cp\phi(u) = pv, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0. \end{cases}$$

Multiplying the above by u and integrating we obtain

$$\int_{R_1}^{R_2} p\phi\left(\frac{u'}{\lambda}\right)u'\, \mathrm{d}r + \int_{R_1}^{R_2} cp\phi(u)u\, \mathrm{d}r = -\int_{R_1}^{R_2} pvu\, \mathrm{d}r.$$

On the other hand, since ϕ is an increasing homeomorphism, for each $\epsilon > 0$, there exists a constant c_ϵ such that

$$|s| \leqslant \epsilon\phi(s)s + c_\epsilon, \quad \text{for } s \in \mathbb{R}.$$

Thus we obtain

$$\left|\int_{R_1}^{R_2} pvu\, \mathrm{d}r\right| \leqslant \|v\| \int_{R_1}^{R_2} p|u|\, \mathrm{d}r \leqslant \|v\| \left\{\epsilon \int_{R_1}^{R_2} p\phi(u)u\, \mathrm{d}r + c_\epsilon(R_2 - R_1)\right\},$$

and, by choosing ϵ appropriately,

$$\left|\int_{R_1}^{R_2} pvu\, \mathrm{d}r\right| \leqslant \frac{1}{2}c \int_{R_1}^{R_2} \big(p\phi(u)u + c_1\big)\, \mathrm{d}r,$$

where c_1 is a constant. Hence we obtain

$$\int_{R_1}^{R_2} p\phi\left(\frac{u'}{\lambda}\right)u'\, \mathrm{d}r \leqslant c_2,$$

for a constant c_2. Therefore

$$\int_{R_1}^{R_2} |u'|\,\mathrm{d}r \leqslant c_3,$$

and hence

$$\|u\| \leqslant c_4.$$

Further

$$\left\| \left(p\phi\left(\frac{u'}{\lambda}\right)\right)' \right\|_{L^1} \leqslant c_5.$$

Since there exists r_0 such that $u'(r_0) = 0$, we obtain from the latter inequality that

$$\left| p\phi\left(\frac{u'}{\lambda}\right)\right| \leqslant c_6$$

and hence

$$\|u'\| \leqslant c_7,$$

where $c_1, \dots, c_7$ are constants independent of λ. The complete continuity of B is easily established and we conclude from these a priori bounds that B has a fixed point $u = T(v)$. If u_1 and u_2 are fixed points, one immediately obtains that

$$\int_{R_1}^{R_2} p\big(\phi(u_1') - \phi(u_2')\big)(u_1' - u_2')\,\mathrm{d}r + \int_{R_1}^{R_2} cp\big(\phi(u_1) - \phi(u_2)\big)(u_1 - u_2)\,\mathrm{d}r = 0,$$

and hence, since ϕ is increasing, that $u_1 = u_2$. Thus the operator T given in the statement of the lemma is well defined. □

Using Lemma 3.1 we obtain a fixed point formulation of problem (3.7) as

$$u = F(u) = T\big(-c\phi(u) - f(u)\big). \tag{3.9}$$

The next lemma is crucial for establishing the existence of nonzero fixed points of (3.9), its proof may be found in [74].

LEMMA 3.2. *If u is the solution of* (3.8) *with $v \leqslant 0$, then*

$$u(r) \geqslant c_0\|u\|k(r), \quad R_1 \leqslant r \leqslant R_2,$$

where

$$k(r) = \frac{1}{R_2 - R_1} \min\{r - R_1, R_2 - r\},$$

and c_0 *is a positive constant.*

3.2. *Existence results*

We next establish the existence of a nontrivial solution of (3.7) by means of a fixed point argument, similar to those used above, (see [65,51]). We need the following auxiliary results.

PROPOSITION 3.3. *Assume that* $f : \mathbb{R} \to \mathbb{R}$ *is continuous and there exists a constant* $c > 0$ *such that*

$$f(s) + c\phi(s) \geqslant 0, \quad s \geqslant 0.$$

Further assume there exists a constant $m > 0$ *such that*

$$u = T\big(\lambda\big(-c\phi\big(|u|\big) - f\big(|u|\big)\big)\big), \quad 0 \leqslant \lambda \leqslant 1, \quad \Rightarrow \quad \|u\| \neq m,$$

and there exists a constant $M > m$ *and an element* $h \in E$, $h \leqslant 0$ *such that*

$$u = T\big(-c\phi(u) - f\big(|u|\big) + \lambda h\big), \quad 0 \leqslant \lambda \leqslant 1, \quad \Rightarrow \quad \|u\| \neq M,$$

further any solution u of

$$u = T\big(-c\phi(u) - f\big(|u|\big) + h\big),$$

satisfies $\|u\| > M$. *Then there exists a fixed point* $u \in E$ *of the operator* F, $F(u) = T(-c\phi(|u|) - f(|u|))$ *such that*

$$m < \|u\| < M.$$

Applying this proposition to the boundary value problem (3.7), we have the following corollary.

COROLLARY 3.4. *Assume that* $f : \mathbb{R} \to \mathbb{R}$ *is continuous and there exists a constant* $c > 0$ *such that*

$$f(u) + c\phi(u) \geqslant 0, \quad u \geqslant 0.$$

Further assume there exists a constant $m > 0$ *such that for* $0 \leqslant \lambda \leqslant 1$ *and any solution* u *of*

$$\begin{cases} \big(p\phi(u')\big)' - cp\phi(u) + \lambda\big(cp\phi\big(|u|\big) + pf\big(|u|\big)\big) = 0, & R_1 < r < R_2, \\ u(R_1) = (R_2) = 0, \end{cases} \tag{3.10}$$

satisfies

$$\|u\| \neq m,$$

and there exists a constant $M > m$ *and an element* $h \in E$, $h \geqslant 0$ *such that any solution* u *of*

$$\begin{cases} \big(p\phi(u')\big)' - p\phi(u) + p\phi\big(|u|\big) + pf\big(|u|\big) \leqslant 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0, \end{cases} \tag{3.11}$$

satisfies $\|u\| \neq M$, *and solutions* u *of*

$$\begin{cases} \big(p\phi(u')\big)' - p\phi(u) + p\phi\big(|u|\big) + pf\big(|u|\big) + ph = 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0, \end{cases} \tag{3.12}$$

satisfy $\|u\| > M$. *Then there exists a solution* u, $u \geqslant 0$ *of* (3.9) (*hence* (3.7)) *such that*

$$m < \|u\| < M.$$

We next impose conditions on f and further conditions on ϕ in order that the above corollary may be applied. These are conditions about the behavior of the functions near zero and near infinity. These conditions in turn will provide the validity of Corollary 3.4 and hence yield the existence of nontrivial solutions.

PROPOSITION 3.5. *Assume that* $f : \mathbb{R} \to \mathbb{R}$ *is continuous and let there exist a constant* $c \geqslant 0$ *such that*

$$f(u) + c\phi(u) \geqslant 0, \quad u \geqslant 0.$$

Further, let ϕ *be an odd increasing homeomorphism of* $\mathbb{R}$ *which satisfies* (3.2) *and assume* (3.5) *holds. Then there exists a positive number* m *such that for* $0 \leqslant \lambda \leqslant 1$ *and any solution* u *of*

$$\begin{cases} \big(p\phi(u')\big)' - cp\phi(u) + \lambda\big(pf\big(|u|\big) + cp\phi\big(|u|\big)\big) = 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0, \end{cases}$$

satisfies $\|u\| \neq m$.

PROOF. First note that solutions are nonnegative. Following an indirect argument, assume the existence of sequences $\{\lambda_n\} \subset [0, 1]$, $\{\epsilon_n\}$, $\epsilon_n \searrow 0$, $\{u_n\}$, $\|u_n\| = \epsilon_n$, such that the triple $(\epsilon_n, \lambda_n, u_n)$ is a solution. Since each $u_n \neq 0$ has a maximum point, say $u_n(r_n^*) = \|u_n\|$, we obtain by integration

$$u_n'(r) \leqslant \phi^{-1}\left(\frac{\lambda_n}{p(r)} \int_r^{r_n^*} p(s) f\big(u_n(s)\big)\,\mathrm{d}s\right)$$

and hence, for given $\epsilon > 0$ we have for large n

$$u_n'(r) \leqslant \phi^{-1}\left(\frac{\lambda_n}{p(r)} \int_r^{r_n^*} \epsilon p(s)\phi\big(u_n(s)\big)\,\mathrm{d}s\right), \quad r \leqslant r_n^*.$$

Integrating once more and using the fact that the integrand is positive and that ϕ^{-1} is monotone, we obtain

$$\|u_n\| \leqslant \int_{R_1}^{R_2} \left(\phi^{-1}\left(\frac{1}{p(r)} \int_{R_1}^{R_2} \epsilon p(s)\phi\big(u_n(s)\big)\,\mathrm{d}s\right)\right)\mathrm{d}t,$$

hence for all large n

$$\|u_n\| \leqslant (R_2 - R_1)\phi^{-1}\big(c_1\epsilon\phi(\epsilon_n)\big),$$

or

$$\phi\left(\frac{\epsilon_n}{R_2 - R_1}\right) \leqslant c_1\epsilon\phi(\epsilon_n).$$

Therefore, if $R_2 - R_1 \leqslant 1$ we obtain a contradiction for ϵ small, on the other hand if $R_2 - R_1 > 1$, then we obtain a contradiction via (3.3). □

The following proposition from [74] cites conditions that guarantee the validity of the second part of Corollary 3.4:

PROPOSITION 3.6. *Assume that* $f : \mathbb{R} \to \mathbb{R}$ *is continuous and let there exist a constant* $c \geqslant 0$ *such that*

$$f(u) + c\phi(u) \geqslant 0, \quad u \geqslant 0.$$

Further, let ϕ *be an odd increasing homeomorphism of* $\mathbb{R}$ *that satisfies* (3.2) *and* (3.6). *Then there exists a positive number* M *such that any solution* u *of*

$$\begin{cases} \big(p\phi(u')\big)' - cp\phi(u) + cp\phi\big(|u|\big) + pf\big(|u|\big) \leqslant 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0 \end{cases}$$

satisfies $\|u\| \neq M$.

In order to apply Corollary 3.4 to Equation (3.4), we need one further result from [74]:

PROPOSITION 3.7. *Assume the hypotheses of Proposition* 3.6 *and let* M *be a constant whose existence is guaranteed there. Then there exists* $h \in E$, $h \geqslant 0$ *such that* $\|u\| > M$ *for any solution of*

$$\begin{cases} \big(p\phi(u')\big)' - cp\phi(u) + cp\phi\big(|u|\big) + pf\big(|u|\big) + ph = 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0. \end{cases}$$

Combining the above propositions and using Corollary 3.4 we have the following theorem.

THEOREM 3.8. *Assume that* $f : \mathbb{R} \to \mathbb{R}$ *is continuous. Further, let* ϕ *be an odd increasing homeomorphism of* $\mathbb{R}$ *which satisfies* (3.2)*, and let the following conditions hold*:

$$-\infty < \liminf_{u \to 0} \frac{f(u)}{\phi(u)} \leqslant \limsup_{u \to 0} \frac{f(u)}{\phi(u)} \leqslant 0,$$

$$\lim_{u \to \infty} \frac{f(u)}{\phi(u)} = \infty.$$

Then the boundary value problem (3.4) *has a positive solution.*

Note that the existence of a constant $c \geqslant 0$ such that

$$f(u) + c\phi(u) \geqslant 0, \quad u \geqslant 0$$

follows from the assumptions of the theorem.

If $f : \mathbb{R} \to \mathbb{R}$ is nonnegative, then the following theorem applies.

THEOREM 3.9. *Assume that* $f : \mathbb{R} \to \mathbb{R}$ *is continuous and nonnegative. Further, let* ϕ *be an odd increasing homeomorphism of* $\mathbb{R}$ *such that* $\frac{\phi(u)}{u}$ *is nondecreasing on* $\mathbb{R}^+$ *and satisfies* (3.2); *also let the following conditions hold*:

$$\liminf_{s \to 0} \frac{F(s)}{\Phi(s)} = 0,$$

$$\lim_{s \to \infty} \frac{f(s)}{\phi(s)} = \infty.$$

Then the boundary value problem (3.4) *has a positive solution.*

We note that in the case $\phi(s) = |s|^{p-2}s$ the conditions above hold if and only if $p \geqslant 2$.

3.3. *Positone problems*

In this section we are interested in the existence and multiplicities of positive solutions of the boundary value problem

$$\begin{cases} \left(\phi(u')\right)' + \lambda f(r,u) = 0, & R_1 < r < R_2, \\ u(R_1) = u(R_2) = 0, \end{cases} \tag{3.13}$$

with $f : [R_1, R_2] \times [0, \infty) \to (0, \infty)$ continuous. As usual, we assume ϕ is an odd increasing homeomorphism on $\mathbb{R}$ such that

$$\limsup_{x\to\infty} \frac{\phi(\sigma x)}{\phi(x)} < \infty, \quad \text{for all } \sigma > 0. \tag{3.14}$$

Concerning f we also assume there exists $[c, d] \subset (R_1, R_2)$, $c < d$, such that

$$\lim_{s\to\infty} \frac{f(r,s)}{\phi(s)} = \infty, \quad \text{uniformly for } r \in [c, d]. \tag{3.15}$$

The main result in this section is:

THEOREM 3.10. *Let* (3.14) *and* (3.15) *hold. Then there exists a positive number* λ^* *such that the problem* (3.13) *has at least two positive solutions for* $0 < \lambda < \lambda^*$*, at least one for* $\lambda = \lambda^*$ *and none for* $\lambda > \lambda^*$.

Note that in the special case where $\phi(u') = u'$, Theorem 3.10 is classical (see, e.g., [30,100]). Related results for the case $\phi(u') = |u'|^{p-2}u'$ can be found in [33,62], and the references in these papers.

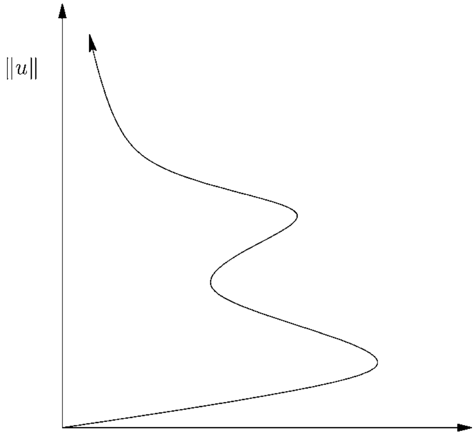


Fig. 3. Continua suggested by Theorem 3.10.

We remark here that the above formulation includes the study of positive radial solutions of equations like

$$\operatorname{div}\big(|\nabla u|^{p-2}\nabla u\big) + \lambda g\big(|x|, u\big) = 0, \quad R_1 < |x| < R_2$$

on annular domains, via the change of variables

$$t = \left(\frac{|N-p|}{(p-1)r}\right)^{\frac{N-p}{p-1}}, \quad r = |x|,$$

and

$$f(t,u) = \left(\frac{(p-1)r}{|N-p|}\right)^{\frac{p(N-1)}{p-1}} g\left(\frac{|N-p|}{(p-1)} t^{-\frac{p-1}{N-p}}, u\right)$$

and thus, higher-dimensional problems, as well (see [31]).

To prove Theorem 3.10, one employs, in addition to continuation methods, also upper and lower solution methods. These methods are, of course, standard for semilinear equations (see [99,39]) and we refer to [12,31,66], where similar types of theorems for the nonlinear case are presented.

Since we are interested in nonnegative solutions we shall make the convention that $f(r,u) = f(r,0)$ if $u < 0$. To prove Theorem 3.10 we first need a lemma which is a special case of Lemma 3.2:

LEMMA 3.11. *Let $v \in C[R_1, R_2]$ with $v \leqslant 0$ and let u satisfy*

$$\big(\phi(u')\big)' = v,$$

$$u(R_1) = u(R_2) = 0.$$

Then

$$u(t) \geqslant \|u\| p(t), \quad t \in [R_1, R_2],$$

where

$$p(t) = \frac{\min(t - R_1, R_2 - t)}{R_2 - R_1}.$$

The next sequence of lemmas allows the use of continuation methods:

LEMMA 3.12. *Suppose that $g : [R_1, R_2] \times \mathbb{R}^+ \to \mathbb{R}^+$ is continuous and there exists a positive number M and an interval $[a_1, b_1] \subset (R_1, R_2)$ such that*

$$g(r,s) \geqslant M(\phi(s) + 1), \quad r \in [a_1, a_2],\ s \geqslant 0.$$

Then there exists a positive number $M_0 = M_0(\phi, a_1, b_1)$ *such that the problem*

$$\begin{aligned}&\big(\phi(u')\big)' = -g(r,u),\\ &u(R_1) = 0 = u(R_2)\end{aligned}$$

has no solution whenever $M \geqslant M_0$.

PROOF. Let u be a solution. Then

$$u(t) = \int_{R_1}^{t} \phi^{-1}\left[c - \int_{R_1}^{s} g(\tau, u)\,\mathrm{d}\tau\right]\mathrm{d}s,$$

where $c = \phi(u'(R_1))$. Let $\|u\| = u(t_0),\ t_0 \in [R_1, R_2]$. Then $u'(t_0) = 0$ and hence

$$u(t) = \int_{R_1}^{t} \phi^{-1}\left[\int_{s}^{t_0} g(\tau, u)\,\mathrm{d}\tau\right]\mathrm{d}s.$$

If $t_0 \geqslant \frac{a_1+b_1}{2}$, then

$$\begin{aligned}\|u\| \geqslant u(a_1) &> \int_{R_1}^{a_1} \phi^{-1}\left[M \int_{a_1}^{\frac{a_1+b_1}{2}} \big(\phi(u)+1\big)\right]\\ &> (a_1 - R_1)\phi^{-1}\left[M\frac{(b_1-a_1)}{2}\big[\phi\big(\|u\|\delta\big)+1\big]\right]\end{aligned}$$

where

$$\delta = \min_{a_1 \leqslant t \leqslant b_1} p(t).$$

This implies

$$\phi\left(\frac{\|u\|}{a_1 - R_1}\right) > M\frac{(b_1-a_1)}{2}\big[\phi\big(\|u\|\delta\big)+1\big].$$

If $t_0 \leqslant \frac{a_1+b_1}{2}$, then since

$$u(t) = \int_{t}^{R_2} \phi^{-1}\left[\int_{t_0}^{s} g(\tau, u)\,\mathrm{d}\tau\right]\mathrm{d}s,$$

we deduce

$$\phi\left(\frac{\|u\|}{R_2 - b_1}\right) > \frac{M(b_1-a_1)}{2}\big[\phi\big(\|u\|\delta\big)+1\big].$$

Combining the above, we obtain

$$\phi\big(\gamma\|u\|\big) > \frac{M(b_1 - a_1)}{2}\big[\phi\big(\|u\|\delta\big) + 1\big],$$

where $\gamma = \max(\frac{1}{R_2 - b_1}, \frac{1}{a_1 - R_1})$. Consequently,

$$\|u\| > \frac{1}{\gamma}\phi^{-1}\left[\frac{M(b_1 - a_1)}{2}\right]$$

and

$$\frac{\phi(\gamma\|u\|)}{\phi(\delta\|u\|)} > \frac{M}{2}(b_1 - a_1)$$

a contradiction, if M is sufficiently large. □

It follows from the above proof, that the problem in Lemma 3.12 has no solution u satisfying

$$g\big(t, u(t)\big) \geqslant M\big(\phi\big(u(t)\big) + 1\big), \quad t \in [a_1, a_2],$$

if $M \geqslant M_0$.

These considerations further imply the following result:

THEOREM 3.13. *There exists a positive number* $\bar{\lambda}$ *such that problem* (3.13) *has no solution for* $\lambda > \bar{\lambda}$.

It follows immediately from (3.15) that there exists a constant $\mu > 0$ such that

$$f(r, s) \geqslant \mu(\phi(s) + 1), \quad s \in \mathbb{R}^+,\ c \leqslant r \leqslant d.$$

Hence the result follows from the previous lemma. We also need the following lemma.

LEMMA 3.14. *For each* $\mu > 0$*, there exists a positive constant* C_μ *such that the problem*

$$\begin{cases} \big(\phi(u')\big)' = -\lambda\theta f(r, u) - (1 - \theta)M_0\big(\big|\phi(u)\big| + 1\big), \\ u(R_1) = u(R_2) = 0 \end{cases}$$

with $\lambda \geqslant \mu$*,* $\theta \in [0, 1]$ *and* M_0 *given by Lemma* 3.12*, has no solution satisfying* $\|u\| > C_\mu$.

Now, let Λ be the set of all $\lambda > 0$ such that (3.13) has a solution and let $\lambda^* = \sup \Lambda$.

LEMMA 3.15. $0 < \lambda^* < \infty$ *and* $\lambda^* \in \Lambda$.

PROOF. A function $u \in C[R_1, R_2]$ is a solution if and only if $u = F(\lambda, u)$, where

$$F : [0, \infty) \times C[R_1, R_2] \to C[R_1, R_2]$$

is the completely continuous mapping given by

$$u = F(\lambda, v),$$

with u the solution of

$$\begin{cases} \left(\phi(u')\right)' = -\lambda f(r, v), \\ u(R_1) = u(R_2) = 0. \end{cases}$$

We note that $F(0, v) = 0$, $v \in C[R_1, R_2]$. Hence it follows from the continuation theorem of Leray–Schauder that there exists a solution continuum $\mathcal{C} \subset [0, \infty) \times C[R_1, R_2]$ of solutions of (3.13) which is unbounded in $[0, \infty) \times C[R_1, R_2]$, and thus, (3.13) has a solution for $\lambda > 0$ sufficiently small, and hence $\lambda^* > 0$. By Theorem 3.13, $\lambda^* < \infty$. We verify that $\lambda^* \in \Lambda$. Let $\{\lambda_n\} \subset \Lambda$ be such that $\lambda_n \to \lambda^*$ and let $\{u_n\}$ be the corresponding solutions. We easily see that $\{u_n\}$ is bounded in $C^1[R_1, R_2]$ and hence $\{u_n\}$ has a subsequence converging to $u \in C[R_1, R_2]$. By standard limiting procedures, it follows that u is a solution of (3.13). □

LEMMA 3.16. *Let* $0 < \lambda < \lambda^*$ *and let* u_{λ^*} *be a solution of* (3.13). *Then there exists* $\epsilon_0 > 0$ *such that* $u_{\lambda^*} + \epsilon$, $0 \leqslant \epsilon \leqslant \epsilon_0$, *is an upper solution of* (3.13).

Theorem 3.10 now follows from these lemmas:

PROOF OF THEOREM 3.10. Let $0 < \lambda < \lambda^*$. Since 0 is a lower solution and u_{λ^*} is an upper solution, there exists a minimum solution u_λ of (3.13) with $0 \leqslant u_\lambda \leqslant u_{\lambda^*}$. We next establish the existence of a second solution.

The mapping

$$\lambda \mapsto u_\lambda, \quad 0 \leqslant \lambda \leqslant \lambda^*,$$

where u_λ is the minimal solution of (3.13), is a continuous mapping $[0, \lambda^*] \to C[R_1, R_2]$. Hence

$$\left\{(\lambda, u_\lambda) \colon 0 \leqslant \lambda \leqslant \lambda^*\right\} \subset \mathcal{C},$$

where $\mathcal{C}$ is the continuum in the proof of Lemma 3.15. Using separation results on closed sets in compact metric spaces (i.e., Whyburn's lemma [1]), one may use the arguments used in the above proof to verify that for each $\lambda \in (0, \lambda^*)$ there are at least two solutions on the continuum $\mathcal{C}$. □

We will discuss consequences of these existence theorems for Gelfand type problems in the next section.

4. The Liouville–Gelfand equation — A case study

4.1. *Introduction*

The classical Liouville–Gelfand problem is concerned with positive solutions of the equation

$$\begin{cases} \Delta u + \lambda e^u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{4.1}$$

for $\lambda > 0$. Equation (4.1) arises in many settings, including the study of combustible gas dynamics [11,49,40], astrophysics [20,38], and prescribed curvature problems [8,63,5].

Let us briefly outline the connection with chemical kinetics. Systems such as

$$\begin{cases} (c_i)_t = k_i \Delta c_i + f_i(c_1, \ldots, c_m, T), \\ T_t = k\Delta T + f(c_1, \ldots, c_m, T), \end{cases} \tag{4.2}$$

arise in the theory of chemical reaction kinetics, where the c_i represent various chemical concentrations with diffusion coefficients k_i, T is the temperature, and f_i represent various nonlinear reactions. This model does not include other important effects such as convection. Assuming a constant concentration c_0 over a short time interval and that the temperature does not exceed a base temperature T_0 the equation for T in (4.2) can be reduced to (see, e.g., [52,11,40,96])

$$u_t = k_0 \Delta u + \hat{\lambda} e^{u/(1-\epsilon u)}, \tag{4.3}$$

for an appropriately nondimensionalized temperature u. Here

$$\hat{\lambda} = \frac{a_0 q_0 c_0}{T_0^2} e^{-a_0/T_0},$$

where a_0 is the activation energy of the substance, q_0 is the latent heat of the reaction, and $\epsilon = T_0/a_0$. Of particular interest is the existence of steady state solutions to (4.3):

$$\Delta u + \lambda e^{u/(1-\epsilon u)} = 0, \tag{4.4}$$

where $\lambda = \hat{\lambda}/k_0$. Equation (4.4) is known as the Arrhenius equation. In the limit of infinitely large activation energy (for a fixed temperature T_0) one obtains the Frank-Kamenetskii approximation

$$\Delta u + \lambda e^u = 0, \tag{4.5}$$

which is also known as the Liouville–Gelfand equation. Of particular importance in this setting is the existence of values of λ for which a solution to (4.5) or (4.4) will exist. It is perhaps not surprising that for large λ (4.5) has no solution (corresponding to blow-up in the parabolic problem). On the other hand, the structure of the solution set to (4.5) for small λ is quite surprising. In the next two sections we discuss this structure as well as its extension to the quasilinear case.

4.2. *Solution set for the semilinear case*

In this section we consider the radial version of the Liouville–Gelfand equation (4.1) with Dirichlet boundary conditions:

$$\begin{cases} u'' + \frac{N-1}{r}u' + \lambda e^u = 0, & r \in (0,1), \\ u'(0) = u(1) = 0. \end{cases} \tag{4.6}$$

To build intuition let us consider carefully the case $N = 1$:

$$\begin{cases} u'' + \lambda e^u = 0, & r \in (0,1), \\ u'(0) = u(1) = 0. \end{cases} \tag{4.7}$$

Multiplying by u' and integrating one finds

$$u'(r) = -\sqrt{2\lambda\left(e^\alpha - e^u\right)},$$

where $\alpha = u(0)$ is a free parameter. A second integration yields the implicit solution

$$\tanh^{-1}\left(\sqrt{1 - e^u/e^\alpha}\right) = \frac{r}{2}\sqrt{2\lambda e^\alpha} + C. \tag{4.8}$$

The condition at $r = 0$ implies $C = \tanh^{-1}(0) = 0$. Thus

$$u(r) = \alpha - 2\ln\cosh\left(\frac{r}{2}\sqrt{2\lambda e^\alpha}\right), \quad 0 \leqslant r < 1. \tag{4.9}$$

From this equation we can determine for each α if there exists a value of λ for which the boundary condition $u(1) = 0$ holds. Indeed from (4.8) at $r = 1$ we conclude

$$\lambda = \frac{1}{2e^\alpha}\left[\ln\left(\frac{1 + \sqrt{1 - e^{-\alpha}}}{1 - \sqrt{1 - e^{-\alpha}}}\right)\right]^2. \tag{4.10}$$

For each α, Equation (4.10) clearly defines a value of λ for which (4.6) has a solution. It is somewhat surprising though that the right-hand side of Equation (4.10) is nonmonotone. Indeed, Figure 4 shows that there exists a unique λ for which (4.6) has exactly one solution and for each smaller λ, Equation (4.6) has precisely two solutions. In terms of the physical setting, it can be shown that for a given λ the solution with larger α value is unstable [11].

The case $N = 2$ can be solved by the change of variables $r = e^{-t}$ and $v(t) = u(r) - 2t$ to obtain $v'' + \lambda e^v = 0$, thus $u(r) = \alpha - 2\ln(1 + \lambda/8e^\alpha r^2)$. In this case

$$\lambda = 8\left(e^{\alpha/2} - e^{-2\alpha}\right).$$

A plot of this relation is qualitatively exactly like Figure 4, now with a maximum at $\lambda = 2$. Thus for both $N = 1$ and $N = 2$ the solution set to Equation (4.6) has the same structure,

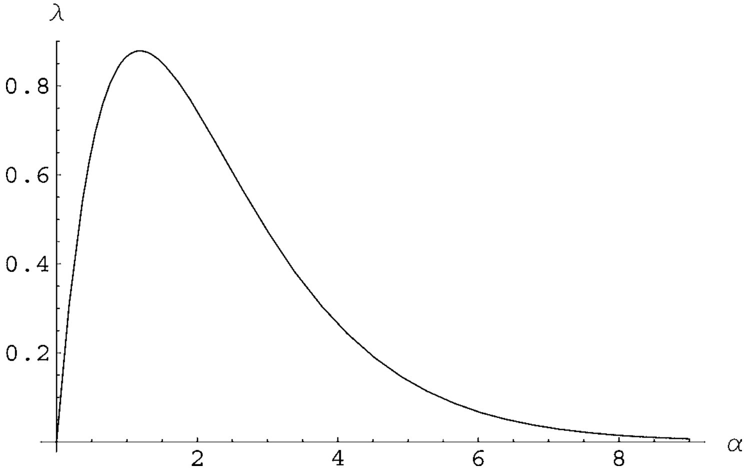


Fig. 4. The dependence of λ on α for (4.10).

there exists a λ^* such that (1.2) has a unique solution for $\lambda = \lambda^*$ and exactly two solutions for each $\lambda \in (0, \lambda^*)$.

Analytic solutions of (4.6) in the physically important case of $N = 3$ are unknown. Based on the results for $N = 1$ and $N = 2$, one might expect a similar solution structure. Historically, Emden [38], Frank-Kamenetskii [40], and Chandrasekhar [20] each numerically integrated (4.6), with Frank-Kamenetskii finding the maximal $\lambda^* \approx 3.32$. However, in [49], Gelfand showed that the structure of the solution set for $N = 3$ is quite different. The key difference is that for $N = 3$, the differential equation in (4.6) has a singular solution $\Phi(r) = -2\ln r$, corresponding to $\lambda = 2$, which also satisfies the boundary condition $u(1) = 0$. Further, as r ranges from 1 to 0, $\Phi(r)$ ranges from 0 to ∞. One can use this structure to transform (4.6) to a state-space where time is represented by $\Phi(r)$. With this motivation, consider the change of variables

$$\begin{cases} s = -\ln r, & 0 < r \leqslant 1, \\ v = \frac{\mathrm{d}u}{\mathrm{d}s}, \\ w = \lambda \mathrm{e}^{-2s}\mathrm{e}^{u}. \end{cases} \tag{4.11}$$

Under this transformation, a solution pair $(\lambda, u_\lambda(r))$ to (4.6) is equivalent to a trajectory $\{(w(s), v(s))\}$ in the phase plane of the system

$$\begin{cases} \frac{\mathrm{d}w}{\mathrm{d}s} = w(v-2), \\ \frac{\mathrm{d}v}{\mathrm{d}s} = v - w, \end{cases} \tag{4.12}$$

with

$$w(0) = \lambda, \tag{4.13}$$

$$v(0) = -u_\lambda'(1), \tag{4.14}$$

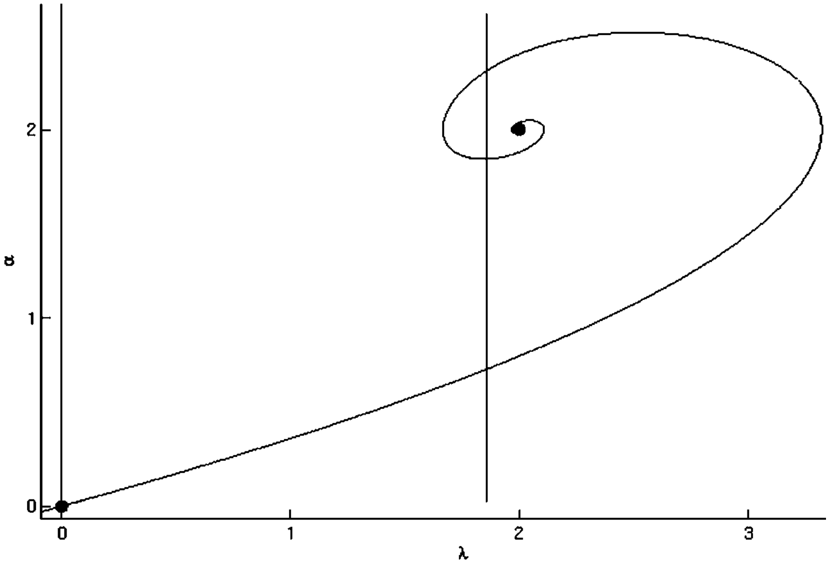


Fig. 5. Heteroclinic orbit for (4.12). The line corresponds to a value of λ for which (4.6) has three solutions.

$$w(\infty) = 0, \tag{4.15}$$

$$v(\infty) = 0. \tag{4.16}$$

An elementary analysis of the system (4.12) shows that there are two equilibrium points $(0, 0)$ and $(2, 2)$, the latter corresponding to the singular solution $\Phi(r)$. Furthermore there is a heteroclinic orbit connecting the spiral node $(2, 2)$ to the origin. Thus each intersection of this orbit with a vertical line $w = \lambda > 0$ corresponds to a solution pair $\{(\lambda, u_\lambda)\}$ of (4.6) (see Figure 5).

Thus, in contrast to the case $N = 1, 2$, for $N = 3$, given any $m \in \mathbb{N}$, there exists values of λ such that (4.6) has exactly m solutions. Moreover, for $\lambda = 2$ it has infinitely many solutions. In the context of Emden and Chandrasekhar's isothermal gas stars, it implies the existence of infinitely many equilibrium positions.

Joseph and Lundgren observed that the change of variables above determines the structure of the solution set to (4.6) for all N. In particular, for $N > 2$, the transformation (4.11) produces the equivalent system

$$\begin{cases} \frac{\mathrm{d}w}{\mathrm{d}s} = w(v-2), \\ \frac{\mathrm{d}v}{\mathrm{d}s} = (N-2)v - w, \end{cases} \tag{4.17}$$

with conditions (4.13)–(4.16). For $2 < N < 10$ the equilibrium point $(2(N-2), 2)$ is a spiral node, thus for this range of N Equation (4.6) has a solution set similar to that of $N = 3$. For $N \geqslant 10$ the critical point becomes a repelling node, which implies that (4.6) has a unique solution for each $\lambda \in (0, 2(N-2))$. Thus we have demonstrated the following theorem:

THEOREM 4.1. *Consider Equation* (4.6). *The following existence results hold*:

- *Case* I: $1 \leqslant N \leqslant 2$. *There exists* $\lambda^* > 0$ *such that* (4.6) *has exactly one solution for* $\lambda = \lambda^*$ *and exactly two solutions for each* $\lambda \in (0, \lambda^*)$.
- *Case* II: $2 < N < 10$. *Equation* (4.6) *has an unbounded continuum of solutions which oscillates around the line* $\lambda = 2(N-2)$, *with the amplitude of oscillations tending to zero, as* $u(0) = \|u\| \to \infty$.
- *Case* III: $N \geqslant 10$. *Equation* (4.6) *has a unique solution for each* $\lambda \in (0, 2(N-2))$ *and no solutions for* $\lambda \geqslant 2(N-2)$. *Moreover,* $\|u\| \to \infty$ *as* $\lambda \to 2(N-2)$.

Finer analysis of the radial solutions to (1.2) may be found in [11,78,82,79]. For results on the Liouville–Gelfand problem on annuli, star-shaped, and more general domains see [11,80,81,107,26,42].

4.3. *Solution set for the quasilinear case*

In this section we discuss extensions of the Liouville–Gelfand problem. The first work in this direction is due to García Azorero and Peral Alonso [44] for the nonlinear diffusive extension:

$$\begin{cases} \Delta_p u + \lambda \mathrm{e}^u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{4.18}$$

for a bounded domain Ω. Using variational methods they prove the following theorem extending the case $p = 2$ (see, e.g., [26,42]):

THEOREM 4.2. *For each* $p > 1$ *there exists a constant* λ_p *such that Equation* (4.18) *has, for* $\lambda \in (0, \lambda_p)$, *at least one solution for* $p < N$ *and at least two solutions for* $p \geqslant N$.

Additional progress for general domains can be found in [6,104].

Studies of (4.18) in the radial case for $p \neq 2$ first appeared in [6] ($1 < p < N$) and [22] ($p = N$). In particular the paper [6] found a "transition" from infinitely many solutions ($p < N < p^2 + 3p/(p-1)$) to unique solutions ($N \geqslant p^2 + 3p/(p-1)$) that extends the Joseph–Lundgren theory. Moreover, in [22] for $p = N$ they find the existence of a $\lambda^* > 0$ such that the radial case of (4.18) has exactly one solution for $\lambda = \lambda^*$ and exactly two solutions for $\lambda \in (0, \lambda^*)$, again extending the Joseph–Lundgren theory.

In [57], the author considers a Monge–Ampère version of the Liouville–Gelfand equation defined by

$$\begin{cases} \det D^2 u + \lambda \mathrm{e}^u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{4.19}$$

on a strictly convex bounded domain Ω. Curiously, unlike the Joseph–Lundgren theory, in the Monge–Ampère case it was shown that there exists a λ^* such that (4.19) has at least two solutions for each $\lambda \in (0, \lambda^*)$, independent of the value of N.

Equation (4.19) is one of a family of equations

$$\begin{cases} S_k\big(D^2u\big) + \lambda \mathrm{e}^u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{4.20}$$

where $S_k(D^2u)$ is the k-Hessian operator discussed in the introduction. The first result for (4.20) ($k > 1$) in the radial case was due to Clément et al. [22] where they considered the case $k = n$ in $\mathbb{R}^{2n}$. They found that in this case (4.20) has the same twofold multiplicity of solutions as in the case $k = 1$, $N = 2$. This motivated the work [58] which determined the precise multiplicity of (4.20) in the radial case. Using the transformation (4.11) with $s = -k \ln r$ one finds the equivalent system:

$$\begin{cases} \frac{\mathrm{d}w}{\mathrm{d}s} = w(v - 2), \\ \frac{\mathrm{d}v}{\mathrm{d}s} = \big[\frac{N-2k}{k^2}\big]v - \frac{1}{k^{k+2}} w v^{1-k}, \end{cases} \tag{4.21}$$

with the additional conditions

$$w(0) = \lambda, \qquad v(0) = -u'(1)/k, \qquad w(\infty) = 0, \qquad v(\infty) = 0.$$

An analysis of the system (4.21) determines the exact multiplicity of solutions to (4.20). Note that we recover the system (4.12) when $k = 1$.

In [59] all of these multiplicity results were extended to the class of quasilinear equations defined by

$$\begin{cases} r^{-\gamma}\big(r^{\alpha}|u'|^{\beta}u'\big)' + \lambda \mathrm{e}^u = 0, & r \in (0, 1), \\ u > 0, & r \in (0, 1), \\ u'(0) = u(1) = 0, \end{cases} \tag{4.22}$$

where the inequalities

$$\alpha \geqslant 0, \tag{4.23}$$

$$\gamma + 1 > \alpha, \tag{4.24}$$

$$\beta + 1 > 0, \tag{4.25}$$

hold. For instance, the following operators are included in this class:

Operator	α	β	γ
Laplacian	$N-1$	0	$N-1$
p-Laplacian ($p > 1$)	$N-1$	$p-2$	$N-1$
k-Hessian	$N-k$	$k-1$	$N-1$

In this setting the transformation

$$\begin{cases} s = \xi \ln r, \\ v = -\frac{\mathrm{d}u}{\mathrm{d}s}, \\ w = \lambda_0 \mathrm{e}^{\delta s} \mathrm{e}^{u}, \end{cases}$$

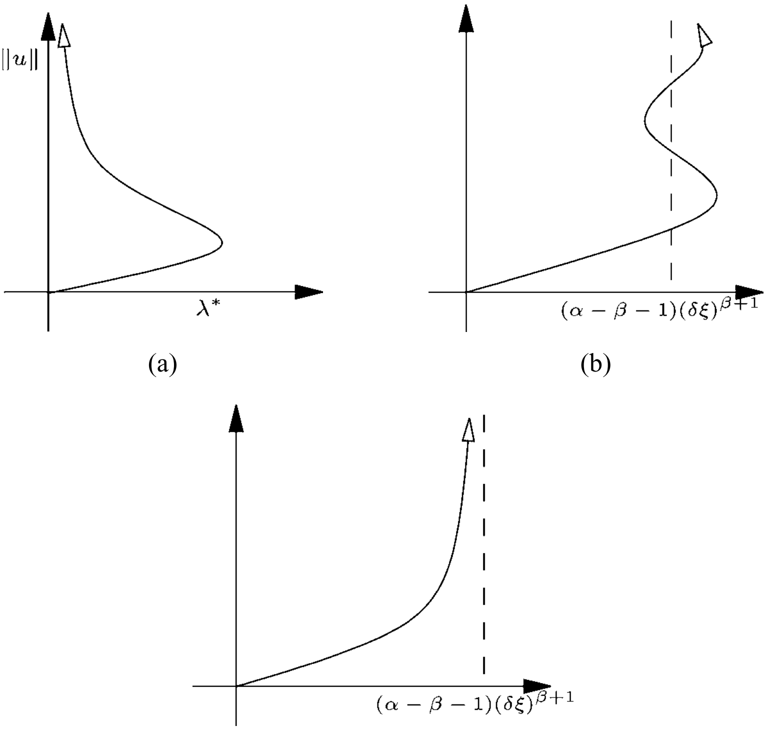
where $\xi = \gamma + 1 - \alpha$ and $\delta = \frac{\gamma+\beta-\alpha+2}{\xi}$ yield the system

$$\begin{cases} \dot{w} = w(\delta - v), \\ \dot{v} = \left(\frac{\beta+1-\alpha}{\xi(\beta+1)}\right) v + \left(\frac{1}{(\beta+1)\xi^{\beta+2}}\right) w v^{-\beta} \end{cases} \tag{4.26}$$

with the additional conditions

$$w(0) = \lambda_0, \qquad v(0) = -u'(1)/\xi, \qquad w(-\infty) = 0, \qquad v(-\infty) = 0.$$

The analysis in [59] yields the following structure theorem for (4.22):



(a) (b)

(c)

Fig. 6. Illustration of Theorem 4.3. (a) $\alpha - \beta - 1 \leqslant 0$, (b) $0 < \alpha - \beta - 1 < \frac{4\delta\xi}{\beta+1}$, (c) $\frac{4\delta\xi}{\beta+1} \leqslant \alpha - \beta - 1$.

THEOREM 4.3. *Consider Equation* (4.22) *with the inequalities* (4.23), (4.24), *and* (4.25) *satisfied. The solution structure to* (4.22) *is as follows* (*see Figure* 6):

- *Case* I: $\alpha - \beta - 1 \leqslant 0$. *There exists a unique* $\lambda^* > 0$ *such that* (4.22) *has exactly one solution for* $\lambda = \lambda^*$ *and exactly two solutions for* $0 < \lambda < \lambda^*$.
- *Case* II: $0 < \alpha - \beta - 1 < \frac{4\delta\xi}{\beta+1}$. *Equation* (4.22) *has a continuum of solutions which oscillates around the line* $\lambda = (\alpha - \beta - 1)(\delta\xi)^{\beta+1}$, *with the amplitude of oscillations tending to zero as* $\|u\| \to \infty$.
- *Case* III: $\alpha - \beta - 1 \geqslant \frac{4\delta\xi}{\beta+1}$. *Equation* (4.22) *has a unique solution for each* $\lambda \in (0, (\alpha - \beta - 1)(\delta\xi)^{\beta+1})$ *and no solutions for* $\lambda \geqslant (\alpha - \beta - 1)(\delta\xi)^{\beta+1}$. *Moreover,* $\|u\| \to \infty$ *as* $\lambda \to (\alpha - \beta - 1)(\delta\xi)^{\beta+1}$.

5. Rellich–Pohozaev identities

In this section we discuss both a nonexistence and a uniqueness result for radial solutions of semilinear elliptic equations defined on N-dimensional balls. These results are based on a well-known theorem due to Pohozaev [89] (see also [77,100]).

The Rellich–Pohozaev identity (cf. [8]) is an important identity (which since has been extended considerably) which has proved useful in establishing that certain nonlinear boundary value problems on starlike domains[1] do not have solutions. The identity is concerned with smooth solutions of problems, like

$$\begin{cases} \Delta u + f(x,u) = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{5.1}$$

where Ω is a starlike domain, which in our case will be assumed to be a ball. Since most of the results are easier to write for partial differential equations, we shall write our equations as partial differential equations, bearing in mind that in the cases of interest positive solutions will be radial solutions and hence solutions of associated ordinary differential equations.

We let

$$F(x,u) = \int_0^u f(x,s)\,\mathrm{d}s$$

and obtain the following identity

$$\begin{aligned} \int_{\partial\Omega} x \cdot |\nabla u|^2 \nu \,\mathrm{d}S = 2N \int_\Omega F(x,u)\,\mathrm{d}x + 2\int_\Omega \nabla F(x,u) \cdot x \,\mathrm{d}x \\ - (N-2)\int_\Omega u f(x,u)\,\mathrm{d}x. \end{aligned} \tag{5.2}$$

We have the following immediate consequence:

---

[1] Ω is star-shaped if there exists $x_0 \in \Omega$ such that $(x - x_0) \cdot \nu \geqslant 0$ for all $x \in \partial\Omega$.

THEOREM 5.1. *Assume that f is independent of x, and Ω is a starlike domain and*

$$2NF(u)-(N-2)uf(u)\leqslant 0,\quad u\geqslant 0.$$

Then (5.1) *has no positive solution.*

A consequence of this theorem when $f(u)=|u|^{p-1}u,\ p\geqslant 1$, is the following result:

COROLLARY 5.2. *Assume that*

$$\frac{N+2}{N-2}\leqslant p,$$

then the problem

$$\begin{cases}\left(r^{N-1}u'\right)'+r^{N-1}f(u)=0, & 0<r<R,\\ u'(0)=u(R)=0,\end{cases}\tag{5.3}$$

has no positive solutions.

Refinements of the original Pohozaev argument for proving the above identity (see [77, 100]) also give uniqueness results for positive solutions of problems like the above containing a parameter. Namely we have the following:

THEOREM 5.3. *Assume that $N\geqslant 3$, that*

$$f(u)>0,\qquad f'(u)>0,\quad u\geqslant 0$$

and there exists $\alpha>0$, such that

$$\sup_{u\geqslant\alpha}\left\{\frac{2N}{N-2}\right\}\left\{\frac{F(u)}{uf(u)}\right\}<1.\tag{5.4}$$

Then

$$\begin{cases}\left(r^{N-1}u'\right)'+\lambda r^{N-1}f(u)=0, & 0<r<R,\\ u'(0)=u(R)=0,\end{cases}\tag{5.5}$$

has a unique positive solution for $\lambda\geqslant 0$ small.

In particular, we can apply the above result for the case that $f(u)=\mathrm{e}^u$ and obtain that for small values of λ the solution found earlier is the unique positive solution. Note that the earlier results also tell us that the condition that $N\geqslant 3$ is needed.

5.1. *More general equations*

Nonexistence results also exist for the class of equations defined in the introduction, i.e.,

$$\begin{cases} r^{-\gamma}\left(r^{\alpha}|u'|^{\beta}u'\right)' = f(r,u), & r \in (0,R), \\ u > 0, & r \in (0,R), \\ u'(0) = u(R) = 0, \end{cases} \tag{5.6}$$

In 1985 Pucci and Serrin [91] extended Pohozaev's identity (5.2) to a larger class of variational equations. Let $L = L(p,z,x)$ denote a Lagrangian which is C^2 on the domain $\mathbb{R}^N \times \mathbb{R} \times \overline{\Omega}$. Smooth critical points of the associated "energy" functional satisfy the Euler–Lagrange equation

$$-\sum_{i=1}^{N}\left(L_{p_i}(Du,u,x)\right)_{x_i} + L_z(Du,u,x) = 0, \quad \text{in } \Omega. \tag{5.7}$$

We assume without loss of generality that $L(0,0,x) = 0$ in Ω. The main identity of Pucci-Serrin is due to the following proposition:

Proposition 5.4. *Let $u \in C^2(\Omega)$ be a solution of the Euler–Lagrange equation* (5.7), *and let a and $\vec{h}$ be, respectively, scalar and vector valued functions of class $C^1(\Omega)$. Then the following relation holds in Ω*:

$$\begin{aligned} &\frac{\partial}{\partial x_i}\left[\vec{h}_i L(Du,u,x) - \vec{h}_j \frac{\partial u}{\partial x_j} L_{p_i}(Du,u,x) - auL_{p_i}(Du,u,x)\right] \\ &\quad = \frac{\partial \vec{h}_i}{\partial x_i} L(Du,u,x) + \vec{h}_i L_{x_i}(Du,u,x) - \left(\frac{\partial u}{\partial x_j}\frac{\partial \vec{h}_j}{\partial x_i} + u\frac{\partial a}{\partial x_i}\right) L_{p_i}(Du,u,x) \\ &\qquad - a\left(\frac{\partial u}{\partial x_i} L_{p_i}(Du,u,x) + uL_z(Du,u,x)\right), \end{aligned} \tag{5.8}$$

where repeated indices i and j are to be summed from 1 *to N.*

The proof is obtained by direct computation, using (5.7).

If $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ solves (5.7) with $u = 0$ on $\partial\Omega$ then $u_{x_i} = (\partial u/\partial \nu)\nu_i$ on $\partial\Omega$ so

$$\vec{h}_j \frac{\partial u}{\partial x_j} L_{p_i}(Du,u,x)\nu_i = \frac{\partial u}{\partial x_i} L_{p_i}(Du,u,x)\vec{h}_j \nu_j \quad \text{on } \partial\Omega. \tag{5.9}$$

Integrating (5.8) over Ω, applying (5.9), $u = 0$ on $\partial\Omega$, and the divergence theorem one obtains the fundamental identity

$$\int_{\partial\Omega}\left[L(Du,0,x) - \frac{\partial u}{\partial x_i} L_{p_i}(Du,0,x)\right](\vec{h}\cdot\nu)\,\mathrm{d}s$$

$$= \int_{\Omega} L(Du, u, x) \operatorname{div} \vec{h} + \vec{h}_i L_{x_i}(Du, u, x)$$
$$- \left(\frac{\partial u}{\partial x_j} \frac{\partial \vec{h}_j}{\partial x_i} + u \frac{\partial a}{\partial x_i} \right) L_{p_i}(Du, u, x)$$
$$- a \left(\frac{\partial u}{\partial x_i} L_{p_i}(Du, u, x) + u L_z(Du, u, x) \right) \mathrm{d}x. \tag{5.10}$$

For example, if $L(p, z) = \frac{1}{2}|p|^2 - F(z)$, $\vec{h} = x$, and a is constant, then (5.10) reduces to

$$-\int_{\partial\Omega} \frac{1}{2} |Du|^2 (x \cdot \nu) \,\mathrm{d}s = \int_{\Omega} \left[\frac{N}{2} - 1 - a \right] |Du|^2 - NF(u) + auf(u) \,\mathrm{d}x. \tag{5.11}$$

The choice of $a(x) = (N-2)/2$ makes the $|Du|^2$ term vanish and reduces (5.11) to the Pohozaev identity (5.2). However, identity (5.10) is applicable to a much larger class of equations. For instance, for the quasilinear equation

$$\begin{cases} \Delta_p u + f(u) = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{5.12}$$

with associated Lagrangian $L(Du, u) = \frac{1}{p}|Du|^p - F(u)$, the choice of $\vec{h} = x$ and constant a yields

$$-\int_{\partial\Omega} \frac{1}{p} |Du|^p (x \cdot \nu) \,\mathrm{d}s = \int_{\Omega} \left[\frac{N}{p} - 1 - a \right] |Du|^p - NF(u) + auf(u) \,\mathrm{d}x. \tag{5.13}$$

Now we see the choice of $a = (N-p)/p$ implies

$$-\int_{\partial\Omega} \frac{1}{p} |Du|^p (x \cdot \nu) \,\mathrm{d}s = \int_{\Omega} \left[\frac{N-p}{p} \right] uf(u) - NF(u) \,\mathrm{d}x, \tag{5.14}$$

from which an appropriate nonexistence result can be stated. To determine the critical exponent we choose $f(u) = |u|^{q-1}u$ and find (5.12) has no nontrivial solutions when $p < n$ and

$$q > \frac{Np}{N-p} - 1 = \frac{(p-1)N + p}{N-p}.$$

Note that $p^* = Np/(N-p)$ is the Sobolev exponent, corresponding to the loss of compactness for the continuous embedding $W^{1,p}(\Omega) \subset L^q(\Omega)$. Many further applications of (5.10) may be found in [91].

Solutions to the k-Hessian equation

$$\begin{cases} S_k\big(D^2u\big) = f(x,u), & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{5.15}$$

correspond to critical points of the functional

$$I_k[u] = -\frac{1}{k+1}\int_\Omega u S_k\big(D^2u\big)\,\mathrm{d}x + \int_\Omega F(x,u)\,\mathrm{d}x, \tag{5.16}$$

where $F(x,u) = \int_0^u f(x,s)\,\mathrm{d}s$. However, Proposition 5.8 does not directly apply to (5.16) since the Lagrangian contains higher order terms, and one needs to derive an appropriate higher order analog of (5.8).

The Euler–Lagrange equation associated with the Lagrangian $L = L(D^2u, Du, u, x) = L(r_{ij}, p_i, z, x)$, where $r_{ij} = r_{ji}$ is

$$\begin{aligned} &\sum_{i,j=1}^N \frac{\partial^2}{\partial x_i \partial x_j} L_{r_{ij}}\big(D^2u, Du, u, x\big) - \sum_{i=1}^N \big(L_{p_i}\big(D^2u, Du, u, x\big)\big)_{x_i} \\ &\quad + L_z\big(D^2u, Du, u, x\big) = 0. \end{aligned} \tag{5.17}$$

In this case the fundamental identity (simplified for our purposes) takes the form (see Equation (29) in [91])

Proposition 5.5. *Let $u \in C^4(\Omega)$ be a solution to the Euler–Lagrange equation* (5.17) *with $L_{p_i} = 0$ and $a \in C^2(\Omega)$ a scalar function. Then*

$$\begin{aligned} &\frac{\partial}{\partial x_i}\left[x_i L + \left(x_l \frac{\partial u}{\partial x_l} + au\right)\frac{\partial L_{r_{ij}}}{\partial x_j} - \frac{\partial}{\partial x_j}\left(x_l \frac{\partial u}{\partial x_l} + au\right) L_{r_{ij}}\right] \\ &\quad = NL + x_i L_{x_i} - auL_z - (a+2)\frac{\partial^2 u}{\partial x_i \partial x_j} L_{r_{ij}}. \end{aligned} \tag{5.18}$$

This identity can be used to determine the critical exponent associated to the operator S_k [109]. For simplicity we assume $F = F(z)$ (e.g., $f(u) = |u|^p$).

Theorem 5.6. *Let Ω be a smooth domain which is star-shaped with respect to the origin. Assume $f : (-\infty, 0] \to [0, \infty)$ is smooth, with $f(s) > 0$ for $s < 0$ and $f(0) = 0$. Then*

$$\begin{cases} S_k(D^2u) = f(u), & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{5.19}$$

has no nontrivial solutions in $C^4(\Omega) \cap C^1(\overline{\Omega})$ when

$$NF(u) - \frac{N-2k}{k+1} u f(u) > 0, \quad \text{for } u < 0. \tag{5.20}$$

PROOF. Applying Proposition 5.5 to the Lagrangian $L = \frac{-zS_k(r_{ij})}{k+1} + F(z)$ with $a = (N - 2k)/(k+1)$ and integrating, one obtains

$$-\frac{1}{k+1}\int_{\partial\Omega}\left[x_l u_{x_l} u_{x_j} S^{ij}\left(D^2 u\right)\right]\nu_i \,\mathrm{d}s = \int_{\Omega}\left(NF(u) - \frac{N-2k}{k+1}uf(u)\right)\mathrm{d}x, \tag{5.21}$$

which simplifies to

$$-\frac{1}{k+1}\int_{\partial\Omega}(x\cdot\nu)|Du|^2 S^{ij}\left(D^2 u\right)\nu_i\nu_j \,\mathrm{d}s = \int_{\Omega}\left(NF(u) - \frac{N-2k}{k+1}uf(u)\right)\mathrm{d}x. \tag{5.22}$$

For $u < 0$ the operator S_k is elliptic [109], thus $S^{ij}(D^2u)\nu_i\nu_j > 0$. Hence the left-hand side of (5.22) is nonpositive and the result follows. □

Note that when $k = 1$, (5.20) is equivalent to the Pohozaev criterion (5.2). If $f(u) = (-u)^p$ then (5.20) reduces to

$$\frac{N-2k}{k+1} > \frac{N}{p+1}. \tag{5.23}$$

If $k \geqslant N/2$, then (5.23) cannot hold and we obtain no a priori obstructions to solution from this method. On the other hand, when $k < N/2$, then (5.23) is true when $p \geqslant \frac{(N+2)k}{N-2k}$. Thus when $k < N/2$ the *critical exponent* $\gamma(k)$ for S_k is defined by

$$\gamma(k) = \frac{(N+2)k}{N-2k}. \tag{5.24}$$

Complementary existence results for radially symmetric solutions for subcritical exponents (and for all exponents when $k \geqslant N/2$) are proved in [109], thus one can extend $\gamma(k)$ to all k via

$$\gamma(k) = \begin{cases} \infty, & k > N/2, \\ \frac{(N+2)k}{N-2k}, & k < N/2. \end{cases} \tag{5.25}$$

In particular, there is no critical exponent for the Monge–Ampère operator. Heuristically, operators "closer" to the Laplace operator have critical exponents, while operators "closer" to Monge–Ampère do not. Note that when $p = k$ one has an eigenvalue problem (see, e.g., [70,110,57]).

5.2. *Critical dimensions*

In 1983 Brezis and Nirenberg [16] observed that lower order perturbations to elliptic equations involving critical exponents recover the lost compactness. More precisely, they proved the equation

$$\begin{cases} \Delta u + u^{\frac{N+2}{N-2}} + \lambda u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{5.26}$$

has a positive solution if $0 < \lambda < \lambda_1$ and $N \geqslant 4$, where λ_1 is the principal eigenvalue for $-\Delta$ on $H_0^1(\Omega)$. Surprisingly, for the case $N = 3$ they observed that there exists $\lambda^* > 0$ such that (5.26) has a solution for $\lambda \in (\lambda^*, \lambda_1)$ and no solution for $\lambda \in (0, \lambda^*)$. If Ω is a ball, then $\lambda^* = \lambda_1/4$. In this context the dimension $N = 3$ is called a *critical dimension.*

Since both Δ_p and S_k have critical exponents (when $p < N$ and $k < N/2$, respectively), it is natural to ask if results similar to the Brezis–Nirenberg result exist for these operators. Several authors have answered this question affirmatively. Rather than treat Δ_p and S_k separately, Clément et al. [22] consider the equation

$$\begin{cases} \left(r^\alpha |u'|^\beta u'\right)' = r^\gamma |u|^{q-2} u, & r \in (0, R), \\ u > 0, & r \in (0, R), \\ u'(0) = u(R) = 0, \end{cases} \tag{5.27}$$

and the perturbed form

$$\begin{cases} \left(r^\alpha |u'|^\beta u'\right)' = r^\gamma |u|^{q-2} u + \lambda r^\delta |u|^\beta u, & r \in (0, R), \\ u > 0, & r \in (0, R), \\ u'(0) = u(R) = 0, \end{cases} \tag{5.28}$$

for various values of exponents α, β, δ and γ. See the table in Section 4.1 for the relevant values of constants for (5.15) or (5.12).

The critical exponent associated with (5.27) is

$$q^* = \frac{(\gamma+1)(\beta+2)}{\alpha - \beta - 1}. \tag{5.29}$$

For the p-Laplacian, $q^* = \frac{Np}{N-p}$ and for S_k, $q^* = \frac{N(k+1)}{N-2k}$.

Concerning the parameters involved, we assume the following inequalities hold:

$$q - 1 > \beta + 1 > 0, \qquad \gamma + 1 > \alpha - \beta - 1, \tag{5.30}$$

$$\alpha - \beta - 1 > 0, \qquad \delta + 1 \geqslant \alpha - \beta - 1, \tag{5.31}$$

$$\gamma, \delta > \alpha - 1, \tag{5.32}$$

$$\alpha - \beta - 2 < \delta. \tag{5.33}$$

When applied to S_k (respectively Δ_p) these inequalities imply $q > k+1$ (respectively $q > p$) and $k < N/2$ (respectively $N < p$), i.e., one is in the realm of critical exponents.

The main theorems concerning (5.28) is as follows:

THEOREM 5.7. *Assume* (5.30), (5.31), (5.32), (5.33) *hold. If* $\lambda \leqslant 0$ *and* $q = q^*$, *then* (5.28) *has no solution.*

THEOREM 5.8. *Assume* (5.30), (5.31), (5.32), (5.33), $\beta \geqslant 0$ *and* $q = q^*$. *If*

$$(\beta+1)(\delta+1) - (\alpha-\beta-1)(\beta+2) > 0, \tag{5.34}$$

then there exists $\lambda^* > 0$ *such that* (5.28) *has no solution for* $\lambda \in (0, \lambda^*)$.

In the case of the operators S_k and Δ_p, their parameters satisfying (5.34) correspond to certain values of the dimension N, called *critical dimensions* by Pucci and Serrin [92]. For the p-Laplace operator, (5.34) corresponds to $N < p^2$, thus the critical dimensions for Δ_p are those n with $p < N < p^2$. Note that for the Laplacian $p = 2$ and we obtain $2 < N < 4$, thus the only critical dimension is $N = 3$, as observed by Brezis and Nirenberg. For the k-Hessian the critical dimensions are those N with $2k < N < 2k(k+1)$.

The proofs are based on the following identity of Pohozaev–Pucci–Serrin type [22]:

PROPOSITION 5.9. *Let* $a, b \in C^1[0,\infty)$. *If* $u \in C^2(0,\infty) \cap C^1[0,\infty)$ *solves*

$$-\left(r^\alpha |u'|^\beta u'\right)' = f(r,u) \quad \text{in } (0,\infty), \tag{5.35}$$

then for $R > 0$ *we have*

$$\begin{aligned}
&\left[-r^\alpha u'|u'|^\beta\left(au + \frac{\beta+1}{\beta+2}bu'\right)\right]_{r=R} + \int_0^R r^\alpha a' u u' |u'|^\beta \,\mathrm{d}r \\
&\qquad + \int_0^R r^\alpha\left(a + \frac{\beta+1}{\beta+2}b' - \frac{\alpha}{\beta+2}\frac{b}{r}\right)|u'|^{\beta+2}\,\mathrm{d}r \qquad (5.36)\\
&\quad = \left[bF(r,u)\right]_{r=R} + \int_0^R auf(r,u) - bF_r(r,u) - b'F(r,u)\,\mathrm{d}r. \qquad (5.37)
\end{aligned}$$

6. Problems linear at infinity

In Section 2 we discussed several bifurcation results for boundary value problems which had a trivial solution. The bifurcation was detected by writing the problems as an equivalent operator equation which was of the form

$$u - F(\lambda, u) = 0,$$

where

$$F:\mathbb{R}\times E\to E$$

is a completely continuous operator and E is the Banach space of continuous functions on $[0, R]$. For the problems to follow such an equivalence will also be the case, except that F will satisfy the following conditions: There exist $a < b$ such solutions of

$$u - F(a,u) = 0 = u - F(b,u)$$

are a priori bounded and

$$\deg_{\mathrm{LS}}\bigl(id - F(a,\cdot), B(0,r), 0\bigr) \neq \deg_{\mathrm{LS}}\bigl(id - F(b,\cdot), B(0,r), 0\bigr),$$

where the latter condition is to hold for all r, sufficiently large. Under these assumptions we obtain the following *bifurcation from infinity* result (see, e.g., [93,101]):

PROPOSITION 6.1. *There exists a continuum $\mathcal{C}$ of solutions of $u - F(\lambda,u) = 0$ which bifurcates from infinity in (a,b).*

The above means that the continuum $\mathcal{C}$ has the property that there exists a sequence $\{(\lambda_n,u_n)\}\subset\mathcal{C}$, such that $\{\lambda_n\}\subset(a,b)$ and $\|u_n\|\to\infty$.

Here we rely mainly on the work [97,98,85–87]. Consider the boundary value problem

$$\begin{cases} \Delta u + \lambda f(u) + h(x) = 0, & x\in\Omega,\\ u = 0, & x\in\partial\Omega,\\ u > 0, & x\in\Omega, \end{cases} \tag{6.1}$$

where f satisfies

$$f(s) = s + g(s),\quad g(s) = o\bigl(|s|\bigr),\ \text{as } |s|\to\infty. \tag{6.2}$$

Problem (6.1) is equivalent to an operator equation

$$u - \lambda F(u) - T(h) = 0 \tag{6.3}$$

in $\mathbb{R}\times E$ where E is the Banach space $C[0, R]$ with the usual norm, and both $F : E\to E$ and $T : E\to E$ are completely continuous.

Using the fact that λ_1, the principal eigenvalue of $-\Delta$ on the Sobolev space $H_0^1(B_R)$, is simple and the above Proposition 6.1, we obtain the following corollary.

COROLLARY 6.2. *There exist continua $\mathcal{C}^{\pm}$ of solutions of* (6.1) *bifurcating from infinity at λ_1.*

It is these continua which we shall next describe for several cases of bounded nonlinear terms g.

6.1. *Landesman–Lazer results*

Regarding the function g, we now introduce the following quantities:

$$\gamma_1 = \limsup_{t\to\infty} g(t),$$
$$\gamma_2 = \liminf_{t\to-\infty} g(t),$$
$$\gamma_3 = \limsup_{t\to-\infty} g(t),$$
$$\gamma_4 = \liminf_{t\to\infty} g(t),$$

and we have the following results.

THEOREM 6.3. *Assume that*

$$\gamma_1 < 0 < \gamma_2.$$

Then there exists a neighborhood I of λ_1, such that problem (6.1) *has*

- *at least one solution for $\lambda \in I$, $\lambda \geqslant \lambda_1$,*
- *at least three solutions for $\lambda \in I$, $\lambda < \lambda_1$.*

THEOREM 6.4. *Assume that*

$$\gamma_3 < 0 < \gamma_4.$$

Then there exists a neighborhood I of λ_1, such that problem (6.1) *has*

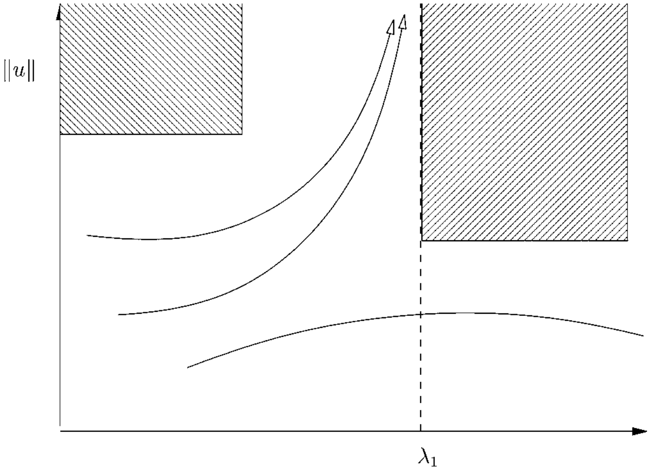


Fig. 7. Illustration of Theorem 6.3. A similar reflected picture holds for Theorem 6.4.

- *at least one solution for* $\lambda \in I$, $\lambda \leqslant \lambda_1$,
- *at least three solutions for* $\lambda \in I$, $\lambda > \lambda_1$.

These results are proved using the existence of the continua above and analyzing their location with respect to the hyperplane $\lambda = \lambda_1$. This may be accomplished by considering the asymptotic behavior of integrals of the form

$$\int_{B_R} g(t\phi + w)\phi \,\mathrm{d}x,$$

with ϕ a principal eigenfunction of $-\Delta$ and w suitably chosen in the orthogonal complement of ϕ. We refer to [75,76,100].

6.2. *Nonlinear terms which oscillate*

We next discuss some multiplicity results from [97,98] regarding the continua which bifurcate from infinity for nonlinear terms g which are bounded and oscillatory. Letting ϕ denote a positive eigenfunction of $-\Delta$ corresponding to λ_1, we assume that

$$\int_\Omega h\phi \,\mathrm{d}x = 0.$$

Let $(\lambda, u) \in \mathcal{C}^+$ be a solution of (6.1), then we obtain

$$\lambda \int_\Omega g(u)\phi \,\mathrm{d}x = \|u\|(\lambda_1 - \lambda) \int_\Omega \frac{u}{\|u\|}\phi \,\mathrm{d}x.$$

Letting $\mu = \|u\|(\lambda_1 - \lambda)$, the following results on the asymptotics of the solution branches hold:

PROPOSITION 6.5. *Assume* $\int_0^\infty g(s)s \,\mathrm{d}s$ *exists, then*

$$\operatorname{sgn}\mu = \operatorname{sgn}\int_0^\infty g(s)s \,\mathrm{d}s, \quad \textit{for } \mu \textit{ large}$$

and

$$\mu\|u\|^2 \to \frac{d(B_R)}{\int_{B_R}\phi^2 \,\mathrm{d}x}\int_0^\infty g(s)s \,\mathrm{d}s,$$

where

$$d(B_R) = \int_{\partial B_R} \frac{1}{|\nabla\phi|} \,\mathrm{d}S.$$

PROPOSITION 6.6. *Assume that*

$$G(\beta) = \int_0^\beta g(s)s\,\mathrm{d}s$$

is bounded and

$$M = \lim_{\beta\to\infty} \frac{1}{\beta} \int_0^\beta G(s)\,\mathrm{d}s$$

exists. Then

- *If* $N = 2$ *then*

  $$\mu\|u\|^2 = \bigl(G\bigl(\|u\|\bigr) - M\bigr)c(B_R) + Md(B_R) + R\bigl(\|u\|\bigr),$$

  with

  $$c(B_R) = \frac{N\omega_N}{2\sqrt{\det A}}, \quad A = -\frac{1}{2}D^2\phi(0).$$

- *If* $N > 2$, *then*

  $$\|u\|^2\mu = Md(B_R) + R\bigl(\|u\|\bigr),$$

  with $R(\alpha) \to 0$ *as* $\alpha \to \infty$.

Concerning the constants $c(B_R)$ and $d(B_R)$ one has the following relationship

$$d(B_R) \geqslant c(B_R)\frac{2\lambda_1^{(N-1)/2}}{N^{N/2}}. \tag{6.4}$$

This relationship will be important in the result below, where interesting phenomena arise which are different for different dimension. Concerning the nonlinear term g, we shall assume that

$$g(t+T) = g(t), \quad \int_0^T g\,\mathrm{d}t = 0$$

and

$$g = g_1' = g_2'' = \cdots, \quad \int_0^T g_i\,\mathrm{d}t = 0.$$

We have the following theorem:

THEOREM 6.7. *Let* $\{(u, \lambda)\}$ *be the solution continuum of* (6.1) *described above. The following hold*:

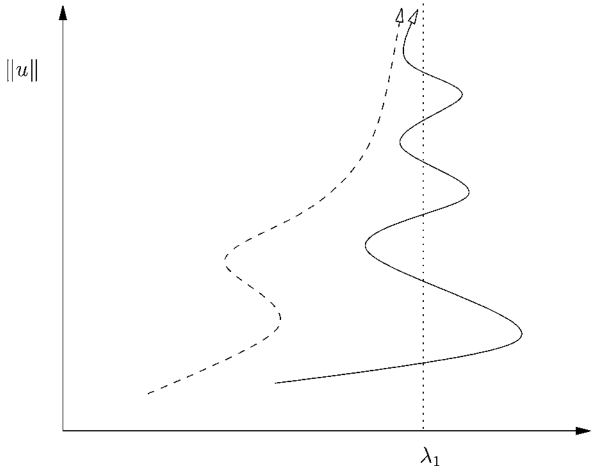


Fig. 8. Possible continua of Theorem 6.7. The the dashed line is possible when $N > 4$.

- *For* $1 \leqslant N \leqslant 3$ *there exist infinitely many solutions for* $\lambda = \lambda_1$.
- *For* $N = 4$

$$\|u\|^2(\lambda - \lambda_1) = \big(g_2\big(\|u\|\big)c(B_R) + g_2(0)d(B_R)\big) + \mathrm{o}\big(\|u\|^{-2}\big)\|u\|^2,$$

as $\|u\| \to \infty$. *Hence infinitely many positive solutions exist, whenever*

$$\min_{s\in[0,T]} g_2(s) < -\frac{d(B_R)}{c(B_R)} g_2(0) < \max_{s\in[0,T]} g_2(s).$$

- *For* $N > 4$

$$\|u\|^2(\lambda - \lambda_1) = g_2(0)d(\Omega) + \mathrm{o}\big(\|u\|^{-2}\big)\|u\|^2$$

as $\|u\| \to \infty$.

We thus conclude that C^+ may lie on one side of the hyperplane $\lambda = \lambda_1$, if $N > 4$ and the nonlinear term g satisfies appropriate conditions. The above results are all valid in case the ball B_R is replaced by a convex set Ω with smooth boundary having the property that the eigenfunction ϕ (corresponding to λ_1) has a single critical point (for details see [98,100]).

7. Symmetry breaking

Our motivation for discussing ordinary differential equations of the form considered earlier has been that they may be considered as equations whose solutions are radial solutions of certain problems for partial differential equations which admit radial solutions because of certain intrinsic symmetries in the equation and the underlying domain. Now, it may be the case that all solutions of the problem at hand must be radial solutions (e.g., if the conditions of the Gidas–Ni–Nirenberg results [50] hold), or it may be case that nonradial solutions bifurcate from radial solutions as parameters are varied (see, e.g., [60,82,105]). It may also happen that nonradial solutions exist because minima of certain energy functionals over a function space of radial functions (which define radial solutions of equations at hand) strictly exceed minima of the same functional considered over a larger space (of not necessarily radial functions) and hence define other solutions of the problem at hand. It is this situation which we shall briefly discuss here.

Variational methods, of course, play an important role in the study of problems of the type that have been considered here and such methods have been used to prove the existence of nonradial solutions of Gelfand type problems on annular domains ([44]). Note that by the results in [50] nonradial solutions cannot exist for such problems if the domain is a ball, and hence the existence of nonradial solutions is due to the geometry of the domain.

Let us consider the boundary value problem

$$\begin{cases} u'' + \frac{N-1}{r}u' + \lambda f(u) = 0, & r \in (a,b), \\ u(a) = u(b) = 0 \end{cases} \tag{7.1}$$

solutions of which are radial solutions of

$$\begin{cases} \Delta u + \lambda f(u) = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{7.2}$$

where

$$\Omega = \{x \in \mathbb{R}^N\colon\ 0 < a < |x| = r < b\}.$$

By changing scale, we may assume that $b = 1$, and changing scale further, we also may assume that $\lambda = 1$ and a is a parameter and we consider

$$\begin{cases} \Delta u + f(u) = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega. \end{cases} \tag{7.3}$$

For various classes of nonlinearities f, we established in Section 3 that these problems have positive solutions.

In case $f(u) = u^p,\ p < \frac{N+2}{N-2},\ N \geqslant 3$, it was proved in [16] that (7.3) has nonradial positive solutions for any a, and if $f(u) = -u + u^p,\ p > 1,\ N = 2$, a similar result was established in [24], where it was also shown that the number of positive nonradial solutions increases without bound as $a \to 1$. These results have been extended by Lin [67,68] who established several bifurcation results (which will be summarized below) motivated by the results cited above. There is also the seminal work of Nagasaki and Suzuki [80,82] concerning nonradial solutions of the Gelfand equation on annuli in $\mathbb{R}^2$, and higher dimensions, which we shall also summarize.

Let us characterize the various nonlinear terms considered and then state the pertinent results. It will always be assumed that f is a smooth function and $f(u) > 0$ for $u > 0$. Furthermore f may belong to one of the following classes of functions:

$$\frac{N}{N-2} f(u) \geqslant f'(u)u \geqslant (1+\delta) f(u), \quad u > 0,\ N \geqslant 3, \tag{7.4}$$

where δ is a positive constant;

$$f(u) = u^p, \quad 1 < p < \frac{N+2}{N-2},\ N \geqslant 2; \tag{7.5}$$

$$f(u) \leqslant \begin{cases} cu^p, & 1 < p < \frac{N+2}{N-2},\ N \geqslant 2 \\ \mathrm{e}^{A(u)}, & A(u) = \mathrm{o}\big(u^2\big),\ u \to \infty,\ N = 2. \end{cases} \tag{7.6}$$

For such nonlinearities the following theorem holds:

THEOREM 7.1. *If* $f(u) = \mathrm{o}(u),\ u \to 0$*, then under any of the above conditions positive nonradial solutions of* (7.3) *exist.*

Consider the Gelfand problem

$$\begin{cases} \Delta u + \lambda \mathrm{e}^u = 0, & x \in \Omega, \\ u = 0, & x \in \partial\Omega, \end{cases} \tag{7.7}$$

with

$$\Omega = \big\{x \in \mathbb{R}^2\colon\ 0 < a < |x| = r < 1\big\}.$$

Using polar coordinates $x = (r\cos\theta, r\sin\theta)$ and the rotation

$$T_m x = \left(r\cos\left(\theta + \frac{\pi}{m}\right), r\sin\left(\theta + \frac{\pi}{m}\right)\right),$$

we consider, besides the space $V = H_0^1(\Omega)$, the spaces

$$V_m = \big\{v \in V\colon\ v(T_m x) = v(x)\big\}$$

and we call m the mode of a function v if m is the largest integer such that $v \in V_m$. For the functional

$$M : V \to \mathbb{R}$$

defined by

$$M(v) = \int_\Omega \mathrm{e}^v\,\mathrm{d}x,$$

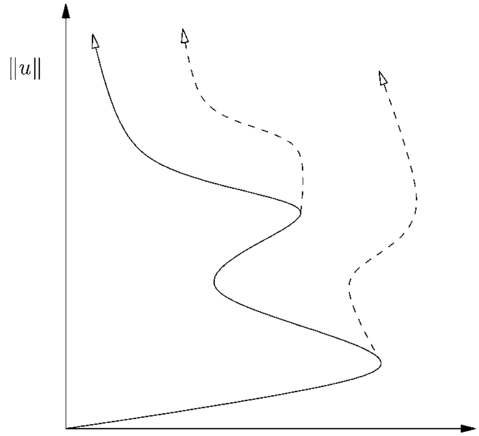


Fig. 9. Possible solution continua for Equation (7.2). The dashed lines represent symmetry breaking bifurcation branches.

the following result was established in [80]:

THEOREM 7.2. *For any positive integer m, there exists a number* $\mu_m > 0$ *such that for any* $\mu > \mu_m$ *problem* (7.7) *has a nonradial solution* (λ, u) *with u of mode m and* $M(u) = \mu$.

This suggests that the solution continua of positive radial solutions of (7.7) whose existence was established in Section 3 undergo secondary bifurcations. However, the proof of the above result uses variational methods and the continua were established using continuation methods based on degree theory. Thus the above suggestion about secondary bifurcations is a suggestion only. On the other hand, in their paper [82] Nagasaki and Suzuki have used a somewhat different method, based on Morse theory and topological degree theory to study symmetry breaking off radial solutions of (7.7) in case $N \geqslant 2$ and they conclude that the solution continuum of radial solutions $\mathcal{C}$ of the Gelfand problem (7.7) undergoes infinitely many symmetry breaking bifurcations (into nonradial solutions) as $\lambda \to 0^+$ and $\|u\| \to \infty$. A similar approach has been used in [106] for semilinear equations involving power nonlinearities.

8. Whole space problems

In this final section we shall consider problems of the type

$$-\Delta u + u - |u|^{p-1}u = 0, \quad x \in \mathbb{R}^N,\ p > 1, \tag{8.1}$$

and more general equations of the form

$$-\Delta u + V(x)u - \bigl(\Delta|u|^2\bigr)u = \lambda|u|^{p-1}u, \quad x \in \mathbb{R}^N, \tag{8.2}$$

where λ is a parameter. Again, we will be interested in radial solutions, i.e., solutions u which only depend upon $r = |x|$. Such solutions, of course, will, as before be solutions of corresponding ordinary differential equations. As in the case when the domain is a ball, positive solutions of (8.1) may be radial solutions by results similar to the results in [50]. In fact very similar results have been obtained, see, e.g., [27,28,102].

We shall call positive solutions of either of the equations ground states and we shall discuss results which guarantee the existence of ground states. The approach used in both cases is variational and we shall rely on results in [83,111] for Equation (8.1) and recent results from [90] for Equation (8.2). The latter equation is a fully nonlinear equation and has many interesting applications in mathematical physics. Some references to such applications may be found in [90].

The important Sobolev space for both problems is the space $W^{1,2}(\mathbb{R}^N)$ with the subspace

$$H_r^1(\mathbb{R}^N) = \{u \in W^{1,2}(\mathbb{R}^N)\colon\ u \text{ is radial}\}.$$

A result of Strauss (see [111]) guarantees that the embedding

$$H_r^1(\mathbb{R}^N) \hookrightarrow L^q(\mathbb{R}^N), \quad 2 < q < \frac{2N}{N-2}, \tag{8.3}$$

is a compact embedding; this fact plays an important role in studying the functionals

$$I(u) = \frac{1}{2}\left(\int_{\mathbb{R}^N} |\nabla u|^2\,\mathrm{d}x + \int_{\mathbb{R}^N} |u|^2\,\mathrm{d}x\right), \tag{8.4}$$

$$J(u) = \frac{1}{2}\left(\int_{\mathbb{R}^N} |\nabla u|^2\,\mathrm{d}x + \int_{\mathbb{R}^N} V(x)|u|^2\,\mathrm{d}x\right) + \int_{\mathbb{R}^N} |\nabla u|^2 u^2\,\mathrm{d}x, \tag{8.5}$$

and

$$K(u) = \frac{1}{p+1}\int_{\mathbb{R}^N} |u|^{p+1}\,\mathrm{d}x, \tag{8.6}$$

on the space $H_r^1(\mathbb{R}^N)$.

The compact embedding result plays an important role in establishing the so-called Palais–Smale condition, a condition which is used in verifying the mountain pass theorem. An alternate approach to such problems is the use of Lagrange multiplier techniques in which case the fact that the level sets

$$\{u \in H_r^1(\mathbb{R}^N)\colon\ K(u) = \text{constant}\}$$

are weakly closed (which follows from the compact embedding result) plays a role. These approaches establish the existence of critical points of the defining functionals (or critical points subject to a constraint) which provide weak solutions of the differential equation. Some classical arguments then show that the weak solutions are actually smooth. Using the well-known mountain pass theorem of Ambrosetti and Rabinowitz the following result has been established (see, e.g., [111]).

THEOREM 8.1. *Assume* $1 < p < \frac{N+2}{N-2}$, *then Equation* (8.1) *has a positive radial solution* $u \in H_r^1(\mathbb{R}^N)$. *This solution belongs to class* $C^2(\mathbb{R}^N)$ *and is a solution of the boundary value problem*

$$\begin{cases} u'' + \frac{N-1}{r}u' - u + |u|^{p-1}u = 0, & r \in (0,\infty), \\ u'(0) = u(\infty) = 0. \end{cases} \tag{8.7}$$

As already mentioned, this problem (for the case $N = 3$) was already studied by Nehari [83] using a variational approach. For Nehari's case the problem becomes

$$\begin{cases} u'' + \frac{2}{r}u' - u + |u|^{p-1}u = 0, & r \in (0,\infty), \\ u'(0) = u(\infty) = 0, \end{cases} \tag{8.8}$$

with $1 < p < 5$. Letting $v = ru$, one finds that v is a solution of the equation

$$v'' - v + r^{1-p}v^p = 0. \tag{8.9}$$

The following example of Nehari [83] shows that for $p = 5$ no such solutions can exist (thus a result similar to the nonexistence result of Pohozaev [89] holds for problems defined in all of space, as well). For, if one assumes that such a solution exists, then v solves

$$v'' - v + r^{-4}v^5 = 0, \tag{8.10}$$

and using the identity

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r(v')^2 - vv' + \frac{v^6}{2r^3}\right) = (2rv' - v)\left(v'' + \frac{v^5}{r^4}\right),$$

one obtains, for any $0 < a < r < b$,

$$\left[r(v')^2 - vv' + \frac{v^6}{2r^3}\right]_a^b = \int_a^b \left(2rvv' - v^2\right)\mathrm{d}r = \left[rv^2\right]_a^b - 2\int_a^b v^2\,\mathrm{d}r. \tag{8.11}$$

One may show that

$$v(r) = \mathrm{O}\left(\mathrm{e}^{-r}\right), \quad r \to \infty, \qquad v(r) = \mathrm{O}(r), \quad r \to 0,$$

and hence conclude from (8.11) that

$$\int_0^\infty v^2\,\mathrm{d}r = 0,$$

contradicting that u is a nontrivial solution of (8.8) for $p = 5$.

The above example is but one of the many results available for the existence of positive radial ground states for equations of the form

$$\operatorname{div}\big(A\big(|\nabla u|\big)\nabla u\big) + f(u) = 0, \quad x \in \mathbb{R}^N.$$

For additional results we refer to the papers [84,88,102] and their references.

We now turn to the problem (8.2), then under various assumptions on the potential V one may establish results similar to the above using a variational approach (critical point theory). The assumptions imposed on $V \in L^1_{\mathrm{loc}}(\mathbb{R}^N)$ and the number p belong to the following class:

(A) There exists a constant $\delta > 0$ such that

$$V(x) \geqslant \delta, \quad \text{a.e.}$$

(B) $V \in L^\infty(\mathbb{R}^N)$, and V is periodic in each coordinate; $p \geqslant 3$.
(C) $V \in L^\infty(\mathbb{R}^N)$, $V(x) \to \|V\|_{L^\infty}$, $|x| \to \infty$; $p \geqslant 3$.
(D) $V(x) \to \infty$, $|x| \to \infty$; $p > 1$.

The following results have been established (see [90]).

THEOREM 8.2. *Assume that* $N = 1$ *and that condition* (A) *and condition* (B), (C), *or* (D) *hold. Then for any* $\alpha > 0$ *there exists* $\lambda_\alpha > 0$ *such that*

$$-u'' + V(x)u - \big(u^2\big)''u = \lambda_\alpha |u|^{p-1}u,$$

has a positive solution $u \in W^{1,2}(\mathbb{R})$. *The function* $\alpha \mapsto \lambda_\alpha$ *is unbounded and lower semicontinuous. If it is the case that* $p \geqslant 3$, *then*

$$\lim_{\alpha \to \infty} \lambda_\alpha = 0, \qquad \lim_{\alpha \to 0^+} \lambda_\alpha = \infty.$$

In this case, furthermore, for each $\lambda > 0$, *there exists an infinite sequence of positive solutions* $\{u_n\} \subset W^{1,2}(\mathbb{R})$ *of*

$$-u'' + V(x)u - \big(u^2\big)''u = \lambda |u|^{p-1}u,$$

with

$$J(u_n) \to \infty,$$

where J *is given by* (8.5).

For higher dimensions the following existence results are valid.

THEOREM 8.3. *Let* $N \geqslant 2$ *and* $1 < p < \frac{N+2}{N-2}$. *Then there exists a sequence* $\{\lambda_n\}$, $\lambda_n \to \infty$, *such that* (8.2) *admits for any* $\lambda = \lambda_n$ *a nontrivial nonnegative radial solution* $u \in H^1_r(\mathbb{R}^N)$.

In case $3 \leqslant p < \frac{N+2}{N-2}$, *i.e.*, $N = 2$ *or* $N = 3$, *there exists a corresponding sequence* $\{\lambda_n\}$, *satisfying* $\lambda_n \to 0$, *as well.*

If $N = 3$ *and* $3 \leqslant p < 5$ *and* $\lambda > 0$, *then* (8.2) *has a nontrivial nonnegative solution* $u \in H_r^1(\mathbb{R}^3)$.

References

[1] J.C. Alexander, *A primer on connectivity*, Fixed Point Theory (Sherbrooke, Que., 1980), Springer, Berlin (1981), 455–483.

[2] W. Allegretto, P. Nistri and P. Zecca, *Positive solutions of elliptic nonpositone problems*, Differential Integral Equations **5** (1992), 95–101.

[3] H. Amann, *On the number of solutions of asymptotically linear two point boundary value problems*, Arch. Rat. Mech. Anal. **55** (1974), 207–213.

[4] A. Anane, *Simplicité et isolation de la première valeur propre du p-Laplacien*, C. R. Acad. Sci Paris **305** (1987), 725–728.

[5] T. Aubin, *Some Nonlinear Problems in Riemannian Geometry*, Springer-Verlag, Berlin (1998).

[6] J.G. Azorero, I.P. Alonso and J.-P. Puel, *Quasilinear problems with exponential growth in the reaction term*, Nonlinear Anal. TMA **22** (1994), 481–498.

[7] M. Badiale and E. Nabana, *A note on radiality of solutions of p-Laplacian equation*, Appl. Anal. **52** (1994), 1–4, 35–43.

[8] C. Bandle, *Isoperimetric Inequalities and Applications*, Pitman, London (1980).

[9] C. Bandle, C.V. Coffman and M. Marcus, *Nonlinear elliptic problems in annular domains*, J. Differential Equations **69** (1987), 322–345.

[10] C. Bandle and L.A. Peletier, *Problèmes de Dirichlet nonlinéaires dans des anneaux*, C. R. Acad. Sci. Paris **305** (1986), 181–184.

[11] J. Bebernes and D. Eberly, *Mathematical Problems from Combustion Theory*, Applied Math. Sci., Vol. 83, Springer-Verlag, New York (1989).

[12] A. ben Naoum and C. De Coster, *On the existence and multiplicity of positive solutions of the p-Laplacian seperated boundary value problems*, Differential Integral Equations **10** (1997), 1093–1112.

[13] M.S. Berger, *Nonlinearity and Functional Analysis*, Academic Press, New York (1977).

[14] T. Bhattacharya, *Radial symmetry of the first eigenfunction for the p-Laplacian in the ball*, Proc. Amer. Math. Soc. **104** (1988), 169–174.

[15] G. Bratu, *Sur les équations intégrales non linéaires*, Bull. Soc. Math. France **42** (1914), 113–142.

[16] H. Brezis and L. Nirenberg, *Positive solutions of nonlinear elliptic equations involving critical Sobolev exponents*, Comm. Pure Appl. Math. **36** (1983), 437–477.

[17] A. Castro and A. Kurepa, *Energy analysis of a nonlinear singular differential equation and applications*, Rev. Colombiana Mat. **21** (1987), 155–166.

[18] A. Castro and A. Kurepa, *Infinitely many radially symmetric solutions to a superlinear Dirichlet problem in a ball*, Proc. Amer. Math. Soc. **101** (1987), 57–64.

[19] A. Castro and R. Shivaji, *Nonnegative solutions for a class of radially symmetric nonpositone problems*, Proc. Amer. Math. Soc. **106** (1989), 735–740.

[20] S. Chandrasekhar, *An Introduction to the Study of Stellar Structure*, Dover, New York (1957).

[21] S.N. Chow and J.K. Hale, *Methods of Bifurcation Theory*, Springer-Verlag, New York (1982).

[22] P. Clément, D. de Figueiredo and E. Mitidieri, *Quasilinear elliptic equations with critical exponents*, Topol. Methods Nonlinear Anal. **7** (1996), 133–170.

[23] P. Clément, M. García-Huidobro, R. Manásevich and K. Schmitt, *Mountain pass type solutions for quasilinear elliptic equations*, Calc. Variations PDE **11** (2000), 33–62.

[24] C.V. Coffman, *A nonlinear boundary value problem with many positive solutions*, J. Differential Equations **54** (1984), 429–437.

[25] C.V. Coffman and M. Marcus, *Existence and uniqueness results for semilinear Dirichlet problems in annuli*, ARCH (1991), 293–307.

[26] M.G. Crandall and P.H. Rabinowitz, *Some continuation and variation methods for positive solutions of nonlinear elliptic eigenvalue problems*, Arch. Rat. Mech. Anal. **58** (1975), 207–218.

[27] L. Damascelli and F. Pacella, *Monotonicity and symmetry of solutions of p-Laplace equations,* $1 < p < 2$, *via the moving plane method*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **9** (1998), 95–100.

[28] L. Damascelli and F. Pacella, *Monotonicity and symmetry results for p-Laplace equations and applications*, Adv. Differential Equations **5** (2000), 1179–1200.

[29] H. Dang, R. Manásevich and K. Schmitt, *Positive radial solutions of some nonlinear partial differential equations*, Math. Nachr. **186** (1997), 101–113.

[30] H. Dang and K. Schmitt, *Existence of positive solutions for semilinear elliptic equations in annular domains*, Differential Integral Equations **7** (1994), 747–758.

[31] H. Dang, K. Schmitt and R. Shivaji, *On the number of solutions of boundary value problems involving the p-Laplacian*, Electr. J. Differential Equations **1996** (1996), 137–148.

[32] K. Deimling, *Nonlinear Functional Analysis*, Springer-Verlag, New York (1987).

[33] M.A. Del Pino, M. Elgueta and R.F. Manásevich, *A homotopic deformation along p of a Leray–Schauder degree result and existence for* $(|u'|^{p-2}u')' + f(t,u) = 0$, $u(0) = u(T) = 0$, $p > 1$, J. Differential Equations **80** (1989), 1–13.

[34] M.A. Del Pino and R.F. Manásevich, *Global bifurcation from the eigenvalues of the p-Laplacian*, J. Differential Equations **92** (1991), 226–251.

[35] P. Delanoë, *Radially symmetric boundary value problems for real and complex elliptic Monge–Ampère equations*, J. Differential Equations **58** (1985), 318–344.

[36] P. Drábek and R. Manásevich, *On the closed solution to some nonhomogeneous eigenvalue problems with p-Laplacian*, Differential and Integral Equations (1999).

[37] A. el Hachimi and F. de Thélin, *Infinitely many radially symmetric solutions for a quasilinear problem in a ball*, J. Differential Equations **128** (1996), 78–102.

[38] V. Emden, *Gaskugeln*, Leipzig (1907).

[39] L. Erbe and K. Schmitt, *On radial solutions of semilinear elliptic problems*, Differential Integral Equations **1** (1988), 71–78.

[40] D.A. Frank-Kamenetskii, *Diffusion and Heat Exchange in Chemical Kinetics*, Princeton University Press, Princeton (1955).

[41] L. Friedlander, *Asymptotic behavior of the eigenvalues of the p-Laplacian*, Comm. Partial Differential Equations **14** (1989), 1059–1069.

[42] T. Gallouët, F. Mignot and J.-P. Puel, *Quelques résultats sur le problème* $-\Delta u = \lambda e^u$, C. R. Acad. Sci. Paris Sér. I Math. **307** (1988), 289–292.

[43] J. García Azorero and I. Peral Alonso, *Multiplicity of solutions for elliptic problems with critical exponent or with a nonsymmetric term*, Trans. Amer. Math. Soc. **323** (1991), 877–895.

[44] J. García Azorero and I. Peral Alonso, *On an Emden–Fowler type equation*, Nonlinear Anal. **18** (1992), 1085–1097.

[45] M. García-Huidobro, R. Manásevich and K. Schmitt, *On the principal eigenvalue of p-Laplacian like operators*, J. Differential Equations **130** (1996), 235–246.

[46] M. García-Huidobro, R. Manásevich and K. Schmitt, *Some bifurcation results for a class of p-Laplacian like operators*, Differential Integral Equations **10** (1997), 51–66.

[47] M. García-Huidobro, R. Manásevich and K. Schmitt, *Positive radial solutions of quasilinear elliptic partial differential equations on a ball*, Nonlinear Anal. TMA **35** (1999), 175–190.

[48] M. García-Huidobro, R. Manásevich and F. Zanolin, *Strongly nonlinear second order ODEs with rapidly growing terms*, J. Math. Anal. Appl. **202** (1996), 1–26.

[49] I.M. Gelfand, *Some problems in the theory of quasilinear equations*, Amer. Math. Soc. Transl. Ser. 2 **29** (1963), 295–381.

[50] B. Gidas, W. Ni and L. Nirenberg, *Symmetry and related properties via the maximum principle*, Comm. Math. Phys. **68** (1979), 209–243.

[51] G.B. Gustafson and K. Schmitt, *Nonzero solutions of boundary value problems for second order ordinary and delay differential equations*, J. Differential Equations **12** (1972), 125–147.

[52] K. Gustafson, *Combustion and explosion equations and their calculation*, Computational Techniques in Heat Transfer (Seattle, Wash., 1983), Vol. 1, Pineridge, Swansea (1985), 161–195.

[53] D. Hai and K. Schmitt, *On radial solutions of quasilinear boundary value problems*, Topics in Nonlinear Analysis, Birkhäuser, Basel (1999), 349–361.

[54] D. Hai, K. Schmitt and R. Shivaji, *Positive solutions of quasilinear boundary value problems*, J. Math. Anal. Appl. **217** (1998), 672–686.

[55] P. Hartman, *Ordinary Differential Equations*, second ed., Birkhäuser, Boston (1982).

[56] J. Heinonen, T. Kilpeläinen and O. Martio, *Nonlinear Potential Theory of Degenerate Elliptic Equations*, Clarendon Press, Oxford University Press, New York (1993).

[57] J. Jacobsen, *Global bifurcation problems associated with k-Hessian operators*, Topol. Methods Nonlinear Anal. **14** (1999), 81–130.

[58] J. Jacobsen, *A Liouville–Gelfand equation for k-Hessian operators*, Rocky Mountain J. Math. **34** (2004), 665–684.

[59] J. Jacobsen and K. Schmitt, *The Liouville–Bratu–Gelfand problem for radial operators*, J. Differential Equations **184** (2002), 283–298.

[60] W. Jäger and K. Schmitt, *Symmetry breaking in semilinear elliptic problems*, Analysis ETC, Moser Anniversary Volume, P. Rabinowitz and E. Zehnder, eds., Academic Press, New York (1990), 451–470.

[61] D.D. Joseph and T.S. Lundgren, *Quasilinear Dirichlet problems driven by positive sources*, Arch. Rat. Mech. Anal. **49** (1973), 241–269.

[62] H. Kaper, M. Knaap and M. Kwong, *Existence theorems for second order boundary value problems*, Differential Integral Equations **4** (1991), 543–554.

[63] J.L. Kazdan and F.W. Warner, *Remarks on some quasilinear elliptic equations*, Comm. Pure Appl. Math. **28** (1975), 567–597.

[64] M. Krasnosel'skii, *Topological Methods in the Theory of Nonlinear Integral Equations*, Pergamon, Oxford (1963).

[65] M. Krasnosel'skii, *Positive Solutions of Operator Equations*, Noordhoff, Groningen (1964).

[66] V. Le and K.Schmitt, *On boundary value problems for degenerate quasilinear elliptic equations and inequalities*, J. Differential Equations **144** (1998), 170–218.

[67] S.S. Lin, *Positive radial solutions and non-radial bifurcation for semilinear elliptic equations in annular domains*, J. Differential Equations **86** (1990), 367–391.

[68] S.S. Lin, *Existence of positive nonradial solutions for nonlinear elliptic equations in annular domains*, Trans. Amer. Math. Soc. **332** (1992), 775–791.

[69] P. Lindqvist, *On the equation* $\operatorname{div}(|\nabla u|^{p-2}\nabla u) + \lambda|u|^{p-2}u = 0$, Proc. Amer. Math. Soc. **109** (1990), 157–164.

[70] P.L. Lions, *Two remarks on Monge–Ampère equations*, Ann. Mat. Pura Appl. **142** (1985), 263–275.

[71] J. Liouville, *Sur l'équation aux différences partielles* $\frac{\mathrm{d}^2 \log\lambda}{\mathrm{d}u\,\mathrm{d}v} \pm \frac{\lambda}{2a^2} = 0$, J. Math. Pures Appl. **18** (1853), 71–72.

[72] N.G. Lloyd, *Degree Theory*, Cambridge Tracts in Mathematics, Vol. 73, Cambridge University Press, Cambridge (1978).

[73] R. Manásevich, F.I. Njoku and F. Zanolin, *Positive solutions for the one-dimensional p-Laplacian*, Differential Integral Equations **8** (1995), 213–222.

[74] R. Manásevich and K. Schmitt, *Boundary value problems for quasilinear second order differential equations*, Non-Linear Analysis and Boundary Value Problems for Ordinary Differential Equations (Udine), F. Zanolin, ed., Springer, Vienna (1996), 79–119.

[75] J. Mawhin and K. Schmitt, *Landesman–Lazer type problems at an eigenvalue of odd multiplicity*, Results Math. **14** (1988), 138–146.

[76] J. Mawhin and K. Schmitt, *Nonlinear eigenvalue problems with the parameter near resonance*, Ann. Polon. Math. **51** (1990), 241–248.

[77] J. S. McGough, *On solution continua of supercritical quasilinear elliptic problems*, Differential Integral Equations **7** (1993), 1453–1471.

[78] F. Mignot and J.-P. Puel, *Solution radiale singulière de* $-\Delta u = \lambda e^{u}$, C. R. Acad. Sci. Paris Sér. I Math. **307** (1988), 379–382.

[79] F. Mignot and J.-P. Puel, *Quelques résultats sur un problème elliptique avec non linéarité exponentielle*, Équations aux dérivées partielles et applications, Gauthier-Villars, Éd. Sci. Méd. Elsevier, Paris (1998), 683–704.

[80] K. Nagasaki and T. Suzuki, *Radial and nonradial solutions for the nonlinear eigenvalue problem* $\Delta u + \lambda e^u = 0$ *on annuli in* R^2, J. Differential Equations **87** (1990), 144–168.

[81] K. Nagasaki and T. Suzuki, *Radial solutions for* $\Delta u + \lambda e^u = 0$ *on annuli in higher dimensions*, J. Differential Equations **100** (1992), 137–161.

[82] K. Nagasaki and T. Suzuki, *Spectral and related properties about the Emden–Fowler equation* $-\Delta u = \lambda e^u$ *on circular domains*, Math. Ann. **299** (1994), 1–15.

[83] Z. Nehari, *On a nonlinear differential equation arising in nuclear physics*, Proc. Roy. Irish Acad. **62** (1963), 117–135.

[84] W.M. Ni and J. Serrin, *Existence and nonexistence theorems for ground states of quasilinear partial differential equations. The anomalous case*, Convegno Celebrativo del Centenario della Nascita di Mauro Picone e di Leonida Tonelli, Atti dei Convegin Lincei, Vol. 77, Academia Nazionale dei Lincei (1986), 231–257.

[85] H.-O. Peitgen and K. Schmitt, *Perturbations globales topologiques des problemes nonlineaires de valeur propre*, Comptes Rendues Acad. Sci. Paris **291** (1980), 271–274.

[86] H.-O. Peitgen and K. Schmitt, *Global topological perturbations of nonlinear elliptic eigenvalue problems*, Math. Methods Appl. Sci. **5** (1983), 376–388.

[87] H.-O. Peitgen and K. Schmitt, *Global analysis of two-parameter elliptic eigenvalue problems*, Trans. Amer. Math. Soc. **283** (1984), 57–95.

[88] L.A. Peletier and J. Serrin, *Ground states for the prescribed mean curvature equation*, Proc. Amer. Math. Soc. **100** (1987), 694–700.

[89] S. Pohozaev, *Eigenfunctions of the equation* $\Delta u + \lambda f(u) = 0$, Sov. Math. Dokl. **165** (1965), 1408–1411.

[90] M. Poppenberg, K. Schmitt and Z. Wang, *On the existence of soliton solutions to quasilinear Schrödinger equations*, Calc. Var. Partial Differential Equations **14** (2002), 329–344.

[91] P. Pucci and J. Serrin, *A general variational identity*, Indiana Univ. Math. J. **35** (1986), 681–703.

[92] P. Pucci and J. Serrin, *Critical exponents and critical dimensions for polyharmonic operators*, J. Math. Pures Appl. (9) **69** (1990), 55–83.

[93] P. Rabinowitz, *On bifurcation from infinity*, J. Differential Equations **14** (1973), 462–475.

[94] P. Rabinowitz, *Some aspects of nonlinear eigenvalue problems*, Rocky Mountain J. Math. **3** (1973), 161–202.

[95] M. Rao and Z. Ren, *Theory of Orlicz Spaces*, Marcel Dekker, New York, 1985.

[96] D.H. Sattinger, *Topics in Stability and Bifurcation Theory*, Lecture Notes in Math., Vol. 309, Springer-Verlag, 1973.

[97] R. Schaaf and K. Schmitt, *A class of nonlinear Sturm–Liouville problems with infintely many solutions*, Trans. Amer. Math. Soc. **306** (1988), 853–859.

[98] R. Schaaf and K. Schmitt, *Asymptotic behavior of positive solution branches of semilinear elliptic problems with linear part at resonance*, Z. Angew. Math. Phys. **43** (1992), 645–676.

[99] K. Schmitt, *Boundary value problems for quasilinear elliptic partial differential equations*, Nonlinear Anal. **2** (1978), 263–309.

[100] K. Schmitt, *Positive solutions of semilinear elliptic boundary value problems*, Topological Methods in Differential Equations and Inclusions (Montreal, PQ, 1994), A. Granas, M. Frigon and G. Sabidussi, eds., Series C: Mathematical and Physical Sciences, Vol. 472, Kluwer Academic, Dordrecht (1995), 447–500.

[101] K. Schmitt and Z.-Q. Wang, *On bifurcation from infinity for potential operators*, Differential Integral Equations **4** (1991), 933–944.

[102] J. Serrin and H. Zou, *Symmetry of ground states of quasilinear elliptic equations*, Arch. Rat. Mech. Anal. **148** (1999), 265–290.

[103] V.L. Shapiro, *Quasilinear ellipticity and the 1st eigenvalue*, Comm. Partial Differential Equations **16** (1991), 1819–1855.

[104] E.A.D.B. Silva and S.H.M. Soares, *Liouville–Gelfand type problems for the* N*-Laplacian on bounded domains of* $\mathbf{R}^N$, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28** (1999), 1–30.

[105] J. Smoller and A. Wasserman, *Symmetry breaking for positive solutions of semilinear elliptic equations*, Arch. Rat. Mech. Anal. **95** (1986), 217–225.

[106] P.N. Srikanth, *Symmetry breaking for a class of semilinear elliptic problems*, Ann. Inst. H. Poincaré **7** (1990), 107–112.

[107] T. Suzuki, *Radial and nonradial solutions for semilinear elliptic equations on circular domains*, Symposia Mathematica, Vol. XXX (Cortona, 1988), Academic Press, London (1989), 153–174.

[108] N.S. Trudinger and X.-J. Wang, *Hessian measures. I*, Topol. Methods Nonlinear Anal. **10** (1997), 225–239.

[109] K. Tso, *Remarks on critical exponents for Hessian operators*, Ann. Inst. H. Poincaré Anal. Non Linéaire **7** (1990), 113–122.

[110] X.-J. Wang, *A class of fully nonlinear elliptic equations and related functionals*, Indiana Univ. Math. J. **43** (1994), 25–54.

[111] M. Willem, *Minimax Theorems*, Birkhäuser, Boston (1996).

This Page Intentionally Left Blank

CHAPTER 5

Integrability of Polynomial Differential Systems

Jaume Llibre
Departament de Matemàtiques, Universitat Autónoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
E-mail: jllibre@mat.uab.es

Contents

1. Basic notions 439
    1.1. Polynomial systems 440
    1.2. First integrals and invariants 441
    1.3. Integrating factors 442
    1.4. Invariant algebraic curves 444
    1.5. Exponential factors 446
2. Darbouxian theory of integrability 448
3. Kapteyn–Bautin theorem 452
4. On the degree of the invariant algebraic curves 454
5. Darboux lemma 460
6. Applications of the Darboux lemma 466
7. Algebraic limit cycles for quadratic systems 475
8. Limit cycles and algebraic limit cycles 480
9. Darbouxian theory of integrability and centres 485
10. Non-existence of limit cycles 487
11. The inverse problem 491
12. Elementary and Liouvillian first integrals 500
13. Liouvillian first integrals for the planar Lotka–Volterra system 503
14. On the integrability of two-dimensional flows 514
15. Darbouxian theory of integrability for polynomial vector fields on surfaces 518
    15.1. Invariant algebraic curves and exponential factors 519
    15.2. The surfaces 520
    15.3. Darbouxian theory 524
References 528

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 1
Edited by A. Cañada, P. Drábek and A. Fonda
© 2004 Elsevier B.V. All rights reserved

This Page Intentionally Left Blank

Nonlinear ordinary differential equations appear in many branches of applied mathematics, physics and, in general, in applied sciences. For a differential system or a vector field defined on the plane $\mathbb{R}^2$ or $\mathbb{C}^2$ the existence of a first integral determines completely its phase portrait; of course, working with real or complex time, respectively. Since for such vector fields the notion of integrability is based on the existence of a first integral the following natural question arises:

> *Given a vector field on $\mathbb{R}^2$ or $\mathbb{C}^2$, how to recognize if this vector field has a first integral?*

The more easiest planar vector fields having a first integral are the Hamiltonian ones. The integrable planar vector fields which are not-Hamiltonian are, in general, very difficult to detect. Many different methods have been used for studying the existence of first integrals for non-Hamiltonian vector fields based on: Noether symmetries [13], the Darbouxian theory of integrability [38], the Lie symmetries [83], the Painlevé analysis [7], the use of Lax pairs [62], the direct method [51] and [55], the linear compatibility analysis method [97], the Carlemann embedding procedure [14] and [2], the quasimonomial formalism [7], etc.

In this chapter we study the existence of first integrals for planar polynomial vector fields through the Darbouxian theory of integrability. The algebraic theory of integrability is a classical one, which is related with the first part of the Hilbert's 16th problem [56]. This kind of integrability is usually called Darbouxian integrability, and it provides a link between the integrability of polynomial vector fields and the number of invariant algebraic curves that they have (see [38,86]).

Darboux [38] showed how can be constructed the first integrals of planar polynomial vector fields possessing sufficient invariant algebraic curves. In particular, he proved that if a planar polynomial vector field of degree m has at least $[m(m+1)/2]$ invariant algebraic curves, then it has a first integral, which can be computed using these invariant algebraic curves. Jouanolou [58] (see also [95] and [31] for an easy proof) shows that if the number of invariant algebraic curves of a planar polynomial vector field of degree m is at least $[m(m+1)/2]+2$, then the vector field has a rational first integral, and consequently all its solutions are invariant algebraic curves. For more details and results on Darbouxian theory of integration for planar polynomial vector fields, see [9,26,28,30,85,87,91–93].

Prelle and Singer [87] using methods of differential algebra, showed that if a polynomial vector field has an elementary first integral, then it can be computed using the Darbouxian theory of integrability. Singer [95] proved that if a polynomial vector field has Liouvillian first integrals, then it has integrating factors given by Darbouxian functions. Some related results can be found in [20].

1. Basic notions

In this section first we present the planar polynomial differential systems that we shall study. Secondly, we introduce the notion of first integral; and finally, we deal with the definition of integrating factor.

1.1. *Polynomial systems*

By definition a *planar polynomial differential system* or simply a *polynomial system* is a differential system of the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \dot{x} = P(x, y), \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = \dot{y} = Q(x, y), \tag{1}$$

or equivalently

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{Q(x, y)}{P(x, y)},$$

where P and Q are polynomials in the variables x and y. Moreover, the dependent variables x and y, the independent variable t, and the coefficients of the polynomials P and Q are either all complex, or all real. In the former case the system is called a *complex* polynomial system and in the later a *real* polynomial system.

The independent variable t will be called the *time* of system (1).

If we have a real polynomial system (1) we will work with it as a complex one, independently if we want to study its real integrability, because as we will see later on, often the real integrability of a real polynomial system is forced by its complex structure.

In this chapter the *degree* m of the polynomial system (1) will be the maximum of the degrees of the polynomials P and Q.

Associated to polynomial system (1) there is either the *polynomial vector field*

$$\mathcal{X} = P(x, y)\frac{\partial}{\partial x} + Q(x, y)\frac{\partial}{\partial y}, \tag{2}$$

or the *polynomial* 1*-form*

$$\omega = P(x, y)\,\mathrm{d}y - Q(x, y)\,\mathrm{d}x.$$

The existence of solutions for the polynomial system (1), or equivalently for the polynomial vector field $\mathcal{X}$ or for the 1-form ω is given by the Existence and Uniqueness theorem of solutions for an ordinary differential system.

If the polynomial system (1) is real, then a solution of it is an analytic function $\varphi : U \to \mathbb{R}^2$ (U is the maximal time open interval in which the solution is defined) such that $\mathrm{d}\varphi(t)/\mathrm{d}t = X(\varphi(t))$, for every $t \in U$.

If the polynomial system (1) is complex, then a solution of it is an analytic or holomorphic function $\varphi : U \to \mathbb{C}^2$ (U is the maximal time simple connected open subset of $\mathbb{C}$ in which the solution is defined) such that $\mathrm{d}\varphi(t)/\mathrm{d}t = X(\varphi(t))$, for every $t \in U$.

In what follows we denote by $\mathbb{F}$ either $\mathbb{R}$, or $\mathbb{C}$, according with system (1) be real or complex, respectively.

Let $\varphi : U \to \mathbb{F}^2$ be a solution of system (1). Then, the set $\{\varphi(t) \in \mathbb{F}^2\colon\ t \in U\}$ is called a *trajectory*, an *integral curve* or an *orbit* of system (1) or of the vector field $\mathcal{X}$.

1.2. *First integrals and invariants*

The vector field $\mathcal{X}$ or equivalently the system (1) is *integrable* on an open subset U of $\mathbb{F}^2$ if there exists a nonconstant analytic function $H : U \to \mathbb{F}$, called a *first integral* of the system on U, which is constant on all solution curves $(x(t), y(t))$ of $\mathcal{X}$ contained in U; i.e., $H(x(t), y(t)) =$ constant for all values of t for which the solution $(x(t), y(t))$ is defined and contained in U. Clearly H is a first integral of $\mathcal{X}$ on U if and only if

$$\mathcal{X}H = 0, \quad \text{or} \quad \omega \wedge \mathrm{d}H = 0,$$

on U.

For system (1) defined on the plane $\mathbb{R}^2$ or $\mathbb{C}^2$ the existence of a first integral determines completely its phase portrait; i.e., the decomposition of the plane as union of the orbits of (1).

EXAMPLE 1.1. The polynomial system

$$\dot{x} = x(ax + c), \qquad \dot{y} = y(2ax + by + c), \tag{3}$$

has the first integral

$$H = \frac{(ax + c)(ax + by)}{x(ax + by + c)}.$$

And the system

$$\dot{x} = -y - b\left(x^2 + y^2\right), \qquad \dot{y} = x, \tag{4}$$

has the first integral

$$H = \exp(2by)\left(x^2 + y^2\right). \tag{5}$$

Of course, once we have a first integral any analytic function of this first integral is another first integral.

Let $U \subset \mathbb{F}^2$ be an open set. We say that an analytic function $H(x, y, t) : U \times \mathbb{F} \to \mathbb{F}$, is an *invariant* of the polynomial vector field $\mathcal{X}$ on U, if $H(x, y, t) =$ constant for all values of t for which the solution $(x(t), y(t))$ is defined and contained in U. If an invariant H is independent on t then, of course, it is a first integral.

The knowledge provided by an invariant is weaker than the one provide by a first integral. The invariant, in general, only gives information about either the α- or the ω-limit set of the orbits of the system.

EXAMPLE 1.2. If $aB - bA = 0$ and $bC - Bc \neq 0$, then the polynomial system

$$\dot{x} = x(ax + by + c), \qquad \dot{y} = y(Ax + By + C), \tag{6}$$

has the invariant

$$H = x^{B/(bC-Bc)} y^{-b/(bC-Bc)} e^t.$$

System (6) is the well-known Lotka–Volterra system, see [72,100].

1.3. *Integrating factors*

The simple fact of associating to system (1) the 1-form ω allows to obtain for the exact systems a first integral. Indeed, system (1) is called *exact* if it satisfies

$$\frac{\partial P}{\partial x} = -\frac{\partial Q}{\partial y}.$$

For these systems the 1-form ω is closed; i.e., $d\omega = 0$. Therefore, the function

$$H(x, y) = \int_{(x_0, y_0)}^{(x,y)} \omega,$$

obtained integrating ω through any path starting at the point (x_0, y_0) and ending at the point (x, y) is well defined, because ω is closed and $\mathbb{F}^2$ is simply connected. Then, $\omega = dH$. Hence, $\omega \wedge dH = 0$, and consequently H is a first integral for system (1).

The exact systems provides us a way to obtain first integrals for a given system (1) which initially is not exact, through the notion of integrating factor.

Let U be an open subset of $\mathbb{F}^2$, and $R : U \to \mathbb{F}$ be an analytic function which is not identically zero on U. The function R is an *integrating factor* of the vector field $\mathcal{X}$, or of the 1-form ω, or of the system (1) on U if one of the following four equivalent conditions holds

$$\begin{aligned}
&\frac{\partial(RP)}{\partial x} = -\frac{\partial(RQ)}{\partial y},\\
&\operatorname{div}(RP, RQ) = 0,\\
&\mathcal{X}R = -R\operatorname{div}(\mathcal{X}),\\
&d(R\omega) = d(RP\,dy - RQ\,dx) = 0,
\end{aligned} \tag{7}$$

on U. As usual the divergence of the vector field $\mathcal{X}$ is defined by

$$\operatorname{div}(\mathcal{X}) = \operatorname{div}(P, Q) = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}.$$

The *first integral H associated to the integrating factor R* is given by

$$H(x, y) = -\int R(x, y)P(x, y)\,dy + h(x),$$

satisfying $\partial H/\partial x = -RQ$. Then

$$\dot{x} = RP = -\frac{\partial H}{\partial y}, \qquad \dot{y} = RQ = \frac{\partial H}{\partial x}. \tag{8}$$

In order that this function H to be single-valued the open set U must be simply connected, if not it can be multi-valued.

Conversely, given a first integral H of the vector field $\mathcal{X}$ we always can find an integrating factor R for which (8) holds.

EXAMPLE 1.3. It is easy to check that system (4) has the integrating factor $R = 2\exp(2by)$, and that using it we get the first integral (5), and vice versa.

EXAMPLE 1.4. Linear differential equations. We consider the polynomial differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = a(x)y + b(x),$$

with a and b polynomials in x. Its associated 1-form is

$$\omega = \mathrm{d}y - \big(a(x)y + b(x)\big)\,\mathrm{d}x.$$

If $R = R(x)$ is an integrating factor it must satisfy

$$\mathrm{d}(R\omega) = \left(\frac{\partial R}{\partial x} + Ra(x)\right)\mathrm{d}x \wedge \mathrm{d}y = 0.$$

Therefore, we get the integrating factor

$$R(x) = \exp\!\left(-\int a(x)\,\mathrm{d}x\right).$$

Later on we shall use the following result.

PROPOSITION 1.5. *If a vector field $\mathcal{X}$ has two integrating factors R_1 and R_2 on the open subset U of $\mathbb{F}^2$, then in the open set $U \setminus \{R_2 = 0\}$ the function R_1/R_2 is a first integral.*

PROOF. Since R_i is an integrating factor, it satisfies that $\mathcal{X}R_i = -R_i\,\mathrm{div}(\mathcal{X})$ for $i = 1, 2$. Therefore, the proposition follows immediately from the next computation:

$$\mathcal{X}\left(\frac{R_1}{R_2}\right) = \frac{(\mathcal{X}R_1)R_2 - R_1(\mathcal{X}R_2)}{R_2^2} = 0.$$

□

1.4. *Invariant algebraic curves*

Let $f \in \mathbb{C}[x, y]$; i.e., f is a polynomial with complex coefficients in the variables x and y. The complex algebraic curve $f(x, y) = 0$ is an *invariant algebraic curve* of the real vector field $\mathcal{X}$ if for some polynomial $K \in \mathbb{C}[x, y]$ we have

$$\mathcal{X} f = P\frac{\partial f}{\partial x} + Q\frac{\partial f}{\partial y} = Kf. \tag{9}$$

The polynomial K is called the *cofactor* of the invariant algebraic curve $f = 0$. We note that since the polynomial system has degree m, then any cofactor has at most degree $m - 1$.

We remark that in the definition of invariant algebraic curve $f = 0$ we always allow this curve to be complex; that is $f \in \mathbb{C}[x, y]$. As we will see this is due to the fact that sometimes for real vector fields the existence of a real first integral can be forced by the existence of complex invariant algebraic curves. Of course, when we look for a complex invariant algebraic curve of a real vector field we are thinking in the real vector field as a complex one.

Since on the points of the algebraic curve $f = 0$ the gradient $(\partial f/\partial x, \partial f/\partial y)$ of the curve is orthogonal to the vector field $\mathcal{X} = (P, Q)$ (see (9)), the vector field $\mathcal{X}$ is tangent to the curve $f = 0$. Hence, the curve $f = 0$ is formed by trajectories of the vector field $\mathcal{X}$. This justifies the name of invariant algebraic curve given to the algebraic curve $f = 0$ satisfying (9) for some polynomial K, because it is *invariant* under the flow defined by $\mathcal{X}$.

EXAMPLE 1.6. It is easy to check that system (3) has the following five invariant algebraic curves, in this case straight lines: $f_1 = x = 0$, $f_2 = ax + c = 0$, $f_3 = y = 0$, $f_4 = ax + by = 0$, and $f_5 = ax + by + c = 0$, having cofactors $K_1 = ax + c$, $K_2 = ax$, $K_3 = 2ax + by + c$, $K_4 = ax + by + c$ and $K_5 = ax + by$, respectively. In fact, all integral curves of system (3) are algebraic because, with the exception of $f_1 = 0$ and $f_5 = 0$, they can be written as $(ax + c)(ax + by) - hx(ax + by + c) = 0$ being h a constant.

EXAMPLE 1.7. System (4) has only the two invariant algebraic curves $x + yi = 0$ and $x - yi = 0$, or equivalently the invariant algebraic curve $x^2 + y^2 = 0$, see Proposition 1.10. This is easy to proof using its first integral $H = \exp(2by)(x^2 + y^2)$.

PROPOSITION 1.8. *For a real polynomial system* (1), $f = 0$ *is an invariant algebraic curve with cofactor* K *if and only if* $\bar{f} = 0$ *is an invariant algebraic curve with cofactor* $\overline{K}$. *Here conjugation denotes conjugation of the coefficients of the polynomials only.*

PROOF. We assume that $f = 0$ is an invariant algebraic curve with cofactor K of the real polynomial system (1.9). Then equality (9) holds. Since P and Q are real polynomials conjugating equality (9) we obtain that

$$P\frac{\partial \bar{f}}{\partial x} + Q\frac{\partial \bar{f}}{\partial y} = \overline{K}\bar{f}.$$

Consequently, $\bar{f} = 0$ is an invariant algebraic curve with cofactor $\overline{K}$ of system (1). The proof in the converse direction is similar. □

LEMMA 1.9. *Let $f, g \in \mathbb{C}[x, y]$. We assume that f and g are relatively prime in the ring $\mathbb{C}[x, y]$. Then, for a polynomial system (1), $fg = 0$ is an invariant algebraic curve with cofactor K_{fg} if and only if $f = 0$ and $g = 0$ are invariant algebraic curves with cofactors K_f and K_g respectively. Moreover, $K_{fg} = K_f + K_g$.*

PROOF. It is clear that

$$\mathcal{X}(fg) = (\mathcal{X} f)g + f(\mathcal{X} g). \tag{10}$$

We assume that $fg = 0$ is an invariant algebraic curve with cofactor K_{fg} of system (1). Then, $\mathcal{X}(fg) = K_{fg} fg$ and from equality (10) we get $K_{fg} fg = (\mathcal{X} f)g + f \mathcal{X} g$. Therefore, since f and g are relatively prime, we obtain that f divides $\mathcal{X} f$, and g divides $\mathcal{X} g$. If we denote by $K_f = \mathcal{X} f/f$ and by $K_g = \mathcal{X} g/g$, then $f = 0$ and $g = 0$ are invariant algebraic curves of system (1) with cofactors K_f and K_g respectively, and $K_{fg} = K_f + K_g$.

The proof in the converse direction follows in a similar way using again (10). □

PROPOSITION 1.10. *We suppose that $f \in \mathbb{C}[x, y]$ and let $f = f_1^{n_1} \cdots f_r^{n_r}$ be its factorization in irreducible factors over $\mathbb{C}[x, y]$. Then, for a polynomial system (1), $f = 0$ is an invariant algebraic curve with cofactor K_f if and only if $f_i = 0$ is an invariant algebraic curve for each $i = 1, \dots, r$ with cofactor K_{f_i}. Moreover $K_f = n_1 K_{f_1} + \cdots + n_r K_{f_r}$.*

PROOF. From Lemma 1.9, we have that $f = 0$ is an invariant algebraic curve with cofactor K_f if and only if $f_i^{n_i} = 0$ is an invariant algebraic curve for each $i = 1, \dots, r$ with cofactor $K_{f_i^{n_i}}$, furthermore $K_f = K_{f_1^{n_1}} + \cdots + K_{f_r^{n_r}}$.

Now for proving the proposition it is sufficient to show, for each $i = 1, \dots, r$, that $f_i^{n_i} = 0$ is an invariant algebraic curve with cofactor $K_{f_i^{n_i}}$ if and only if $f_i = 0$ is an invariant algebraic curve with cofactor K_{f_i}, and that $K_{f_i^{n_i}} = n_i K_{f_i}$. We assume that $f_i^{n_i} = 0$ is an invariant algebraic curve with cofactor $K_{f_i^{n_i}}$. Then

$$K_{f_i^{n_i}} f_i^{n_i} = \mathcal{X}\big(f_i^{n_i}\big) = n_i f_i^{n_i - 1} \mathcal{X}(f_i),$$

or equivalently

$$\mathcal{X}(f_i) = \frac{1}{n_i} K_{f_i^{n_i}} f_i.$$

So defining $K_{f_i} = K_{f_i^{n_i}}/n_i$. We obtain that $f_i = 0$ is an invariant algebraic curve with cofactor K_{f_i} such that $K_{f_i^{n_i}} = n_i K_{f_i}$. The proof in the converse direction follows in a similar way. □

An *irreducible* invariant algebraic curve $f = 0$ will be an invariant algebraic curve such that f is an irreducible polynomial in the ring $\mathbb{C}[x, y]$.

A natural question in this subject is whether a polynomial vector field has or not invariant algebraic curves. The answer is not easy, see the large section in Jouanolou's book [58], or the long paper [79] devoted to show that one particular polynomial system has no invariant algebraic solutions. Even one of the more studied limit cycles, the limit cycle of the van der Pol system, until 1995 it was unknown that it is not algebraic [82]. Some results about the algebraic limit cycles of quadratic polynomial vector fields can be found in [21].

1.5. *Exponential factors*

In this section we introduce the notion of an exponential factor due to Christopher [28]. We shall see that an exponential factor appears when an invariant algebraic curve has in some sense multiplicity larger than 1. Therefore, as we shall show an exponential factor will play the same role than an invariant algebraic curve in order to obtain a first integral for the polynomial system (1).

We assume that we have two invariant algebraic curves $h = 0$ and $h + \varepsilon g = 0$ with cofactors K_h and $K_{h+\varepsilon g}$ and that ε is in a neighborhood of zero. Using that $\mathcal{X}h = K_h h$ and that $\mathcal{X}(h + \varepsilon g) = K_{h+\varepsilon g}(h + \varepsilon g)$, if we expand in power series of ε the cofactor $K_{h+\varepsilon g}$ we obtain that $K_{h+\varepsilon g} = K_h + \varepsilon K + \mathrm{O}(\varepsilon^2)$, where K is some polynomial of degree at most $m - 1$.

Since

$$\begin{aligned}
\mathcal{X}\left(\frac{h+\varepsilon g}{h}\right) &= \frac{\mathcal{X}(h+\varepsilon g)h - (h+\varepsilon g)\mathcal{X}h}{h^2} \\
&= \frac{K_{h+\varepsilon g}(h+\varepsilon g)h - (h+\varepsilon g)K_h h}{h^2} \\
&= \frac{(K_h + \varepsilon K + \mathrm{O}(\varepsilon^2))(h+\varepsilon g)h - (h+\varepsilon g)K_h h}{h^2} \\
&= \varepsilon K + \mathrm{O}\big(\varepsilon^2\big),
\end{aligned}$$

we have

$$\begin{aligned}
\mathcal{X}\left(\left(\frac{h+\varepsilon g}{h}\right)^{1/\varepsilon}\right) &= \frac{1}{\varepsilon}\left(\frac{h+\varepsilon g}{h}\right)^{1/\varepsilon}\left(1+\varepsilon\frac{h}{g}\right)^{-1}\mathcal{X}\left(\frac{h+\varepsilon g}{h}\right) \\
&= \frac{1}{\varepsilon}\left(\frac{h+\varepsilon g}{h}\right)^{1/\varepsilon}\big(1+\mathrm{O}(\varepsilon)\big)\big(\varepsilon K + \mathrm{O}\big(\varepsilon^2\big)\big) \\
&= \big(K + \mathrm{O}(\varepsilon)\big)\left(\frac{h+\varepsilon g}{h}\right)^{1/\varepsilon}. \qquad (11)
\end{aligned}$$

Therefore the function

$$\left(\frac{h+\varepsilon g}{h}\right)^{1/\varepsilon}$$

has cofactor $K+\mathrm{O}(\varepsilon)$. As ε tends to zero, the above expression tends to

$$\exp\left(\frac{g}{h}\right), \tag{12}$$

and from (11) we obtain that

$$\mathcal{X}\left(\exp\left(\frac{g}{h}\right)\right) = K\exp\left(\frac{g}{h}\right). \tag{13}$$

Therefore, function (12) satisfies the same Equation (9) that the invariant algebraic curves, with a cofactor of degree at most $m-1$.

Let $h, g \in \mathbb{C}[x, y]$ and assume that h and g are relatively prime in the ring $\mathbb{C}[x, y]$. Then the function $\exp(g/h)$ is called an *exponential factor* of the $\mathbb{F}$-polynomial system (1) if for some polynomial $K \in \mathbb{C}[x, y]$ of degree at most $m-1$ it satisfies Equation (13). As before we say that K is the *cofactor* of the exponential factor $\exp(g/h)$.

As we will see from the point of view of the integrability of polynomial systems (1) the importance of the exponential factors is double. On one hand, they verify Equation (13), and on the other hand, their cofactors are polynomials of degree at most $m-1$. These two facts will allow that they play the same role that the invariant algebraic curves in the integrability of a polynomial system (1). We note that the exponential factors do not define invariant curves for the flow of system (1).

We remark that in the definition of exponential factor $\exp(g/h)$ we always allow that this function is complex; that is $h, g \in \mathbb{C}[x, y]$. This is due to the same reason that in the case of invariant algebraic curves. That is, sometimes for real polynomial systems the existence of a real first integral can be forced by the existence of complex exponential factors. Again when we look for a complex exponential factor of a real polynomial system we are thinking in the real polynomial vector field as a complex one.

PROPOSITION 1.11. *For a real polynomial system* (1.9) *the function* $\exp(g/h)$ *is an exponential factor with cofactor K if and only if the function* $\exp(\bar{g}/\bar{h})$ *is an exponential factor with cofactor $\overline{K}$. Again conjugation denotes conjugation of the coefficients of the polynomials only.*

PROOF. We assume that $\exp(g/h)$ is an exponential factor with cofactor K of the real polynomial system (1). Then equality (13) holds. Since P and Q are real polynomials conjugating equality (13) we obtain that

$$P\frac{\partial\exp(\bar{g}/\bar{h})}{\partial x} + Q\frac{\partial\exp(\bar{g}/\bar{h})}{\partial y} = \overline{K}\exp(\bar{g}/\bar{h}).$$

Consequently, $\exp(\bar{g}/\bar{h})$ is an exponential factor with cofactor $\overline{K}$ of system (1). The proof in the converse direction is similar. □

PROPOSITION 1.12. *If $F = \exp(g/h)$ is an exponential factor for the polynomial system* (1)*, then $h = 0$ is an invariant algebraic curve, and g satisfies the equation*

$$\mathcal{X}g = gK_h + hK_F,$$

where K_h and K_F are the cofactors of h and F respectively.

PROOF. Since $F = \exp(g/h)$ is an exponential factor with cofactor K_F, we have

$$K_F \exp\left(\frac{g}{h}\right) = \mathcal{X} \exp\left(\frac{g}{h}\right) = \exp\left(\frac{g}{h}\right)\mathcal{X}\left(\frac{g}{h}\right) = \exp\left(\frac{g}{h}\right)\frac{(\mathcal{X}g)h - g(\mathcal{X}h)}{h^2},$$

or equivalently

$$(\mathcal{X}g)h - g(\mathcal{X}h) = h^2 K_F.$$

Hence, since h and g are relatively prime, we obtain that h divides $\mathcal{X}h$. So $h = 0$ is an invariant algebraic curve with cofactor $K_h = \mathcal{X}h/h$. Now substituting $\mathcal{X}h$ by $K_h h$ in the last equality, we have that $\mathcal{X}g = gK_h + hK_F$. □

We remark that the exponential factors of the form $\exp(g/h)$ with $h =$ constant appear when the straight line at infinity is a solution with multiplicity higher than 1 for the projectivized version of the vector field, for additional details see [34].

2. Darbouxian theory of integrability

As far as we know, the problem of integrating a polynomial system by using its invariant algebraic curves was started to be considered by Darboux in [38]. The version that we present improves Darboux's one essentially because here we also take into account the exponential factors (see [28]), the independent singular points (see [26]), and the invariants (see [9,11]).

Before stating the main results of the Darboux theory of integrability we need some definitions. If $S(x, y) = \sum_{i+j=0}^{m-1} a_{ij} x^i y^j$ is a polynomial of degree $m - 1$ with $m(m + 1)/2$ coefficients in $\mathbb{C}$, then we write $S \in \mathbb{C}_{m-1}[x, y]$. We identify the linear vector space $\mathbb{C}_{m-1}[x, y]$ with $\mathbb{C}^{m(m+1)/2}$ through the isomorphism $S \to (a_{00}, a_{10}, a_{01}, \dots, a_{m-1,0}, a_{m-2,1}, \dots, a_{0,m-1})$.

We say that r points $(x_k, y_k) \in \mathbb{C}^2$, $k = 1, \dots, r$, are *independent* with respect to $\mathbb{C}_{m-1}[x, y]$ if the intersection of the r hyperplanes

$$\left\{(a_{ij}) \in \mathbb{C}^{m(m+1)/2} \colon \sum_{i+j=0}^{m-1} x_k^i y_k^j a_{ij} = 0,\ k = 1, \dots, r\right\},$$

is a linear subspace of $\mathbb{C}^{m(m+1)/2}$ of dimension $m(m+1)/2 - r > 0$.

We recall that (x_0, y_0) is a singular point of system (1) if $P(x_0, y_0) = Q(x_0, y_0) = 0$.

We remark that the maximum number of isolated singular points of the polynomial system (1) is m^2 (by Bezout theorem), that the maximum number of independent isolated singular points of the system is $m(m+1)/2 - 1$, and that $m(m+1)/2 < m^2$ for $m \geqslant 2$.

A singular point (x_0, y_0) of system (1) is called *weak* if the divergence, $\operatorname{div}(P, Q)$, of system (1) at (x_0, y_0) is zero.

THEOREM 2.1 (Darbouxian theory of integrability). *Suppose that a $\mathbb{F}$-polynomial system* (1) *of degree m admits p irreducible invariant algebraic curves $f_i = 0$ with cofactors K_i for $i = 1, \dots, p$, q exponential factors* $\exp(g_j/h_j)$ *with cofactors L_j for $j = 1, \dots, q$, and r independent singular points (x_k, y_k) such that $f_i(x_k, y_k) \neq 0$ for $i = 1, \dots, p$ and for $k = 1, \dots, r$. Note that the irreducible factors of the polynomials h_j are some f_i's.*

(a) *There exist $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = 0$, if and only if the* (*multi-valued*) *function*

$$f_1^{\lambda_1} \cdots f_p^{\lambda_p} \left(\exp\left(\frac{g_1}{h_1}\right)\right)^{\mu_1} \cdots \left(\exp\left(\frac{g_q}{h_q}\right)\right)^{\mu_q} \tag{14}$$

is a first integral of system (1.9), *real if* $\mathbb{F} = \mathbb{R}$.

(b) *If $p + q + r = [m(m+1)/2] + 1$, then there exist $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = 0$.*

(c) *If $p + q + r \geqslant [m(m+1)/2] + 2$, then system* (1) *has a rational first integral, and consequently all trajectories of the system are contained in invariant algebraic curves.*

(d) *There exist $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = -\operatorname{div}(P, Q)$, if and only if the function* (14) *is an integrating factor of system* (1), *real if* $\mathbb{F} = \mathbb{R}$.

(e) *If $p + q + r = m(m+1)/2$ and the r independent singular points are weak, then function* (14) *for convenient $\lambda_i, \mu_j \in \mathbb{C}$ not all zero is a first integral if $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = 0$, or an integrating factor if $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = -\operatorname{div}(P, Q)$.*

(f) *There exist $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = -s$ for some $s \in \mathbb{F}\backslash\{0\}$, if and only if the function*

$$f_1^{\lambda_1} \cdots f_p^{\lambda_p} \left(\exp\left(\frac{g_1}{h_1}\right)\right)^{\mu_1} \cdots \left(\exp\left(\frac{g_q}{h_q}\right)\right)^{\mu_q} \exp(st) \tag{15}$$

is an invariant of system (1), *real if* $\mathbb{F} = \mathbb{R}$.

PROOF. We prove the theorem when the polynomial system (1) is complex. For the real case, we only do some minor comments.

We denote $F_j = \exp(g_j/h_j)$. By hypothesis we have p invariant algebraic curves $f_i = 0$ with cofactors K_i, and q exponential factors F_j with cofactors L_j. That is, the f_i's satisfy $\mathcal{X} f_i = K_i f_i$, and the F_j's satisfy $\mathcal{X} F_j = L_j F_j$.

The first part of statement (a) follows easily from the next computations

$$\begin{aligned}&\mathcal{X}\big(f_1^{\lambda_1}\cdots f_p^{\lambda_p}F_1^{\mu_1}\cdots F_q^{\mu_q}\big)\\&\quad=\big(f_1^{\lambda_1}\cdots f_p^{\lambda_p}F_1^{\mu_1}\cdots F_q^{\mu_q}\big)\left(\sum_{i=1}^{p}\lambda_i\frac{\mathcal{X}f_i}{f_i}+\sum_{j=1}^{q}\mu_j\frac{\mathcal{X}F_j}{F_j}\right)\\&\quad=\big(f_1^{\lambda_1}\cdots f_p^{\lambda_p}F_1^{\mu_1}\cdots F_q^{\mu_q}\big)\left(\sum_{i=1}^{p}\lambda_iK_i+\sum_{j=1}^{q}\mu_jL_j\right).\end{aligned}$$

In statement (a) we claim that the function (14) is real if $\mathcal{X}$ is real. This follows from the following fact. Since $\mathcal{X}$ is real, it is well known that if a complex invariant algebraic curve or exponential factor appears, then its conjugate must appear simultaneously (see Propositions 1.8 and 1.11). If among the invariant algebraic curves of $\mathcal{X}$ a complex conjugate pair $f=0$ and $\bar{f}=0$ occurs, the function (14) has a real factor of the form $f^{\lambda}\bar{f}^{\bar{\lambda}}$, which is the multi-valued real function

$$\big[(\operatorname{Re} f)^2+(\operatorname{Im} f)^2\big]^{\operatorname{Re}\lambda}\exp\big(-2\operatorname{Im}\lambda\arg(\operatorname{Re} f+i\operatorname{Im} f)\big),$$

if $\operatorname{Im}\lambda\operatorname{Im} f\not\equiv 0$. If among the exponential factors of $\mathcal{X}$ a complex conjugate pair $F=\exp(h/g)$ and $\overline{F}=\exp(\bar{h}/\bar{g})$ occurs, the first integral (14) has a real factor of the form

$$\left(\exp\left(\frac{h}{g}\right)\right)^{\mu}\left(\exp\left(\frac{\bar{h}}{\bar{g}}\right)\right)^{\bar{\mu}}=\exp\left(2\operatorname{Re}\left(\mu\frac{h}{g}\right)\right).$$

(b) Since the cofactors K_i and L_j are polynomials of degree $\leqslant m-1$, we have that $K_i, L_j\in\mathbb{C}_{m-1}[x,y]$. We note that the dimension of $\mathbb{C}_{m-1}[x,y]$ as a vector space over $\mathbb{C}$ is $m(m+1)/2$.

Since (x_k,y_k) is a singular point of system (1), $P(x_k,y_k)=Q(x_k,y_k)=0$. Then, from $\mathcal{X}f_i=P(\partial f_i/\partial x)+Q(\partial f_i/\partial y)=K_if_i$, it follows that $K_i(x_k,y_k)f_i(x_k,y_k)=0$. By assumption $f_i(x_k,y_k)\neq 0$, therefore $K_i(x_k,y_k)=0$ for $i=1,\dots,p$. Again, from $\mathcal{X}F_j=P(\partial F_j/\partial x)+Q(\partial F_j/\partial y)=L_jF_j$, it follows that $L_j(x_k,y_k)F_j(x_k,y_k)=0$. Since $F_j=\exp(g_j/h_j)$ does not vanish, $L_j(x_k,y_k)=0$ for $j=1,\dots,q$. Consequently, since the r singular points are independent, all the polynomial K_i and L_j belong to a linear subspace S of $\mathbb{C}_{m-1}[x,y]$ of dimension $[m(m+1)/2]-r$. We have $p+q$ polynomials K_i and L_j and since from the assumptions $p+q>[m(m+1)/2]-r$, we obtain that the $p+q$ polynomials must be linearly dependent on S. So, there are $\lambda_i,\mu_j\in\mathbb{C}$ not all zero such that $\sum_{i=1}^{p}\lambda_iK_i+\sum_{j=1}^{q}\mu_jL_j=0$. Hence statement (b) is proved.

(c) Since the number of independent singular points $r<m(m+1)/2$, it follows that $p+q>2$. Under the assumptions of statement (c) we apply statement (b) to two subsets of $p+q-1>0$ functions defining invariant algebraic curves or exponential factors. Thus, we get two linear dependencies between the corresponding cofactors, which after some linear algebra and relabeling, we can write into the following form

$$M_1+\alpha_3M_3+\cdots+\alpha_{p+q-1}M_{p+q}=0,$$

$$M_2 + \beta_3 M_3 + \cdots + \beta_{p+q-1} M_{p+q} = 0,$$

where M_l are the cofactors K_i and L_j, and the α_l and β_l are complex numbers. Then, by statement (a), it follows that the two functions

$$\log\bigl(G_1 G_3^{\alpha_3} \cdots G_{p+q}^{\alpha_{p+q}}\bigr), \qquad \log\bigl(G_2 G_3^{\beta_3} \cdots G_{p+q}^{\beta_{p+q}}\bigr),$$

are first integrals of system (1), where G_l is the polynomial defining an invariant algebraic curve or the exponential factor having cofactor M_l for $l = 1, \ldots, p+q$. Then, taking logarithms to the above two first integrals, we obtain that

$$\begin{aligned} H_1 &= \log(G_1) + \alpha_3 \log(G_3) + \cdots + \alpha_{p+q-1} \log(G_{p+q}), \\ H_2 &= \log(G_2) + \beta_3 \log(G_3) + \cdots + \beta_{p+q-1} \log(G_{p+q}), \end{aligned}$$

are first integrals of system (1). Each provides an integrating factor R_i such that

$$R_i P = \frac{\partial H_i}{\partial y}, \qquad R_i Q = -\frac{\partial H_i}{\partial x}.$$

Therefore, we obtain that

$$\frac{R_1}{R_2} = \frac{\partial H_1/\partial x}{\partial H_2/\partial x}.$$

Since the functions G_l are polynomials or exponentials of a quotient of polynomials, it follows that the functions $\partial H_i/\partial x$ are rational for $i = 1, 2$. So, from the last equality, we get that the quotient between the two integrating factors R_1/R_2 is a rational function. From Proposition 1.5 it follows statement (c).

(d) We have $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = -\operatorname{div}(P, Q)$. Then, from the computations of the proof of statement (a), we obtain

$$\begin{aligned} &\mathcal{X}\bigl(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q}\bigr) \\ &\quad = \bigl(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q}\bigr) \left(\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j \right) \\ &\quad = -\bigl(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q}\bigr) \operatorname{div}(P, Q). \end{aligned}$$

So, from (7), statement (d) follows.

(e) Let $K = \operatorname{div}(P, Q)$, clearly $K \in \mathbb{C}_{m-1}[x, y]$. By assumption the r singular points (x_k, y_k) are weak, therefore $K(x_k, y_k) = 0$ for $k = 1, \ldots, r$. So K belongs to the linear subspace S of the proof of statement (b).

On the other hand, since $\dim S = p + q = [m(m+1)/2] - r \geqslant 0$ and we have $p + q + 1$ polynomials $K_1, \ldots, K_p, L_1, \ldots, L_q, K$ in S (we are using the same arguments that in the proof of statement (b), it follows that these polynomials are linearly dependent on S. Therefore, we obtain $\lambda_i, \mu_j, \alpha \in \mathbb{C}$ not all zero such that

$$\left(\sum_{i=1}^{p} \lambda_i K_i\right) + \left(\sum_{j=1}^{q} \mu_j L_j\right) + \alpha K = 0. \tag{16}$$

If $\alpha = 0$ then, as in the proof of statement (a), we obtain that function (14) is a first integral of system (1).

We assume now that $\alpha \neq 0$. Dividing by α the equality (16) (if necessary), we can assume without loss of generality that $\alpha = 1$. So we have that

$$\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = -\operatorname{div}(P, Q).$$

Then, by using statement (d) it follows (e).

(f) We have $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = -s$. Then, from

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q} \mathrm{e}^{st}\right) \\
&\quad = \left(\mathcal{X} + \frac{\partial}{\partial t}\right)\left(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q} \mathrm{e}^{st}\right) \\
&\quad = \left(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q} \mathrm{e}^{st}\right)\left(\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j + s\right) = 0,
\end{aligned}$$

it follows statement (f). □

A (multi-valued) function of the form (14) is called a *Darbouxian* function. The associated first integral to a Darbouxian integrating factor is called a *Liouvillian* first integral.

3. Kapteyn–Bautin theorem

An interesting application of the Darbouxian theory of integrability allows us to present a new and shorter proof of the sufficient conditions for the classification theorem of centers of quadratic polynomial differential systems due to Kapteyn [59,60] and Bautin [6]. The first proof of this fact was due to Kapteyn in the 1910s. The proof that we present here can be found in [9].

KAPTEYN–BAUTIN THEOREM. *Any quadratic system candidate to have a center can be written in the form*

$$\dot{x} = -y - bx^2 - Cxy - dy^2, \qquad \dot{y} = x + ax^2 + Axy - ay^2. \tag{17}$$

This system has a center at the origin if and only if one of the following conditions holds

(I) $A - 2b = C + 2a = 0$,

(II) $C = a = 0$,

(III) $b + d = 0$,

(IV) $C + 2a = A + 3b + 5d = a^2 + bd + 2d^2 = 0$.

The following result gives a short proof of the sufficient conditions of Kapteyn–Bautin theorem.

THEOREM 4. *If system* (17) *satisfies one of the four conditions of the Kapteyn–Bautin theorem, then it has a center at the origin.*

PROOF. Since system (17) has a linear center at the origin, this is a center or a focus. Consequently, to prove that system (17) satisfying one of the four conditions of the Kapteyn–Bautin theorem has a center at the origin, it is sufficient to show that it has a first integral in a neighborhood of the origin.

Assume that system (17) *satisfies condition* (I). Then, it is easy to check that the system is Hamiltonian, i.e., $\dot{x} = -\partial H/\partial y$, $\dot{y} = \partial H/\partial x$ with $H = \frac{1}{2}(x^2 + y^2) + \frac{a}{3}x^3 + bx^2y - axy^2 + \frac{d}{3}y^3$. Therefore H is a first integral defined in a neighborhood of the origin.

Suppose that system (17) *satisfies condition* (II). Then, the system can be written in the form

$$\dot{x} = -y - bx^2 - dy^2, \qquad \dot{y} = x + Axy.$$

If $A \neq 0$ this system has the invariant straight line $f_1 = 1 + Ay = 0$ with cofactor $K_1 = Ax$. The divergence of the system is $(A - 2b)x$. Then, if $A(A - 2b) \neq 0$ we have the divergence of the system is equal to $(1 - \frac{2b}{A})K_1$. Hence, by Theorem 2.1(d) we obtain that $(1 + Ay)^{2b/A-1}$ is an integrating factor of system (17). Since this integrating factor is not zero at the origin, the associated first integral is defined in a neighborhood of the origin, and consequently the origin is a center.

We can assume that $A - 2b \neq 0$, otherwise we would be under the assumptions of condition (I). So, it remains only to study the case $A = 0$ and $b \neq 0$. Then, the system becomes $\dot{x} = -y - bx^2 - dy^2$, $\dot{y} = x$. This system has the algebraic solution $f_1 = 2b^2(bx^2 + dy^2) + (b - d)(2by - 1) = 0$ with cofactor $K_1 = -2bx$, which is equal to the divergence of the system. Therefore, by Theorem 2.1(d) we obtain that f_1^{-1} is an integrating factor. Hence the first integral associated to this integrating factor is defined at the origin if $b - d \neq 0$, and consequently the origin is a center.

Now we suppose that in addition $b - d = 0$. Then the system goes over to $\dot{x} = -y - b(x^2 + y^2)$, $\dot{y} = x$. From Examples 1.1 and 1.3 we know that $H = \exp(2by)(x^2 + y^2)$ is a first integral, which is defined at the origin, and therefore the origin must be a center.

Assume that system (17) *satisfies condition* (III). As Frommer observed in [44] (see also [92]) the form of system (17) with $b + d = 0$ is preserved under a rotation of axes. After performing a rotation of axes of an angle θ, the new coefficient a' of x^2 in the second equation of system (13) becomes of the form $a' = a\cos^3\theta + \alpha\cos^2\theta\sin\theta + \beta\cos\theta\sin^2\theta + d\sin^3\theta$.

Therefore, if $a \neq 0$ we can find θ such that $a' = 0$. So we can assume that $a = 0$, and consequently $C \neq 0$; otherwise we would be under the assumptions of condition (II).

The system

$$\dot{x} = -y - bx^2 - Cxy + by^2, \qquad \dot{y} = x + Axy,$$

has the algebraic solutions $f_1 = 1 + Ay = 0$ if $A \neq 0$ with cofactor $K_1 = Ax$, and $f_2 = (1 - by)^2 + C(1 - by)x - b(A + b)x^2 = 0$ with cofactor $K_2 = -2bx - Cy$. Since the divergence of the system is equal to $K_1 + K_2$, by Theorem 2.1 (d) we obtain that $f_1^{-1} f_2^{-1}$ is an integrating factor. Hence, again the first integral associated to the integrating factor is defined at the origin, and consequently the origin is a center.

We remark that if $A = 0$ then f_1 is not an algebraic solution of the system, but then the divergence of the system is equal to K_2 and the integrating factor of the system is f_2^{-1}, and using the same arguments we obtain that the origin is a center.

Suppose that system (17) *satisfies condition* (IV). Then, if $d \neq 0$ the system becomes

$$\begin{aligned} \dot{x} &= -y + \frac{a^2 + 2d^2}{d} x^2 + 2axy - dy^2, \\ \dot{y} &= x + ax^2 + \frac{3a^2 + d^2}{d} xy - ay^2. \end{aligned}$$

We note that if $d = 0$ then we are under the assumptions of condition (II). This system has the algebraic solution $f_1 = (a^2 + d^2)[(dy - ax)^2 + 2dy] + d^2 = 0$ with cofactor $K_1 = 2(a^2 + d^2)x/d$. Therefore the divergence of the system is equal to $\frac{5}{2}K_1$. Hence, by Theorem 2.1(d) the function $f_1^{-5/2}$ is an integrating factor of the system. Since $d \neq 0$, its associated first integral is defined in a neighborhood of the origin, and consequently the origin is a center. □

In order to see the explicit expression for the first integrals of system (17) satisfying conditions (I)–(IV), see Schlomiuk [92].

4. On the degree of the invariant algebraic curves

From Jouanolou's result (see Theorem 2.1(c))) it follows that for a given polynomial differential system of degree m the maximum degree of its irreducible invariant algebraic curves is bounded, since either it has a finite number $p < [m(m + 1)/2] + 2$ of invariant algebraic curves, or all its trajectories are contained in invariant algebraic curves and the system admits a rational first integral. Thus for each polynomial system there is a natural number N which bounds the degree of all its *irreducible* invariant algebraic curves. A natural question, going back to Poincaré [86], is to give an effective procedure to find N. Partial answers to this question were given by Cerveau and Lins Neto [16], Carnicer [15], Campillo and Carnicer [12], and Walcher [101]. These results depend on either restricting the nature of the polynomial differential system, or more specifically on the singularities of its invariant algebraic curves.

Of course, given such a bound, it is then easy to compute the algebraic curves of the system and also describe its elementary or Liouvillian first integrals (modulo any exponential factors) see for instance [73,28,84].

Unfortunately, for the class of polynomial systems with fixed degree m, there does not exist a uniform upper bound $N(m)$ for N as shown by the system:

$$\dot{x} = rx, \qquad \dot{y} = sy,$$

with r and s be positive integers. This system has a rational first integral

$$H = \frac{y^r}{x^s}$$

and hence invariant algebraic curves $x^s - hy^r = 0$ for all $h \in \mathbb{C}$.

A common suggestion was that the following question would have a positive answer:

> *Given $m \geqslant 2$, is there a positive integer $M(m)$ such that if a polynomial vector field of degree m has an irreducible invariant algebraic curve of degree $\geqslant M(m)$, then it has a rational first integral.*

See for instance the open question 2 in [31], or the question at the end of the introduction of [65].

The purpose of this section is to present two families of polynomial differential systems of degree 2 without rational first integrals but with irreducible invariant algebraic curves of arbitrarily high degree. Thus we show that no such function $M(m)$ exists.

THEOREM 4.1. *We consider the quadratic polynomial differential system*

$$\begin{aligned} \dot{x} &= x(1-x), \\ \dot{y} &= -\lambda y + Ax^2 + Bxy + y^2, \end{aligned} \tag{18}$$

where

$$\lambda = c - 1, \qquad A = \frac{ab(c-a)(c-b)}{c^2}, \qquad B = a + b - 1 - \frac{2ab}{c}.$$

If, for any positive integer k, we choose $a = 1 - k$, $b \leqslant a$, and c irrational, then the polynomial system (18) *has*:

(a) *no rational first integrals*;

(b) *an irreducible invariant algebraic curve*

$$\left(y - \frac{ab}{c}x\right)F + x(1-x)F' = 0,$$

of degree k, where $F = F(a, b; c; x)$ is the hypergeometric function;

(c) *a Darbouxian integrating factor.*

PROOF. We first prove statement (a). We shall use the following result known by Poincaré [86]: *If a polynomial system* (1.9) *has a rational first integral, then the eigenvalues* λ_1 *and* λ_2 *associated to any singular point of the system must be resonant in the following sense*: *there exist nonnegative integers* m_1 *and* m_2 *with* $m_1 + m_2 \geqslant 1$ *such that* $m_1\lambda_1 + m_2\lambda_2 = 0$. For a proof see, for instance, [46] or [64]. Now it is easy to check that the origin of system (18) is a singular point whose ratio of eigenvalues is $1 - c$. Since c is irrational, the system can have no rational first integral.

The system (18) has the invariant solutions $x = 0$ and $x = 1$ together with the following two explicit solutions:

$$f_1 = \left(y - \frac{ab}{c}x\right)F_1 + x(1-x)F_1' = 0,$$
$$f_2 = \left(c - 1 - y + \left(1 - c + \frac{ab}{c}\right)x\right)F_2 - x(1-x)F_2' = 0,$$

where $F_1 = F(a, b; c; x)$, $F_2 = F(1 + a - c, 1 + b - c; 2 - c; x)$, and $F(a, b; c; x)$ is the hypergeometric function

$$F(a, b; c; x) = \sum_{k=0}^{\infty} \frac{(a)_k(b)_k}{(c)_k}\frac{x^k}{k!}.$$

Here we have used the notation

$$(a)_k = \begin{cases} 1, & \text{if } k = 0, \\ a(a+1)(a+2)\cdots(a+k-1), & \text{if } k > 0. \end{cases}$$

The hypergeometric function $F(a, b; c; x)$ is a solution of the hypergeometric differential equation

$$x(1-x)y'' + \left[c - (a+b+1)x\right]y' - aby = 0,$$

and so F and F' can have a common zero only at $x = 1$.

If we take $a = 1 - k$, $b \leqslant a$ and $c > 0$ irrational, with k a positive integer, then $F(a, b; c; x)$ is a polynomial of degree $k - 1$, with all coefficients positive. Hence $F(1) > 0$, and so F and F' can have no common roots at all. Thus $f_1 = 0$ defines an irreducible algebraic curve of degree k, which proves statement (b).

Finally, after some calculation it can be shown that

$$H = \frac{x^{c-1} f_1}{f_2}$$

is a first integral of the system, with (reciprocal) integrating factor

$$x^c f_1^2 (1-x)^{a+b+1-c}$$

which proves statement (c). ☐

Theorem 4.1 can be found in [32].

Moulin Ollagnier in [75] showed that the quadratic spatial homogeneous Lotka–Volterra systems in $\mathbb{C}^3$:

$$\begin{aligned}\dot{x} &= x(Cy+z),\\ \dot{y} &= y(x+Az),\\ \dot{z} &= z(Bx+y),\end{aligned} \tag{19}$$

with

$$(A,B,C) = \bigl(-(2r+1)/(2r-1), 1/2, 2\bigr)$$

for $r = 1, 2, \ldots$ has an invariant algebraic curve $f(x,y,z) = 0$ of degree $2r$. So these systems present invariant algebraic curves of unbounded degree. He added: *It is not easy to write f in "closed form"*.

This previous example is interesting because it was used by Moulin Ollagnier in [78] to provide another negative answer to the question about the existence of the number $M(m)$.

Now we want to understand the algebraic structure of the invariant algebraic curves of system (19), and at the same time we present a new proof of the existence of such invariant algebraic curves. The proof of the existence presented here is essentially analytical, while the proof of the existence given in [77,78] is algebraic.

The homogeneous Lotka–Volterra systems (19) in $\mathbb{C}^3$ can be thought as the planar projective model of the following planar Lotka–Volterra systems in $\mathbb{C}^2$:

$$\begin{aligned}\dot{x} &= x\left(1-\frac{x}{2}+y\right),\\ \dot{y} &= y\left(-\frac{2r+1}{2r-1}+\frac{x}{2}-y\right).\end{aligned} \tag{20}$$

For more details between the affine and the projective model of a planar polynomial vector fields, see for instance [77].

Let $f(x,y) \in \mathbb{C}[x,y]$. If $f = 0$ is an *invariant algebraic curve* of the polynomial differential system (20), then f satisfies the following linear partial differential equation

$$x\left(1-\frac{x}{2}+y\right)\frac{\partial f}{\partial x} + y\left(-\frac{2r+1}{2r-1}+\frac{x}{2}-y\right)\frac{\partial f}{\partial y} = Kf, \tag{21}$$

where the polynomial K of degree $\leqslant 1$ is the cofactor of $f = 0$.

THEOREM 4.2. *For any positive integer r system* (20) *has an invariant algebraic curve $f(x,y) = 0$ of degree $2r$ of the following form*

$$\begin{aligned}f(x,y) = {} & x^r y^r + \bigl(a_{r-1,r-1} + a_{r,r-1}x + a_{r+1,r-1}x^2\bigr)x^{r-1}y^{r-1}\\ & + \bigl(a_{r-2,r-2} + a_{r-1,r-2}x + a_{r,r-2}x^2\bigr)x^{r-2}y^{r-2}\\ & + \cdots\end{aligned}$$

$$+\big(a_{1,1}+a_{2,1}x+a_{3,1}x^2\big)xy$$
$$+\big(a_{0,0}+a_{1,0}x+a_{2,0}x^2\big), \tag{22}$$

with cofactor $-x$. Moreover, an effective algorithm for the computation of $f(x,y)$ is given in the proof.

PROOF. For $r=1,2$ the following two polynomials

$$f(x,y)=xy-\frac{1}{2}\big(4-4x+x^2\big),$$
$$f(x,y)=x^2y^2+\frac{1}{6}\big(4-8x+3x^2\big)xy-\frac{1}{9}\big(4-4x+x^2\big),$$

provide, respectively, the invariant algebraic curves stated in the theorem. So, in the rest of the proof we assume that $r\geqslant 3$.

We consider the polynomial $f(x,y)$ written in the form $f(x,y)=f_0(x)+f_1(x)y+\cdots+f_r(x)y^r$. Since $f=0$ must be an invariant algebraic curve it must satisfy

$$\begin{aligned}&\left[x\left(1-\frac{x}{2}\right)+xy\right]\big(f_0'(x)+f_1'(x)y+\cdots+f_r'(x)y^r\big)\\&\quad+\left[y\left(-\frac{2r+1}{2r-1}+\frac{x}{2}\right)-y^2\right]\big(f_1(x)+2f_2(x)y+\cdots+rf_r(x)y^{r-1}\big)\\&\quad=-x\big(f_0(x)+f_1(x)y+\cdots+f_r(x)y^r\big),\end{aligned} \tag{23}$$

where we are forcing that the cofactor of $f=0$ is $-x$.

Equaling to zero the coefficients of the different powers y^k in the expression (23) for $k=r+1,r,\ldots,1,0$, we obtain that the polynomials $f_j(x)$ for $j=0,1,\ldots,r$ must satisfy the following ordinary differential system

$$\begin{aligned}&xf_r'(x)-rf_r(x)=0,\\&xf_j'(x)-jf_j(x)=p_j(x),\quad \text{for } j=r-1,r-2,\ldots,1,0,\\&\left(1-\frac{x}{2}\right)f_0'(x)+f_0(x)=0,\end{aligned} \tag{24}$$

where

$$p_j(x)=-x\left(1-\frac{x}{2}\right)f_{j+1}'(x)-\left[(j+1)\left(-\frac{2r+1}{2r-1}+\frac{x}{2}\right)+x\right]f_{j+1}(x).$$

Using the first $r+1$ equations of system (24) we can determine recursively $f_j(x)$ starting with $f_r(x)$ and ending with $f_0(x)$ as follows. After, we must verify that the function $f_0(x)$ obtained in this way satisfies the last equation of (24). Now, we will do these computations carefully.

The general solution of the first equation of system (24) is $f_r(x) = c_r x^r$, where c_r is the integration constant. Since we are looking for invariant algebraic curves $f(x, y)$ of the form (22), we take $c_r = 1$. Then, it easy to see that the general solution of the differential equation $xf_j'(x) - jf_j(x) = p_j(x)$ is

$$f_j(x) = x^j \int \frac{p_j(x)}{x^{j+1}} \, \mathrm{d}x. \tag{25}$$

So, solving (25) for $j = r - 1$, we get

$$f_{r-1} = \left[c_{r-1} + \frac{2r}{2r-1} x - \frac{1}{2} x^2 \right] x^{r-1},$$

where, in what follows, the integration constant which appears solving (25) is denoted by c_j. Solving again (25) but now for $j = r - 2$, we obtain

$$f_{r-2} = \left[c_{r-2} + \frac{2(r-1)}{2r-1} c_{r-1} x - \left[\frac{r}{(2r-1)^2} + \frac{1}{2} c_{r-1} \right] x^2 \right] x^{r-2}.$$

Now, we solve (8) for $j = r - 3$ and we have that f_{r-3} is equal to

$$c_{r-3} + \frac{2(r-2)c_{r-2}}{2r-1} x - \left[\frac{3(r-1)c_{r-1}}{(2r-1)^2} + \frac{c_{r-2}}{2} \right] x^2 \\ + \frac{2}{3} \left[\frac{r(r+1)}{(2r-1)^3} + \frac{c_{r-1}}{2r-1} \right] x^3$$

times x^{r-3}. In order to obtain for the polynomial $f(x, y)$ the structure given in the statement of the theorem, we choose the constant $c_{r-1} = -r(r+1)/(2r-1)^2$. Therefore, we get that

$$f_{r-3} = \left[c_{r-3} + \frac{2(r-2)c_{r-2}}{2r-1} x + \left[\frac{3r(r^2-1)}{(2r-1)^4} - \frac{c_{r-2}}{2} \right] x^2 \right] x^{r-3}.$$

We shall see that this structure that we have found computing $f_{r-3}(x)$, repeats until $f_0(x)$. More precisely, integrating Equation (25) for $j \leqslant r - 3$ and assuming that the expression of f_{j+1} is of the form

$$f_{j+1}(x) = \left[c_{j+1} + \alpha_{j+1} c_{j+2} x + \left[\beta_{j+1} - \frac{1}{2} c_{j+2} \right] x^2 \right] x^{j+1},$$

where the constants α and β depend on r, we have that

$$f_j(x) = \left[c_j + \frac{2(j+1)c_{j+1}}{2r-1} x + \left[\frac{(3 - 2(r-j))\alpha_{j+1} c_{j+2}}{2(2r-1)} - \frac{c_{j+1}}{2} \right] x^2 \right. \\ \left. + \frac{4\beta_{j+1}(2(1-r) + j) + (4(r-1) - (2r-1)\alpha_{j+1} - 2j)c_{j+2}}{6(2r-1)} x^3 \right] x^j.$$

For obtaining the structure of the polynomial given in the statement of the theorem, we choose the value for the constant c_{j+2} in order that the coefficient of x^{j+3} in the last equality becomes zero; i.e.,

$$c_{j+2} = -\frac{4\beta_{j+1}(2(1-r)+j)}{4(r-1)-(2r-1)\alpha_{j+1}-2j}.$$

We remark that the denominator of this expression never vanishes, otherwise j would be equal to $2r-2$. Hence, we can write $f_j(x)$ into the form

$$f_j(x) = \left[c_j + \alpha_j c_{j+1} x + \left[\beta_j - \frac{1}{2}c_{j+1}\right]x^2\right]x^j$$

for $j = r-1, r-2, \ldots, 1, 0$. Notice that $\alpha_j = 2(j+1)/(2r-1)$. Clearly, when we reach $f_0(x)$ we have chosen values for the integrating constants c_j for $j = r-1, r-2, \ldots, 2$, only the constants c_0 and c_1 remain arbitrary.

On the other hand, if we solve the last differential equation of system (24), we get that $f_0(x) = c(x-2)^2$, where c is the integration constant. The polynomial $f(x,y)$ that we are determining solving the differential system (24) provides an invariant algebraic curve of system (20), if the two expressions obtained by $f_0(x)$ coincide; i.e.,

$$c_0 + \alpha_0 c_1 x + \left[\beta_0 - \frac{1}{2}c_1\right]x^2 = c(x-2)^2.$$

Taking

$$c = \frac{\alpha_0\beta_0}{\alpha_0 - 2}, \quad c_0 = \frac{4\alpha_0\beta_0}{\alpha_0 - 2}, \quad \text{and} \quad c_1 = \frac{4\beta_0}{\alpha_0 - 2},$$

both expressions are equal. Notice $\alpha_0 - 2 = 4(1-r)/(2r-1) \neq 0$. Hence, the proof of the theorem is completed. □

In the light of these two theorems, we have:

OPEN QUESTION 4.3. *There is some number $D(m)$ for which any polynomial differential system of degree m having some irreducible invariant algebraic curve of degree $> D(m)$ has a Darbouxian first integral or Darbouxian integrating factor.*

5. Darboux lemma

Darboux was the first to give the following relation of enumerative geometry [38, pp. 83–84]:

> On peut rattacher cette recherche à un lemme relatif à six polynômes A, A', B, B', C, C', de degrés l, l', m, m', n, n' satisfaisant à l'identité déjà considérée
>
> (48) $\qquad AA' + BB' + CC' = 0;$

> il est évident que les degrés des produits AA', BB', CC' sont égaux.
> On a donc déjà
>
> $$l + l' = m + m' = n + n' = \lambda.$$
>
> Cela posé, je dis que *la somme du nombre des points communs aux trois courbes*
>
> $$A = 0, \qquad B = 0, \qquad C = 0,$$
>
> *et du nombre des points communs aux trois courbes*
>
> $$A' = 0, \qquad B' = 0, \qquad C' = 0,$$
>
> *est égale à*
>
> $$\frac{lmn + l'm'n'}{\lambda}.$$

We will refer to this result as the *Darboux lemma*; it can be stated more precisely as follows.

DARBOUX LEMMA. *Let $\mathbb{K}$ be an algebraically closed field and let A, B, C, A', B', C' be six homogeneous polynomials of degrees l, m, n, l', m', n' in three variables with coefficients in $\mathbb{K}$ such that*:

(i) *A, B, C are relatively prime and so are A', B', C',*

(ii) *$l + l' = m + m' = n + n' = r$ and the* orthogonality *relation holds*:

$$AA' + BB' + CC' = 0, \tag{26}$$

(iii) *the homogeneous ideal (A, B, C, A', B', C') generated by all six polynomials has no zero in the projective plane $\mathbb{P}_2(\mathbb{K})$.*

Then, the homogeneous ideals generated by the triples (A, B, C) and (A', B', C') have only finitely many zeroes in the projective plane.

Denoting by h and h' the total multiplicities $I(A, B, C)$ and $I(A', B', C')$ of these homogeneous ideals in the projective plane, there is a relation between h, h' and the degrees:

$$h + h' = \frac{lmn + l'm'n'}{r} = r^2 - r(l + m + n) + (lm + mn + nl). \tag{27}$$

Darboux started the proof of his result as follows

> En effet, soient h le nombre des points communs aux trois courbes A, B, C; h' celui des points communs aux trois courbes $A', B', C', \ldots$

This original proof is wrong; in particular, Darboux paid little attention to the last hypothesis (no common zeroes) and a counterexample is easy to find. Jouanolou noticed that Darboux's result was wrong and established clearly the formula (27) in his book [58, pp. 183–184], but his proof is far from being elementary. Jouanolou uses Chern's classes. Here, we present one of the two simple proofs given in [24]. Later on we shall apply Darboux lemma to polynomial differential equations.

In the proof that we present here we use standard facts about the intersection index of plane algebraic curves. This proof is divided into two steps. The first step deals with a special case of the statement under extra assumptions and it follows the ideas of Darboux. The second step consists in reducing the general case to the special case in order to get the complete result.

In this section $\mathbb{K}$ will denote an algebraically closed field. It will be convenient to denote an ideal generated in some polynomial ring over $\mathbb{K}$ by the elements $A_1, \dots, A_k$ simply by $(A_1, \dots, A_k)$. Other notations and definitions will be given when they become necessary.

We first recall some standard facts about homogeneous ideals of the polynomial ring $\mathbb{A} = \mathbb{K}[x_0, \dots, x_t]$, that have only a finite number of zeroes in the projective space $\mathbb{P}_t(\mathbb{K})$.

Let P be a point in the projective space $\mathbb{P}_t(\mathbb{K})$. The *local ring* $\mathcal{O}_P$ can be defined in two ways. First, it is the subring of the field $\mathbb{K}(x_0, \dots, x_t)_0$ of homogeneous rational fractions of degree 0 consisting of all those with a denominator that does not vanish at P. Second, one of the projective coordinates of P does not vanish and there is no restriction in supposing $x_0(P) \neq 0$ to fix matters. Then the polynomial ring $\mathbb{K}[x_1, \dots, x_t]$ is isomorphic to the quotient ring of $\mathbb{K}[x_0, \dots, x_t]$ by its ideal generated by $x_0 - 1$, and $\mathcal{O}_P$ is the local ring $S^{-1}\mathbb{K}[x_1, \dots, x_t]$, where S is the multiplicative set of all t-variable polynomials that do not vanish at P.

Let P be a point of $\mathbb{P}_t(\mathbb{K})$ and let $\mathcal{I}$ be a homogeneous ideal of $\mathbb{A}$. The ideal $\mathcal{I}_P$ is the ideal of the local ring $\mathcal{O}_P$ generated by $\mathcal{I}$. If the quotient ring $\mathcal{O}_P/\mathcal{I}_P$ is a finite-dimensional vector space over $\mathbb{K}$, its dimension is called the *multiplicity* or the *intersection index* of $\mathcal{I}$ at P; $I_P(\mathcal{I})$ is a convenient notation for this number. In particular $I_P(\mathcal{I}) \neq 0$ means that P is a zero of $\mathcal{I}$. Thus, if $\mathcal{I}$ is a homogeneous ideal with a finite number of zeroes in $\mathbb{P}_t(\mathbb{K})$, the sum $I(\mathcal{I}) = \sum_P I_P(\mathcal{I})$ over all zeroes of $\mathcal{I}$ is well defined. It is called the *total multiplicity*, or the *total intersection index* or the *degree* of $\mathcal{I}$.

In the case of two or more *three-variable* homogeneous polynomials over $\mathbb{K}$, $A_1, \dots, A_k$ the intersection index $I_P(A_1, \dots, A_k)$ at some point P of $\mathbb{P}_2(\mathbb{K})$ can be defined as the corresponding index for the homogeneous ideal generated by $A_1, \dots, A_k$. We note $I_P(A_1, \dots, A_k) = I_P((A_1, \dots, A_k))$.

In particular, if $A_1, \dots, A_k$ (with $k \geqslant 2$) are relatively prime, then $I_P(A_1, \dots, A_k)$ is defined at every point P. On the other hand if they have a nontrivial greatest common divisor D, $I_P(A_1, \dots, A_k)$ is defined at all points P of $\mathbb{P}_2(\mathbb{K})$ where $D(P) \neq 0$.

Here are some standard properties of the intersection index whose proof can be found in the book of Fulton [45]. The first two are general:

(i) $I_P(A_1, \dots, A_k)$ only depends on the ideal $(A_1, \dots, A_k)$ of $\mathbb{A}$,

(ii) in fact, $I_P(A_1, \dots, A_k)$ only depends on the ideal generated by $A_1, \dots, A_k$ in $\mathcal{O}_P$: if $B(P) \neq 0$, B is invertible in $\mathcal{O}_P$ and $I_P(BA_1, A_2, \dots, A_k) = I_P(A_1, A_2, \dots, A_k)$.

The next two ones are specific to the three-variable case:

(iii) If B has no nontrivial common factor with CC', $I_P(B, CC') = I_P(B, C) + I_P(B, C')$ (Addition formula).

(iv) If F and G are two homogeneous polynomials without nontrivial common factor, they have a finite number of common projective zeroes and $I(F, G) = \deg(F) \cdot \deg(G)$ (Bézout's theorem).

We will call a family $[A, B, C, A', B', C']$ of homogeneous polynomials in $\mathbb{K}[x, y, z]$ an *orthogonal system of polynomials* if

(i) A, B, C are relatively prime, and so are A', B', C',

(ii) $\deg(A) + \deg(A') = \deg(B) + \deg(B') = \deg(C) + \deg(C') = \rho(A, B, C, A', B', C')$, in which case $\rho(A, B, C, A', B', C')$ is called the *degree* of the system,

(iii) the *orthogonality* condition (26) $AA' + BB' + CC' = 0$ holds.

We will say that an orthogonal system of polynomials $[A, B, C, A', B', C']$ is *without projective zero* if A, B, C, A', B', C' have no common zero in the projective plane.

If $[A, B, C, A', B', C']$ is an orthogonal system of polynomials, A, B, C are relatively prime and the total intersection index $I(A, B, C)$ is well-defined and so is $I(A', B', C')$.

Now denote the degrees of the polynomials A, B, C, A', B', C' by l, m, n, l', m', n', respectively, and the degree of the orthogonal system $\rho(A, B, C, A', B', C')$ by r, to simplify the discussion.

The ratio

$$\frac{lmn + l'm'n'}{r} = r^2 - r(l + m + n) + (lm + mn + nl)$$

is a well-defined positive integer, which is 0 when $r = 0$. We then denote

$$\begin{aligned} \Delta(A, B, C, A', B', C') &= I(A, B, C) + I(A', B', C') - \frac{lmn + l'm'n'}{r} \\ &= h + h' - \frac{lmn + l'm'n'}{r}, \end{aligned}$$

and call this difference the *gap* of the orthogonal system.

With these definitions, the Darboux lemma can be stated as

The gap is zero for an orthogonal system of polynomials without projective zero.

We first give the result under the additional assumption that all six polynomials are pairwise relatively prime except maybe A and A', B and B', C and C'; in other words, we suppose that there is no nontrivial common factor to AA', BB', CC', and we then say that the orthogonal system is *irreducible*. This proof follows the ideas of Darboux. We will afterwards reduce the general case to this special case.

In our opinion, it is convenient and nonconfusing to identify a homogeneous nonzero three-variable polynomial F with the projective planar curve $F = 0$ it defines. We thus follow the free intuitive notations of Darboux.

For instance, the notation "$P \in A \cap B$" means that the point P of $\mathbb{P}_2(\mathbb{K})$ is a common zero of the two homogeneous polynomials A and B and belongs to the intersection of the two curves $A = 0$ and $B = 0$, as well as the alternative notation "$A(P) = 0$ and $B(P) = 0$".

PROPOSITION 5.1. *Let $[A, B, C, A', B', C']$ be an irreducible orthogonal system of polynomials without projective zero. Then $\Delta(A, B, C, A', B', C') = 0$.*

PROOF. We first notice that $I_P(A, B, CC') = I_P(A, B, -AA' - BB') = I_P(A, B)$ at every point P of $\mathbb{P}_2(\mathbb{K})$. We want to prove the following equality at $P \in A \cap B$:

$$I_P(A, B) = I_P(A, B, CC') = I_P(A, B, C) + I_P(A, B, C'). \tag{28}$$

From the orthogonality relation (26), $P \in CC'$ and, by exchanging A, B, C by A', B', C', there is no restriction in supposing $C(P) = 0$.

If $C'(P) \neq 0$, $I_P(A, B, C') = 0$. According to Section 2, $I_P(A, B, C) = I_P(A, B, CC') = I_P(A, B)$ and equality (28) holds in this case.

We now suppose that $C'(P) = 0$. Since $A \cap A' \cap B \cap B' \cap C \cap C' = \emptyset$, either $A'(P) \neq 0$ or $B'(P) \neq 0$ and there is no restriction in supposing $A'(P) \neq 0$. Then, $I_P(A, B, C) = I_P(A'A, B, C) = I_P(B, C)$, $I_P(A, B, C') = I_P(A'A, B, C') = I_P(B, C')$, $I_P(A, B, CC') = I_P(A'A, B, CC') = I_P(B, CC')$.

As B has no nontrivial common factor with C or C', equality (28) follows from the addition formula of the intersection index: $I_P(B, CC') = I_P(B, C) + I_P(B, C')$. The use of the addition formula is the only place where the extra assumption that the orthogonal system is irreducible plays a role.

Summing intersection indices at all points $P \in A \cap B$, relation (28) leads to

$$\begin{aligned} lm = I(A, B) &= \sum_{P \in A \cap B} I_P(A, B) = \sum_{P \in A \cap B} I_P(A, B, CC') \\ &= \sum_{P \in A \cap B} I_P(A, B, C) + \sum_{P \in A \cap B} I_P(A, B, C') = h + (lm - h), \end{aligned} \tag{29}$$

which means that the total intersection index $I(A, B, C')$ is $lm - h$.

Similar considerations lead to the next two equalities.

The first one,

$$\begin{aligned} ln' = I(A, C') &= \sum_{P \in A \cap C'} I_P(A, C') = \sum_{P \in A \cap C'} I_P(A, BB', C') \\ &= \sum_{P \in A \cap C'} I_P(A, B, C') + \sum_{P \in A \cap C'} I_P(A, B', C') \\ &= (lm - h) + \big(ln' - (lm - h)\big), \end{aligned} \tag{30}$$

means that the total intersection index $I(A, B', C')$ is $ln' - lm + h$.

The second one,

$$\begin{aligned} m'n' = I(B', C') &= \sum_{P \in B' \cap C'} I_P(B', C') = \sum_{P \in B' \cap C'} I_P(AA', B', C') \\ &= \sum_{P \in B' \cap C'} I_P(A, B', C') + \sum_{P \in B' \cap C'} I_P(A', B', C') \\ &= ln' - lm + h + h', \end{aligned} \tag{31}$$

means that the total intersection index $I(A', B', C')$ is $m'n' - ln' + lm - h$. From this last result, we deduce $\Delta(A, B, C, A', B', C') = 0$. ☐

In order to prove the Darboux lemma in the general case, we need a way to reduce non-irreducible orthogonal systems to irreducible ones. Let $[A, B, C, A', B', C']$ be a nonirreducible orthogonal system of polynomials. The orthogonality relation easily implies that two polynomials of the same triple have a nontrivial great common divisor (gcd) and there is no restriction in supposing that $D = \gcd(A, B) \notin \mathbb{K}$ to describe what a reduction is.

According to the orthogonality relation, D also divides C' as it is coprime with C and we have $A = DA_1$, $B = DB_1$, $C' = DC'_1$, with $\gcd(A_1, B_1) = 1$. Then $[A_1, B_1, C, A', B', C'_1]$ is another orthogonal system of polynomials. If $[A, B, C, A', B', C']$ is without projective zero, so is $[A_1, B_1, C, A', B', C'_1]$. We say that $[A_1, B_1, C, A', B', C'_1]$ is a *one-step reduction* of $[A, B, C, A', B', C']$.

There are as many possible one-step reductions of an orthogonal system of polynomials as pairs of noncoprime polynomials of the same triple. Thus, after at most six successive one-step reductions, we get an irreducible orthogonal system that can be called *the complete reduction* of the original one.

The following lemma is then the key to deduce the general case of the Darboux lemma from the special case of irreducible orthogonal systems.

LEMMA 5.2. *Let* $[A, B, C, A', B', C']$ *be a nonirreducible orthogonal system of polynomials without projective zero and such that* $D = \gcd(A, B) \notin \mathbb{K}$. *If* $[A_1, B_1, C, A', B', C'_1]$ *is the corresponding one-step reduction of* $[A, B, C, A', B', C']$ *then* $\Delta(A, B, C, A', B', C') = \Delta(A_1, B_1, C, A', B', C'_1)$.

PROOF. Let us denote by s the degree of D, so that $\deg(A_1) = l_1 = l - s$, $\deg(B_1) = m_1 = m - s$, $\deg(C'_1) = n'_1 = n' - s$, $\rho(A_1, B_1, C, A', B', C'_1) = r_1 = r - s$. h_1 will stand for $I(A_1, B_1, C)$ and h'_1 for $I(A_1, B_1, C)$.

With these notations, after straightforward cancellations, proving the announced result amounts to proving

$$(h - h_1) + (h' - h'_1) = ns. \tag{32}$$

This relation (32) will come from $h - h_1 = ns$ and $h' - h'_1 = 0$.

We first show $h - h_1 = ns$ by proving the following equality for all $P \in A \cap B \cap C$:

$$I_P(A, B, C) = I_P(A_1, B_1, C) + I_P(D, C). \tag{33}$$

Let P belong to $D \cap C$. Since $A \cap A' \cap B \cap B' \cap C \cap C' = \emptyset$, either $A'(P) \neq 0$ or $B'(P) \neq 0$ and there is no restriction in supposing $A'(P) \neq 0$.

Let then D' be the greatest common divisor of B and C: $B = D'B_2, C = D'C_2$. As A, B, C are relatively prime, D' is relatively prime with D and it divides B_1: $B_1 = D'B_3$; it also divides A', thus $D'(P) \neq 0$. Then, $I_P(B, C)$ and $I_P(B_1, C)$ are well-defined and the following equalities hold:

$$I_P(A, B, C) = I_P(AA', B, C) = I_P(B, C),$$
$$I_P(A_1, B_1, C) = I_P(A_1A', B_1, C) = I_P(B_1, C).$$

Now, the addition formula of the intersection index gives

$$I_P(B,C)=I_P(B_2,C_2)=I_P(B_3,C_2)+I_P(D,C_2)=I_P(B_1,C)+I_P(D,C).$$

Thus, for a P in $D\cap C$, equality (33) holds.

Consider now a point P in $A\cap B\cap C$ that does not belong to D. Then $I_P(D,C)=0$ whereas $I_P(A,B,C)=I_P(A_1,B_1,C)$ and equality (33) also holds.

Summing now equality (33) over all $P\in A\cap B\cap C$ gives

$$h=I(A,B,C)=I(A_1,B_1,C)+I(D,C)=h_1+sn, \tag{34}$$

according to Bézout's theorem.

Now we show that $h'-h'_1=0$ by proving $I_P(A',B',C')=I_P(A',B',C'_1)$ at all points P of $A'\cap B'\cap C'$.

If P belongs to $A'\cap B'\cap C'$ without being in D, then, we have that $I_P(A',B',C')=I_P(A',B',DC'_1)=I_P(A',B',C'_1)$.

If P belongs to $A'\cap B'\cap D$, then P does not belong to C and

$$\begin{aligned}I_P(A',B',C')&=I_P(A',B',CC')=I_P(A',B')\\&=I_P(A',B',CC_1)=I_P(A',B',C'_1).\end{aligned}$$

□

We can now conclude.

COROLLARY 5.3 (Darboux lemma). *Let* $[A,B,C,A',B',C']$ *be an orthogonal system of polynomials without projective zero. Then* $\Delta(A,B,C,A',B',C')=0$.

PROOF. If $[A_1,B_1,C_1,A'_1,B'_1,C'_1]$ is the complete reduction of $[A,B,C,A',B',C']$, then, according to Lemma 5.2, $\Delta(A,B,C,A',B',C')$ is the same as $\Delta(A_1,B_1,C_1,A'_1,B'_1,C'_1)$ and, according to Proposition 5.1, $\Delta(A_1,B_1,C_1,A'_1,B'_1,C'_1)=0$ for an irreducible orthogonal system of polynomials. □

6. Applications of the Darboux lemma

Ideas in the remaining part of this section go back to Darboux's work [38]. Let $p(x,y)$ and $q(x,y)$ be polynomials with complex coefficients. For the vector field

$$p\frac{\partial}{\partial x}+q\frac{\partial}{\partial y}, \tag{35}$$

or equivalently for the differential system

$$\dot x=p(x,y),\qquad \dot y=q(x,y), \tag{36}$$

we consider the associated differential 1-form $\omega_1 = q(x, y)\,\mathrm{d}x - p(x, y)\,\mathrm{d}y$, and the differential equation

$$\omega_1 = 0. \tag{37}$$

Clearly, Equation (37) defines a foliation with singularities on $\mathbb{C}^2$. The affine plane $\mathbb{C}^2$ is compactified on the complex projective space $\mathbb{CP}^2 = (\mathbb{C}^3 \setminus \{0\})/\sim$, where $(X, Y, Z) \sim (X', Y', Z')$ if and only if $(X, Y, Z) = \lambda(X', Y', Z')$ for some complex $\lambda \neq 0$. The equivalence class of (X, Y, Z) will be denoted by $[X : Y : Z]$.

The foliation defined by equation (37) on $\mathbb{C}^2$ can be extended to a singular foliation on $\mathbb{CP}^2$ and the 1-form ω_1 can be extended to a meromorphic 1-form ω on $\mathbb{CP}^2$ which yields an equation $\omega = 0$, i.e.,

$$A(X, Y, Z)\,\mathrm{d}X + B(X, Y, Z)\,\mathrm{d}Y + C(X, Y, Z)\,\mathrm{d}Z = 0, \tag{38}$$

whose coefficients A, B, C are homogeneous polynomials and satisfy the relation:

$$A(X, Y, Z)X + B(X, Y, Z)Y + C(X, Y, Z)Z = 0. \tag{39}$$

Indeed, consider the map $i : \mathbb{C}^3 \setminus \{Z = 0\} \to \mathbb{C}^2$, given by $i(X, Y, Z) = (X/Z, Y/Z) = (x, y)$ and suppose that $\max\{\deg(p), \deg(q)\} = m > 0$. Since $x = X/Z$ and $y = Y/Z$ we have:

$$\mathrm{d}x = (Z\,\mathrm{d}X - X\,\mathrm{d}Z)/Z^2, \qquad \mathrm{d}y = (Z\,\mathrm{d}Y - Y\,\mathrm{d}Z)/Z^2,$$

the pull-back form $i^*(\omega_1)$ has poles at $Z = 0$ and Equation (37) can be written as

$$\begin{aligned} i^*(\omega_1) &= q(X/Z, Y/Z)(Z\,\mathrm{d}X - X\,\mathrm{d}Z)/Z^2 - p(X/Z, Y/Z)(Z\,\mathrm{d}Y - Y\,\mathrm{d}Z)/Z^2 \\ &= 0. \end{aligned}$$

Then the 1-form $\omega = Z^{m+2} i^*(\omega_1)$ in $\mathbb{C}^3 \setminus \{Z \neq 0\}$ has homogeneous polynomial coefficients of degree $m + 1$, and for $Z = 0$ the equations $\omega = 0$ and $i^*(\omega_1) = 0$ have the same solutions. Therefore the differential equation $\omega = 0$ can be written as (38) where

$$\begin{aligned} A(X, Y, Z) &= ZQ(X, Y, Z) = Z^{m+1} q(X/Z, Y/Z), \\ B(X, Y, Z) &= -ZP(X, Y, Z) = -Z^{m+1} p(X/Z, Y/Z), \\ C(X, Y, Z) &= YP(X, Y, Z) - XQ(X, Y, Z). \end{aligned} \tag{40}$$

Clearly A, B and C are homogeneous polynomials of degree $m + 1$ satisfying (39).
Singular points in $\mathbb{P}_2(\mathbb{C})$ are the points satisfying $A = B = C = 0$.
A homogeneous 1-form

$$\omega = A\,\mathrm{d}X + B\,\mathrm{d}Y + C\,\mathrm{d}Z = 0,$$

where A, B and C are homogeneous polynomials of degree $m+1$ is called *projective* if $XA+YB+ZC=0$; that is, if there exist three homogeneous polynomials L, M and N of degree m such that

$$A=ZM-YN, \qquad B=XN-ZL, \qquad C=YL-XM;$$

or equivalently

$$(A,B,C)=(P,Q,R)\wedge(X,Y,Z).$$

Then we can write

$$\omega=P(Y\,\mathrm{d}Z-Z\,\mathrm{d}Y)+Q(Z\,\mathrm{d}X-X\,\mathrm{d}Z)+R(X\,\mathrm{d}Y-Y\,\mathrm{d}X). \tag{41}$$

The vector field

$$\mathcal{X}=P\frac{\partial}{\partial X}+Q\frac{\partial}{\partial Y}+R\frac{\partial}{\partial Z}, \tag{42}$$

can be thought as a *homogeneous polynomial vector field* of $\mathbb{C}^3$ of degree m associated to $\omega=0$.

In short, we have seen that the homogeneous vector field (42) of degree m in $\mathbb{C}^3$ with

$$P=Z^m p(X/Z,Y/Z), \qquad Q=Z^m q(X/Z,Y/Z), \qquad R=0, \tag{43}$$

is associated to the 1-form $\omega=0$, and consequently to the vector field (35).

PROPOSITION 6.1 (Darboux proposition). *For any homogeneous polynomial vector field* (42) *of degree m in $\mathbb{C}^3$ having finitely many singular points and satisfying* (43)*, we have that its number of singular points taking into account their multiplicities or numbers of intersection satisfies*

$$\sum_p I(p,P\cap Q\cap R)=m^2+m+1.$$

PROOF. Take $A=P$, $B=Q$, $C=R$, $A'=X$, $B'=Y$ and $C'=Z$. Then $AA'+BB'+CC'\equiv 0$. Since there are no common points to the curves $A'=0$, $B'=0$ and $C'=0$ it follows that

$$h'=\sum_p I(p,A'\cap B'\cap C')=0.$$

Therefore, from the Darboux lemma we obtain

$$h=\sum_p I(p,A\cap B\cap C)=\frac{(m+1)^3+1}{m+2}=m^2+m+1.$$

□

Let $f \in \mathbb{C}[x, y]$. The algebraic curve $f(x, y) = 0$ is an invariant algebraic curve of the affine polynomial vector field $\mathcal{X}$ given by (35) if for some polynomial $k \in \mathbb{C}[x, y]$ (the cofactor of $f = 0$), we have

$$\mathcal{X}f = \frac{\partial f}{\partial x}p + \frac{\partial f}{\partial y}q = kf.$$

It is easy to verify that if $f(x, y) = 0$ is an invariant algebraic curve of degree r for the polynomial vector field $\mathcal{X}$ with cofactor $k(x, y)$, then $F(X, Y, Z) = Z^r f(X/Z, Y/Z) = 0$ is an invariant algebraic curve of degree r for its projective vector field with cofactor $K(X, Y, Z) = Z^{m-1}k(X/Z, Y/Z)$; i.e.,

$$\mathcal{X}F = \frac{\partial F}{\partial X}P + \frac{\partial F}{\partial Y}Q + \frac{\partial F}{\partial Z}R = KF, \tag{44}$$

here $R = 0$.

If $F(X, Y, Z) = 0$ is an algebraic curve of $\mathbb{P}_2(\mathbb{C})$ of degree n. Let $p = (X_0, Y_0, Z_0)$ be a point of $\mathbb{P}_2(\mathbb{C})$. Since the three coordinates of p cannot be zero, without loss of generality we can assume that $p = (0, 0, 1)$. Then suppose that the expression of $F(X, Y, Z)$ restricted to $Z = 1$ is

$$F(X, Y, 1) = F_i(X, Y) + F_{i+1}(X, Y) + \cdots + F_n(X, Y),$$

where $0 \leqslant i \leqslant n$ and $F_j(X, Y)$ denotes a homogeneous polynomial of degree j in the variables X and Y for $j = i, \ldots, n$, with F_i different from the zero polynomial. We say that $i = m_p(F)$ is the *multiplicity* of the curve $F = 0$ at the point p. If $i = 0$ then the point p does not belong to the curve $F = 0$. If $i = 1$ we say that p is a *simple* point for the curve $F = 0$. If $i > 1$ we say that p is a *multiple* point.

PROPOSITION 6.2. *Let $f(x, y) = 0$ be an irreducible invariant algebraic curve of degree $n \geqslant 1$ without multiple points for the affine polynomial vector field $\mathcal{X}$ of degree m. Then $n \leqslant m + 1$.*

PROOF. Since $F = 0$ is an invariant algebraic curve of $\mathcal{X}$ with cofactor K we have that

$$\frac{\partial F}{\partial X}P + \frac{\partial F}{\partial Y}Q = KF$$

in $\mathbb{P}_2(\mathbb{C})$. By using Euler theorem for the homogeneous function F of degree n, this equation goes over to

$$\frac{\partial F}{\partial X}\left(P - \frac{1}{n}XK\right) + \frac{\partial F}{\partial Y}\left(Q - \frac{1}{n}YK\right) + \frac{\partial F}{\partial Z}\left(-\frac{1}{n}ZK\right) = 0. \tag{45}$$

Now we take in the Darboux lemma

$$A = \frac{\partial F}{\partial X}, \qquad B = \frac{\partial F}{\partial Y}, \qquad C = \frac{\partial F}{\partial Z},$$

$$A' = P - \frac{1}{n} XK, \qquad B' = Q - \frac{1}{n} YK, \qquad C' = -\frac{1}{n} ZK,$$

and

$$h = I(A \cap B \cap C), \qquad h' = I(A' \cap B' \cap C').$$

We note that by the assumptions h and h' are finite. Moreover, as $A \cap B \cap C = \emptyset$, $h = 0$.

Since A, A', B, B', C and C' satisfy equality (45), Darboux lemma can be used to get

$$h + h' = \frac{m^3 + (n-1)^3}{m + n - 1} = m^2 + (n-1)(n-m+1). \tag{46}$$

By the Bézout theorem, the number of intersection points of the curves $A' = 0$, $B' = 0$ and $C' = 0$ is at most m^2 taking into account their multiplicities; i.e., $h' \leqslant m^2$. Whence a lower bound for h: $0 = h \geqslant (n-1)(n-m-1)$, and $1 \leqslant n \leqslant m+1$. □

In [23] it has been proved the following result that provides sufficient conditions for the existence of a rational first integral.

THEOREM 6.3. *Let $f(x, y) = 0$ be an irreducible algebraic curve of degree $n > 1$, which is invariant with cofactor $k \neq 0$, for the affine polynomial vector field $\mathcal{X}$ of degree $m > 1$. If m^2 is the total number of solutions of the system*

$$nP - XK = 0, \qquad nQ - YK = 0, \qquad ZK = 0, \tag{47}$$

in the projective plane, taken into account their multiplicities or numbers of intersection, then $\mathcal{X}$ has a rational first integral.

PROOF. Since m^2 is the number of solutions of system (47), by using Bezout theorem it follows that the number of solutions of systems

$$nP - XK = 0, \qquad nQ - YK = 0, \qquad Z = 0; \tag{48}$$

and

$$P = 0, \qquad Q = 0, \qquad K = 0; \tag{49}$$

are m and $m^2 - m$ respectively. We note that always we take into account the number of solutions with their multiplicities.

We write

$$\begin{aligned}
P(X, Y, Z) &= p_0 Z^m + p_1(X, Y) Z^{m-1} + \cdots + p_m(X, Y), \\
Q(X, Y, Z) &= q_0 Z^m + q_1(X, Y) Z^{m-1} + \cdots + q_m(X, Y), \\
K(X, Y, Z) &= k_0 Z^{m-1} + k_1(X, Y) Z^{m-2} + \cdots + k_{m-1}(X, Y),
\end{aligned}$$

where $A_i(x,y)$ denotes a homogeneous polynomial of degree i and $A \in \{p,q,k\}$. Then, for $Z=0$ system (48) becomes

$$np_m(X,Y) - Xk_{m-1}(X,Y) = 0, \qquad nq_m(X,Y) - Yk_{m-1}(X,Y) = 0.$$

Since this system of homogeneous polynomials of degree m in the variables X and Y intersect at m points taking into account their multiplicities, there exists a homogeneous polynomial $A(X,Y)$ of degree m such that

$$\begin{aligned} p_m(X,Y) &= \lambda_1 A(X,Y) + \frac{1}{n} X k_{m-1}(X,Y), \\ q_m(X,Y) &= \lambda_2 A(X,Y) + \frac{1}{n} Y k_{m-1}(X,Y), \end{aligned}$$

where $\lambda_1, \lambda_2 \in \mathbb{C}$ are not zero.

The polynomial system associated to the vector field $\mathcal{X}$ can be written as

$$\begin{aligned} \dot{x} &= p_0 + p_1(x,y) + \cdots + p_{m-1}(x,y) + \lambda_1 A(x,y) + \frac{x}{n} k_{m-1}(x,y) \\ &= P(x,y,1), \\ \dot{y} &= q_0 + q_1(x,y) + \cdots + q_{m-1}(x,y) + \lambda_2 A(x,y) + \frac{1y}{n} k_{m-1}(x,y) \\ &= Q(x,y,1), \end{aligned}$$

or equivalently

$$\begin{aligned} \dot{x} &= p_0 + p_1 + \cdots + p_{m-1} - \frac{1}{n} x(k_0 + k_1 + \cdots + k_{m-2}) + \lambda_1 A + \frac{1}{n} xk, \\ \dot{y} &= q_0 + q_1 + \cdots + q_{m-1} - \frac{1}{n} y(k_0 + k_1 + \cdots + k_{m-2}) + \lambda_2 A + \frac{1}{n} yk. \end{aligned}$$

Since $\lambda_1^2 + \lambda_2^2 \neq 0$ (otherwise system (48) does not have m intersection points), without loss of generality we can assume that $\lambda_1 \neq 0$. We change from the variables (x,y) to the variables (x,z) where $z = \lambda_2 x - \lambda_1 y$. In the new variables the second equation of the polynomial system goes over to

$$\dot{z} = \lambda_2 P(x,y,1) - \lambda_1 Q(x,y,1) = b(x,y) + \frac{1}{n} zk(x,y),$$

with $y = (\lambda_2 x - z)/\lambda_1$ and where

$$\begin{aligned} b(x,y) = \lambda_2 &\left(p_0 + p_1 + \cdots + p_{m-1} - \frac{1}{n} x(k_0 + k_1 + \cdots + k_{m-2}) \right) \\ &- \lambda_1 \left(q_0 + q_1 + \cdots + q_{m-1} - \frac{1}{n} y(k_0 + k_1 + \cdots + k_{m-2}) \right). \end{aligned}$$

Since, from (49), the curves $\dot{x} = P(x, y, 1) = 0$, $\dot{y} = Q(x, y, 1) = 0$ and $K(x, y, 1) = k(x, y) = 0$ have $m^2 - m$ intersection points taking into account their multiplicities, it follows that the curves $P(x, y, 1) = 0$, $b(x, y) + k(x, y)z/n = 0$ and $k(x, y) = 0$ have the same number of intersection points, where $y = (\lambda_2 x - z)/\lambda_1$. Hence, the number of intersection points of the curves $b(x, y) = 0$ and $k(x, y) = 0$ is at least $m^2 - m$ points taking into account their multiplicities. But, from Bézout theorem, if these last two curves do not have a common component, then they have $(m - 1)^2$ intersection points taking into account their multiplicities. Since $m^2 - m > (m - 1)^2$ if $m > 1$, it follows that the curves $b = 0$ and $k = 0$ have a maximal common component $c = 0$ of degree $r \geqslant 1$. Therefore, $b = \bar{b}c$ and $k = \bar{k}c$ where $\bar{b}$ and $\bar{k}$ are polynomials of degree $m - r - 1$.

From (47) the number of intersection points of the curves $P(x, y, 1) = 0$ and $Q(x, y, 1) = 0$ with $k(x, y) = 0$ is maximal, i.e., $m^2 - m$, then the number of intersection points of the curves $P(x, y, 1) = 0$, $Q(x, y, 1) = 0$ and $\bar{k}(x, y) = 0$ is $m(m - r - 1)$, and the number of intersection points of the curves $P(x, y, 1) = 0$, $Q(x, y, 1) = 0$ and $c(x, y) = 0$ is mr. Of course, always we compute the number of intersection points taking into account their multiplicities. Since the number of intersection points of the curves $P(x, y, 1) = 0$, $b(x, y) + k(x, y)z/n = 0$ and $\bar{k}(x, y) = 0$ is $m(m - r - 1)$, it follows that the number of intersection points of the curves $b(x, y) = 0$ and $\bar{k}(x, y) = 0$ is at least $m(m - r - 1)$. On the other hand, since the curves $b = 0$ and $\bar{k} = 0$ has no common components, by Bézout theorem they intersect at $(m - 1)(m - r - 1)$ points taking into account their multiplicities. Hence, since $m > 1$ and $m(m - r - 1) > (m - 1)(m - r - 1)$ except if $r = m - 1$, we have that $r = m - 1$. Therefore $b = ak$ with $a \in \mathbb{C}$. So, $\dot{z} = k(a + z/n)$, and consequently $z + an = \lambda_2 x - \lambda_1 y + an = 0$ is an invariant straight line with cofactor k/n. Then, by statement (a) of Theorem 2.1, we obtain that $H = f(x, y)(\lambda_2 x - \lambda_1 y + an)^{-n}$ is a rational first integral of $\mathcal{X}$. We note that since $f = 0$ is different by a straight line, we have that $f \neq (\lambda_2 x - \lambda_1 y + an)^n$, and consequently H is a first integral. □

From the proof of Theorem 6.3 it follows that any polynomial vector field $\mathcal{X}$ in the assumptions of Theorem 6.3 has an invariant straight line $ax + by + c = 0$ such that a rational first integral of $\mathcal{X}$ is of the form

$$\frac{f(x, y)}{(ax + by + c)^n}.$$

Now we go back to study the number of multiple points that an invariant algebraic curve of degree n of a polynomial vector field of degree m can have in function of m and n.

PROPOSITION 6.4. *Let $f(x, y) = 0$ be an invariant algebraic curve of degree n of the polynomial vector field*

$$\mathcal{X} = p(x, y)\frac{\partial}{\partial x} + q(x, y)\frac{\partial}{\partial y}$$

of degree m with cofactor k. Assume that:

If

$$P(X,Y,Z)=Z^m p\left(\frac{X}{Z},\frac{Y}{Z}\right),$$
$$Q(X,Y,Z)=Z^m q\left(\frac{X}{Z},\frac{Y}{Z}\right),$$
$$K(X,Y,Z)=Z^{m-1} k\left(\frac{X}{Z},\frac{Y}{Z}\right),$$

then

(1) *the curves* $nP-XK=0$, $nQ-YK=0$, $ZK=0$, *do not have a common component, and*
(2) *the curve* $F(X,Y,Z)=Z^n f(X/Z,Y/Z)=0$ *has finitely many multiple points in* $\mathbb{P}_2(\mathbb{C})$ *taking into account their multiplicities, namely* h.

Then

$$(n-1)(n-m-1)\leqslant h\leqslant m^2+(n-1)(n-m-1).$$

PROOF. Since $F=0$ is an invariant algebraic curve of (44) with cofactor K we have that

$$\frac{\partial F}{\partial X}P+\frac{\partial F}{\partial Y}Q=KF$$

in $\mathbb{P}_2(\mathbb{C})$. By using Euler theorem for the homogeneous function F of degree n, this equation goes over to

$$\frac{\partial F}{\partial X}\left(P-\frac{1}{n}XK\right)+\frac{\partial F}{\partial Y}\left(Q-\frac{1}{n}YK\right)+\frac{\partial F}{\partial Z}\left(-\frac{1}{n}ZK\right)=0. \tag{50}$$

Now we take

$$A=\frac{\partial F}{\partial X},\qquad B=\frac{\partial F}{\partial Y},\qquad C=\frac{\partial F}{\partial Z},$$
$$A'=P-\frac{1}{n}XK,\qquad B'=Q-\frac{1}{n}YK,\qquad C'=-\frac{1}{n}ZK,$$

and

$$h=\sum_p I(p,A\cap B\cap C),\qquad h'=\sum_p I(p,A'\cap B'\cap C').$$

We note that by assumptions h and h' are finite.

Since A, A', B, B', C and C' satisfy equality (50), by Darboux lemma we obtain that

$$h+h'=\frac{m^3+(n-1)^3}{m+n-1}=m^2+(n-1)(n-m+1).$$

Therefore the upper bound for h given in the statement of the theorem is proved. By Bézout theorem, the number of intersection points of the curves $A' = 0$, $B' = 0$ and $C' = 0$ is at most m^2 taking into account their multiplicities; i.e., $h' \leqslant m^2$. Therefore, $h \geqslant (n-1)(n-m-1)$, and the proposition is proved. □

From Proposition 6.4 it follows as a corollary Proposition 6.2.

COROLLARY 6.5. *Under the assumptions of Proposition* 6.4*, if $f(x,y) = 0$ is an invariant algebraic curve of degree $n = m+1$ for the polynomial vector field $\mathcal{X}$ of degree $m > 1$ such that its projectivization $F(X,Y,Z) = Z^n f(X/Z, Y/Z) = 0$ has no multiple points, then $\mathcal{X}$ has a rational first integral.*

PROOF. Using the same notation than in the proof of Proposition 6.4 we have that $h' = m^2$, because from the assumptions we have that $n = m+1$ and $h = 0$. Now, since we are in the hypotheses of Theorem 6.3 the statement of the corollary follows. □

The last result in this section about rational first integrals is the following one.

PROPOSITION 6.6. *Under the assumptions of Proposition* 6.4*, if $f(x,y)$ is irreducible in $\mathbb{C}[x,y]$ and all the multiple points of $F(X,Y,Z) = 0$ are double and ordinary, then $n \leqslant 2m$. Moreover, if $n = 2m$ then $\mathcal{X}$ has a rational first integral.*

PROOF. We use the notation and the results introduced in the proof of Proposition 6.4. Since every multiple point p of $F(X,Y,Z) = 0$ is double and ordinary, it follows that

$$I\left(p, \frac{\partial F}{\partial X} \cap \frac{\partial F}{\partial Y} \cap \frac{\partial F}{\partial Z}\right) = 1. \tag{51}$$

If $F = 0$ is an irreducible algebraic curve of degree n, we know that

$$\sum_p \frac{1}{2} m_p(m_p - 1) \leqslant \frac{1}{2}(n-1)(n-2), \tag{52}$$

where p runs over the multiple points of $F = 0$ and m_p denotes the multiplicity of p, for a proof of this result see Section 4 of Chapter 5 of [46]. Since every multiple point p of $F = 0$ is double, $m_p = 2$. If h is the number of multiple points of $F = 0$ taking into account their multiplicity, from (51) and (52), it follows that $h \leqslant (n-1)(n-2)/2$. By Proposition 6.4, we have $(n-1)(n-m-1) \leqslant h$. Consequently

$$(n-1)(n-m-1) \leqslant \frac{1}{2}(n-1)(n-2).$$

Hence $n \leqslant 2m$, and the first part of the proposition is proved.

Now we assume that $n = 2m$. From the proof of Proposition 6.4 we have that $h + h' = m^2 + (n-1)(n-m-1) = m^2 + (2m-1)(m-1)$.Then since $h \leqslant (n-1)(n-2)/2 =$

$(2m-1)(m-1)$, we obtain that $h' \geqslant m^2$. But, by definition, $h' \leqslant m^2$. Hence, we are in the assumptions of Theorem 6.3, and consequently $\mathcal{X}$ has a rational first integral. ☐

7. Algebraic limit cycles for quadratic systems

We recall that a *limit cycle* of a real affine polynomial vector fields is an isolated periodic orbit in the set of all periodic orbits of the system. An *algebraic limit cycle* of degree r is an oval of an irreducible invariant algebraic curve $f(x,y)=0$ of degree r which is a limit cycle of the system.

In 1958 Qin Yuan-Xun [88] (see also [104]) proved that quadratic systems can have algebraic limit cycles of degree 2, moreover when such a limit cycle exists then it is the unique limit cycle of the system. Evdokimenco in [40–42] proved that quadratic systems do not have algebraic limit cycles of degree 3, for two different shorter proofs see [22,25]. We provide one of these proofs in what follows.

THEOREM 7.1. *Quadratic systems have no algebraic limit cycles of degree* 3.

PROOF. Let $f=0$ be an invariant algebraic curve of degree 3 of a real affine polynomial vector field of degree 2. If the cubic curve $f=0$ has multiple points, then it is rational (its genus is 0) and there is no oval in it. If $f=0$ has no multiple points, Equation (46) in the proof of Proposition 6.2 implies $h'=2^2=4$. According to Theorem 6.3, the system has a rational first integral and thus no limit cycle. ☐

The first class of algebraic limit cycles of degree 4 was given in 1966 by Yablonskii [103]. The second class was found in 1973 by Filiptsov [43]. Recently, two new classes has been found and in [25] the authors proved that there are no other algebraic limit cycles of degree 4 for quadratic systems. The uniqueness of these limit cycles has been proved in [21]. We summarize all these results into the following theorem, for a proof see the mentioned papers.

THEOREM 7.2. *The following statements hold.*

(a) *After an affine change of variables and a rescaling of the time variable the only quadratic systems having an algebraic limit cycle of degree* 2 *are*

$$\begin{aligned}\dot{x} &= -y(ax+by+c)-\left(x^2+y^2-1\right),\\ \dot{y} &= x(ax+by+c),\end{aligned} \tag{53}$$

with $c^2+4(b+1)<0$ *and* $a^2+b^2<c^2$. *This system possesses the irreducible invariant algebraic curve*

$$x^2+y^2-1=0,$$

of degree 2. *This algebraic limit cycle is the unique limit cycle of system* (18).

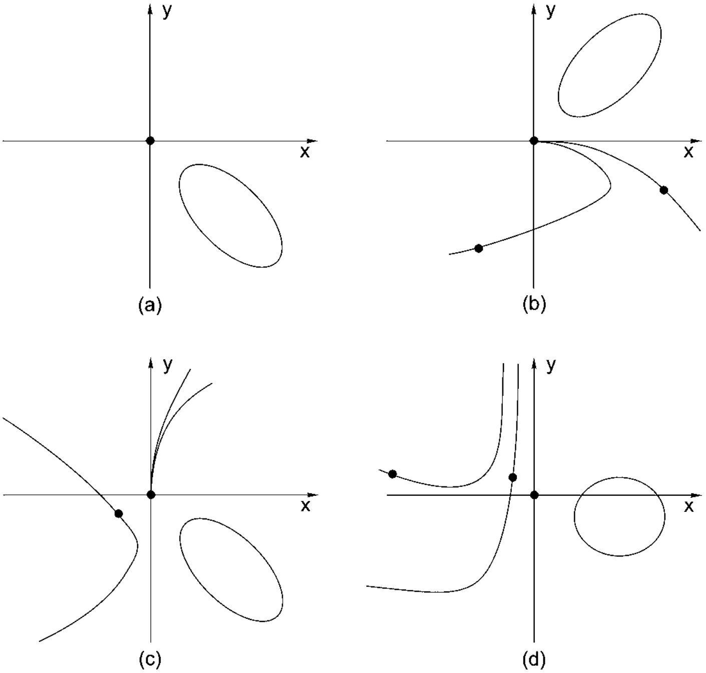


Fig. 1. Algebraic limit cycles of degree 4.

(b) *There are no quadratic systems having algebraic limit cycles of degree* 3.

(c) *After an affine change of variables the only quadratic systems having an algebraic limit cycle of degree* 4 *are*

(c.1) *Yablonskii's system*

$$\begin{aligned} \dot{x} &= -4abcx - (a+b)y + 3(a+b)cx^2 + 4xy, \\ \dot{y} &= (a+b)abx - 4abcy + \big(4abc^2 - 3(a+b)^2/2 + 4ab\big)x^2 \\ &\quad + 8(a+b)cxy + 8y^2, \end{aligned} \tag{54}$$

with $abc \neq 0$, $a \neq b$, $ab > 0$ *and* $4c^2(a-b)^2 + (3a-b)(a-3b) < 0$. *This system possesses the irreducible invariant algebraic curve*

$$\big(y + cx^2\big)^2 + x^2(x-a)(x-b) = 0, \tag{55}$$

of degree 4 *having two components, an oval* (*the algebraic limit cycle*) *and an isolated point* (*a singular point*), *see Figure* 1(a).

(c.2) *Filiptsov's system*

$$\begin{aligned} \dot{x} &= 6(1+a)x + 2y - 6(2+a)x^2 + 12xy, \\ \dot{y} &= 15(1+a)y + 3a(1+a)x^2 - 2(9+5a)xy + 16y^2, \end{aligned} \tag{56}$$

with $0 < a < 3/13$. *This system possesses the irreducible invariant algebraic curve*

$$3(1+a)\big(ax^2+y\big)^2+2y^2\big(2y-3(1+a)x\big)=0,$$

of degree 4 *having two components, one is an oval and the other is homeomorphic to a straight line. This last component contains three singular points of the system, see Figure* 1(b).

(c.3) *The system*

$$\begin{aligned}\dot x &= 5x+6x^2+4(1+a)xy+ay^2,\\ \dot y &= x+2y+4xy+(2+3a)y^2,\end{aligned} \tag{57}$$

with $(-71+17\sqrt{17})/32 < a < 0$ *possesses the irreducible invariant algebraic curve*

$$x^2+x^3+x^2y+2axy^2+2axy^3+a^2y^4=0,$$

of degree 4 *having three components, one is an oval and each of the other two is homeomorphic to a straight line. Each one of these last two components contains one singular point of the system, see Figure* 1(c).

(c.4) *The system*

$$\begin{aligned}\dot x &= 2\big(1+2x-2ax^2+6xy\big),\\ \dot y &= 8-3a-14ax-2axy-8y^2,\end{aligned} \tag{58}$$

with $0 < a < 1/4$ *possesses the irreducible invariant algebraic curve*

$$\frac{1}{4}+x-x^2+ax^3+xy+x^2y^2=0, \tag{59}$$

of degree 4 *having three components, one is an oval and each of the other two is homeomorphic to a straight line. Each one of these last two components contains one singular point of the system, see Figure* 1(d).

(d) *Quadratic systems* (53), (54), (56), (57) *and* (58) *have a unique limit cycle, the algebraic one.*

We note that the algebraic limit cycle of Filiptsov's system is born in a Hopf bifurcation at the singular point $(4, 48/13)$ when $a = 3/13$. Then, when a decreases the algebraic limit cycle increases its size and ends having infinite size at the curve $y^2(3-6x+4y)=0$ when $a=0$.

We note that the algebraic limit cycle of system (c.3) is born in a Hopf bifurcation at the singular point $((9-\sqrt{17})/8, -(5+3\sqrt{17})/8)$ when $a = (-71+17\sqrt{17})/32$. Then, when a increases the algebraic limit cycle increases its size and ends having infinite size at the curve $x^2(1+x+y)=0$ when $a=0$.

We note that the algebraic limit cycle of system (c.4) is born in a Hopf bifurcation at the singular point $(2, -1/4)$ when $a = 1/4$. Then, when a decreases the algebraic limit cycle increases its size and ends having infinite size at the irreducible curve $1/4 + x - x^2 + xy + x^2y^2 = 0$ when $a = 0$.

The uniqueness of limit cycles for planar differential systems is a classical problem which, in general, does not have an easy solution, for more details see the books of Ye Yanqian [104] and Zhang Zhifen [105].

Now following the paper [35] we apply a change of variables to the known quadratic systems having an algebraic limit cycle, which preserves the degree of the system, but increases the degree of the algebraic curve. For this purpose we use the birational transformation

$$(x, y) \to \big(x/y^2, 1/y\big), \tag{60}$$

after an appropriate translation. In fact, this transformation is an involution.

If the system is of the form

$$\begin{aligned} \dot{x} &= \alpha x + \beta y + 2ex^2 + bxy + cy^2, \\ \dot{y} &= \gamma x + \delta y + exy + fy^2, \end{aligned} \tag{61}$$

then it is easy to see that we can apply the transformation above and still remain in the class of quadratic systems.

As a simple example, we show that the example of Yablonskii with an algebraic limit cycle of degree 4 can be obtained from the well-known example of an algebraic limit cycle of degree 2 due to Qin Yuan-Xun [88].

PROPOSITION 7.3. *The system of Yablonskii* (54) *with its irreducible invariant algebraic curve* (55) *can be transformed into the system*

$$\begin{aligned} \dot{x} &= -3(a+b)cx + 4abcx^2 - 4y + (a+b)xy, \\ \dot{y} &= \big(a\big(b + 4bc^2\big) - \big(3a^2 + 3b^2\big)/2\big)x + 2(a+b)cy + ab(a+b)x^2 \\ &\quad + 4abcxy + 2(a+b)y^2, \end{aligned}$$

with the invariant algebraic curve

$$-\frac{(a-b)^2}{4ab} + ab\left(x - \frac{a+b}{2ab}\right)^2 + (y+c)^2 = 0, \tag{62}$$

by the transformation $(x, y) \to (1/x, y/x^2)$.

The algebraic curves (55) *and* (62) *both give limit cycles when* $abc \neq 0$, $a \neq b$, $ab > 0$, *and* $4c^2(a-b)^2 + (3a-b)(a-3b) < 0$, *and the transformation maps the one limit cycle onto the other.*

We now apply the birational transformation to system (58) having the irreducible invariant algebraic curve (59) which defines an algebraic limit cycle of degree 4 for $0 < a < 1/4$, and we get a quadratic system having an algebraic limit cycle of degree 5.

THEOREM 7.4. *System*

$$\begin{aligned}\dot{x} &= 28x - \frac{12}{\alpha+4}y^2 - 2\big(\alpha^2-16\big)(12+\alpha)x^2 + 6(3\alpha-4)xy,\\ \dot{y} &= \big(32-2\alpha^2\big)x + 8y - (\alpha+12)(\alpha^2-16)xy + (10\alpha-24)y^2,\end{aligned} \tag{63}$$

has an irreducible algebraic invariant curve of degree 5 *given by*

$$\begin{aligned}&x^2 + (16-\alpha^2)x^3 + (\alpha-2)x^2y + \frac{1}{(4+\alpha)^2}y^4 - \frac{6}{(4+\alpha)^2}y^5 - \frac{2}{4+\alpha}xy^2\\ &\quad + \frac{(\alpha-4)(12+\alpha)}{4}x^2y^2 + \frac{(12+\alpha)}{4+\alpha}xy^4 + \frac{8-\alpha}{4+\alpha}xy^3 = 0.\end{aligned} \tag{64}$$

For $\alpha \in (3\sqrt{7}/2, 4)$ *the curve* (64) *contains an algebraic limit cycle of degree* 5.

PROOF. Let $a = 16-\alpha^2$. When we make the change of coordinates

$$(x,y) = \left(\frac{u}{v^2} - \frac{1}{\alpha+4}, \frac{1}{v} + \frac{\alpha-2}{2}\right), \tag{65}$$

multiply by v, and replace (u,v) again with (x,y), system (58) becomes (63). The curve (64) is obtained from (59) by means of the same change of coordinates and multiplication by v^6. The irreducibility of (64) follows from the irreducibility of (59).

Since the curve (59) contains an algebraic limit cycle for $a \in (0, 1/4)$, one may easily check, that the above oval does not intersect the singular line of the transformation (65), so the theorem follows. □

In a similar way we have the following result.

THEOREM 7.5. *System*

$$\begin{aligned}\dot{x} &= 28(\beta-30)\beta x + y + 168\beta^2x^2 + 3xy,\\ \dot{y} &= 16\beta(\beta-30)\big(14(\beta-30)\beta x + 5y + 84\beta^2x^2\big)\\ &\quad + 24(17\beta-6)\beta xy + 6y^2,\end{aligned} \tag{66}$$

has an irreducible algebraic invariant curve of degree 6 *given by*

$$\begin{aligned}&-7y^3 + 3(\beta-30)^2\beta y^2 + 18(\beta-30)(-2+\beta)\beta xy^2 + 27(\beta-2)^2\beta x^2y^2\\ &\quad + 24(\beta-30)^3\beta^2xy + 144(\beta-30)(\beta-2)^2\beta^2x^3y + 48(\beta-30)^4\beta^3x^2\end{aligned}$$

$$
\begin{aligned}
&+576(\beta-30)^2(-2+\beta)^2\beta^3x^4-432(\beta-2)^2\beta^2(3+2\beta)x^4y\\
&-3456(\beta-30)(-2+\beta)^2\beta^3(3+2\beta)x^5\\
&+3456(\beta-2)^2\beta^3(12+\beta)(3+2\beta)x^6\\
&+24(\beta-30)^2\beta^2(9\beta-4)x^2y+64(\beta-30)^3\beta^3(9\beta-4)x^3=0.
\end{aligned}
\tag{67}
$$

For $\beta\in(3/2,2)$ *the curve* (67) *contains an algebraic limit cycle of degree* 6.

PROOF. Let $a=(4-\beta^2)/7$. When we make the change of coordinates

$$
(x,y)=\left(\frac{v+4u\beta(-30+3u(-2+\beta)+\beta)}{12u^2\beta(\beta^2-4)},\frac{30-\beta-u(8+3\beta)}{14u}\right),
\tag{68}
$$

multiply by $-21\beta u/2$, and replace (u,v) again with (x,y), system (58) becomes (66). The curve (67) is obtained from (59) by means of the same change of coordinates and multiplication by $2016\beta^2(\beta^2-4)^2u^6$. The irreducibility of (67) is now obvious.

Since the curve (59) contains an algebraic limit cycle for $a\in(0,1/4)$, the theorem follows in a way similar to the last part of Theorem 7.4. ☐

After these results some natural questions are:

OPEN QUESTION 7.6. *Does there exist a chain of rational transformations like the ones above which give examples of quadratic polynomial systems with algebraic limits cycles of arbitrary degree*?

OPEN QUESTION 7.7. *What is the maximum degree of all algebraic limit cycles for quadratic systems*?

8. Limit cycles and algebraic limit cycles

In this section we follow the paper [69]. In 1900 Hilbert [55] in the second part of its 16th problem proposed to find an estimation of the uniform upper bound for the number of limit cycles of all polynomial vector fields of a given degree, and also to study their distribution or configuration in the plane. This has been one of the main problems in the qualitative theory of planar differential equations in the XX century. The contributions of Bamon [5] for the particular case of quadratic vector fields, and mainly of Écalle [39] and Ilyashenko [57] proving that any polynomial vector field has finitely many limit cycles have been the best results in this area. But until now it is not proved the existence of an uniform upper bound. This problem remains open even for the quadratic polynomial vector fields.

A *configuration of limit cycles* is a finite set $C=\{C_1,\dots,C_n\}$ of disjoint simple closed curves of the plane such that $C_i\cap C_j=\emptyset$ for all $i\neq j$.

Given a configuration of limit cycles $C=\{C_1,\dots,C_n\}$ the curve C_i is *primary* if there is no curve C_j of C contained into the bounded region limited by C_i.

Two configurations of limit cycles $C = \{C_1, \dots, C_n\}$ and $C' = \{C'_1, \dots, C'_m\}$ are (*topologically*) *equivalent* if there is a homeomorphism $h: \mathbb{R}^2 \to \mathbb{R}^2$ such that $h(\bigcup_{i=1}^n C_i) = (\bigcup_{i=1}^m C'_i)$. Of course, for equivalent configurations of limit cycles C and C' we have that $n = m$.

We say that the vector field $\mathcal{X}$ *realizes* the configuration of limit cycles C if the set of all limit cycles of $\mathcal{X}$ is equivalent to C.

THEOREM 8.1. *Let $C = \{C_1, \dots, C_n\}$ be a configuration of limit cycles, and let r be its number of primary curves. Then the following statements hold.*

(a) *The configuration C is realizable by a polynomial vector field.*

(b) *The configuration C is realizable as algebraic limit cycles by a polynomial vector field of degree $\leqslant 2(n + r) - 1$.*

In the proof of Theorem 8.1 we shall provide an explicit expression for the polynomial differential system of degree at most $2(n + r) - 1$ satisfying statement (b) of Theorem 8.1. Of course, statement (a) of Theorem 8.1 follows immediately from statement (b).

The problem of given a configuration of limit cycles realize it by a polynomial differential system has been studied by several authors. Thus, for C^r vector fields the problem has been solved by Al'mukhamedov [1], Balibrea and Jimenez [4] and Valeeva [99]. Statement (a) of Theorem 8.1 has been solved by Schecter and Singer [90] and Sverdlove [98], but they do not provide an explicit polynomial vector field satisfying the given configuration of limit cycles. The result presented in statement (b) of Theorem 8.1 appears in [69], and its proof provides simultaneously the shortest and easiest proof of statement (a) of Theorem 8.1.

We consider C^1 the vector field

$$\mathcal{X} = P\frac{\partial}{\partial x} + Q\frac{\partial}{\partial y}$$

defined in the open subset U of $\mathbb{R}^2$. Then, $\mathcal{X}$ is *exact* in U if

$$\frac{\partial P}{\partial x} = -\frac{\partial Q}{\partial y},$$

for all $(x, y) \in U$. Furthermore, if U is simply connected, then there exists a function $H: U \to \mathbb{R}$ satisfying

$$P = -\frac{\partial H}{\partial y}, \qquad Q = \frac{\partial H}{\partial x}.$$

Therefore, the function H is the *Hamiltonian* of the *Hamiltonian vector field* $\mathcal{X}$. Clearly, the Hamiltonian function is a first integral of $\mathcal{X}$.

A C^1 function $R: U \to \mathbb{R}$ such that

$$\frac{\partial (RP)}{\partial x} = -\frac{\partial (RQ)}{\partial y} \tag{69}$$

is an *integrating factor* of the vector field $\mathcal{X}$. We know that R is an integrating factor of $\mathcal{X}$ in U if and only if R is a solution of the partial differential equation

$$P\frac{\partial R}{\partial x}+Q\frac{\partial R}{\partial y}=-\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}\right)R \tag{70}$$

in U.

A function $V:U\to\mathbb{R}$ is an *inverse integrating factor* of the vector field $\mathcal{X}$ if V verifies the partial differential equation

$$P\frac{\partial V}{\partial x}+Q\frac{\partial V}{\partial y}=\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}\right)V \tag{71}$$

in U. We note that V satisfies (71) in U if and only if $R=1/V$ satisfies (70) in $U\setminus\{(x,y)\in U\colon\ V(x,y)=0\}$.

The following result due to Giacomini, Llibre and Viano [47] will be essential in our proof of statement (b) of Theorem 8.1. Here, we provide an easier and direct proof.

THEOREM 8.2. *Let X be a C^1 vector field defined in the open subset U of $\mathbb{R}^2$. Let $V:U\to\mathbb{R}$ be an inverse integrating factor of X. If γ is a limit cycle of X, then γ is contained in $\boldsymbol{\Sigma}=\{(x,y)\in U\colon\ V(x,y)=0\}$.*

PROOF. Due to the existence of the inverse integrating factor V defined in U, we have that the vector field X/V is Hamiltonian in $U\setminus\boldsymbol{\Sigma}$. Since the flow of a Hamiltonian vector field preserves the area and in a neighborhood of a limit cycle a flow does not preserve the area, the theorem follows. □

PROOF OF THEOREM 8.1. Let $C=\{C_1,\dots,C_n\}$ be the configuration of limit cycles given in the statement of Theorem 8.1. For every primary curve C_j we select a point p_j in the interior of the bounded component limited by C_j. Since we will work with an equivalent configuration of limit cycles, without loss of generality we can assume that

(i) each curve C_i is a circle defined by

$$f_i(x,y)=(x-x_i)^2+(y-y_i)^2-r_i^2=0,$$

for $i=1,\dots,n$; and that

(ii) the primary curves of the configuration C are the curves C_j, and the selected points p_j have coordinates (x_j,y_j), for $j=1,\dots,r$.

For every selected point p_j we define

$$f_{n+2j-1}(x,y)=(x-x_j)+i(y-y_j),$$
$$f_{n+2j}(x,y)=(x-x_j)-i(y-y_j).$$

Now, we consider the function

$$\widetilde{H} = f_1^{\lambda_1} \cdots f_n^{\lambda_n} f_{n+1}^{\lambda_{n+1}} f_{n+2}^{\lambda_{n+2}} \cdots f_{n+2r-1}^{\lambda_{n+2r-1}} f_{n+2r}^{\lambda_{n+2r}} = \prod_{k=1}^{n+2r} f_k,$$

with $\lambda_1 = \cdots = \lambda_n = 1$, and $\lambda_{n+2j-1} = 1 + i$ and $\lambda_{n+2j} = 1 - i$, for $j = 1, \ldots, r$. After an easy computation, we have that

$$\widetilde{H}(x, y) = A(x, y)B(x, y)C(x, y),$$

where

$$A(x, y) = \prod_{i=1}^{n} \big[(x - x_i)^2 + (y - y_i)^2 - r_i^2\big],$$

$$B(x, y) = \prod_{j=1}^{r} \big[(x - x_j)^2 + (y - y_j)^2\big],$$

$$C(x, y) = \exp\left(-2 \sum_{j=1}^{r} \arg\big[(x - x_j) + i\,(y - y_j)\big]\right).$$

Clearly $\widetilde{H}(x, y)$ is a real function. Therefore, the function

$$H = \log \widetilde{H} = \sum_{k=1}^{n+2r} \lambda_k \log f_k$$

is also real.

We claim that the vector field

$$\begin{aligned} \mathcal{X} &= P(x, y)\frac{\partial}{\partial x} + Q(x, y)\frac{\partial}{\partial y} \\ &= -\sum_{k=1}^{n+2r} \lambda_k \left(\prod_{\substack{l=1 \\ l \neq k}}^{n+2r} f_l\right) \frac{\partial f_k}{\partial y} \frac{\partial}{\partial x} + \sum_{k=1}^{n+2r} \lambda_k \left(\prod_{\substack{l=1 \\ l \neq k}}^{n+2r} f_l\right) \frac{\partial f_k}{\partial x} \frac{\partial}{\partial y}, \end{aligned}$$

satisfies the conclusion of statement (b) of Theorem 8.1. Now we shall prove the claim.

First, we note that we have the equalities

$$\frac{\partial H}{\partial x} = \frac{Q}{\prod_{k=1}^{n+2r} f_k}, \qquad \frac{\partial H}{\partial y} = -\frac{P}{\prod_{k=1}^{n+2r} f_k}. \tag{72}$$

Therefore, since H and $\prod_{k=1}^{n+2r} f_k$ are real functions, we get that P, Q, and consequently $\mathcal{X}$ are real.

Clearly, from the definition of $\mathcal{X}$ it follows that P and Q are polynomials of degree at most $n+2r-1$. So, $\mathcal{X}$ is a real polynomial vector field of degree at most $n+2r-1$.

From (72) it follows that $V=\prod_{k=1}^{n+2r} f_k$ is a polynomial inverse integrating factor of $\mathcal{X}$, and that H is a Hamiltonian for the Hamiltonian vector field

$$\frac{1}{V}X=\frac{P}{V}\frac{\partial}{\partial x}+\frac{Q}{V}\frac{\partial}{\partial y},$$

defined in $\mathbb{R}^2\setminus\{V=0\}$.

Since V is polynomial, V is defined in the whole $\mathbb{R}^2$. Therefore, by Theorem 8.2 and since $V(x,y)=0$ if and only if $(x,y)\in(\bigcup_{i=1}^{n} C_i)\cup\{p_1,\dots,p_r\}$, if the vector field $\mathcal{X}$ has limit cycles, these must be the circles C_i for $i=1,\dots,n$. Now, we shall prove that all these circles are limit cycles. Hence, the polynomial vector field $\mathcal{X}$ will realize the configuration of limit cycles $\{C_1,\dots,C_n\}$ and the theorem will be proved.

We note that since $\widetilde{H}=\exp(H)$ is a first integral of the vector field $\mathcal{X}$ in $\mathbb{R}^2\setminus\{V=0\}$, the circles are formed by solutions because they are contained in the level curve $\tilde{V}=0$, and $V=0$ is formed by solutions. Now we shall prove that on every circle C_i there is no singular points of $\mathcal{X}$ and, therefore, C_i will be a periodic orbit. Assume that (x_0,y_0) is a singular point of $\mathcal{X}$ contained into the circle C_i; i.e., $P(x_0,y_0)=Q(x_0,y_0)=f_i(x_0,y_0)=0$. From the definition of P and Q we have that

$$P(x_0,y_0)=-\lambda_i\left(\prod_{\substack{l=1\\ l\neq i}}^{n+2r} f_l(x_0,y_0)\right)\frac{\partial f_i}{\partial y}(x_0,y_0)=0,$$

$$Q(x_0,y_0)=\lambda_i\left(\prod_{\substack{l=1\\ l\neq i}}^{n+2r} f_l(x_0,y_0)\right)\frac{\partial f_i}{\partial x}(x_0,y_0)=0.$$

Since $f_l(x_0,y_0)\neq 0$ for $l\neq i$, we obtain that $\frac{\partial f_i}{\partial x}(x_0,y_0)=0$ and $\frac{\partial f_i}{\partial y}(x_0,y_0)=0$. Therefore, the point (x_0,y_0) is the center of the circle C_i in contradiction that $f_i(x_0,y_0)=0$. Hence, every circle C_i is a periodic orbit of the vector field $\mathcal{X}$. Now, we shall prove that C_i will be a limit cycle, and this will complete the proof of Theorem 8.1.

We note that all circles C_i and all points p_j are in the level $\widetilde{H}(x,y)=0$, and that they are the unique orbits of $\mathcal{X}$ in this level. Now suppose that C_i is not a limit cycle. Then, there is a periodic orbit $\gamma=\{(x(t),y(t))\colon t\in\mathbb{R}\}$ different from $C_1,\dots,C_n$ and sufficiently close to C_i such that in the bounded component B limited by γ there are the same points of $\{p_1,\dots,p_r\}$ than in the bounded component limited by C_i. Without loss of generality we can assume that these points are $p_1,\dots,p_s$.

As γ is different from $C_1,\dots,C_n$, there exists $h\neq 0$ such that

$$\widetilde{H}\big(x(t),y(t)\big)=A\big(x(t),y(t)\big)B\big(x(t),y(t)\big)\exp\left(-2\sum_{j=1}^{r}\theta_j(t)\right)=h,\tag{73}$$

where $\theta_j(t) = \arg[(x(t) - x_j) + i(y(t) - y_j)]$. The function $A(x(t), y(t))\, B(x(t), y(t))$ is bounded on γ. Clearly, the angles $\theta_1(t), \dots, \theta_s(t)$ tend all simultaneously (due to its definition) to either $+\infty$ or $-\infty$, when $t \to +\infty$, while the angles $\theta_{s+1}(t), \dots, \theta_r(t)$ remain bounded when $t \to +\infty$. These facts are in contradiction with equality (73). Consequently, we have proved that C_i is a limit cycle. In short, Theorem 8.1 is proved. □

9. Darbouxian theory of integrability and centres

In this section we follow the paper [30]. One of the main applications of the Darbouxian theory of integrability is proving the existence of a centre, see for instance Section 3.

In the elementary theory of qualitative differential equations we identify three main types of behavior at a nondegenerate singular point: a node, a focus, or a saddle. All three are stable, in that a small perturbation will not change the stability of the singular point. Moreover, topologically, we can read off their behavior from just their linear terms. However, there is also the possibility that the singular point is a *fine focus* or a *centre*. That is, the divergence vanishes at that point. In this case the linear terms give a centre: a neighborhood of the origin which consists of closed trajectories. In this case, the nonlinear terms must be examined in order to determine whether the point is stable or unstable. If it is neither and the system is analytic, then the singular point is also a centre for the nonlinear system. Without loss of generality, we can consider the singular point to be at the origin and to be in the form

$$\dot{x} = \lambda x - y + p(x, y), \qquad \dot{y} = x + \lambda y + q(x, y), \tag{74}$$

where p and q represent the nonlinear terms. The case of a fine focus or centre corresponds to $\lambda = 0$.

We can distinguish between a centre and a (fine) focus in a number of ways; we follow the most direct first. It can be shown that for any N there is a change of coordinates which brings the origin of (74) to the polar form

$$\begin{aligned} \dot{r} &= c_3 r^3 + c_5 r^5 + \cdots + \mathrm{O}\big(r^N\big), \\ \dot{\theta} &= 1 + d_2 r^2 + d_4 r^4 + \cdots + \mathrm{O}\big(r^N\big). \end{aligned} \tag{75}$$

If all the c_i are zero up to c_{2k+1}, then the system is said to have a *fine focus of order k*. The stability is given by the sign of c_{2k+1}.

It can also be shown that perturbations of the nonlinear terms of (74) can produce in this case at most k limit cycles bifurcating from the origin. Furthermore, if the class of systems (74) are sufficiently general, there are perturbations which produce this number of limit cycles in a multiple Hopf bifurcation. We call c_{2k+1} the kth *Liapunov quantity*.

If all the c_i are zero, then it can be shown that there is an analytic change of coordinates which brings the system into the polar form

$$\dot{r} = 0, \qquad \dot{\theta} = 1 + d_2 r^2 + d_4 r^4 + \cdots.$$

The singular point is obviously a centre in this case.

Given a class of polynomial equations in the form (74), we are often interested in the subclass with centres at the origin. The problem is that showing that we have a centre requires an infinite number of conditions. However, if we could find an analytic first integral in a neighborhood of the singular point, then the singular point must be a centre. In fact, the linear terms of (74) implies that the first terms of such an integral are $a+b(x^2+y^2)^s+\cdots$, and therefore trajectories close to the origin are closed.

Alternatively, we could find an integrating factor which is well-defined and nonzero in a neighborhood of the singular point. In either case an obvious method for constructing such functions is the Darbouxian theory of integrability. The surprising thing is that this method is so successful.

THEOREM 9.1. *All the nondegenerate centres of systems* (74) *with homogeneous quadratic or cubic* p *and* q *are integrable with Darbouxian first integrals. The same is true if* p *and* q *are of the form*

$$p = p_2 + xf, \qquad q = q_2 + yf,$$

where p_2, q_2 *and* f *are all homogeneous quadratics.*

For a proof of this theorem see [92,89,9].

The last system is the projective version of the quadratic systems and in fact was the system studied by Darboux.

As an example of this, consider the system

$$\begin{aligned}\dot{x} &= y + a_1x^2 + (a_2+2b_1)xy - a_1y^2 + x^2y,\\ \dot{y} &= -x + b_1x^2 + (b_2-2a_1)xy - b_1y^2 + xy^2,\end{aligned}$$

generically this has 4 invariant lines, whose cofactors are all of the form

$$\alpha x + \beta y + xy.$$

Hence we can conclude that the system has a centre at the origin.

Another method for distinguishing between a focus and a centre is to use a *Liapunov function*. This is a function $V = k + x^2 + y^2 + \mathrm{O}((x^2+y^2)^3/2)$ which satisfies

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \eta_4\bigl(x^2+y^2\bigr)^4 + \eta_6\bigl(x^2+y^2\bigr)^6 + \cdots + \mathrm{O}\bigl(\bigl(x^2+y^2\bigr)^{N/2}\bigr). \tag{76}$$

Such a function can always be found in the neighborhood of a fine focus. The origin is a centre if all the η_i vanish. If η_{2k+2} is the first nonzero term, then the origin is a fine focus of order k. Computationally, this method is easier to handle than the normal form (75). The coefficients η_{2i+2} are essentially positive multiples of the c_{2i+1} if we assume that the previous η_{2j+1}, $j<i$ vanish.

There is also a third method, closely related to the second. Here we seek a function which is almost an integrating factor. That is we look for a function $R = 1 + \mathrm{O}(x,y)$

which satisfies

$$\begin{aligned}&\frac{\partial}{\partial x}\left(\frac{P}{R}\right)+\frac{\partial}{\partial y}\left(\frac{Q}{R}\right)\\&\quad=\zeta_2\left(x^2+y^2\right)+\zeta_4\left(x^2+y^2\right)^4+\cdots+\mathrm{O}\left(\left(x^2+y^2\right)^{N/2-1}\right).\end{aligned}\tag{77}$$

Here the origin is a centre if all the ζ_i vanish, and a fine focus of order k if ζ_{2k} is the first nonzero term. The advantage here is that the degrees of the polynomials required here is less than (76). Again the coefficients ζ_{2i} are essentially positive multiples of the c_{2i+1} modulo the previous ζ_{2j}'s.

Now, suppose we are seeking conditions for a centre at the origin. Consider a Darbouxian function B which is composed of invariant algebraic curves which do not pass through the origin and exponential factors. Such a B is well-defined at the origin and

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}B=B\left[\sum l_iL_i\right],\\&\frac{\partial}{\partial x}\left(\frac{P}{B}\right)+\frac{\partial}{\partial y}\left(\frac{Q}{B}\right)=\frac{1}{B}\left[\operatorname{div}(P,Q)-\sum l_iL_i\right].\end{aligned}$$

Since the invariant algebraic curves do not pass through the origin, then the definition of invariant algebraic curve implies that the cofactors L_i must vanish there. We also assume that the divergence is zero at the origin, or there would be no centre. If we have at least $m(m+1)/2-1-q$ invariant algebraic curves and exponential factors, $0<q\leqslant(m-1)/2$, then we can choose the l_i nontrivially so that the square brackets of one of the expressions above lies in the vector space generated by the polynomials $(x^2+y^2)^j$, $j=1,\dots,q$. Comparing these expressions with (76) and (77), we obtain the following result.

THEOREM 9.2. *Suppose the origin is a fine focus, and that the first q Liapunov quantities at the origin vanish, $0<q\leqslant(m-1)/2$. If there are at least $m(m+1)/2-1-q$ invariant algebraic curves or exponential factors not passing through the origin, then there is a local Darbouxian integrating factor. If there are at least $m(m+1)/2-q$, then there is a Darbouxian first integral. In either case the origin is a centre.*

The result was first noticed by Cozma and Şubă [37,94] using different methods. Another related result is Chavarriga, Giacomini and Giné [18].

10. Non-existence of limit cycles

This section follows the paper [30]. We rename a *Liapunov function* a function $\phi(x,y)$ for which

$$D\phi>0$$

in the region of interest. Clearly the existence of such a function in a region precludes the existence of periodic solutions and, in particular, limit cycles.

In the same way, we define a *Dulac function* to be a function R such that

$$\frac{\partial}{\partial x}\left(\frac{P}{R}\right)+\frac{\partial}{\partial y}\left(\frac{Q}{R}\right)>0.$$

By applying the divergence criterion, we can see that such a function also precludes the existence of periodic solutions or limit cycles in any simply connected region where R is well-defined and nonzero.

The analogy between these functions and the first integrals and integrating factors examined up to now is obvious. In particular, given a collection of invariant algebraic curves or exponential factors with cofactors L_i, if we can find constants l_i such that

$$\sum l_i L_i>0, \quad \text{or} \quad \operatorname{div}(P,Q)-\sum l_i L_i>0,$$

then there are no limit cycles in any simply connected region where the Darbouxian function is well defined.

For example, the Lokta–Volterra equations are quadratic with two invariant lines. There is also a singular point which does not lie on either line around which any limit cycle must lie. Thus the cofactors of the two lines must vanish at this point. If the divergence also vanishes at the singular point, then we can find a linear dependency between the cofactors and the divergence which means that the system must be integrable and we have a family of closed orbits. If the divergence does not vanish then we can construct a Dulac function of Liapunov function as above. Thus there are no limit cycles in either case.

The example above shows in a simple way how detailed calculations can be reduced to simple geometric arguments instead. We could have replaced the lines above by invariant hyperbolas or parabolas with no increase in difficulty.

If one of the curves was an ellipse, however, we might have problems. First, the ellipse may be a limit cycle in its own right. Second, if the polynomial representing the ellipse appeared to a negative power in the Dulac function, then we cannot apply Green's theorem since the region surrounding the ellipse is not simply connected. This can be overcome in certain cases by considering line integrals around the loop itself.

In order to get some deeper results, we need to allow some curves which are almost invariant. Rather than abstract things more, we give an example which is very representative of other results.

THEOREM 10.1. *Suppose a quadratic system has an invariant algebraic curve and a singular point not on this curve where the divergence vanishes, then the system has no limit cycles in any simply connected region of the complement of the curve.*

PROOF. Any limit cycle in a quadratic system surrounds only one singular point which must be a focus (see [104]). Suppose a limit cycle surrounds a singular point with nonzero divergence. Let $C=0$ be the invariant algebraic curve with cofactor L. Thus

$$\left(PC^r\right)_x+\left(QC^r\right)_y=C^r(rL+\Delta).$$

In this section we denote by Δ the divergence of the system, i.e., $\Delta = P_x + Q_y$. Since both L and Δ must pass through the other singular point, and limit cycles of quadratic system must be convex, we can chose r so that $rL + \Delta$ does not pass through the limit cycle. Hence we have a Dulac function in this case, which contradicts our assumption of a limit cycle.

Suppose now that there is a limit cycle which surrounds the divergence free singular point. Since this point must be a focus we transform the system to the form

$$P = -y + ax^2 + bxy + cy^2, \qquad Q = x + dx^2 + exy + fy^2.$$

A further rotation allows us to set $c = 0$ without loss of generality. Consider the function $G = x - 1/b$ for $b \neq 0$ and $G = e^x$ if $b = 0$. We calculate

$$\frac{\mathrm{d}}{\mathrm{d}t}G = \begin{cases} byG + ax^2, & b \neq 0, \\ -yG + ax^2G, & b = 0. \end{cases}$$

Thus the line $G = 0$ is a transversal, and no limit cycle can cross it. Now, we calculate that

$$\left(PC^rG^s\right)_x + \left(QC^rG^s\right)_y = \begin{cases} C^rG^{s-1}(G[rL + sby + \Delta] + sax^2), & b \neq 0, \\ C^rG^{s-1}(G[rL - sy + \Delta] + sGax^2), & b = 0. \end{cases}$$

In either case we can find values of r and s to eliminate the term in square brackets. Once again we have a Dulac function. □

It seems that algebraic curve methods are the natural ones for proving nonexistence of limit cycles. In Coppel's survey paper [36], for example, all the nonexistence results are obtained this way except one which uses a Liénard system argument. We reprove this here using algebraic curves.

THEOREM 10.2. *A quadratic system* (2) *with* $rP + sQ = \Delta M$ *for some polynomial* M *and real numbers* r *and* s *can have no limit cycles. In particular, a quadratic system with two singular points with zero divergence has no limit cycles.*

PROOF. If Δ is a constant we have finished. If $\Delta = a(rx + sy + t)$ for some a and t, then $\Delta = 0$ would be an invariant line. All limit cycles would have to lie in one of the regions $\Delta > 0$ or $\Delta < 0$ which is not possible. Hence the linear terms of Δ must be linearly independent from $rx + sy$ for limit cycles to exist. We can therefore write $M = a\Delta + b(rx + sy) + c$ for some a, b and c.

If $b = 0$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}(rx + sy) = a\Delta^2 \quad (c = 0),$$

or

$$\left(\mathrm{e}^{-(rx+sy)/c}P\right)_x + \left(\mathrm{e}^{-(rx+sy)/c}Q\right)_y = -\frac{a}{c}\Delta^2\mathrm{e}^{-(rx+sy)/c} \quad (c \neq 0).$$

Hence there are no limit cycles. When $b \neq 0$ then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(rx + sy + \frac{c}{b}\right) = a\Delta^2 \quad \text{on } rx + sy + \frac{c}{b} = 0,$$

so $rx + sy + c/b = 0$ is a transversal. Furthermore

$$(BP)_x + (BQ)_y = -\frac{a}{b(rx + sy + c/b)} B\Delta^2,$$

where

$$B = \left(rx + sy + \frac{c}{b}\right)^{-1/b},$$

and so we have a Dulac function in this case too. □

Is it possible to extend these methods to prove the uniqueness of limit cycles if we allow the Dulac function to vanish at the singular point which the limit cycle surrounds? It would be very nice if this was true as many of the uniqueness results for quadratic systems have a nice algebraic content. For example a quadratic system with an invariant line or an invariant parabola have at most one limit cycle. However, no such methods are known at the moment and other less direct methods need to be used.

We finish with a simple example of such a proof. We show that the van der Pol oscillator has a unique limit cycle for small values of the nonlinear terms.

THEOREM 10.3. *The system*

$$\dot{x} = y, \qquad \dot{y} = -x - \mu\left(1 - x^2\right)y$$

has at most one limit cycle for $|\mu| < \sqrt{3}$.

PROOF. We first let $Y = y + \mu(x - x^3/3)$ to transform the system to the Liénard plane:

$$\dot{x} = Y - \mu\left(x - x^3/3\right), \qquad \dot{Y} = -x.$$

Now, taking $B = (x^2 - \mu xY + Y^2)^{-1}$, we have

$$(B\dot{x})_x + (B\dot{y})_y = \mu B^2 x^2\left(x^2/3 - 2\mu xY/3 + Y^2\right).$$

For $|\mu| < \sqrt{3}$, B is positive definite and we have a Dulac function defined in the whole of the plane except at the origin. Since each limit cycle must surround the origin, a simple application of Green's theorem shows that there can be at most one limit cycle. □

11. The inverse problem

In this section we are mainly interested in the polynomial differential systems which have a given set of invariant algebraic curves, independently if they are integrable or not. Thus, first we study the normal forms of planar polynomial vector fields having a given set of generic invariant algebraic curves. That is, in some sense we are interested in a kind of inverse theory of the Darboux theory of integrability. We follow the paper [33]. The main result of this section is the following one.

THEOREM 11.1. *Let* $C_i = 0$ *for* $i = 1, \ldots, p$, *be irreducible invariant algebraic curves in* $\mathbb{C}^2$, *and set* $r = \sum_{i=1}^{p} \deg C_i$. *We assume that all* C_i *satisfy the following generic conditions*:

(i) *There are no points at which* C_i *and its first derivatives are all vanish.*
(ii) *The highest order terms of* C_i *have no repeated factors.*
(iii) *If two curves intersect at a point in the finite plane, they are transversal at this point.*
(iv) *There are no more than two curves* $C_i = 0$ *meeting at any point in the finite plane.*
(v) *There are no two curves having a common factor in the highest order terms.*

Then any polynomial vector field $\mathcal{X}$ *of degree* m *tangent to all* $C_i = 0$ *satisfies one of the following statements.*

(a) *If* $r < m + 1$ *then*

$$\mathcal{X} = \left(\prod_{i=1}^{p} C_i\right)\mathcal{Y} + \sum_{i=1}^{p} h_i \left(\prod_{\substack{j=1 \\ j\neq i}}^{p} C_j\right)\mathcal{X}_{C_i}, \tag{78}$$

where $\mathcal{X}_{C_i} = (-C_{iy}, C_{ix})$ *is a Hamiltonian vector field, the* h_i *are polynomials of degree no more than* $m - r + 1$, *and the* $\mathcal{Y}$ *is a polynomial vector field of degree no more than* $m - r$.

(b) *If* $r = m + 1$ *then*

$$\mathcal{X} = \sum_{i=1}^{p} \alpha_i \left(\prod_{\substack{j=1 \\ j\neq i}}^{p} C_j\right)\mathcal{X}_{C_i}, \tag{79}$$

with $\alpha_i \in \mathbb{C}$.

(c) *If* $r > m + 1$ *then* $\mathcal{X} = 0$.

This theorem due to Christopher [28] was stated in several papers without proof like [28] and [61], and used in other papers as [9] and [66]. The proof that we present here of Theorem 11.1 essentially circulated as the preprint [29] but was never published. Żołądek in [106] (see also Theorem 3 of [107]) stated a similar result to Theorem 11.1, but as far as we know the paper [106] has not been published. In any case Żołądek's approach to Theorem 11.1 is analytic, while our approach is completely algebraic.

Statement (b) of this theorem has a corollary due to Christopher and Kooij [61] showing that system (79) has the integrating factor $\mathcal{R} = (\prod_{i=1}^{p} C_i)^{-1}$, and consequently the system is Darbouxian integrable.

The following result shows that the generic conditions of Theorem 11.1 are necessary.

THEOREM 11.2. *If one of the conditions* (i)–(v) *of Theorem* 11.1 *is not satisfied, then the statements of Theorem* 11.1 *do not hold.*

We have mentioned that system (1) satisfying the five assumptions of Theorem 11.1 with $r = m + 1$ is Darbouxian integrable. Now we provide two examples of polynomial systems satisfying all assumptions of Theorem 11.1 with $r = m + 1$ except either (ii) or (iii) and which are not Darbouxian integrable. Until now there are very few proofs of polynomial systems which are not Darbouxian integrable, see for instance Jouanolou [58], Moulin Ollagnier [75–77] and [10], see Section 13.

Consider the following quadratic systems:

$$\begin{aligned} \dot{x} &= y(ax - by + b) + x^2 + y^2 - 1, \\ \dot{y} &= bx(y - 1) + a\big(y^2 - 1\big), \end{aligned} \tag{80}$$

which has the invariant circle $C_1 = x^2 + y^2 - 1 = 0$ with cofactor $K_1 = 2(x + ay)$ and the invariant straight line $C_2 = y - 1 = 0$ with cofactor $K_2 = bx + ay + a$. We note that C_1 and C_2 are tangent at the point (0, 1).

THEOREM 11.3. *There are values of the parameters a and b for which system* (80) *is not Darbouxian integrable.*

As a corollary the following result shows that there are polynomial systems with an invariant algebraic curve whose highest order term have repeated factors such that they are not Darbouxian integrable. Consider the following quadratic system:

$$\begin{aligned} \dot{x} &= (1 - b)\big(x^2 + 2y - 1\big) - (ax - b)(y - 1) = P(x, y), \\ \dot{y} &= -(bx + 2ay - a)(y - 1) = Q(x, y), \end{aligned} \tag{81}$$

which has the invariant algebraic curves $C_1 = x^2 + 2y - 1 = 0$ with cofactor $K_1 = 2[(1 - b)x - ay + a]$ and $C_2 = y - 1 = 0$ with cofactor $K_2 = -(bx + 2ay - a)$. We note that the highest order term of C_1 has a repeated factor x.

COROLLARY 11.4. *There exist values of the parameters a and b for which system* (81) *is not Darbouxian integrable.*

In the rest of the section we shall prove the first of these stated four results, for the proof of the others see [33].

PROOF OF THEOREM 11.1. In the proof of this theorem we will use intensively the Hilbert's Nullstellensatz (see for instance, [46]):

Set $A, B_i \in \mathbb{C}[x, y]$ *for* $i = 1, \dots, r$. *If* A *vanishes in* $\mathbb{C}^2$ *whenever the polynomials* B_i *vanish simultaneously, then there exist polynomials* $M_i \in \mathbb{C}[x, y]$ *and a nonnegative integer* n *such that* $A^n = \sum_{i=1}^{r} M_i B_i$. *In particular, if all* B_i *have no common zero, then there exist polynomial* M_i *such that* $\sum_{i=1}^{r} M_i B_i = 1$.

In what follows if we have a polynomial A we will denotes its degree by a. If we do not say anything we denote by C^c the homogeneous part of degree c for the polynomial C. We shall need the following result.

LEMMA 11.5. *If* C^c *has no repeated factors, then* $(C_x, C_y) = 1$.

PROOF. Suppose that $(C_x, C_y) \neq 1$. Then there exists a polynomial A nonconstant such that $A|C_x$ and $A|C_y$. Here $A|C_x$ means that the polynomial A divides the polynomial C_x. Therefore, $A^a|(C^c)_x$ and $A^a|(C^c)_y$. By the Euler theorem for homogeneous polynomials we have that $x(C^c)_x + y(C^c)_y = cC^c$. So $A^a|C^c$. Since A^a, $(C^c)_x$, $(C^c)_y$ and C^c are homogeneous polynomials of $\mathbb{C}[x, y]$ and A^a divides $(C^c)_x$, $(C^c)_y$ and C^c, the linear factors of A^a having multiplicity m, must be linear factors of C^c having multiplicity $m + 1$. This last statement follows easily identifying the linear factors of the homogeneous polynomial $C^c(x, y)$ in two variables with the roots of the polynomial $C^c(1, z)$ in the variable z. Hence, A^a is a repeated factor of C^c. It is in contradiction with the assumption. ☐

We first consider the case that system (1) has a given invariant algebraic curve.

LEMMA 11.6. *Assume that polynomial system* (1) *of degree* m *has an invariant algebraic curve* $C = 0$ *of degree* c, *and that* C *satisfies condition* (i) *of Theorem* 11.1.

(a) *If* $(C_x, C_y) = 1$, *then system* (1) *has the following normal form*:

$$\dot{x} = AC - DC_y, \qquad \dot{y} = BC + DC_x, \tag{82}$$

where A, B *and* D *are suitable polynomials.*

(b) *If* C *satisfies condition* (ii) *of Theorem* 11.1, *then system* (1) *has the normal form* (82) *with* $a, b \leqslant m - c$ *and* $d \leqslant m - c + 1$. *Moreover, if the highest order term* C^c *of* C *does not have the factors* x *and* y, *then* $a \leqslant p - c$, $b \leqslant q - c$ *and* $d \leqslant \min\{p, q\} - c + 1$.

PROOF. (a) Since there are no points at which C, C_x and C_y vanish simultaneously, from Hilbert's Nullstellensatz we obtain that there exist polynomials E, F and G such that

$$EC_x + FC_y + GC = 1. \tag{83}$$

As C satisfies the definition of invariant algebraic curve, we get from (83) that

$$K = (KE + GP)C_x + (KF + GQ)C_y.$$

Substituting K into the definition of invariant algebraic curve $C = 0$, we get

$$\big[P - (KE + GP)C\big]C_x = -\big[Q - (KF + GQ)C\big]C_y.$$

Since $(C_x, C_y) = 1$, there exists a polynomial D such that

$$P - (KE + GP)C = -DC_y, \qquad Q - (KF + GQ)C = DC_x.$$

This proves that system (1) has the form (82) with $A = KE + GP$ and $Q = KF + GQ$.

(b) From (a) and Lemma 11.5 we get that system (1) has the normal form (82). Without loss of generality we can assume that $p \leqslant q$.

We first consider the case that C^c has neither factor x nor y. So we have $(C^c, (C^c)_x) = 1$ and $(C^c, (C^c)_y) = 1$, where $(C^c)_x$ denotes the derivative of C^c with respect to x. In (82) we assume that $a > p - c$, otherwise the statement follows. Then $d = a + 1$. Moreover, from the highest order terms of (82) we get $A^a C^c = D^{a+1} C_y^{c-1}$, where C_y^{c-1} denotes the homogeneous part with degree $c - 1$ of C_y. Since $(C^c, C_y^{c-1}) = 1$, there exists a polynomial F such that $A^a = F C_y^{c-1}$, $D^{a+1} = F C^c$. In (82) we replace A by $A - FC_y$ and D by $D - FC$, so the degrees of polynomials under consideration reduce by one. We continue this process and do the same for $\dot{y}$ until we reach a system of the form

$$\dot{x} = AC - DC_y, \qquad \dot{y} = BC + EC_x, \tag{84}$$

with $a \leqslant p - c$, $d \leqslant p - c + 1$, $b \leqslant q - c$ and $e \leqslant q - c + 1$. Since $C = 0$ is an invariant algebraic curve of (84), from the definition of invariant algebraic curve we get

$$C(AC_x + BC_y) + C_x C_y (E - D) = KC.$$

This implies that there exists a polynomial R such that $E - D = RC$, because C with C_x and C_y are coprime.

If $e \geqslant d$, then $r = e - c$. We write $BC + EC_x = (B + RC_x)C + DC_x$ and denote $B + RC_x$ again by B, then system (84) has the form (82) where A, B and D have the required degrees.

If $e < d$, then $r = d - c$. We write $AC - DC_y = (A + RC_y)C - EC_y$ and denote $A + RC_y$ again by A, then system (84) has the form (82) where A, B and E instead of D have the required degrees. This proves the second part of (b).

Now we prove the first part of (b). We note that even though C^c has no repeated factor, C^c with C_x^{c-1} or C_y^{c-1} may have a common factor in x or y (for example $C^3 = x(x^2 + y^2)$, $C^3 = y(x^2 + y^2)$ or $C^4 = xy(x^2 + y^2)$). In order to avoid this difficulty we rotate system (1) slightly such that C^c has no factors in x and y. Then, applying the above method to the new system we get that the new system has a normal form (82) with the degrees of A, B and D as those of the second part of (b).

We claim that under affine changes system (82) preserves its form and the upper bound of the polynomials, i.e., $a, b \leqslant m - c$ and $d \leqslant m - c + 1$. Indeed, using the affine change of variables $u = a_1 x + b_1 y + c_1$ and $v = a_2 x + b_2 y + c_2$ with $a_1 b_2 - a_2 b_1 \neq 0$, system (82) becomes

$$\begin{aligned} \dot{u} &= (a_1 A + b_1 B)C - (a_1 b_2 - a_2 b_1) DC_v, \\ \dot{v} &= (a_2 A + b_2 B)C + (a_1 b_2 - a_2 b_1) DC_u. \end{aligned}$$

Hence, the claim follows. This completes the proof of (b), and consequently we have the proof of the lemma. □

LEMMA 11.7. *Assume that* $C = 0$ *and* $D = 0$ *are different irreducible invariant algebraic curves of system* (1) *of degree* m*, and that they satisfy conditions* (i) *and* (iii) *of Theorem* 11.1.

(a) *If* $(C_x, C_y) = 1$ *and* $(D_x, D_y) = 1$*, then system* (1) *has the normal form*

$$\dot{x} = ACD - EC_yD - FCD_y, \qquad \dot{y} = BCD + EC_xD + FCD_x. \tag{85}$$

(b) *If* C *and* D *satisfy conditions* (ii) *and* (v)*, then system* (1) *has the normal form* (85) *with* $a, b \leqslant m - c - d$ *and* $e, f \leqslant m - c - d + 1$.

PROOF. Since $(C, D) = 1$, the curves C and D have finitely many intersection points. By assumption (i) at each of such points there is at least one nonzero first derivative of both C and D. In a similar way to the proof of the claim inside the proof of Lemma 11.6, we can prove that under an affine change of the variables, system (85) preserves its form and the bound for the degrees of A, B, E and F. So, we rotate system (1) slightly such that all first derivatives of C and D are not equal to zero at the intersection points.

From the Hilbert's Nullstellensatz, there exist polynomials M_i, N_i and R_i, $i = 1, 2$, such that

$$M_1C + N_1D + R_1D_y = 1, \qquad M_2C + N_2D + R_2C_y = 1. \tag{86}$$

By Lemma 11.6 we get that

$$P = A_1C - E_1C_y = G_1D - F_1D_y, \tag{87}$$

for some polynomials A_1, E_1, G_1 and F_1. Moreover, using the first equation of (86) we have $F_1 = SC + TD + UC_y$ for some polynomials S, T and U. Substituting F_1 into (87) we obtain that

$$(A_1 + SD_y)C + (-G_1 + TD_y)D + (-E_1 + UD_y)C_y = 0. \tag{88}$$

Using the second equation of (86) and (88) to eliminate C_y we get

$$-E_1 + UD_y = VC + WD, \tag{89}$$

for some polynomials V and W. Substituting (89) into (88), we have

$$(A_1 + SD_y + VC_y)C = (G_1 - TD_y - WC_y)D.$$

Since $(C, D) = 1$, there exists a polynomial K such that

$$A_1 + SD_y + VC_y = KD, \qquad G_1 - TD_y - WC_y = KC. \tag{90}$$

Substituting E_1 of (89) and A_1 of (90) into (87), then we have

$$P = KCD - SCD_y + WC_yD - UC_yD_y. \tag{91}$$

Similarly, we can prove that there exist some polynomials K', S', W' and U' such that

$$Q = K'CD + S'CD_x - W'C_xD + U'C_xD_x. \tag{92}$$

Since C is an invariant algebraic curve of (1), we have that $PC_x + QC_y = K_CC$ for some polynomial K_C. Using (91) and (92) we get

$$\begin{aligned} K_CC = C\big[&D(KC_x + K'C_y) - SC_xD_y + S'C_yD_x\big] \\ &+ C_xC_y\big[D(W - W') - UD_y + U'D_x\big]. \end{aligned}$$

As C, C_x and C_y are coprime, there exists a polynomial Z such that

$$D(W - W') - UD_y + U'D_x = ZC. \tag{93}$$

Substituting the expression $DW - UD_y$ into (91), we get

$$P = KCD - SCD_y + W'C_yD - U'C_yD_x + ZCC_y. \tag{94}$$

Since $D = 0$ is an invariant algebraic curve of system (1), we have $PD_x + QD_y = K_DD$ for some polynomial K_D. Using (92) and (94) we get

$$\begin{aligned} K_DD = D\big[&C(KD_x + K'D_y) + W'(C_yD_x - C_xD_y)\big] \\ &+ D_x\big[CD_y(-S + S') + U'(C_xD_y - C_yD_x) + ZCC_y\big]. \end{aligned}$$

As D and D_x are coprime, there exists a polynomial M such that

$$CD_y(-S + S') + U'(C_xD_y - C_yD_x) + ZCC_y = MD. \tag{95}$$

The curves C and D are transversal implies that C, D and $C_xD_y - C_yD_x$ have no common zeros. From Hilbert's Nullstellensatz, there exist some polynomials M_3, N_3 and R_3 such that

$$M_3C + N_3D + R_3(C_xD_y - C_yD_x) = 1. \tag{96}$$

Eliminating the term $C_xD_y - C_yD_x$ from (95) and (96), we obtain that $U' = IC + JD$ for some polynomials I and J. Hence, Equation (95) becomes

$$\begin{aligned} &C\big[I(C_xD_y - C_yD_x) + D_y(-S + S') + ZC_y\big] \\ &\quad + D\big[J(C_xD_y - C_yD_x) - M\big] = 0. \end{aligned}$$

Since $(C, D) = 1$, there exists a polynomial G such that

$$M = J(C_x D_y - C_y D_x) + GC,$$
$$I(C_x D_y - C_y D_x) + D_y(-S + S') + ZC_y = GD.$$

Substituting $ZC_y - SD_y$ and U' into (94) we obtain that

$$P = (K + G)CD - (IC_x + S')CD_y + (W' - JD_x)DC_y.$$

This means that P can be expressed in the form (91) with $U = 0$.

Working in a similar way, we can express Q in the form (92) with $U' = 0$. Thus, (93) is reduced to $D(W - W') = ZC$. Hence, we have $W = W' + HC$ for some polynomial H. Consequently, $Z = HD$. Therefore, from (95) we obtain that $CD_y(-S + S') = D(M - HCC_y)$. Since $(C, D) = 1$ and $(D, D_y) = 1$, we have $S = S' + LD$ for some polynomial L. Substituting W and S into (91) we obtain that P and Q have the form (85). This proves statement (a).

As in the proof of Lemma 11.6 we can prove that under suitable affine change of variables the form of system (85) and the bound of the degrees of the polynomials A, B, E and F are invariant. So, without loss of generality we can assume that the highest order terms of C and D are neither divisible by x nor y.

By the assumptions, the conditions of statement (a) hold, so we get that system (1) has the form (85). If the bounds of the degrees of A, B, E and F are not satisfied, we have by (85) that

$$\begin{aligned} &A^a C^c D^d - E^e C_y^{c-1} D^d - F^f C^c D_y^{d-1} = 0, \\ &B^b C^c D^d + E^e C_x^{c-1} D^d + F^f C^c D_x^{d-1} = 0. \end{aligned} \tag{97}$$

We remark that if one of the numbers $a + c + d$, $e + c - 1 + d$ and $f + c + d - 1$ is less than the other two, then its corresponding term in the first equation of (97) is equal to zero. The same remark is applied to the second equation of (97). From the hypotheses it follows that C^c and C_y^{c-1} are coprime, and also D^d and D_y^{d-1}, and C^c and D^d, respectively. Hence, from these last two equations we obtain that there exist polynomials K and L such that $E^e = KC^c$, $F^f = LD^d$, and

$$A^a = KC_y^{c-1} + LD_y^{d-1}, \qquad B^b = -KC_x^{c-1} - LD_x^{d-1}.$$

We rewrite Equation (85) as

$$\dot{x} = (A - KC_y - LD_y)CD - (E - KC)C_y D - (F - LD)CD_y,$$
$$\dot{y} = (B + KC_x + LD_x)CD + (E - KC)C_x D + (F - LD)CD_x.$$

Thus, we reduce the degrees of A, B, E and F in (85) by one. We can continue this process until the bounds are reached. This completes the proof of statement (b). ☐

LEMMA 11.8. *Let $C_i = 0$ for $i = 1, \dots, p$ be different irreducible invariant algebraic curves of system* (1) *with* $\deg C_i = c_i$. *Assume that C_i satisfy conditions* (i), (iii) *and* (iv) *of Theorem* 11.1. *Then*

(a) *If $(C_{ix}, C_{iy}) = 1$ for $i = 1, \dots, p$, then system* (1) *has the normal form*

$$\dot{x} = \left(B - \sum_{i=1}^{p} \frac{A_i C_{iy}}{C_i}\right) \prod_{i=1}^{p} C_i, \qquad \dot{y} = \left(D + \sum_{i=1}^{p} \frac{A_i C_{ix}}{C_i}\right) \prod_{i=1}^{p} C_i, \tag{98}$$

where B, D and A_i are suitable polynomials.

(b) *If C_i satisfy conditions* (ii) *and* (v) *of Theorem* 11.1, *then system* (1) *has the normal form* (98) *with $b, d \leqslant m - \sum_{i=1}^{p} c_i$ and $a_i \leqslant m - \sum_{i=1}^{p} c_i + 1$.*

PROOF. We use induction to prove this lemma. By Lemmas 11.6 and 11.7 we assume that for any l with $2 \leqslant l < p$ we have

$$P = \sum_{i=1}^{l} \left(B_i - \frac{A_i C_{iy}}{C_i}\right) \prod_{i=1}^{l} C_i, \qquad Q = \sum_{i=1}^{l} \left(D_i + \frac{A_i C_{ix}}{C_i}\right) \prod_{i=1}^{l} C_i,$$

where $\sum_{i=1}^{l} B_i = B$ and $\sum_{i=1}^{l} D_i = D$. Since $C_{l+1} = 0$ is an invariant algebraic curve, from Lemma 11.6 we get that there exist some polynomials E, G and H such that

$$\begin{aligned} P &= \sum_{i=1}^{l} \left(B_i - \frac{A_i C_{iy}}{C_i}\right) \prod_{i=1}^{l} C_i = E C_{l+1} - G C_{l+1,y}, \\ Q &= \sum_{i=1}^{l} \left(D_i + \frac{A_i C_{ix}}{C_i}\right) \prod_{i=1}^{l} C_i = H C_{l+1} + G C_{l+1,x}. \end{aligned} \tag{99}$$

Now we consider the curves

$$K_j = \prod_{\substack{i=1 \\ i \neq j}}^{l} C_i = 0, \quad j = 1, \dots, l.$$

From the assumptions we obtain that there is no points at which all the curves $K_i = 0$ and $C_{l+1} = 0$ intersect. Otherwise, at least three of the curves $C_i = 0$ for $i = 1, \dots, l + 1$ intersect at some point. Hence, there exist polynomials U and V_i for $i = 1, \dots, l$ such that

$$U C_{l+1} + \sum_{i=1}^{l} V_i K_i = 1. \tag{100}$$

Using this equality, we can rearrange (99) as

$$\begin{aligned}(E - GUC_{l+1,y})C_{l+1} &= \sum_{i=1}^{l}(B_iC_i - A_iC_{iy} + GV_iC_{l+1,y})K_i,\\ (H + GUC_{l+1,x})C_{l+1} &= \sum_{i=1}^{l}(D_iC_i + A_iC_{ix} - GV_iC_{l+1,x})K_i.\end{aligned} \tag{101}$$

Using (100) and (101) to eliminate C_{l+1} we obtain that

$$E - GUC_{l+1,y} = \sum_{i=1}^{l} I_iK_i, \qquad H + GUC_{l+1,x} = \sum_{i=1}^{l} J_iK_i,$$

for some polynomials I_i and J_i. Substituting these last equalities into (101), we have

$$\begin{aligned}&\sum_{i=1}^{l}(B_iC_i - A_iC_{iy} + GV_iC_{l+1,y} - I_iC_{l+1})K_i = 0,\\ &\sum_{i=1}^{l}(D_iC_i + A_iC_{ix} - GV_iC_{l+1,x} - J_iC_{l+1})K_i = 0.\end{aligned} \tag{102}$$

It is easy to check that the expression multiplying K_i in the two summations of (102) are divisible by C_i. Hence, there exist polynomials L_i and F_i for $i = 1, \dots, l$ such that

$$\begin{aligned}&B_iC_i - A_iC_{iy} + GV_iC_{l+1,y} - I_iC_{l+1} = L_iC_i,\\ &D_iC_i + A_iC_{ix} - GV_iC_{l+1,x} - J_iC_{l+1} = F_iC_i.\end{aligned} \tag{103}$$

So, from (102) we get that $\sum_{i=1}^{l} L_i = 0$ and $\sum_{i=1}^{l} F_i = 0$. This implies that (99) can be rewritten as

$$\begin{aligned}P &= \sum_{i=1}^{l}\big((B_i - L_i)C_i - A_iC_{iy}\big)K_i,\\ Q &= \sum_{i=1}^{l}\big((C_i - F_i)C_i + A_iC_{ix}\big)K_i.\end{aligned} \tag{104}$$

Moreover, we write (103) in the form

$$\begin{aligned}&(B_i - L_i)C_i - A_iC_{iy} = I_iC_{l+1} - GV_iC_{l+1,y} = P_i,\\ &(D_i - F_i)C_i + A_iC_{ix} = J_iC_{l+1} + GV_iC_{l+1,x} = Q_i.\end{aligned} \tag{105}$$

It is easy to see that C_i and C_{l+1} are invariant algebraic curves of the system $\dot{x} = P_i$, $\dot{y} = Q_i$. So, from statement (a) of Lemma 11.7 we can obtain that

$$P_i = (B_i - L_i)C_i - A_iC_{iy} = X_iC_iC_{l+1} - Y_iC_{iy}C_{l+1} - N_iC_iC_{l+1,y},$$
$$Q_i = (D_i - F_i)C_i + A_iC_{ix} = Z_iC_iC_{l+1} + Y_iC_{ix}C_{l+1} + N_iC_iC_{l+1,x}.$$

Substituting these last two equations into (104), we obtain that system (1) has the form (98) with the $l + 1$ invariant algebraic curves $C_1, \dots, C_{l+1}$. From induction we have finished the proof of statement (a).

The proof of statement (b) is almost identical with those of Lemma 11.7(b), so we shall omit it here. Hence, this ends the proof of the lemma. ☐

PROOF OF THEOREM 11.1. From Lemma 11.8 it follows statement (a) of Theorem 11.1. By checking the degrees of polynomials A_i, B and D in statement (b) of Lemma 11.8 we obtain statement (b) of Theorem 11.1.

From statement (a) of Lemma 11.8, we can rearrange system (1) such that it has the form (98). But from statement (b) of Lemma 11.8 we must have $B = 0$, $D = 0$ and $A_i = 0$. This proves statement (c) of Theorem 11.1. ☐

12. Elementary and Liouvillian first integrals

We now examine the effectiveness of the Darbouxian theory of integrability, i.e., what sort of integrals does it capture. The surprising result is that in some sense it captures every "closed form solution". However, we clearly need to make this idea precise before we can explain the known results. Here, we follow the paper [30].

The idea of calculating what sort of functions can arise as the result of evaluating an indefinite integral or solving a differential equation goes back to Liouville. The modern formulation of these ideas is usually done through differential algebra. The advantage over an analytic approach is first that the messy details of branch points etc., is hidden completely, and second that the way is open to apply these methods to symbolic computation.

We assume that the set of functions we are interested in form a field together with a number of *derivations*. We call such an object a *differential field*. The process of adding more functions to a given set of functions is described by a tower of such fields:

$$F_0 \subset F_1 \subset \cdots \subset F_n.$$

Of course, we must also specify how the derivations of F_0 are extended to derivations on each F_i.

The fields we are interested in arise by adding exponentials, logarithms or the solutions of algebraic equations based on the previous set of functions. That is we take

$$F_i = F_0(\theta_1, \dots, \theta_i),$$

where one of the following holds:

(i) $\delta\theta_i = \theta_i \delta g$, for some $g \in F_{i-1}$ and for each derivation δ.
(ii) $\delta\theta_i = g^{-1}\delta g$, for some $g \in F_{i-1}$ and for each derivation δ.
(iii) θ_i is algebraic over F_{i-1}. If we have such a tower of fields, F_n is called an *elementary* extension of F_0.

This is essentially what we mean by a function being expressible in closed form. We call the set of all elements of a differential field which are annihilated by all the derivations of the field the *field of constants*. We shall always assume that the field of constants is algebraically closed.

THEOREM 12.1 (Liouville theorem). *If an element f in a differential field F is the derivative of an element g in an elementary extension field G with the same field of constants, then we must have*

$$f = h_0 + \sum c_i \ln(h_i),$$

where c_i are constants and all the h_i lie in F.

We say that our system (1) has an *elementary* first integral if there is an element u in an elementary extension field of the field of rational functions $\mathbb{C}(x, y)$ with the same field of constants such that $Du = 0$. The derivations on $\mathbb{C}(x, y)$ are of course $\mathrm{d}/\mathrm{d}x$ and $\mathrm{d}/\mathrm{d}y$.

THEOREM 12.2 (Prelle and Singer [87]). *If the system* (2) *has an elementary first integral, then there is also an elementary first integral of the form*

$$v_0 + \sum c_i \ln(v_i),$$

where the c_i are constants and the v_i are algebraic functions over $\mathbb{C}(x, y)$.

It is known that we cannot strengthen this theorem to make all the v_i rational functions in x and y. By manipulating this formula and taking traces we obtain the following corollary.

COROLLARY 12.3. *In the situation above there is always an integrating factor of the form $R^{1/N}$, with $R \in \mathbb{C}(x, y)$ and N an integer.*

Thus the method of Darboux finds all elementary first integrals. For a direct proof of this corollary, see [20].

Another class of integrals we are interested in are the *Liouvillian* ones. Here we say that an extension F_n is a *Liouvillian* extension of F_0 if there is a tower of differential fields as above which satisfies conditions (i), (iii) or

(ii)′ $\delta_\alpha \theta_i = h_\alpha$ for some elements $h_\alpha \in F_{i-1}$ such that $\delta_\alpha h_\beta = \delta_\beta h_\alpha$.

This last condition, mimics the introduction of line integrals into the class of functions. Clearly (ii) is included in (ii)′.

This class of functions represents those functions which are obtainable “by quadratures”. An element u of a Liouvillian extension field of $\mathbb{C}(x, y)$ with the same field of constants is said to be a *Liouvillian* first integral.

THEOREM 12.4 (Singer [95]). *If the system* (2) *has a Liouvillian first integral, then it has an integrating factor of the form*

$$e^{\int U\,\mathrm{d}x+V\,\mathrm{d}y}, \qquad U_y = V_x,$$

where U and V are rational functions.

It can be shown that this last expression can always be integrated to get a Darbouxian function. More specifically,

THEOREM 12.5. *Let U, V be two rational functions with*

$$U_y = V_x,$$

then

$$\int U\,\mathrm{d}x + V\,\mathrm{d}y = w_0 + \sum c_i \ln(w_i),$$

for some constants c_i and w_i rational functions.

Hence we have

COROLLARY 12.6. *If system* (1) *has a Liouvillian first integral, then there is a Darbouxian integrating factor.*

For a direct proof of this result see Pereira [85].

Thus the method of Darboux finds all Liouvillian solutions. However, there is another surprising result.

THEOREM 12.7 (Singer [95]). *Suppose that a trajectory of* (1) *can be described by a function in a Liouvillian extension of $\mathbb{C}(x, y)$. Then either this function is a first integral of the system, or the function is a polynomial.*

Thus a system has a Darbouxian integrating factor, or the only trajectories that can be described by closed form solutions or quadratures are the polynomial ones.

What does the general first integral of a system with a Darbouxian integrating factor look like? Generically we can show that the first integral is also Darbouxian [27], but stranger things can happen. A reasonable conjecture which embodies all the cases which we know is that it is a sum of a Darbouxian function and terms of the form

$$\int^{R(x,y)} e^{s(u)} \prod k_i(u)^{l_i}\,\mathrm{d}u,$$

where R, s and the k_i are rational functions. Several examples of these have been given by Żołądek [106].

13. Liouvillian first integrals for the planar Lotka–Volterra system

The aim of this section is to present the complete classification of the Liouvillian first integrals for the quadratic Lotka–Volterra systems in $\mathbb{C}^2$:

$$\begin{aligned} \dot{x} &= x(ax + by + c), \\ \dot{y} &= y(Ax + By + C). \end{aligned} \tag{106}$$

We say that system (106) has a *Liouvillian first integral* if it has a first integral or an integrating factor given by a Darbouxian function, see for more details Section 12 or Singer [95]. We note that our Darbouxian theory of integrability takes into account the invariant algebraic curves and their multiplicity through the exponential factors, see Section 1.5 or for more details [34]. We emphasize that a complete characterization of Liouvillian integrable polynomial differential systems has been made for very few families of differential systems. We follows the paper [10].

As we shall see, systems (106) can be formulated as the following quadratic homogeneous Lotka–Volterra systems in $\mathbb{C}^3$:

$$\begin{aligned} \dot{x} &= x\big((b - B)y + cz\big), \\ \dot{y} &= y\big((A - a)x + Cz\big), \\ \dot{z} &= -z(ax + By). \end{aligned} \tag{107}$$

Several authors and mainly Moulin Ollagnier [77] have studied the Liouvillian first integrals of the system

$$\begin{aligned} \dot{x} &= x(\omega Cy + z) = x\Phi_x, \\ \dot{y} &= y(x + \omega Az) = y\Phi_y, \\ \dot{z} &= z(\omega Bx + y) = z\Phi_z. \end{aligned} \tag{108}$$

In fact, these homogeneous Lotka–Volterra systems in $\mathbb{C}^3$ can be thought as the planar projective version of the following planar Lotka–Volterra systems in $\mathbb{C}^2$:

$$\begin{aligned} \dot{x} &= x\big(-\omega Bx + (\omega C - 1)y + 1\big), \\ \dot{y} &= y\big((1 - \omega B)x - y + \omega A\big). \end{aligned} \tag{109}$$

For more details between the affine and the projective version of a planar polynomial vector field, see Section 6, or for instance [77] or [70].

We note that if $c(a - A)B \neq 0$, then system (106) becomes system (109) with the following rescaling of the variables:

$$(x, y, t) \to \left(\frac{c}{A - a}x, -\frac{c}{B}y, \frac{1}{c}t\right), \tag{110}$$

where

$$\omega A = \frac{C}{c}, \qquad \omega B = \frac{a}{a-A}, \qquad \omega C = \frac{B-b}{B}. \tag{111}$$

The Darbouxian theory of integrability says that if a planar polynomial differential system of degree $m = 2$ has at least $m(m+1)/2 = 3$ invariant algebraic curves or exponential factors, then the system has a Liouvillian first integral (see again Section 12 or Singer [95]). This result will play a main role in the classification of all Liouvillian first integrals of systems (106).

We want to show that the Darbouxian theory of integrability is one of the best methods for finding first integrals of polynomial ordinary differential equations. For showing this we choose the 2-dimensional Lotka–Volterra systems (106) as paradigmatic systems for the study of the integrability, and we complete the classification of their Liouvillian first integrals. This model, introduced by Lotka [72] and Volterra [100], appears in ecology where it models two species in competition, and it has been widely used in applied mathematics, in chemistry and in a large variety of problems in physics.

Moulin Ollagnier in [77] classified the irreducible systems (108) (and his results provide almost the classification of the nonirreducible ones) having a Liouvillian first integral. In the paper [10] the authors, first, provide a new tool for studying the existence of Liouvillian first integrals, and second they use this tool for completing the classification of systems (107).

System (107) is called *irreducible* if the 1-form

$$\begin{aligned}\omega &= (\Phi_z - \Phi_y)yz\,\mathrm{d}x + (\Phi_x - \Phi_z)zx\,\mathrm{d}y + (\Phi_y - \Phi_x)xy\,\mathrm{d}z \\ &= -(Ax + By + Cz)yz\,\mathrm{d}x + (ax + by + cz)zx\,\mathrm{d}y \\ &\quad + \big((A-a)x + (B-b)y + (C-c)z\big)xy\,\mathrm{d}z\end{aligned}$$

has no nontrivial common factor in its three components, for more details see [77].

Systems (109) always have two invariant algebraic curves $x = 0$ and $y = 0$. From the Darbouxian theory of integrability (see Theorem 2.1), it is known that if they have another invariant algebraic curve or an exponential factor, then the system has a Liouvillian first integral. We remark that almost all the Liouvillian integrable systems (109) given by Moulin Ollagnier in [77] have a third invariant algebraic curve. For system (107) it is shown in [10] that in seven of the new Liouvillian integrable cases the integrability is due to the existence of an exponential factor.

We associate to a given 3-dimensional Lotka–Volterra system (108) two if $\omega A\omega B \times \omega C = 0$, or five if $\omega A\omega B\omega C \neq 0$, *equivalent* 3-dimensional Lotka–Volterra systems. The first two are obtained doing circular permutation of the variables x, y, z and of the parameters $\omega A, \omega B, \omega C$; i.e.,

$$(x, y, z, \omega A, \omega B, \omega C) \to (y, z, x, \omega B, \omega C, \omega A);$$

and the remainder three systems are obtained doing the transformation

$$(x, y, z, \omega A, \omega B, \omega C) \to (\omega B x, \omega A z, \omega C y, 1/\omega C, 1/\omega B, 1/\omega A),$$

and the two new transformations obtained from this one doing the above circular permutations. Sometimes the conjugated systems to these previous ones are different, obtaining in all 12 related systems. We say that all these Lotka–Volterra systems are E *equivalent*. All the results of this paper are stated modulo these E equivalences. Of course, the analogous E equivalences exist for 2-dimensional Lotka–Voterra systems (109).

The main results of [77] and [10] are the following two theorems, respectively.

THEOREM 13.1. *Consider an irreducible system* (109). *Then the following statements are equivalent.*

(a) *System* (109) *has a Liouvillian first integral.*
(b) *System* (109) *has an invariant algebraic curve* $f = 0$ *different from* $x = 0$ *and* $y = 0$.
(c) *The triple* $[\omega A, \omega B, \omega C]$ *of parameters of system* (109) *falls, up to E equivalences, in one of the cases of the following list.*

We introduce the following parameters

$$p = -\omega A - \frac{1}{\omega B}, \qquad q = -\omega B - \frac{1}{\omega C}, \qquad r = -\omega C - \frac{1}{\omega A}.$$

The numbers p, q and r are related with the Kowalevskaya exponents of system (108), see [52] for more details.

In the first two cases ωA, ωB and ωC are related by a unique condition and f has degree 1.

1. $\omega A\omega B\omega C + 1 = 0$, $f = x - \omega Cy + \omega A\omega C$.

2. $\omega B = 1$ where $\omega C = 0$ is possible, $f = y - \omega A$.

Case 2 is formed by three E equivalent systems. The other two are: $\omega C = 1$ where $\omega A = 0$ is possible, $f = 1 - \omega Bx$; and $\omega A = 1$ where $\omega B = 0$ is possible, $f = x - \omega Cy$.

In the next two cases ωA, ωB and ωC are related by two conditions and f has degree 2.

3. $p = 1$, $q = 1$, consequently $\omega A\omega B\omega C = 1$ and $r = 1$; $f = \omega A^2(\omega Bx - 1)^2 - 2\omega A(\omega Bx + 1)y + y^2$.

4. $\omega A = 2$, $q = 1$; $f = (x - \omega Cy)^2 - 2\omega C^2 y$.

Case 4 is formed by five other E equivalent systems, namely:
$\omega B = 2, r = 1, f = (y - \omega A)^2 - 2\omega A^2 x$;
$\omega C = 2, p = 1, f = (1 - \omega Bx)^2 - 2\omega B^2 xy$;
$\omega C = 1/2, p = 1, f = (Bx - 1)^2 - 2y\omega C/\omega A$;
$\omega A = 1/2, q = 1, f = (\omega Cy - x)^2 - 2x\omega A/\omega B$;
$\omega B = 1/2, r = 1, f = (\omega A - y)^2 - 2xy\omega B/\omega C$.

There is a finite number of isolated triples of complex numbers providing an invariant algebraic curve of system (21) with degree 3, 4 or 6. Here i stands for the square root of -1 and $j = (-1 + i\sqrt{3})/2$ is a third root of 1. Firstly, these cases were found in [53] using the so-called Painlevé analysis, see [52].

5. $[\omega A, \omega B, \omega C] = [(j - 1)/3, j - 1, j]$ or equivalently $[p, q, r] = [1, 2, 2]$, here $(\omega A\omega B\omega C)^3 = -1$; $f = 9x^3 - 27jx^2y + 9(2 + j)x^2 - 27(1 + j)xy^2 + 9(2 + j)xy + 9(1 + j)x - 9y^3 - 9(1 - j)y^2 + 9jy + 1 + 2j$.

6. $[\omega A,\omega B,\omega C]=[(i-2)/5,(i-3)/2,i-1]$ or equivalently $[p,q,r]=[1,2,3]$, here $(\omega A\omega B\omega C)^2=-1$; $f=625x^4+2500(1-i)x^3y+500(3+i)x^3-7500ix^2y^2+5000x^2y+300(4+3i)x^2-5000(1+i)xy^3+1000(1-3i)xy^2+200(9+13i)xy+40(9+13i)x-2500y^4+2000(-2+i)y^3+600(-3+4i)y^2+80(-2+11i)y+4(7+24i)$.

7. $[\omega A,\omega B,\omega C]=[-1+j,(-2+j)/7,(-4+j)/3]$ or equivalently $[p,q,r]=[4,1,2]$, here $(\omega A\omega B\omega C)^3=1$; $f=6(-397-683j)y^5+405(8+5j)x^4+540(19+18j)x^3+135(-323+37j)y^2+45(-360-323j)y^4+162(3+j)x^5+6(25-211j)xy^5+45(16-39j)x^2y^4+135(5-3j)x^4y^2+162(62+149j)x+54(-286+397j)y+60(17-20j)x^3y^3+54(4-j)x^5y+27x^6+(-37-360j)y^6+27(37+360j)+324(-87+62j)xy+54(94+71j)x^2y+378(-149-87j)xy^2+54(179-29j)x^3y+27(-104-947j)x^2y^2+18(-1427-2074j)xy^3+162(18-j)x^4y+108(52-41j)x^3y^2+36(73-328j)x^2y^3+126(-23-94j)xy^4+60(-683-286j)y^3+405(39+55j)x^2$.

Case 5 is formed by six E equivalent systems which can be obtained doing circular permutations to $[\omega A,\omega B,\omega C]$ and to $[1/\omega C,1/\omega B,1/\omega A]$. Cases 6 and 7 are formed by twelve E equivalent systems which are obtained as in Case 5, and additionally conjugating all the parameters.

There are some isolated triples of rational numbers providing an irreducible invariant algebraic curve of degree 3, 4 or 6.

8. $[\omega A,\omega B,\omega C]=[-7/3,3,-4/7]$, $f=-259308x^3-185220x^2y+259308x^2+567xy^3-13230xy^2-71001xy-86436x+324y^4+3024y^3+10584y^2+16464y+9604$.

9. $[\omega A,\omega B,\omega C]=[-3/2,2,-4/3]$, $f=108x^2+6xy^2+180xy-108x+8y^3+36y^2+54y+27$.

10. $[\omega A,\omega B,\omega C]=[2,4,-1/6]$, $f=216x^3+108x^2y-54x^2+18xy^2-36xy+y^3-4y^2+4y$.

11. $[\omega A,\omega B,\omega C]=[2,-8/7,1/3]$, $f=216x^3+189x^2+882xy-343y^2+686y$.

12. $[\omega A,\omega B,\omega C]=[6,1/2,-2/3]$, $f=9x^2y+12xy^2-144xy+432x+4y^3-72y^2+432y-864$.

13. $[\omega A,\omega B,\omega C]=[-6,1/2,1/2]$, $f=3x^2y+24xy+144x-8y^2-96y-288$.

14. $[\omega A,\omega B,\omega C]=[3,1/5,-5/6]$, $f=1296x^4+4320x^3y-6480x^3+5400x^2y^2-18900x^2y+3000xy^3-18000xy^2+27000xy+625y^4-5625y^3+16875y^2-16875y$.

15. $[\omega A,\omega B,\omega C]=[2,-13/7,1/3]$, $f=648x^4-216x^3y-252x^2y+1176xy^2-343y^3+686y^2$.

16. $[\omega A,\omega B,\omega C]=[2,2,2]$, $f=x^2+xy^2-3xy+y$.

17. $[\omega A,\omega B,\omega C]=[2,3,-3/2]$, $f=8x^2+16xy-y^3+4y^2-4y$.

18. $[\omega A,\omega B,\omega C]=[2,2,-5/2]$, $f=8x^2-4xy^2+24xy+y^4-6y^3+12y^2-8y$.

19. $[\omega A,\omega B,\omega C]=[-4/3,3,-5/4]$, $f=576x^2+864xy-384x+27y^3+108y^2+144y+64$.

20. $[\omega A,\omega B,\omega C]=[-9/4,4,-5/9]$, $f=419904x^3+279936x^2y-314928x^2+15552xy^2+69984xy+78732x-256y^4-2304y^3-7776y^2-11664y-6561$.

21. $[\omega A,\omega B,\omega C]=[-3/2,2,-7/3]$, $f=324x^2+72xy^2+864xy-324x+16y^4+96y^3+216y^2+216y+81$.

22. $[\omega A,\omega B,\omega C]=[-5/2,2,-8/5]$, $f=125000x^3-5000x^2y^2+225000x^2y-187500x^2-1600xy^4-6000xy^3+15000xy^2+87500xy+93750x-64y^6-960y^5-6000y^4-20000y^3-37500y^2-37500y-15625$.

23. $[\omega A,\omega B,\omega C]=[-10/3,3,-7/10]$, $f=81000000x^4+64800000x^3y-108000000x^3-243000x^2y^3+3240000x^2y^2+29700000x^2y+54000000x^2-97200xy^4-1296000xy^3-6480000xy^2-14400000xy-12000000x+729y^6+14580y^5+121500y^4+540000y^3+1350000y^2+1800000y+1000000$.

There is a special family where the degree of f is unbounded, see Section 4.

24. $[\omega A,\omega B,\omega C]=[-(2l+1)/(2l-1),1/2,2]$, $l=1,2,\ldots$. In this case, it is not easy to provide the explicit expression of f.

Cases 8 to 24 are formed by six E equivalent systems, with the exception of Case 16 which is formed by only two E equivalent systems.

From (110) and (111) it follows that systems (106) studied by Moulin Ollagnier are those which satisfy $a(B-b)C\neq 0$ and $c(a-A)B\neq 0$. Therefore, it remains to determine:

(i) Which are the Liouvillian integrable systems (106) with $a(B-b)C=0$ and $c(a-A)B\neq 0$?

(ii) Which are the Liouvillian integrable systems (106) that cannot be written into the form (109)? Or equivalently, what are the Liouvillian integrable systems (106) having $c(a-A)B=0$?

Now we shall determine the normal forms of the Lotka–Volterra systems (106) which can have a Liouvillian first integral and which have not been studied by Moulin Ollagnier. According with the previous paragraph those systems must be systems (106) satisfying

(i) either $a(B-b)C=0$ and $c(a-A)B\neq 0$,

(ii) or $c(a-A)B=0$.

In fact, case (ii) can be subdivided into the following two subcases:

(ii.1) $a(B-b)C\neq 0$ and $c(a-A)B=0$,

(ii.2) $a(B-b)C=0$ and $c(a-A)B=0$.

Doing the change of variables $(x,y)\to(y,x)$ it is immediate to check that the expressions $a(B-b)C$ and $c(a-A)B$ are interchanged. Therefore, this change of variables interchanges cases (i) and (ii.1). In other words, case (i) is contained in case (ii). Finally, again due to the change of variables $(x,y)\to(y,x)$ to study case (ii) is equivalent to study the case $a(B-b)C=0$.

In the next proposition we reduce the study of systems (106) with $a(B-b)C=0$ to analyze the Liouvillian integrability of some subclasses of systems (106).

PROPOSITION 13.2. *To complete the study of the Liouvillian integrable Lotka–Volterra systems* (106) *it is sufficient to consider the Liouvillian integrability of the following subclasses of systems* (106):

(lv1) $\dot{x} = cx, \qquad \dot{y} = y(Ax + By + C).$

(lv2) $\dot{x} = x(y + c), \qquad \dot{y} = y(By + 1).$

(lv3) $\dot{x} = xy, \qquad \dot{y} = y(x + By + 1).$

(lv4) $\dot{x} = xy, \qquad \dot{y} = y(Ax + By).$

(lv5) $\dot{x} = x(ax + by + c), \qquad \dot{y} = By^2.$

(lv6) $\dot{x} = x(y + 1), \qquad \dot{y} = y(x + By).$

(lv7) $\dot{x} = x(ax + by + c), \qquad \dot{y} = xy.$

(lv8) $\dot{x} = x(ax + by), \qquad \dot{y} = y(x + y).$

(lv9) $\dot{x} = x(ax + by + 1), \qquad \dot{y} = y(x + y).$

(lv10) $\dot{x} = x(x + c), \qquad \dot{y} = y(Ax + 1)$ *with* $cA = 0.$

(lv11) $\dot{x} = x(x + y + 1), \qquad \dot{y} = y(Ax + y + 1).$

PROOF. We divide the class of systems (106) satisfying $a(B - b)C = 0$ into the following four subclasses.

Case 1: $a = 0$, $C \neq 0$, $B - b$ arbitrary. Then, doing the rescaling $(x, y, t) \to (\alpha x, \beta y, t/C)$ in system (106) we obtain the system

$$\dot{x} = x\left(\frac{b\beta}{C}y + \frac{c}{C}\right), \qquad \dot{y} = y\left(\frac{A\alpha}{C}x + \frac{B\beta}{C}y + 1\right). \tag{112}$$

Now we consider the following three subcases:

(1.1) $b = 0$. Then system (106) is contained into system (lv1) after a redefinition of the parameters.

(1.2) $b \neq 0$, $A = 0$. Taking $\beta = C/b$ system (112) becomes $\dot{x} = x(y + c/C)$, $\dot{y} = y(By/b + 1)$; i.e., system (lv2) after a redefinition of the parameters.

(1.3) $bA \neq 0$. Taking $\alpha = C/A$ and $\beta = C/b$ system (112) becomes $\dot{x} = x(y + c/C)$, $\dot{y} = y(x + By/b + 1)$; or equivalently $\dot{x} = x(y + c)$, $\dot{y} = y(x + By + 1)$. Now, if $c \neq 0$, then doing the changes of variables $x = -c/Y$, $y = cX/Y$ and the rescaling of the independent variable $t \to -Yt/c$, we get system (lv9). If $c = 0$, then we obtain system (lv3).

Case 2: $a = 0, C = 0, B - b$ arbitrary. Then, doing the rescaling $(x, y, t) \to (\alpha x, \beta y, \gamma t)$ in system (106) we obtain the system

$$\dot{x} = x(b\beta\gamma y + c\gamma), \qquad \dot{y} = y(A\alpha\gamma x + B\beta\gamma y). \tag{113}$$

We deal with the following three subcases:

(2.1) $b = 0$. We obtain a particular case of system (lv1).

(2.2) $b \neq 0, c = 0$. Taking $\beta = 1/b$ and $\gamma = 1$ in (113) we get system (lv4).

(2.3) $bc \neq 0$. Taking $\beta = c/b$ and $\gamma = 1/c$ in (113) we get a particular case of system (lv5) if $A = 0$. Additionally, if $A \neq 0$ taking $\alpha = c/A$ we obtain system (lv6).

Case 3: $a \neq 0, C = 0, B - b$ arbitrary. Then, doing the rescaling $(x, y, t) \to (\alpha x, \beta y, \gamma t)$ in system (1) we obtain the system

$$\dot{x} = x(a\alpha\gamma x + b\beta\gamma y + c\gamma), \qquad \dot{y} = y(A\alpha\gamma x + B\beta\gamma y).$$

We consider four subcases:

(3.1) $A=0$. Then we obtain system (lv5).
(3.2) $A\neq 0$, $B=0$. Taking $\alpha\gamma=1/A$, we get system (lv7).
(3.3) $AB\neq 0$, $c=0$. Taking $\alpha\gamma=1/A$ and $\beta\gamma=1/B$, we have system (lv8).
(3.4) $ABc\neq 0$. Taking $\gamma=1/c$, $\alpha=c/A$ and $\beta=c/B$, we obtain system (lv9).

Case 4: $a\neq 0$, $C\neq 0$, $B-b=0$. Then, doing the rescaling $(x,y,t)\to(Cx/a,\beta y,t/C)$ in system (106) we obtain the system

$$\dot{x}=x\left(x+\frac{b\beta}{C}y+\frac{c}{C}\right),\qquad \dot{y}=y\left(\frac{A}{a}x+\frac{b\beta}{C}y+1\right).\tag{114}$$

We distinguish the two subcases:

(4.1) $b=0$. Then we obtain system $\dot{x}=x(x+c)$, $\dot{y}=y(Ax+1)$. If $cA\neq 0$, then doing the change of variables $x=Y$, $y=X$, and a rescaling of the independent variable this system becomes system (lv2). If $cA=0$, then we get system (lv10).

(4.2) $b\neq 0$. Taking $\beta=C/b$ system (114) goes over to system $\dot{x}=x(x+y+c)$, $\dot{y}=y(Ax+y+1)$. If $c\neq 1$ and $A\neq 1$, then doing the change of variables $x=(1-c)Y/((A-1)X)$, $y=(1-c)/X$, and the rescaling of the independent variable $t\to Xt/(c-1)$, we obtain system (lv9). If $c\neq 1$ and $A=1$, then doing the change of variables $x=1/X$, $y=Y/X$, and the rescaling of the independent variable $t\to Xt/(1-c)$, we obtain system (lv7). If $c=1$, we get system (lv11).

□

In the next proposition we characterize which systems (lvi), for $i=1,\dots,11$, are irreducible.

PROPOSITION 13.3. *The following statements hold.*

(a) *The next systems are irreducible*: (lv1) *with* $c(A^2+B^2)(B^2+C^2)\neq 0$, (lv2) *with* $c(B-1)(cB-1)\neq 0$, (lv5) *with* $(a^2+c^2)(a^2+(B-b)^2)\neq 0$, (lv6) *with* $B\neq 0$, (lv9) *with* $a\neq 1$ *or* $b\neq 1$, (lv10) *with* $c^2+(A-1)^2\neq 0$, *and* (lv11).
(b) *Systems* (lv1)–(lv11) *which do not appear in* (a) *are reducible.*

PROOF. The proposition follows easily from the definition of irreducible system. □

From Theorem 2.1 and Propositions 13.2 and 13.3 we obtain the following new Liouvillian integrable Lotka–Volterra systems (106). These new cases are easy to obtain, many of them come from [9]. The difficult problem will be to prove that the Lotka–Volterra systems of Proposition 13.2 which not appear in the following list are not Liouvillian integrable.

25. System (lv1) has the exponential factor $\exp(x)$ if $c\neq 0$; and it has the function x as a first integral if $c=0$.

26. System (lv2) has the invariant straight line $By+1=0$ if $B\neq 0$, and the exponential factor $\exp(y)$ if $B=0$.

27. System (lv3) has the integrating factor $1/(x^{B+1}y)$.

28. System (lv4) has the straight line $y=0$ formed by singular points. So, in $\mathbb{R}^2\setminus\{y=0\}$ the system has the first integral of the linear system $\dot{x}=x$, $\dot{y}=Ax+By$.

29. System (lv5) has the exponential factor $\exp(1/y)$ if $B \neq 0$, and the function y is a first integral if $B = 0$.

30. System (lv6) has the integrating factor $1/(xy)$ if $B = 0$.

31. System (lv7) has the straight line $x = 0$ formed by singular points. So, in $\mathbb{R}^2 \setminus \{x = 0\}$ the system has the first integral of the linear system $\dot{x} = ax + by + c$, $\dot{y} = y$.

32. System (lv8) is a quadratic homogeneous system and consequently it is integrable. In fact, if $a \neq 1$ or $b \neq 1$, then the function $1/(xy((a-1)x + (b-1)y))$ is an integrating factor of the system. If $a = b = 1$ then the straight line $x + y = 0$ is formed by singular points. So, in $\mathbb{R}^2 \setminus \{x + y = 0\}$ the system has the first integral of the linear system $\dot{x} = x$, $\dot{y} = y$.

33. System (lv9) for $b = 0$ has the invariant straight line $ax + 1 = 0$ if $a \neq 0$, and the exponential factor $\exp(x)$ if $a = 0$. If $a = 1$ it has the integrating factor $1/(xy^2)$. If $(a, b) = (1/2, -1)$ then the system has the invariant algebraic curve $xy + 1 + x + x^2/4 = 0$. If $(a, b) = (1/2, 1/2)$ then the system has the exponential factor $\exp((2 + x)^2/y)$.

34. System (lv10) has the invariant straight line $x + c = 0$ if $c \neq 0$, and the exponential factor $\exp(1/x)$ if $c = 0$.

35. System (lv11) for $A = 1$ has the straight line $x + y + 1 = 0$ formed by singular points. So, in $\mathbb{R}^2 \setminus \{x + y + 1 = 0\}$ the system has the first integral of the linear system $\dot{x} = x$, $\dot{y} = y$. If $A \neq 1$ it has the exponential factor $\exp(y/x)$.

We note that the exponential factors which appear in cases **33, 34** and **35** (respectively **29**) correspond to the fact that the geometric multiplicity of the invariant algebraic curve $x = 0$ (respectively $y = 0$) is higher than 1. The exponential factors of the cases **25, 26** and **33** correspond to the fact that the infinite straight line has geometric multiplicity higher than 1 as invariant algebraic curve of the projectivized vector field. For more details see [34].

In order to show that there are no additional Lotka–Volterra systems to those described in the list which are Liouvillian integrable (modulo the corresponding equivalencies or the reductions to the normal forms of Proposition 13.2), we must prove that the following two systems

$$\text{(lv6)} \quad \text{with } B \neq 0.$$

$$\text{(lv9)} \quad \text{with } b \neq 0,\ a \neq 1 \text{ and } (a, b) \notin \big\{(1/2, -1), (1/2, 1/2)\big\}$$

have no Liouvillian first integrals. We shall prove this for the system (lv6), similarly can be made for the system (lv9), for more details see [10]. Then the classification of the Liouvillian integrable Lotka–Volterra systems is completed by the following result.

THEOREM 13.4. *The following statements hold.*

(a) *After Theorem* 13.1 *to complete the study of the Liouvillian integrable Lotka–Volterra systems* (106) *it is sufficient to consider the Liouvillian integrability of the eleven subclasses* (lvi) *of systems* (106) *for* $i = 1, \ldots, 11$.

(b) *System* (lvi) *for* $i = 1, \ldots, 11$ *has a Liouvillian first integral if and only if it is one of the cases described from cases* **25** *to* **35**.

Of course, statement (a) is equivalent to Proposition 13.2.

Now in order to show how to prove Theorem 13.4 we shall study system (lv6) with $B \neq 0$. Systems (lv9) with $b \neq 0$, $a \neq 1$ and $(a, b) \notin \{(1/2, -1), (1/2, 1/2)\}$ can be proved using similar arguments, for more details see [10].

PROPOSITION 13.5. *System* (lv6) *with* $B \neq 0$ *has no Liouvillian first integrals.*

PROOF. From Singer results (see Section 12) we know that if system (lv6) with $B \neq 0$ (simply (lv6) in what follows) has a Liouvillian first integral, then it has an integrating factor of the form (14) where the $f_i = 0$ are invariant algebraic curves and the F_j are exponential factors. Now we shall prove that the unique irreducible invariant algebraic curves of (lv6) are $x = 0$ and $y = 0$. So from Proposition 1.10, it follows that the unique invariant algebraic curves of (lv6) are $x = 0$ and $y = 0$. Additionally, we shall also prove that (lv6) has no exponential factors. Since $B \neq 0$ a linear combination of the cofactors of $x = 0$ and $y = 0$ cannot be equal to minus the divergence of system (lv6). Then, by Theorem 2.1 it follows that system (lv6) has no integrating factors of the form (14) with $f_1 = x$ and $f_2 = y$.

We start by proving that system (lv6) has no irreducible invariant algebraic curves different from $x = 0$ and $y = 0$. Suppose that $f = 0$ is an irreducible invariant algebraic curve of (lv6) different from $x = 0$ and $y = 0$. Then f satisfies (9), i.e.,

$$x(y+1)\frac{\partial f}{\partial x} + y(x+By)\frac{\partial f}{\partial y} = (k_0 + k_1 x + k_2 y) f, \tag{115}$$

where $k_0 + k_1 x + k_2 y$ is the cofactor of $f = 0$.

Since f is a polynomial we can write

$$\begin{aligned} f &= f_0(x) + f_1(x)y + \cdots + f_m(x)y^m \\ &= \omega f_0(y) + \omega f_1(y)x + \cdots + \omega f_n(y)x^n, \end{aligned} \tag{116}$$

with $f_0(x)$, $f_m(x)$, $\omega f_0(y)$ and $\omega f_n(y)$ nonzero. We note that $f(x, y)$ cannot be independent on x or y, otherwise it factorizes and its factors would correspond to invariant straight lines (real or complex) which do not exist.

Taking $y = 0$ in (115) we get the following ordinary differential equation

$$x f_0'(x) = (k_0 + k_1 x) f_0(x).$$

Its general solution is $f_0(x) = D x^{k_0} \exp(k_1 x)$, where D is a constant which cannot be zero, otherwise y would be a factor of f in contradiction with the fact that f is irreducible and different from y. Since $f_0(x)$ is a polynomial, we get that $k_1 = 0$ and $k_0 \in \mathbb{Z}^+$, where $\mathbb{Z}^+$ denotes the nonnegative integers.

Now, taking $x = 0$ in (115) and since $B \neq 0$, we get the following ordinary differential equation

$$B y^2 \omega f_0'(y) = (k_0 + k_2 y) \omega f_0(y).$$

Its general solution is $\omega f_0(y) = Ey^{k_2/B}\exp(-k_0/(By))$, where E is a constant which cannot be zero, otherwise x would be a factor of f in contradiction with the fact that f is irreducible and different from x. Since $\omega f_0(y)$ must be a polynomial, we get that $k_0 = 0$, and that $k_2/B \in \mathbb{Z}^+$.

In short, we have $f_0(x) = D$ and $\omega f_0(y) = Ey^{k_2/B}$. Hence, if $k_2 \neq 0$ then $f_0(x) = 0$, because $f(0,0) = 0$. Therefore, y is a factor of f, a contradiction because f is irreducible and different from y. Consequently, we must have that $k_2 = 0$ and $f_0(x) = \omega f_0(y) = D = E$. Since the cofactor of $f = 0$ is zero, it follows that if there is another irreducible algebraic invariant curve $f = 0$ different from $x = 0$ and $y = 0$, then the polynomial f is a first integral of system (lv6).

Suppose that f is an irreducible polynomial first integral of system (lv6) of minimum degree. Then $H = f - f(0,0)$ is another polynomial first integral such that $H(0,0) = 0$ and consequently, $H = x^l y^k g$ with l and k positive integers, g coprime with x and y, and the polynomial g cannot be constant because it is easy to check that system (lv6) has no first integrals of the form $x^l y^k$. Since $H = 0$ is formed by invariant algebraic curves, it follows that $g = 0$ is formed by invariant algebraic curves different from $x = 0$ and $y = 0$. Hence, there exists an irreducible invariant algebraic curve $h = 0$ with h dividing to g (eventually h can be equal to g) different from $x = 0$ and $y = 0$, and having degree smaller than the degree of f. Since h also is an irreducible polynomial first integral (because its cofactor is zero) and its degree is smaller than the degree of f, we have a contradiction with the minimality of the degree of f as a polynomial first integral of system (lv6). In short, we have proved that there are no irreducible invariant algebraic curves of system (lv6) different from $x = 0$ and $y = 0$.

Now we shall prove that system (lv6) has no exponential factors. According to Proposition 1.12, if system (lv6) has exponential factors, then they must be of the form

$$\exp(h), \qquad \exp\bigl(h/x^l\bigr), \qquad \exp\bigl(h/y^k\bigr), \qquad \exp\bigl(h/\bigl(x^l y^k\bigr)\bigr), \tag{117}$$

where h is a polynomial, and l and k are positive integers.

The singular points of system (lv6) are $(0,0)$ and $(B,-1)$. Let $k_0 + k_1x + k_2y$ be a cofactor of an exponential factor of system (lv6). From the definition of exponential factor, the left-hand side of equality (13) is zero on the singular point $(B,-1)$, then from the right-hand side we obtain that $k_2 = k_0 + Bk_1$. So, a cofactor of any exponential factor of system (lv6) can be written as $k_0 + k_1x + (k_0 + Bk_1)y$.

First we shall see that there are no exponential factors of the form $\exp(h/x^l)$. Without loss of generality we can assume that h and x are coprime, and l is positive.

Suppose that $\exp(h/x^l)$ is an exponential factor of system (lv6). Then it satisfies (13), i.e.,

$$x(y+1)\frac{x\partial h/\partial x - lh}{x^{l+1}} + y(x+By)\frac{\partial h/\partial y}{x^l} = k_0 + k_1x + (k_0 + Bk_1)y,$$

where we have simplified the common factor $\exp(h/x^l)$. This last equality can be written as

$$(y+1)\left(x\frac{\partial h}{\partial x}-lh\right)+y(x+By)\frac{\partial h}{\partial y}=\big(k_0+k_1x+(k_0+Bk_1)y\big)x^l. \tag{118}$$

Since h is a polynomial we can write

$$h=\omega h_0(y)+\omega h_1(y)x+\omega h_2(y)x^2+\cdots. \tag{119}$$

Taking $x=0$ in (118) and since $B\neq 0$, we get the following ordinary differential equation

$$-(y+1)l\omega h_0(y)+By^2\omega h_0'(y)=0.$$

Its general solution is $\omega h_0(y)=Dy^{l/B}\exp(-l/(By))$ where D is a nonzero constant, otherwise x would be a factor of h. The fact that $D\neq 0$ gives a contradiction, because $\omega h_0(y)$ must be a polynomial and l is a positive integer. Hence, there are no exponential factors of the form $\exp(h/x^l)$.

We shall prove that there are no exponential factors of the form $\exp(h)$. Suppose that $\exp(h)$ is an exponential factor of system (lv6). Then it satisfies (13), i.e.,

$$x(y+1)\frac{\partial h}{\partial x}+y(x+By)\frac{\partial h}{\partial y}=k_0+k_1x+(k_0+Bk_1)y,$$

where we have simplified the common factor $\exp(h)$. Taking $(x,y)=(0,0)$ in this equality, it follows that $k_0=0$. So, we can write it as follows

$$x(y+1)\frac{\partial h}{\partial x}+y(x+By)\frac{\partial h}{\partial y}=k_1(x+By). \tag{120}$$

Now, taking $x=0$ in (120) we obtain the following ordinary differential equation

$$y\omega h_0'(y)=k_1,$$

here we have used once more that $B\neq 0$. Its general solution is $\omega h_0(y)=E+k_1\log y$, where E is a nonzero constant. Since $\omega h_0(y)$ is a polynomial, we get that $k_1=0$. However k_1 cannot be zero, otherwise h will be a polynomial first integral, in contradiction with the fact that the unique irreducible invariant algebraic curves are $x=0$ and $y=0$. Hence, there are no exponential factors of the form $\exp(h)$.

Suppose that $\exp(h/y^k)$ is an exponential factor of system (lv6). Without loss of generality we can assume that h and y are coprime, and k is positive. Then it satisfies (13), i.e.,

$$x(y+1)\frac{\partial h}{\partial x}+(x+By)\left(y\frac{\partial h}{\partial y}-kh\right)=y^k\big(k_0+k_1x+(k_0+Bk_1)y\big), \tag{121}$$

where we have simplified the common factor $\exp(h/y^k)$ and multiplied by y^k. Since h is a polynomial we can write

$$h = h_0(x) + h_1(x)y + h_2(x)y^2 + \cdots. \tag{122}$$

Taking $y = 0$ in Equation (121) we get the following ordinary differential equation

$$h_0'(x) - kh_0(x) = 0.$$

Its general solution is $h_0(x) = D\exp(kx)$ where D is a nonzero constant, otherwise y would be a factor of h. The fact that $D \neq 0$ gives a contradiction, because $h_0(x)$ must be a polynomial and k is a positive integer. Hence, there are no exponential factors of the form $\exp(h/y^k)$.

We shall prove now that there are no exponential factors of the form $\exp(h/(x^l y^k))$ with l and k positive integers. Without loss of generality we can assume that h, x and y are coprime. Then this exponential factor must satisfy (13), i.e.,

$$\begin{aligned}(y+1)\left(x\frac{\partial h}{\partial x} - lh\right) + (x+By)\left(y\frac{\partial h}{\partial y} - kh\right)\\ = x^l y^k\big(k_0 + k_1 x + (k_0 + Bk_1)y\big),\end{aligned}$$

where we have simplified the common factor $\exp(h/(x^l y^k))$ and multiplied by $x^l y^k$. Taking $y = 0$ in this equation we get the following ordinary differential equation

$$xh_0'(x) - (l + kx)h_0(x) = 0.$$

Its general solution is $h_0(x) = Dx^l \exp(kx)$ where D is a nonzero constant, otherwise y would be a factor of h. The fact that $D \neq 0$ gives a contradiction, because $h_0(x)$ must be a polynomial and k is a positive integer. Hence, there are no exponential factors of the form $\exp(h/(x^l y^k))$. This completes the proof of the proposition. □

14. On the integrability of two-dimensional flows

By definition a *planar differential system* is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x' = P(x,y), \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = y' = Q(x,y), \tag{123}$$

where P and Q are C^r maps with $r \geqslant 1$ from an open subset U of $\mathbb{R}^2$ to $\mathbb{R}$. We say that U is the *domain of definition* of the differential system (123), and that

$$X = P\frac{\partial}{\partial x} + Q\frac{\partial}{\partial y} \tag{124}$$

is the C^r *vector field* defined on U associated to differential system (123).

A C^k function $H : U \to \mathbb{R}$ with $k \geqslant 0$ is a *strong first integral* of the differential system (123) defined in U if H is constant on each solution of this system, and H is nonconstant on any open subset of U. Here $k \geqslant 0$ means that $k = 0, 1, 2, \ldots, \infty, \omega$. More precisely, $k = 0$ means that H is continuous, $k = 1, 2, \ldots, \infty$ means that H is C^k, $k = \omega$ means that H is analytic. If $k \geqslant 1$ then the previous definition of integrability implies that the derivative of H following the direction of the vector field X is zero; i.e., if $XH = 0$ on U.

This definition of strong first integral is the usual definition of first integral which appears in the major part of books on differential equations (see for instance [3] and [96]). With this definition the linear differential system

$$x' = x, \qquad y' = y, \tag{125}$$

defined on $\mathbb{R}^2$ has no strong first integrals. This is due to the fact that every strong first integral of system (125) must be a continuous function on $\mathbb{R}^2$ that must take a constant value on each straight line through the origin, because these straight lines are formed by orbits of the system. Hence it must be constant on the whole $\mathbb{R}^2$, consequently there are no strong first integrals for system (125). By using this argument it follows that if system (123) has a strong first integral, then it cannot have nodes, foci, center–foci, singular points having some parabolic or elliptic sectors, limit cycles and separatrix cycles that be the α- or ω-limit set of some orbit of the system (see [96] for definitions). Since we do not like that differential systems so easy as system (125) have no first integrals, we will introduce the notion of weak first integral.

Let Σ be a set of orbits of system (123) such that $U \setminus \Sigma$ is open. We say that a C^k function $H : U \setminus \Sigma \to \mathbb{R}$ with $k \geqslant 0$ is a *weak first integral* of the differential system (1.9) defined in U if H is constant on each solution of system (123) contained in $U \setminus \Sigma$, and H is nonconstant on any open subset of $U \setminus \Sigma$. If $k \geqslant 1$ this definition implies that the derivative of H following the direction of the vector field X is zero on $U \setminus \Sigma$.

We remark that the unique difference between the notions of strong and weak first integral is that a weak first integral does not need to be defined in the whole domain of definition U of the differential system (123). This difference has been noted by many authors. Thus, the first integrals computed by Darboux [38] in 1878 for polynomial differential systems possessing sufficient algebraic solutions are in general weak first integrals. If we study the integrability of the linear differential systems

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x' = ax + by, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = y' = cx + dy, \tag{126}$$

with $ad - bc \neq 0$ (nondegenerate), then with the notion of strong first integral only the centers and the saddles are integrable, but with the notion of weak first integral all linear systems (126) are integrable. In particular we show that system (125) has the weak first integral $H : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$ where $H(x, y) = xy/(x^2 + y^2)$.

We claim that the notion of weak first integral is the natural notion of integrability for two-dimensional differential systems instead of the more usual notion of strong first integral. Theorem 14.1 will confirm this claim. In order to present it we need some preliminary definitions and results. In this section we follow the papers [19,63].

For differential system (123) the following three properties are well known, see for more details [96].

(a) (*Existence and uniqueness of maximal solutions for a differential system*) For all $p \in U$ there exists an open interval I_p of the real line where it is defined the unique maximal solution $\varphi_p : I_p \to U$ of (123) such that $\varphi_p(0) = p$.

(b) (*Group property*) If $q = \varphi_p(t)$ and $t \in I_p$, then $I_q = I_p - t = \{r - t\colon\ r \in I_p\}$ and $\varphi_q(s) = \varphi_p(t + s)$ for all $s \in I_q$.

(c) (*Differentiability with respect to initial conditions*) The set $D = \{(t, p)\colon\ p \in U, t \in I_p\}$ is open in $\mathbb{R}^3$ and the map $\varphi : D \to U$ defined by $\varphi(t, p) = \varphi_p(t)$ is C^r.

The map $\varphi : D \to U$ is a *local flow* of class C^r with $r \geqslant 1$ on U associated to system (123), which verifies

(i) $\varphi(0, p) = p$ for all $p \in U$;

(ii) $\varphi(t, \varphi(s, p)) = \varphi(t + s, p)$ for all $p \in U$, and all s and t such that $s, t + s \in I_p$;

(iii) $\varphi_p(-t) = \varphi_p^{-1}(t)$ for all $p \in U$ such that $t, -t \in I_p$.

Let φ be a local flow on the two-dimensional manifold M, and let Σ be a subset of M formed by orbits of φ such that $M \setminus \Sigma$ is open. We say that a C^k function $H : M \setminus \Sigma \to \mathbb{R}$ with $k \geqslant 0$ is a *weak first integral* of φ if $H \circ \varphi_p$ is constant for each $p \in M \setminus \Sigma$ and H is not constant on any open subset of $M \setminus \Sigma$. Of course, when the local flow is the local flow associated to a C^r differential system (123) with $r \geqslant 1$, the above definition of weak first integral for system (123) and the definition of weak first integral for its associated local flow coincide.

In this paper we consider C^r local flows with $r \geqslant 0$ on an arbitrary two-dimensional manifold M (separable metric, but not necessarily compact nor orientable and possibly with boundary). Of course, when $r = 0$ the flow is only continuous. Two such flows, (M, φ) and (M', φ'), are *C^k-equivalent* with $k \geqslant 0$ if there is a C^k diffeomorphism of M onto M' which takes orbits of φ onto orbits of φ' preserving sense (but not necessarily the parametrization). Of course, a C^0 diffeomorphism is a homeomorphism.

Let φ be a C^r local flow with $r \geqslant 0$ on the two-dimensional manifold M. We call (M, φ) *C^k-parallel* if it is C^k-equivalent to one of the following flows:

(1) $\mathbb{R}^2$ with the flow defined by $x' = 1$, $y' = 0$;

(2) $\mathbb{R}^2 \setminus \{0\}$ with the flow defined (in polar coordinates) by $r' = 0$, $\theta' = 1$;

(3) $\mathbb{R}^2 \setminus \{0\}$ with the flow defined by $r' = r$, $\theta' = 0$;

(4) $\mathbb{S}^1 \times \mathbb{S}^1$ with rational flow (e.g., the flow induced by (123) above under the usual covering map; note in particular that all rational flows on the torus are equivalent). We call these flows as *strip*, *annular*, *spiral*, and *toral*, respectively.

Let $p \in M$. We denote by $\gamma(p)$ the *orbit* of the flow φ on M through p, more precisely $\gamma(p) = \{\varphi_p(t)\colon\ t \in I_p\}$. The *positive half-orbit* of $p \in M$ is $\gamma^+(p) = \{\varphi_p(t)\colon\ t \in I_p, t \geqslant 0\}$. In a similar way it is defined the *negative half-orbit* $\gamma^-(p)$ of $p \in M$.

We define the *α-limit* and the *ω-limit* of p as $(\gamma^{\pm}(p))$ and let

$$\alpha(p) = \mathrm{cl}\big(\gamma^-(p)\big) - \gamma^-(p), \qquad \omega(p) = \mathrm{cl}\big(\gamma^+(p)\big) - \gamma^+(p),$$

respectively. Here, as usual, cl denotes the closure.

Let $\gamma(p)$ be an orbit of the flow φ defined on M. A *parallel neighborhood* of the orbit $\gamma(p)$ is an open neighborhood N of $\gamma(p)$ such that (N, φ) is C^k-equivalent to a parallel flow for some $k \geqslant 0$.

We say that $\gamma(p)$ is a *separatrix* of φ if $\gamma(p)$ is not contained in a parallel neighborhood N satisfying the following two assumptions:

(1) for any $q \in N$, $\alpha(q) = \alpha(p)$ and $\omega(q) = \omega(p)$, and

(2) $\mathrm{cl}(N) \setminus N$ consists of $\alpha(p)$, $\omega(p)$ and exactly two orbits $\gamma(a)$, $\gamma(b)$ of φ, with $\alpha(a) = \alpha(p) = \alpha(b)$ and $\omega(a) = \omega(p) = \omega(b)$.

We denote by Σ the union of all separatrices of φ. Then Σ is a closed invariant subset of M. A component of the complement of Σ in M, with the restricted flow, is called a *canonical region* of φ.

The main result of this section is the following one.

THEOREM 14.1. *Let φ be a C^r flow on a two-dimensional manifold M for some $r \in \{0, 1, \dots, \infty, \omega\}$, and let Σ be the union of all separatrices of ϕ. Then*

(1) *Every canonical region of (M, φ) is C^r parallel.*

(2) *The flow φ restricted to every canonical region has a C^r (respectively C^∞, C^ω) first integral for $r \in \mathbb{N}$ (respectively $r = \infty, \omega$).*

Statement (1) in the case C^0 parallel was proved by Markus [74] and Neumann [81], the rest of the theorem is proved in [63]. See [19] for a version C^0 of statement (2). Later on, here we will provide a proof of Theorem 14.1 for the C^0 version.

In the next two theorems we use the fact that any analytic vector field on $\mathbb{S}^2$ has finitely many limit cycles as it was proved by Il'Yashenko [57] and Écalle [39]. The following two results improve Theorem 14.1 for planar polynomial differential systems. For a proof of them see [63].

THEOREM 14.2. *For every planar polynomial system there exist finitely many invariant curves in $\mathbb{R}^2$ and singular points γ_i, $i = 1, 2, \dots, l$, such that $\mathbb{R}^2 \backslash (\bigcup_{i=1}^{l} \gamma_i)$ has finitely many connected open sets, and on each of these connected sets the system has an analytic first integral.*

COROLLARY 14.3. *Let $\mathbf{X} = (P, Q)$ be a polynomial vector field in $\mathbb{R}^2$ such that P and Q are coprime. Then, using the notations of Theorem* 14.2, *the set $\bigcup_{i=1}^{l} \gamma_i$ is formed by all the separatrices of $\mathbf{X}$ in $\mathbb{R}^2$, and the open components of $\mathbb{R}^2 \backslash (\bigcup_{i=1}^{l} \gamma_i)$ are the canonical regions of $\mathcal{X}$.*

Statement (1) of Theorem 14.1 follows directly from the following result.

LEMMA 14.4 (Neumann lemma [81]). *Let φ be a local flow on the two-dimensional manifold M. Then every canonical region of (M, φ) is C^0-parallel.*

PROOF. Let $(R, \varphi' = \varphi|R)$ be a canonical region. There are no separatrices in R, so the set consisting of orbits homeomorphic with $\mathbb{S}^1$ is open, and similarly the set consisting of orbits homeomorphic with $\mathbb{R}$ is open. Hence, R consists entirely of closed orbits or entirely of line orbits.

Also, two orbits of φ' can be separated with disjoint parallel neighborhoods. To prove this we suppose $\gamma(p)$ and $\gamma(q)$ are distinct orbits (closed or not) which cannot be separated. Then, for any parallel neighborhood N_p of p, we have $q \in \mathrm{cl}(N_p)$; i.e.,

$$q \in \bigcap_{N_p} \mathrm{cl}(N_p) = \alpha(p) \cup \gamma(p) \cup \omega(p).$$

But then $q \in \alpha(p)$ (or $q \in \omega(p)$) and this is impossible because $q \in N_q \subset R$ and $\alpha(p) \cup \omega(p) \subset \mathrm{cl}(N_q) \setminus N_q$.

It follows that the quotient space R/φ', obtained by collapsing orbits of (R, φ') to points, is a (Hausdorff) one-dimensional manifold. Hence the natural projection $\pi : R \to R/\varphi'$ is a locally trivial fibering of R over $\mathbb{R}$ or $\mathbb{S}^1$. Since the flow provides a natural orientation on the fibers, there are only four possibilities, the four classes of parallel flows described above. □

PROOF OF STATEMENT (2) OF THEOREM 14.1 FOR C^0 FLOWS. From Lemma 14.4 it follows that every canonical region R of (M, φ) is C^0-parallel; i.e., there is a homeomorphism h of R onto M' which takes orbits of φ onto orbits of φ' preserving the sense (but not necessarily the parametrization), and (M', φ') is one of the following flows:

(1) $M' = \mathbb{R}^2$ with the flow defined by $x' = 1$, $y' = 0$;
(2) $M' = \mathbb{R}^2 \setminus \{0\}$ with the flow defined by $r' = 0$, $\theta' = 1$;
(3) $M' = \mathbb{R}^2 \setminus \{0\}$ with the flow defined by $r' = r$, $\theta' = 0$;
(4) $M' = \mathbb{S}^1 \times \mathbb{S}^1$ with the rational flow.

Clearly $H(x, y) = y = C$ is a first integral for the flows (1) and (4), $H(r, \theta) = r = C$ a first integral for the flow (2), and $H(r, \theta) = \theta = C$ a first integral for the flow (3). Hence, $H \circ h$ is a continuous valued first integral for the flow φ on the canonical region R. □

15. Darbouxian theory of integrability for polynomial vector fields on surfaces

In this section we follow the paper [68]. A *polynomial vector field* $\mathcal{X}$ in $\mathbb{R}^3$ is a vector field of the form

$$\mathcal{X} = P(x, y, z)\frac{\partial}{\partial x} + Q(x, y, z)\frac{\partial}{\partial y} + R(x, y, z)\frac{\partial}{\partial z},$$

where P, Q and R are polynomials in the variables x, y and z with real coefficients. In all this section $m = \max\{\deg P, \deg Q\ \deg R\}$ will denote the *degree* of the polynomial vector field $\mathcal{X}$.

Let $G : \mathbb{R}^3 \to \mathbb{R}$ be a C^r map with $r \geqslant 1$. The *gradient* of G is defined by

$$\nabla G(x, y, z) = \left(\frac{\partial G}{\partial x}, \frac{\partial G}{\partial y}, \frac{\partial G}{\partial z}\right)(x, y, z).$$

If $\nabla G(x, y, z) \neq 0$ for all points $(x, y, z) \in \mathbb{R}^3$, then

$$\boldsymbol{\Sigma} = \left\{(x, y, z) \in \mathbb{R}^3 \colon G(x, y, z) = 0\right\} \tag{127}$$

is called a 2-dimensional *regular surface*. A polynomial vector field $\mathcal{X}$ in $\mathbb{R}^3$ defines a *polynomial vector field on* $\boldsymbol{\Sigma}$ if

$$\mathcal{X}(x, y, z) \cdot \nabla G(x, y, z) = 0, \tag{128}$$

for all the points (x, y, z) of the surface $\boldsymbol{\Sigma}$; i.e., for all $(x, y, z) \in \boldsymbol{\Sigma}$ the vector $\mathcal{X}(x, y, z)$ belongs to the tangent plane to $\boldsymbol{\Sigma}$ at the point (x, y, z).

In what follows sometimes we will consider real regular surfaces extended to complex ones; that is, being G a real map we consider its natural extension to $\mathbb{C}^3$. Thus, we continue denoting by $\boldsymbol{\Sigma} = \{(x, y, z) \in \mathbb{C}^3\colon\ G(x, y, z) = 0\}$ the complexification of (127). Always we will work with real regular surfaces (127) such that its complexification also satisfies (128).

Let U be an open subset of $\mathbb{R}^3$. A polynomial vector field $\mathcal{X}$ on the regular surface $\boldsymbol{\Sigma}$ is *integrable* on the open subset $U \cap \boldsymbol{\Sigma}$ if there exists a nonconstant analytic function $H : U \to \mathbb{R}$, called a *first integral* of $\mathcal{X}$ on $U \cap \boldsymbol{\Sigma}$, which is constant on all solution curves $(x(t), y(t), z(t))$ of the vector field $\mathcal{X}$ on $U \cap \boldsymbol{\Sigma}$; i.e., $H(x(t), y(t), z(t)) =$ constant for all values of t for which the solution $(x(t), y(t), z(t))$ is defined on $U \cap \boldsymbol{\Sigma}$. Clearly H is a first integral of the polynomial vector field $\mathcal{X}$ on $U \cap \boldsymbol{\Sigma}$ if and only if $XH = 0$ on all points (x, y, z) of $U \cap \boldsymbol{\Sigma}$. We note that the curves $\{H(x, y, z) = \text{constant}\} \cap (U \cap \boldsymbol{\Sigma})$ are formed by trajectories of $\mathcal{X}$.

15.1. *Invariant algebraic curves and exponential factors*

In a similar way as we did for studying the invariant algebraic curves of real planar polynomial vector fields, also for the real polynomial vector fields on real surfaces we will look for complex invariant algebraic curves and we will think both the real polynomial vector field on a surface and the real surface as complex ones.

Let $f \in \mathbb{C}[x, y, z]$, where as it is usual $\mathbb{C}[x, y, z]$ denotes the ring of the polynomials in the variables x, y and z with complex coefficients. The algebraic surface $f = 0$ defines an *invariant algebraic curve* $\{f = 0\} \cap \boldsymbol{\Sigma}$ of the polynomial vector field $\mathcal{X}$ on the regular surface $\boldsymbol{\Sigma} = \{(x, y, z) \in \mathbb{C}^3\colon\ G(x, y, z) = 0\}$ if

(i) for some polynomial $K \in \mathbb{C}[x, y, z]$ of degree at most $m - 1$ we have

$$\mathcal{X}f = P\frac{\partial f}{\partial x} + Q\frac{\partial f}{\partial y} + R\frac{\partial f}{\partial z} = Kf, \tag{129}$$

on all the points (x, y, z) of the surface $\boldsymbol{\Sigma}$;

(ii) the intersection of the two surfaces $f = 0$ and $\boldsymbol{\Sigma}$ is *transversal*; i.e., for all the points $(x, y, z) \in \{f = 0\} \cap \boldsymbol{\Sigma}$ we have that

$$\nabla G(x, y, z) \wedge \nabla f(x, y, z) \neq 0,$$

where $\wedge$ denotes the vectorial product of two vectors of $\mathbb{C}^3$.

The polynomial K is called the *cofactor* of the invariant algebraic curve $\{f = 0\} \cap \boldsymbol{\Sigma}$.

Since on the points of the invariant algebraic curve $\{f = 0\} \cap \boldsymbol{\Sigma}$ the gradient ∇f of the surface $f = 0$ is orthogonal to the polynomial vector field $\mathcal{X} = (P, Q, R)$ (see (129)), and the vector field $\mathcal{X}$ is tangent to the surface $\boldsymbol{\Sigma}$, it follows that the vector field $\mathcal{X}$ is tangent to the curve $\{f = 0\} \cap \boldsymbol{\Sigma}$. Hence, the curve $\{f = 0\} \cap \boldsymbol{\Sigma}$ is formed by trajectories of the vector field $\mathcal{X}$. This justifies the name of invariant algebraic curve given to the curve $\{f = 0\} \cap \boldsymbol{\Sigma}$ satisfying (129) for some polynomial K, because it is *invariant* under the flow defined by $\mathcal{X}$. We note that our definition of invariant algebraic curve $\{f = 0\} \cap \boldsymbol{\Sigma}$ does not need that the regular surface to be algebraic, only needs that the surface $\{f = 0\}$ be algebraic.

Let h and g polynomials of $\mathbb{C}[x, y, z]$. Then the function $F = \exp(g/h)$ is called an *exponential factor* of the polynomial vector field $\mathcal{X}$ on $\boldsymbol{\Sigma}$ if for some polynomial $K \in \mathbb{C}[x, y, z]$ of degree at most $m - 1$ it satisfies the following equality

$$\mathcal{X}F = KF, \tag{130}$$

on all the points (x, y, z) of $\boldsymbol{\Sigma}$. As before we say that K is the *cofactor* of the exponential factor F.

As we will see from the point of view of the integrability of polynomial vector fields on regular surfaces $\boldsymbol{\Sigma}$ the importance of the exponential factors is double. On one hand, they verify equation (130), and on the other hand, their cofactors are polynomials of degree at most $m - 1$. These two facts will allow that they play the same role that the invariant algebraic curves in the integrability of a polynomial vector field $\mathcal{X}$ on $\boldsymbol{\Sigma}$.

The following proposition has essentially the same proof as Proposition 1.12.

PROPOSITION 15.1. *Let $F = \exp(g/h)$ be an exponential factor for the polynomial vector field $\mathcal{X}$ on the regular surface $\boldsymbol{\Sigma}$, satisfying that the polynomials g and h are relatively prime, and that the intersection of the surfaces $\{h = 0\}$ with $\boldsymbol{\Sigma}$ is transversal. If h is non-constant, then $\{h = 0\} \cap \boldsymbol{\Sigma}$ is an invariant algebraic curve of $\mathcal{X}$ on $\boldsymbol{\Sigma}$, and g satisfies the equation*

$$\mathcal{X}g = gK_h + hK_F,$$

on the points (x, y, z) of $\boldsymbol{\Sigma}$, where K_h and K_F are the cofactors of h and F respectively.

15.2. *The surfaces*

Now we describe the 2-dimensional regular surfaces on which we will study the Darbouxian theory of integrability for the polynomial vector fields defined on them. These surfaces will be the quadrics and the 2-torus.

It is well known that the quadrics can be classified into seventeen types. There are five types of imaginary quadrics (namely the imaginary ellipsoid, cone, cylinder, two planes that intersect, and two parallel planes). Since in this section we only consider real polynomial vector fields on real surfaces, we omit these previous five imaginary quadrics. From the twelve real quadrics there are three which are formed by planes (namely two planes that intersect, two parallel planes, and a double plane). Since the study of the Darbouxian theory of integrability of the polynomial vector fields on these surfaces is reduced to the classical study of the Darbouxian theory of integrability of planar polynomial vector fields, we omit these three types of quadrics. In short we only deal with the remainder nine quadrics.

We remark that the reduction of every quadric to each canonical form is made through a linear isomorphism, and that a such isomorphism and its inverse are polynomial diffeomorphisms. Therefore, the study of the Darbouxian theory of integrability of polynomial vector fields on these nine quadrics can be restricted to take for each of these quadrics its canonical form. In what follows we describe the canonical forms for these nine quadrics and their local charts, that we shall need later on. In fact we only present one local chart for each surface, the full atlas would be described in a similar way, but it is not necessary for studying the existence of first integrals.

Parabolic cylinder. Its canonical form is given by $z^2 - x = 0$.

Elliptic or circular paraboloid. Its canonical form is given by $y^2 + z^2 - x = 0$.

Hyperbolic paraboloid. Its canonical form is given by $y^2 - z^2 - x = 0$.

Let $h(x, y)$ be equal to x, $x - y^2$ and $y^2 - x$ for the parabolic cylinder, the circular paraboloid and the hyperbolic paraboloid, respectively. Then for each one of these cases we have that there is an open subset Ω of $\mathbb{R}^2$ and a diffeomorphism $\lambda : \Omega \to \boldsymbol{\Sigma}$ defined by

$$\lambda(r, s) = (x, y, z) = \left(r, s, \sqrt{h(r, s)}\right),$$

whose inverse is $\lambda^{-1}(x, y, z) = (r, s) = (x, y)$. The entries of the Jacobian matrix $D\lambda(r, s)$ are polynomial functions in the variables r, s and $\sqrt{h(r, s)}$.

Ellipsoid or sphere. Its canonical form is given by

$$x^2 + y^2 + z^2 - 1 = 0.$$

We identify $\mathbb{R}^2$ as the tangent plane to the 2-dimensional sphere $\boldsymbol{\Sigma}$ at the point $(0, 0, -1)$. Then we have $\lambda : \mathbb{R}^2 \to \boldsymbol{\Sigma} \setminus \{(0, 0, 1)\}$ the diffeomorphism given by

$$\begin{aligned}\lambda(r, s) &= (x, y, z)\\ &= \frac{1}{1 + r^2 + s^2}(2r, 2s, r^2 + s^2 - 1).\end{aligned}$$

That is, λ is the inverse map of the stereographic projection $\lambda^{-1} : \boldsymbol{\Sigma} \setminus \{(0, 0, 1)\} \to \mathbb{R}^2$ defined by

$$\lambda^{-1}(x, y, z) = (r, s) = \left(\frac{x}{1 - z}, \frac{y}{1 - z}\right).$$

The entries of the Jacobian matrix $D\lambda(r, s)$ are rational functions in the variables r and s.

Hyperboloid of two sheets. Its canonical form is given by

$$x^2 + y^2 - z^2 + 1 = 0.$$

Without loss of generality we restrict our attention to the sheet $\boldsymbol{\Sigma}$ having $z > 0$. Then we have the diffeomorphism $\lambda : \mathbb{R}^2 \to \boldsymbol{\Sigma}$ defined by

$$\lambda(r, s) = (x, y, z) = \left(r, s, \sqrt{r^2 + s^2 + 1}\right),$$

whose inverse is $\lambda^{-1}(x, y, z) = (r, s) = (x, y)$. The entries of the Jacobian matrix $D\lambda(r, s)$ are rational functions in the variables r, s and $\sqrt{r^2 + s^2 + 1}$.

Cone. Its canonical form is given by

$$x^2 + y^2 - z^2 = 0.$$

Since the cone is not a regular surface at the origin, without loss of generality we restrict our attention to the sheet $\boldsymbol{\Sigma}$ having $z > 0$. Then we have the diffeomorphism $\lambda : \mathbb{R}^2 \to \boldsymbol{\Sigma}$ defined by

$$\lambda(r, s) = (x, y, z) = \left(r, s, \sqrt{r^2 + s^2}\right),$$

whose inverse is $\lambda^{-1}(x, y, z) = (r, s) = (x, y)$. The entries of the Jacobian matrix $D\lambda(r, s)$ are rational functions in the variables r, s and $\sqrt{r^2 + s^2}$.

Hyperbolic cylinder. Its canonical form is given by

$$x^2 - z^2 - 1 = 0.$$

Without loss of generality we restrict our attention to the sheet $\boldsymbol{\Sigma}$ having $x > 0$. Then we have the diffeomorphism $\lambda : \mathbb{R}^2 \to \boldsymbol{\Sigma}$ defined by

$$\lambda(r, s) = (x, y, z) = \left(\sqrt{s^2 + 1}, r, s\right),$$

whose inverse is $\lambda^{-1}(x, y, z) = (r, s) = (y, z)$. The entries of the Jacobian matrix $D\lambda(r, s)$ are rational functions in the variables r, s and $\sqrt{s^2 + 1}$.

Hyperboloid of one sheet. Its canonical form is given by

$$x^2 + y^2 - z^2 - 1 = 0.$$

Without loss of generality we restrict our attention to the surface $\boldsymbol{\Sigma}$ having $z > 0$. Then we have the diffeomorphism $\lambda : \{(r, s) \in \mathbb{R}^2 \colon r^2 + s^2 > 1\} \to \boldsymbol{\Sigma} \cap \{z > 0\}$ defined by

$$\lambda(r, s) = (x, y, z) = \left(r, s, \sqrt{r^2 + s^2 - 1}\right),$$

whose inverse is $\lambda^{-1}(x, y, z) = (r, s) = (x, y)$. The entries of the Jacobian matrix $D\lambda(r, s)$ are rational functions in the variables r, s and $\sqrt{r^2 + s^2 - 1}$.

Elliptic or circular cylinder. Its canonical form is given by

$$x^2 + z^2 - 1 = 0.$$

We identify $\mathbb{R}^2$ with the tangent plane to the cylinder $\boldsymbol{\Sigma}$ at the generatrix through the point $(0, 0, -1)$. Let l be the generatrix to the cylinder through the point $(0, 0, 1)$. Then we consider the diffeomorphism $\pi : \mathbb{R}^2 \to \boldsymbol{\Sigma} \setminus l$ defined by

$$\pi(r, s) = (x, y, z) = \left(\frac{4r}{r^2+4}, s, \frac{r^2-4}{r^2+4}\right),$$

with inverse

$$\pi^{-1}(x, y, z) = (r, s) = \left(\frac{2x}{1-z}, y\right).$$

The entries of the Jacobian matrix $D\pi(r, s)$ are rational functions in the variables r and s.

After the presentation of the nine quadrics for which we shall study the Darbouxian theory of integrability of their polynomial vector fields, we present the 2-torus.

The 2-*torus*. We choose for the 2-torus $\mathbf{T}^2 = \mathbb{R}^2/\mathbf{Z}^2$ the following canonical form. For $a > 1$ let $\boldsymbol{\Sigma} \approx \mathbf{T}^2$ be the regular surface of $\mathbf{R}^3$ defined by

$$z^2 + \left(\sqrt{x^2+y^2} - a\right)^2 = 1.$$

The intersection of this torus with the (x, y)-plane is formed by the two circles γ_1 and γ_2 with center at the origin and radius $a - 1$ and $a + 1$, respectively. Let C be the circular cylinder parallel to the z-axis and tangent to the torus $\boldsymbol{\Sigma}$ at the circle γ_2. Let $\mu : C \to \boldsymbol{\Sigma} \setminus \gamma_1$ be the diffeomorphism

$$\mu(u, v, w) = (x, y, z) = \left(\frac{uv}{(a+1)^2}\left[a - 1 + \frac{8}{w^2+4}\right]^2, \frac{4w}{w^2+4}\right).$$

On each plane P containing the z-axis, the diffeomorphism $\mu|_{C\cap P}$ is the inverse of the stereographic projection at the two points $P \cap \gamma_1$ from $(\boldsymbol{\Sigma} \setminus \gamma_1) \cap P$ onto the two generatrices $C \cap P$. Its inverse μ^{-1} is given by

$$\mu^{-1}(x, y, z) = (u, v, w) = \left(\frac{(a+1)x}{\sqrt{x^2+y^2}} \frac{(a+1)y}{\sqrt{x^2+y^2}}, \frac{2z}{1-a+\sqrt{x^2+y^2}}\right).$$

Then we have the diffeomorphism $\lambda = \mu \circ \pi : \mathbb{R}^2 \to \boldsymbol{\Sigma} \setminus (\gamma_1 \cup \mu(l))$. Here π corresponds to the local chart of a circular cylinder parallel to the z-axis, while the π presented previously corresponds to a cylinder parallel to the y-axis. Here, we have that

$$\lambda^{-1}(x, y, z) = (r, s) = \left(\frac{2(a+1)x}{\sqrt{x^2+y^2} - (a+1)y}, \frac{2z}{1-a+\sqrt{x^2+y^2}}\right).$$

The entries of the Jacobian matrix $D\lambda(r, s)$ are rational functions in the variables r and s.

We remark that the nine quadrics that we consider have canonical form of the type $z^2 = h(x, y)$ where h is a polynomial and its square root is not a polynomial. Similarly, the 2-torus that we consider has canonical form $z^2 = h(x, y) = 1 - (\sqrt{x^2+y^2} - a)^2$.

15.3. *Darbouxian theory*

Now we will extend the results of the Darbouxian theory of integrability for planar polynomial vector fields (see for instance Theorem 2.1) to polynomial vector fields on a regular surface $\boldsymbol{\Sigma}$. Here the regular surface will be one of the nine quadrics described in the previous subsection, or the 2-torus. In order to develop this theory we need some preliminary algebraic results.

Let $\mathbb{R}_m[x, y, z]$ denote the real linear vector space of all polynomials of $\mathbb{R}[x, y, z]$ having degree $\leqslant m$. Let $\mathbb{R}_m[x, y, z = \alpha(x, y)]$ be the real linear vector space of all the functions obtained from the polynomials of $\mathbb{R}_m[x, y, z]$ substituting z by the function $\alpha(x, y)$. Here $\alpha(x, y)$ is equal to $\sqrt{h(x, y)}$ for the quadrics and for the 2-torus, where $h(x, y)$ is a polynomial such that its square root is not a polynomial if $\boldsymbol{\Sigma}$ is a quadric, and $h(x, y) = 1 - (\sqrt{x^2 + y^2} - a)^2$ if $\boldsymbol{\Sigma}$ is the 2-torus.

We define $d(m)$ as $(m + 1)^2$ or $2(m^2 + 1)$ for the quadrics or the 2-torus, respectively.

LEMMA 15.2. *The dimension of the linear vector space* $\mathbb{R}_m[x, y, z = \alpha(x, y)]$ *is* $d(m)$.

PROOF. We denote by $\mathcal{B}_k$ a basis of the real vector space formed by all polynomials in the variables x and y having degree $\leqslant k$. It is well known that the cardinal of $\mathcal{B}_k$ is $(k + 1)(k + 2)/2$.

Assume that the function $\alpha(x, y)$ is associated to one of our nine quadrics. Then it is easy to check that $\mathcal{B}_m \cup \alpha\mathcal{B}_{m-1}$ is a basis for the real vector field $\mathbb{R}_m[x, y, z = \alpha(x, y)]$. Since the cardinal of this basis is

$$\frac{(m + 1)(m + 2)}{2} + \frac{m(m + 1)}{2} = (m + 1)^2,$$

it follows the lemma when the surface is a quadric.

Assume that

$$\alpha(x, y) = \sqrt{1 - \left(\sqrt{x^2 + y^2} - a\right)^2} = \sqrt{1 - (\beta - a)^2},$$

is the function associated to the 2-torus. Then it is easy to verify that

$$\mathcal{B}_m \cup \alpha\mathcal{B}_{m-1} \cup \beta\mathcal{B}_{m-2} \cup \alpha\beta\mathcal{B}_{m-3},$$

is a basis for the real vector field $\mathbb{R}_m[x, y, z = \alpha(x, y)]$. Since the cardinal of this basis is

$$\frac{(m + 1)(m + 2)}{2} + \frac{m(m + 1)}{2} + \frac{(m - 1)m}{2} + \frac{(m - 2)(m - 1)}{2} = 2(m^2 + 1),$$

it follows the lemma when the surface is the 2-torus. ☐

The functions of the linear vector space $\mathbb{R}_m[x, y, z = \alpha(x, y)]$ are polynomials in the variables x, y and α when $\boldsymbol{\Sigma}$ is a quadric, and polynomials in the variables x, y, α, β and

$\alpha\beta$ when $\boldsymbol{\Sigma}$ is the 2-torus. We note that such polynomials always have degree one when they are thought as polynomials in the variable either α, β or $\alpha\beta$.

Since the dimension of the linear vector space $\mathbb{R}_m[x, y, z = \alpha(x, y)]$ is $d(m)$, this space is linearly isomorphic to $\mathbb{R}^{d(m)}$ by identifying each polynomial with the $d(m)$-tuple of its coefficients.

Let $\boldsymbol{\Sigma}$ be a quadric. For $k = 1, \ldots, r$ we say that r points $(x_k, y_k, z_k) \in \boldsymbol{\Sigma}$ are *independent* with respect to the linear vector space $\mathbb{R}_{m-1}[x, y, z = \alpha(x, y)] \equiv \mathbb{R}^{d(m)}$ if the intersection of the r hyperplanes

$$\sum_{i+j=0}^{m(m+1)/2} x_k^i y_k^j a_{ij} + \sum_{i+j=0}^{(m-1)m/2} x_k^i y_k^j \alpha(x_k, y_k) b_{ij} = 0$$

(in the variables a_{ij} and b_{ij} of $\mathbb{R}^{d(m)}$), for $l = 1, \ldots, r$, defines a linear subspace of dimension $m^2 - r$.

Let $\boldsymbol{\Sigma}$ be the 2-torus. We say that r points $(x_k, y_k, z_k) \in \boldsymbol{\Sigma}$, for $k = 1, \ldots, r$, are *independent* with respect to the linear vector space $\mathbb{R}_{m-1}[x, y, z = \alpha(x, y)] \equiv \mathbb{R}^{d(m)}$ if the intersection of the r hyperplanes

$$\sum_{i+j=0}^{m(m+1)/2} x_k^i y_k^j a_{ij} + \sum_{i+j=0}^{(m-1)m/2} x_k^i y_k^j \alpha(x_k, y_k) b_{ij} + \sum_{i+j=0}^{(m-2)(m-1)/2} x_k^i y_k^j \beta(x_k, y_k) c_{ij}$$
$$+ \sum_{i+j=0}^{(m-3)(m-2)/2} x_k^i y_k^j \alpha(x_k, y_k) \beta(x_k, y_k) d_{ij} = 0,$$

for $l = 1, \ldots, r$ (in the variables a_{ij}, b_{ij}, c_{ij} and d_{ij} of $\mathbb{R}^{d(m)}$) defines a linear subspace of dimension $2m^2 - r$.

THEOREM 15.3. *Let $\boldsymbol{\Sigma}$ be one of the nine quadrics of the previous subsection or the 2-torus. Suppose that the real polynomial vector field $\mathcal{X}$ on the real regular surface $\boldsymbol{\Sigma}$ of degree m admits p invariant algebraic curves $\{f_i = 0\} \cap \boldsymbol{\Sigma}$ with cofactors K_i for $i = 1, \ldots, p$, q exponential factors $F_j = \exp(g_j/h_j)$ with cofactors L_j for $j = 1, \ldots, q$, and r independent singular points (x_k, y_k, z_k) of $\mathcal{X}$ in $\boldsymbol{\Sigma}$ such that $f_i(x_k, y_k, z_k) \neq 0$ for $i = 1, \ldots, p$ and for $k = 1, \ldots, r$. Of course, every h_j is equal to some f_i except if h_j is constant. The following statements hold.*

(a) *There exist $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = 0$ on all the points (x, y, z) of $\boldsymbol{\Sigma}$, if and only if the function $f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_1}$ substituting $f_i^{\lambda_i}$ by $|f_i|^{\lambda_i}$ if $\lambda_i \in \mathbb{R}$, is a first integral of the vector field $\mathcal{X}$ on $\boldsymbol{\Sigma}$.*

(b) *If $p + q + r = d(m-1) + 1$, then there exist $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = 0$ on all the points (x, y, z) of $\boldsymbol{\Sigma}$.*

(c) *The vector field $\mathcal{X}$ on $\boldsymbol{\Sigma}$ has a rational first integral if and only if $p + q + r \geqslant d(m-1) + 2$ and $p + q \geqslant 3$.*

PROOF. By hypothesis we have p invariant algebraic curves $\{f_i = 0\} \cap \boldsymbol{\Sigma}$ with cofactors K_i, and q exponential factors F_j with cofactors L_j. That is, the polynomials f_i's satisfy $\mathcal{X} f_i = K_i f_i$, and the F_j's satisfy $\mathcal{X} F_j = L_j F_j$, on all the points (x, y, z) of the regular surface $\boldsymbol{\Sigma}$.

(a) We have $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that $\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j = 0$ on all the points (x, y, z) of $\boldsymbol{\Sigma}$. On the other hand, we have that

$$\begin{aligned}
&\mathcal{X}\big(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q}\big) \\
&\quad = \big(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q}\big)\left(\sum_{i=1}^{p} \lambda_i \frac{Xf_i}{f_i} + \sum_{j=1}^{q} \mu_j \frac{XF_j}{F_j}\right) \\
&\quad = \big(f_1^{\lambda_1} \cdots f_p^{\lambda_p} F_1^{\mu_1} \cdots F_q^{\mu_q}\big)\left(\sum_{i=1}^{p} \lambda_i K_i + \sum_{j=1}^{q} \mu_j L_j\right) = 0,
\end{aligned}$$

on all the points (x, y, z) of $\boldsymbol{\Sigma}$. Hence statement (a) follows.

(b) Since (x_k, y_k, z_k) is a singular point of the vector field $\mathcal{X} = (P, Q, R)$ on $\boldsymbol{\Sigma}$, we have that $P(x_k, y_k, z_k) = Q(x_k, y_k, z_k) = R(x_k, y_k, z_k) = 0$. Then, since $\mathcal{X} f_i = P(\partial f_i/\partial x) + Q(\partial f_i/\partial y) + R(\partial f_i/\partial z) = K_i f_i$ on all the points (x, y, z) of $\boldsymbol{\Sigma}$, it follows that $K_i(x_k, y_k, z_k)\ f_i(x_k, y_k, z_k) = 0$. By assumption $f_i(x_k, y_k, z_k) \neq 0$, therefore $K_i(x_k, y_k, z_k) = 0$ for $i = 1, \ldots, p$. Again, since $\mathcal{X} F_j = P(\partial F_j/\partial x) + Q(\partial F_j/\partial y) + R(\partial F_j/\partial z) = L_j F_j$ on all the points (x, y, z) of $\boldsymbol{\Sigma}$, it follows that $L_j(x_k, y_k, z_k) F_j(x_k, y_k, z_k) = 0$. Since $F_j = \exp(g_j/h_j)$ does not vanish, $L_j(x_k, y_k, z_k) = 0$ for $j = 1, \ldots, q$. Consequently, since the r singular points are independent, all the vectors $K_i(x, y, z = \alpha(x, y))$ and $L_j(x, y, z = \alpha(x, y))$ belong to a linear vector subspace S of $\mathbb{C}_{m-1}[x, y, z = \alpha(x, y)]$ of dimension $d(m-1) - r$. We have $p + q$ vectors $K_i(x, y, z = \alpha(x, y))$ and $L_j(x, y, z = \alpha(x, y))$, and since from the assumptions $p + q > d(m-1) - r$, we obtain that these $p + q$ vectors must be linearly dependent on S. So, there are $\lambda_i, \mu_j \in \mathbb{C}$ not all zero such that

$$\sum_{i=1}^{p} \lambda_i K_i\big(x, y, z = \alpha(x, y)\big) + \sum_{j=1}^{q} \mu_j L_j\big(x, y, z = \alpha(x, y)\big) = 0.$$

Hence statement (b) is proved.

(c) Under the assumptions of statement (c) we can apply statement (b) to two subsets of $p + q - 1 \geqslant 2$ functions defining invariant algebraic curves or exponential factors. Therefore, we get two linear dependencies between the corresponding cofactors, which after some linear algebra and relabeling, we can write into the following form

$$\begin{aligned}
&M_1 + \alpha_3 M_3 + \cdots + \alpha_{p+q} M_{p+q} = 0, \\
&M_2 + \beta_3 M_3 + \cdots + \beta_{p+q} M_{p+q} = 0,
\end{aligned}$$

where M_l are the cofactors K_i and L_j, and the α_l and β_l are real numbers. Of course, these two equalities must be satisfied only on all the points (x, y, z) of $\boldsymbol{\Sigma}$. Then, by statement (a), it follows that the two functions

$$|G_1||G_3|^{\alpha_3}\cdots|G_{p+q}|^{\alpha_{p+q}},$$
$$|G_2||G_3|^{\beta_3}\cdots|G_{p+q}|^{\beta_{p+q}},$$

are first integrals of the vector field $\mathcal{X}$ on $\boldsymbol{\Sigma}$, where for $l = 1, \dots, p+q$ the function G_l is the polynomial defining an invariant algebraic curve or the exponential factor having cofactor M_l. Then, taking logarithms to the above two first integrals, we obtain that

$$H_1 = \log|G_1| + \alpha_3 \log|G_3| + \cdots + \alpha_{p+q} \log|G_{p+q}|,$$
$$H_2 = \log|G_2| + \beta_3 \log|G_3| + \cdots + \beta_{p+q} \log|G_{p+q}|,$$

are first integrals of the vector field $\mathcal{X}$ on $\boldsymbol{\Sigma}$.

For each one of the regular surfaces $\boldsymbol{\Sigma}$ that we consider, we have defined in the previous subsection a local chart $\lambda : \mathbb{R}^2 \to \boldsymbol{\Sigma}$ given by $\lambda(r,s) = (x,y,z)$. We continue denoting by λ the natural extension of λ from $\mathbb{C}^2$ to the complexified $\boldsymbol{\Sigma}$. We note that when $\boldsymbol{\Sigma}$ is the circular cylinder the map here denoted by λ was denoted by ϕ in the previous subsection. Now, the vector field $\mathcal{X}$ on $\boldsymbol{\Sigma}$ induces a vector field $\omega\mathcal{X}$ on $\mathbb{C}^2$ through $\omega\mathcal{X}^{\mathrm{T}}(r,s) = D\lambda^{-1}(\lambda(r,s))X^{\mathrm{T}}$ associated to $\omega\mathcal{X}$ is denoted by

$$\dot{r} = \omega P(r,s), \qquad \dot{s} = \omega Q(r,s). \tag{131}$$

Since $H_i(x,y,z)$ is a first integral of $\mathcal{X}$ on $\boldsymbol{\Sigma}$, it follows that $\omega H_i(r,s) = H_i(\lambda(r,s))$ is a first integral of $\omega\mathcal{X}$ on $\mathbb{C}^2$ (or $\mathbb{R}^2$ when we restrict on the real domain), for $i = 1, 2$.

Each first integral ωH_i provides an integrating factor $R_i(r,s)$ for system (131) such that

$$\omega P R_i = \frac{\partial \omega H_i}{\partial r}, \qquad \omega Q R_i = -\frac{\partial \omega H_i}{\partial s},$$

for $i = 1, 2$. Therefore, we obtain that

$$\frac{R_1}{R_2} = \frac{\frac{\partial \omega H_2}{\partial r}}{\frac{\partial \omega H_1}{\partial r}}. \tag{132}$$

Since the functions G_l are polynomials or exponentials of a quotient of polynomials, from the expression of H_i it follows that the entries of the Jacobian matrix $DH_i(x,y,z)$ are rational functions in the variables x, y and z. From the chain rule we have that

$$D\omega H_i(r,s) = DH_i(\lambda(r,s)) \circ D\lambda(r,s).$$

In the previous subsection we have studied that the functions $\lambda(r,s)$ and $D\lambda(r,s)$ are rational in the variables r, s and, sometimes $\delta(x,y)$, a convenient square root depending on r and s, according with the type of the surface $\boldsymbol{\Sigma}$. So the partial derivatives of (132) are rational functions in the variables r, s and perhaps $\delta(r,s)$.

Now using for each surface $\boldsymbol{\Sigma}$ the expression $(r,s)=\lambda^{-1}(x,y,z)$, we can write the function R_1/R_2 of (132) in the variables (x,y,z). From Proposition 15.1 it follows that the function $(R_1/R_2)(r,s)$ is a first integral of system (131), and consequently that the function $(R_1/R_2)(x,y,z)$ is a first integral of the vector field $\mathcal{X}$ on $\boldsymbol{\Sigma}$. This completes the proof of statement (c). □

We note that in statement (b) of Theorem 15.3 the number $d(m-1)+1$ is equal to m^2+1 for the nine quadratics and $2m^2-4m+5$ for the torus.

We remark that in Theorem 15.3(c) once we know a rational first integral in the variables x, y, z and $\mu=\sqrt{x^2+y^2}$, it is always possible to construct a rational first integral in the variables x, y and z.

For an extension of the results of Theorem 15.3 to vector fields on a regular algebraic hypersurface of $\mathbb{R}^n$ see [71].

References

[1] M.I. Al'mukhamedov, *On the construction of a differential equation having given curves as limit cycles*, Izv. Vyss. Ucebn. Zaved. Mat. **1** (44) (1963), 3–6 (in Russian).

[2] R.F.S. Andrade and A. Rauh, *The Lorenz model and the method of Carleman embedding*, Phys. Lett. A **82** (1981), 276–278.

[3] V. Arnold, *Équations Différentielles Ordinaires*, Éditions Mir, Moscou (1974).

[4] F. Balibrea and V. Jimenez, *A characterization of the ω-limit sets of planar continuous dynamical systems*, J. Differential Equations **145** (1998), 469–488.

[5] R. Bamon, *Quadratic vector fields in the plane have a finite number of limit cycles*, Int. Hautes Études Sci. Publ. Math. **64** (1986), 111–142.

[6] N.N. Bautin, *On the number of limit cycles which appear with the variation of the coefficients from an equilibrium position of focus or center type*, Math. USSR-Sb. **100** (1954), 397–413.

[7] T.C. Bountis, A. Ramani, B. Grammaticos and B. Dorizzi, *On the complete and partial integrability of non-Hamiltonian systems*, Phys. A **128** (1984), 268–288.

[8] L. Brenig and A. Goriely, *Quasimonomial transformations and integrability*, Partially Integrable Evolution Equations in Physics (Les Houches, 1989), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., Vol. 310, Kluwer Academic, Dordrecht (1990), 571–572.

[9] L. Cairó, M.R. Feix and J. Llibre, *Integrability and algebraic solutions for planar polynomial differential systems with emphasis on the quadratic systems*, Resenhas de Universidade de Sao Paulo **4** (2000), 127–161.

[10] L. Cairó, H. Giacomini and J. Llibre, *Liouvillian first integrals for the planar Lotka–Volterra system*, Rend. Circ. Mat. Palermo **52** (2003), 389–418.

[11] L. Cairó and J. Llibre, *Darbouxian first integrals and invariants for real quadratic systems having an invariant conic*, J. Phys. A **35** (2002), 589–608.

[12] A. Campillo and M.M. Carnicer, *Proximity inequalities and bounds for the degree of invariant curves by foliations of* $\mathbf{P}^2_{\mathbb{C}}$, Trans. Amer. Math. Soc. **349** (1997), 2211–2228.

[13] F. Cantrijn and W. Sarlet, *Generalizations of Noether's theorem in classical mechanics*, SIAM Rev. **23** (1981), 467–494.

[14] T. Carleman, *Application de la théorie des équations intégrales linéaires aux systémes d'équations différentielles non linéaires*, Acta Math. **59** (1932), 63.

[15] M.M. Carnicer, *The Poincaré problem in the nondicritical case*, Ann. Math. **140** (1994), 289–294.

[16] D. Cerveau and A. Lins Neto, *Holomorphic foliations in CP(2) having an invariant algebraic curve*, Ann. Inst. Fourier **41** (1991), 883–903.

[17] J. Chavarriga, *Integrable systems in the plane with a center type linear part*, Appl. Math. **22** (1994), 285–309.

[18] J. Chavarriga, H. Giacomini and J. Giné, *An improvement to Darboux integrability theorem for systems having a center*, Appl. Math. Lett. **12** (1999), 85–89.

[19] J. Chavarriga, H. Giacomini, J. Giné and J. Llibre, *On the integrability of two-dimensional flows*, J. Differential Equations **157** (1999), 163–182.

[20] J. Chavarriga, H. Giacomini, J. Giné and J. Llibre, *Darboux integrability and the inverse integrating factor*, J. Differential Equations, **194** (2003), 116–139.

[21] J. Chavarriga, H. Giacomini and J. Llibre, *Uniqueness of algebraic limit cycles for quadratic systems*, J. Math. Anal. Appl. **261** (2001), 85–99.

[22] J. Chavarriga and J. Llibre, *On the algebraic limit cycles of quadratic systems*, Proceedings of the IV Catalan Days of Applied Mathematics (Tarragona, 1998), Univ. Rovira Virgili, Tarragona (1998), 17–24.

[23] J. Chavarriga and J. Llibre, *Invariant algebraic curves and rational first integrals for planar polynomial vector fields*, J. Differential Equations **169** (2001), 1–16.

[24] J. Chavarriga, J. Llibre and J. Moulin Ollagnier, *On a result of Darboux*, J. Comput. Math. **4** (2001), 197–210.

[25] J. Chavarriga, J. Llibre and J. Sorolla, *Algebraic limit cycles of quadratic systems*, J. Differential Equations, to appear

[26] J. Chavarriga, J. Llibre and J. Sotomayor, *Algebraic solutions for polynomial systems with emphasis in the quadratic case*, Exposition. Math. **15** (1997), 161–173.

[27] C. Christopher, *Polynomial systems with invariant algebraic curves*, Preprint, University of Wales, Aberystwyth (1991).

[28] C.J. Christopher, *Invariant algebraic curves and conditions for a center*, Proc. Roy. Soc. Edinburgh A **124** (1994), 1209–1229.

[29] C. Christopher, *Polynomial systems with invariant algebraic curves*, Preprint (1993).

[30] C. Christopher and J. Llibre, *Algebraic aspects of integrability for polynomial systems*, Qualitative Theor. Planar Differential Equations **1** (1999), 71–95.

[31] C. Christopher and J. Llibre, *Integrability via invariant algebraic curves for planar polynomial differential systems*, Ann. Differential Equations **16** (2000), 5–19.

[32] C. Christopher and J. Llibre, *A family of quadratic polynomial differential systems with invariant algebraic curves of arbitrarily high degree without rational first integrals*, Proc. Amer. Math. Soc. **130** (2002), 2025–2030.

[33] C. Christopher, J. Llibre, C. Pantazi and X. Zhang, *Darboux integrability and invariant algebraic curves for planar polynomial systems*, J. Phys. A **35** (2002), 2457–2476.

[34] C. Christopher, J. Llibre and J.V. Pereira, *Multiplicity of invariant algebraic curves in polynomial vector fields*, Preprint (2003).

[35] C. Christopher, J. Llibre and G. Świrszcz, *Invariant algebraic curves of large degree for quadratic systems*, J. Math. Anal. Appl., to appear.

[36] W.A. Coppel, *Some quadratic systems with at most one limit cycle*, Dynam. Report. **2** (1989), 61–88.

[37] D. Cozma and A. Şubă, *Partial integrals and the first focal value in the problem of centre*, Nonlinear Differential Equations Appl. **2** (1995), 21–34.

[38] G. Darboux, *Mémoire sur les équations différentielles algébriques du premier ordre et du premier degré (Mélanges)*, Bull. Sci. Math. 2ème Série **2** (1878), 60–96, 123–144, 151–200.

[39] J. Écalle, *Introduction aux fonctions analysables et preuve constructive de la conjecture de Dulac*, Hermann (1992).

[40] R.M Evdokimenco, *Construction of algebraic paths and the qualitative investigation in the large of the properties of integral curves of a system of differential equations*, Differential Equations **6** (1970), 1349–1358.

[41] R.M. Evdokimenco, *Behavior of integral curves of a dynamic system*, Differential Equations **9** (1974), 1095–1103.

[42] R.M. Evdokimenco, *Investigation in the large of a dynamic system with a given integral curve*, Differential Equations **15** (1979), 215–221.

[43] V.F. Filiptsov, *Algebraic limit cycles*, Differential Equations **9** (1973), 983–986.

[44] M. Frommer, *Über das Auftreten von Wirbeln und Strudeln (geschlossener und spiraliger Integralkurven) in der Umgebung rationaler Unbestimmheitsstellen*, Math. Ann. **109** (1934), 395–424.

[45] W. Fulton, *Algebraic Curves*, Mathematics Lecture Note Series, Benjamin–Cummings (1969). Reprinted by Addison–Wesley (1989).

[46] S.D. Furta, *On non-integrability of general systems of differential equations*, Z. Angew. Math. Phys. **47** (1996), 112–131.

[47] H. Giacomini, J. Llibre and M. Viano, *On the nonexistence, existence and uniqueness of limit cycles*, Nonlinearity **9** (1996), 501–516.

[48] H. Giacomini, J. Llibre and M. Viano, *On the shape of limit cycles that bifurcate from Hamiltonian centers*, Nonlinear Anal. **41** (2000), 523–537.

[49] H. Giacomini, J. Llibre and M. Viano, *On the shape of limit cycles that bifurcate from non-Hamiltonian centers*, Nonlinear Anal. **43** (2001), 837–859.

[50] H. Giacomini, J. Llibre and M. Viano, *Arbitrary order bifurcations for perturbed Hamiltonian planar systems via the reciprocal of an integrating factor*, Nonlinear Anal. **48** (2002), 117–136.

[51] H.J. Giacomini, C.E. Repetto and O.P. Zandron, *Integrals of motion of three-dimensional non-Hamiltonian dynamical systems*, J. Phys. A **24** (1991), 4567–4574.

[52] A. Goriely, *Integrability and Non-Integrability of Dynamical Systems*, Advanced Series in Nonlinear Dynamics, Vol. 19, World Scientific (2001).

[53] B. Grammaticos, J. Moulin Ollagnier, A. Ramani, J.M. Strelcyn and S. Wojciechowski, *Integrals of quadratic ordinary differential equations in* $\mathbb{R}^3$*: the Lotka–Volterra system*, Physica A **163** (1990), 683–722.

[54] C. Gutierrez and J. Llibre, *Darboux integrability for polynomial vector fields on the* 2*-dimensional sphere*, Extracta Math. **17** (2002), 289–301.

[55] J. Hietarinta, *Direct methods for the search of the second invariant*, Phys. Rep. **147** (1987), 87–154.

[56] D. Hilbert, *Mathematische Probleme*, Lecture, Second Internat. Congr. Math. (Paris, 1900), Nachr. Ges. Wiss. Göttingen Math. Phys. KL. (1900), 253–297; English transl.: Bull. Amer. Math. Soc. **8** (1902), 437–479.

[57] Yu. Ilyashenko, *Finiteness Theorems for Limit Cycles*, Translations of Math. Monographs, Vol. 94, Amer. Math. Soc. (1991).

[58] J.P. Jouanolou, *Equations de Pfaff Algébriques*, Lectures Notes in Mathematics, Vol. 708, Springer-Verlag, New York/Berlin (1979).

[59] W. Kapteyn, *On the midpoints of integral curves of differential equations of the first degree*, Nederl. Akad. Wetensch. Verslag. Afd. Natuurk. Konikl. Nederland (1911), 1446–1457 (in Dutch).

[60] W. Kapteyn, *New investigations on the midpoints of integrals of differential equations of the first degree*, Nederl. Akad. Wetensch. Verslag Afd. Natuurk. **20** (1912), 1354–1365; **21** (1912), 27–33 (in Dutch).

[61] R. Kooij and C. Christopher, *Algebraic invariant curves and the integrability of polynomial systems*, Appl. Math. Lett. **6** (1993), 51–53.

[62] P.D. Lax, *Integrals of nonlinear equations of evolution and solitary waves*, Commun. Pure Appl. Math. **21** (1968), 467–490.

[63] Weigu Li, J. Llibre, M. Nicolau and X. Zhang, *On the differentiability of first integrals of two-dimensional flows*, Proc. Amer. Math. Soc. **130** (2002), 2079–2088.

[64] Weigu Li, J. Llibre and Xiang Zhang, *Planar analytic vector fields with generalized rational first integrals*, Bull. Sci. Math. **125** (2001), 341–361.

[65] A. Lins Neto, *Some examples for the Poincaré and Painlevé problems*, Ann. Sci. École Norm. Sup. **35** (2002), 231–266.

[66] J. Llibre, J.S. Pérez del Río and J.A. Rodríguez, *Phase portraits of a new class of integrable quadratic vector fields*, Dynam. Contin. Discrete Impuls. Systems **7** (2000), 595–616.

[67] J. Llibre and G. Rodríguez, *Invariant hyperplanes and Darbox integrability for* d*-dimensional polynomial differential systems*, Bull. Sci. Math. **124** (2000), 1–21.

[68] J. Llibre and G. Rodríguez, *Darboux integrability of polynomial vector fields on* 2*-dimensional surfaces*, J. Bifurcation Chaos **12** (2002), 2821–2833.

[69] J. Llibre and G. Rodríguez, *Configurations of limit cycles and planar polynomial vector fields*, J. Differential Equations **198** (2004), 374–380.

[70] J. Llibre and D. Schlomiuk, *The geometry of differential quadratic systems with a weak focus of third order*, Canadian J. Math., to appear.

[71] J. Llibre and X. Zhang, *Darboux integrability of real polynomial vector fields on regular algebraic hypersurfaces*, Rend. Circ. Mat. Palermo, Serie II **LI** (2002), 109–126.

[72] A.J. Lotka, *Analytical note on certain rhythmic relations in organic systems*, Proc. Nat. Acad. Sci. **6** (1920), 410.

[73] Yiu-Kwong Man and M.A.H. Maccallum, *A rational approach to the Prelle–Singer algorithm*, J. Symbolic Comput. **24** (1997), 31–43.

[74] L. Markus, *Global structure of ordinary differential equations in the plane*, Trans. Amer. Math. Soc. **76** (1954), 127–148.

[75] J. Moulin Ollagnier, *Polynomial first integrals of the Lotka–Volterra system*, Bull. Sci. Math. **121** (1997), 463–476.

[76] J. Moulin Ollagnier, *Rational integration of the Lotka–Volterra system*, Bull. Sci. Math. **123** (1999), 437–466.

[77] J. Moulin Ollagnier, *Liouvillian Integration of the Lotka–Volterra system*, Qualitative Theor. Dynamical Systems **2** (2001), 307–358.

[78] J. Moulin Ollagnier, *About a conjecture on quadratic vector fields*, J. Pure Appl. Algebra **165** (2001), 227–234.

[79] J. Moulin Ollagnier, A. Nowicki and J.M. Strelcyn, *On the non-existence of constants of derivations: the proof of a theorem of Jouanolou and its development*, Bull. Sci. Math. **119** (1995), 195–233.

[80] V.V. Nemytskii and V.V. Stepanov, *Qualitative Theory of Differential Equations*, Dover, New York (1989).

[81] D.A. Neumann, *Classification of continuous flows on 2-manifolds*, Proc. Amer. Math. Soc. **48** (1975), 73–81.

[82] K. Odani, *The limit cycle of the van der Pol equation is not algebraic*, J. Differential Equations **115** (1995), 146–152.

[83] P.J. Olver, *Applications of Lie Groups to Differential Equations*, Springer, New York (1986).

[84] J.M. Pearson, N.G. Lloyd and C.J. Christopher, *Algorithmic derivation of centre conditions*, SIAM Rev. **38** (1996), 619–636.

[85] J.V. Pereira, *Integrabilidade de equaçoes diferenciais no plano complexo*, Monografias del IMCA, Vol. 25, Lima, Peru (2002).

[86] H. Poincaré, *Sur l'intégration des équations différentielles du premier ordre et du premier degré I and II*, Rend. Circ. Mat. Palermo **5** (1891), 161–191; **11** (1897), 193–239.

[87] M.J. Prelle and M.F. Singer, *Elementary first integrals of differential equations*, Trans. Amer. Math. Soc. **279** (1983), 613–636.

[88] Qin Yuan-Xun, *On the algebraic limit cycles of second degree of the differential equation* $\mathrm{d}y/\mathrm{d}x = \sum_{0\leqslant i+j\leqslant 2} a_{ij}x^i y^j / \sum_{0\leqslant i+j\leqslant 2} b_{ij}x^i y^j$, Acta Math. Sinica **8** (1958), 23–35.

[89] C. Rousseau and D. Schlomiuk, *Cubic vector fields symmetric with respect to a center*, J. Differential Equations **123** (1995), 388–436.

[90] S. Schecter and M.F. Singer, *A class of vectorfields on S^2 that are topologically equivalent to polynomial vectorfields*, J. Differential Equations **57** (1985), 406–435.

[91] D. Schlomiuk, *Elementary first integrals and algebraic invariant curves of differential equations*, Exposition. Math. **11** (1993), 433–454.

[92] D. Schlomiuk, *Algebraic particular integrals, integrability and the problem of the center*, Trans. Amer. Math. Soc. **338** (1993), 799–841.

[93] D. Schlomiuk, *Algebraic and geometric aspects of the theory of polynomial vector fields*, Bifurcations and Periodic Orbits of Vector Fields, D. Schlomiuk, ed. (1993), 429–467.

[94] A.S. Shubé, *Partial integrals, integrability, and the center problem*, Differential Equations **32** (1996), 884–892.

[95] M.F. Singer, *Liouvillian first integrals of differential equations*, Trans. Amer. Math. Soc. **333** (1992), 673–688.

[96] J. Sotomayor, *Liçóes de Equaçóes Diferenciais Ordinárias*, IMPA, Rio de Janerio (1979).

[97] J.M. Strelcyn and S. Wojciechowski, *A method of finding integrals for three-dimensional dynamical systems*, Phys. Lett. A **133** (1988), 207–212.

[98] R. Sverdlove, *Inverse problems for dynamical systems*, J. Differential Equations **42** (1981), 72–105.

[99] R.T. Valeeva, *Construction of a differential equation with given limit cycles and critical points*, Volz Math. Sb. **5** (1966), 83–85 (in Russian).

[100] V. Volterra, *Leçons sur la théorie Mathematique de la lutte pour la vie*, Gauthier–Villars, Paris (1931).

[101] S. Walcher, *On the Poincaré problem*, J. Differential Equations **166** (2000), 51–78.

[102] J.A. Weil, *Constantes et polynômes de Darboux en algèbre différentielle : applications aux systèmes différentiels linéaires*, Ph.D., École Polytecnique (1995).

[103] A.I. Yablonskii, *Limit cycles of a certain differential equations*, Differential Equations **2** (1966), 335–344 (in Russian).

[104] Ye Yanqian, *Theory of Limit Cycles*, Translations of Math. Monographs, Vol. 66, Amer. Math. Soc., Providence (1986).

[105] Zhang Zhifen, Ding Tongren, Huang Wenzao and Dong Zhenxi, *Qualitative Theory of Differential Equations*, Translations of Math. Monographs, Vol. 101, Amer. Math. Soc., Providence (1992).

[106] H. Żołądek, *The solution of the center-focus problem*, Preprint (1992).

[107] H. Żołądek, *On algebraic solutions of algebraic Pfaff equations*, Studia Math. **114** (1995), 117–126.

CHAPTER 6

Global Results for the Forced Pendulum Equation

Jean Mawhin

Département de mathématique, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium
E-mail: mawhin@amm.ucl.ac.be

Contents

1. Introduction . . . 535
2. Autonomous pendulum equations . . . 537
    2.1. The free conservative pendulum . . . 537
    2.2. The free damped pendulum . . . 539
    2.3. The pendulum with constant torque . . . 540
3. Periodic solutions of the forced pendulum . . . 541
    3.1. The problems . . . 541
    3.2. The possibly dissipative case $c \geqslant 0$ . . . 543
    3.3. The conservative case $c = 0$ . . . 546
    3.4. The dissipative case $c > 0$ . . . 551
    3.5. The degeneracy problem . . . 554
    3.6. Multiplicity and stability results under restrictions of the coefficients . . . 558
    3.7. Many T-periodic solutions for special forcings . . . 560
    3.8. Many T-periodic solutions for small length pendulum . . . 562
    3.9. Subharmonic solutions when $c = 0$ . . . 563
4. Rotating solutions and Mather sets . . . 565
    4.1. Periodic solutions of the second kind . . . 565
    4.2. Solutions with an arbitrary rotation number . . . 566
    4.3. Mather set . . . 570
5. KAM theory and Lagrange stability . . . 571
    5.1. Twist mappings . . . 571
    5.2. Chaotic dynamics . . . 574
6. Bounded forcing . . . 577
    6.1. Bounded functions and their averages . . . 577
    6.2. A necessary condition for the existence of bounded solutions . . . 577
    6.3. Sufficient conditions for the existence of bounded solutions . . . 578
    6.4. Local uniqueness of bounded solutions . . . 580
7. Almost periodic forcings . . . 582
    7.1. Almost periodic functions . . . 582
    7.2. Almost periodic solutions . . . 583
References . . . 585

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 1
Edited by A. Cañada, P. Drábek and A. Fonda
© 2004 Elsevier B.V. All rights reserved

This Page Intentionally Left Blank

1. Introduction

The obtention of global (i.e., nonperturbational) results for the forced pendulum equation started with the rigorous mathematical study of the periodic solutions of the periodically forced pendulum equation

$$y'' + \alpha^2 \sin y = \beta \sin t \tag{1}$$

initiated in 1922 by Georg Hamel, in a paper of the special issue of the *Mathematische Annalen* dedicated to Hilbert's sixtieth birthday anniversary [55].

Hamel's contribution starts by proving the existence of a 2π-periodic solution of Equation (1) through the *direct method of the calculus of variations*. He shows that the corresponding *action integral*

$$A(y) := \int_0^{2\pi} \left(\frac{y'^2(t)}{2} + \alpha^2 \cos y(t) + \beta y(t) \sin t \right) \mathrm{d}t$$

has a minimum over the space of 2π-periodic C^1-functions, and his argument easily extends to the case where $b \sin t$ is replaced by a continuous 2π-periodic function $h(t)$ with mean value zero. The *Ritz method* is then used to find a first approximation of the amplitude of the periodic solution.

Hamel then notices that the symmetries of the equation imply that any solution of Equation (1) satisfying the boundary conditions

$$y(0) = y(\pi) = 0 \tag{2}$$

can be extended as an *odd* 2π-periodic solution. The problem (1)–(2) is reduced to the *nonlinear integral equation*

$$y(t) = -\alpha^2 \int_0^{2\pi} K(t,\tau) \sin y(\tau)\, \mathrm{d}\tau - \beta \sin t := F(y)(t), \tag{3}$$

where $K(t,\tau)$ is the Green function of

$$y'' = h(t), \quad y(0) = y(\pi) = 0.$$

The *method of successive approximations*

$$y_{n+1} = F(y_n), \quad y_0(t) = -b \sin t,$$

is shown to converge to a (unique) solution when $\alpha^2 < 1$. Hamel's argument is equivalent to proving the existence of a sufficiently large integer m, for which the mth iterate F^m of F is a contraction in the space $C([0,\pi])$.

When $\alpha^2 \geqslant 1$, Hamel uses *Lyapunov–Schmidt's method* to study the 2π-periodic solutions of (1) when $|b|$ is sufficiently small, and mentions the possibility of the existence of more than one 2π-periodic solution.

So, Hamel anticipates or uses several of the fundamental methods of nonlinear analysis, in a work which will not be essentially superseded during some sixty years.

An important role in renewing the interest to the forced pendulum equation was played by Fučik in the late nineteen seventies, when he wrote, in the Introduction of Chapter 26 of his monograph [49]:

> Finally we shall present here one attempt to obtain the existence of a T-periodic solution of the mathematical pendulum equation
>
> $$-u''(t) + \sin u(t) = f(t). \tag{4}$$
>
> The result is not final since the necessary and sufficient condition obtained for T-periodic solvability of (4) is not useful.

After describing very partial results in this direction and mentioning extensions personally communicated by Dancer, Fučik concluded that the

> description of the set $\mathcal{P}$ of f for which the equation $u'' + \sin u = f(t)$ has a T-periodic solution seems to remain a *terra incognita.*

Motivated by Fučík's remarks, but unaware of the existence of Hamel's paper, Castro [32] (for $\alpha \leqslant 1$), Dancer [36] and Willem [129], independently (for arbitrary α), reintroduced in the early eighties the use of the direct method of the calculus of variations, in the setting of Sobolev spaces. The time was ripe for the obtention by Mawhin and Willem [85], some sixty years after the first one, of a *second periodic solution,* using a refinement of the *mountain pass lemma*.

At the same time, the forced pendulum equation also became a paradigm for the *theory of chaos*, and appeared in the description of *Josephson type junctions*. We only describe global results in this direction here, and refer to the bibliography of [82,117] for the numerous perturbational and numerical aspects.

A fundamental role in the development of the qualitative theory of nonlinear differential equations and its applications to engineering, was also played by the *pendulum equation with a constant torque* or *Tricomi's equation*

$$y'' + cy' + a \sin y = b,$$

introduced by Tricomi [125,126] in his studies of synchronous electrical machines, and widely developed since (see [83] for references).

To keep the size of the survey reasonable and facilitate the comparison between results obtained through different methods, we only state the theorems for the standard (possibly dissipative) forced pendulum equation

$$y'' + cy' + a \sin y = h(t).$$

Most of the assertions remain valid if $a \sin y$ is replaced by an arbitrary continuous S-periodic function $g(y)$ with mean value zero. Recent results depend upon the fact that $a \sin y$ is replaced by a S-periodic function whose Fourier series contains higher harmonics [63]. Also, some conclusions survive when the friction term cy' is replaced by a more general one of Liénard type $f(y)y'$ or of Rayleigh type $f(y')$. See the references in [82] and the recent papers [52,54].

For the same reason of brevity, we will not describe the possible generalizations to *systems of pendulum-type,* and in particular to the equations of the *forced multiple pendulum,* and to *higher order pendulum-type equations.* The reader can consult the bibliography in [82] and the recent paper [116]. We shall also leave aside the existence of forced oscillations of the *spherical pendulum* (which depends upon methods of a quite different nature), considered by Furi, Pera and Spadini, and the pendulum-type equations describing the *libration of satellites*. Again, references to the corresponding literature for those questions can be found in [82].

Let us mention also that the corresponding problem for the case of *Dirichlet boundary conditions,* namely

$$y'' + y + a\sin y = h(t), \quad y(0) = 0 = y(\pi),$$

and its analog for partial differential equations, has been the object, since the pioneering papers of Ward [128] using critical point theory, and of Schaaf and Schmitt [118] using global bifurcation, of a number of studies by Arcoya, Cañada, Lupo, Roca, Ruiz, Solimini, Ureña and others. References can be found in [82] and in the more recent contributions [28–31]. This problem, which has both deep analogies and strong differences with the periodic boundary value problem for the forced pendulum, will not be considered here.

Finally, in order to keep the number of references at a reasonable level, we have only quoted the papers whose results are directly mentioned in this survey. Further bibliographical information can be found in [82].

2. Autonomous pendulum equations

2.1. *The free conservative pendulum*

To allow the evaluation of the subsequent results, let us briefly recall the structure of the solution set of the *free conservative pendulum equation*

$$y'' + a\sin y = 0, \tag{5}$$

where $a > 0$. This information can easily be obtained from the *energy integral*

$$\frac{y'^2}{2} - a\cos y = C, \tag{6}$$

by studying the orbits (y, y') in the phase-plane (or in the phase-cylinder $S^1 \times \mathbb{R}$). Equation (5) admits two geometrically distinct *equilibria*, namely

$$\big(y(t), y'(t)\big) \equiv (0,0), \qquad \big(y(t), y'(t)\big) \equiv (\pi,0).$$

The orbits through $(A, 0)$, with $0 < y(0) = A < \pi$ are closed and surround the equilibrium $(0, 0)$, which is therefore a *center*. They correspond to *periodic solutions* of (5) with maximum amplitude A and maximum velocity $\sqrt{2a(1 - \cos A)}$. Their period

$$T(A) = 2\sqrt{2} \int_0^A \frac{\mathrm{d}y}{\sqrt{a(\cos y - \cos A)}}$$

is an increasing function such that

$$\lim_{A \to 0+} T(A) = \frac{2\pi}{\sqrt{a}}, \qquad \lim_{A \to \pi-} T(A) = +\infty. \tag{7}$$

Physically, those periodic solutions correspond to *oscillations* of the pendulum around its stable equilibrium position.

The orbits through $(0, B)$ with $B > 2\sqrt{a}$ (respectively $B < -2\sqrt{a}$) are the periodic curves

$$y' = \pm\bigl[B^2 - 2a(1 - \cos y)\bigr]^{1/2} \tag{8}$$

in the upper (respectively lower) phase-plane (or closed curves surrounding the phase-cylinder). So $y(t)$ is strictly monotone and its inverse function $t(y)$, given by

$$t(y) = \pm \int_0^y \frac{\mathrm{d}s}{[B^2 - 2a(1 - \cos y)]^{1/2}},$$

is the indefinite integral of a 2π-periodic function. Thus, if we define

$$\tau(B) := \int_0^{2\pi} \frac{\mathrm{d}y}{\sqrt{B^2 - 2a(1 - \cos y)}}, \tag{9}$$

$t(y)$ has the form

$$t(y) = \pm\left(\frac{\tau(B)}{2\pi} y + P(y)\right),$$

with P a 2π-periodic function. Therefore,

$$\frac{\tau(B)}{2\pi}(y + 2\pi) + P(y + 2\pi) = \tau(B) + \frac{\tau(B)}{2\pi} y + P(y) = \tau(B) + t,$$

which shows that

$$y\bigl(t + \tau(B)\bigr) = y(t) + 2\pi \quad (t \in \mathbb{R}),$$

and hence, by (8),

$$y'\bigl(t + \tau(B)\bigr) = y'(t) \quad (t \in \mathbb{R}).$$

Such a y is called a *periodic solution of the second kind* of Equation (5). The function $\tau(B)$ is such that

$$\lim_{B\to 2\sqrt{a}+} \tau(B) = +\infty, \qquad \lim_{B\to+\infty} \tau(B) = 0.$$

Physically, those solutions correspond to *rotations* of the pendulum.

Finally, the orbit through $(0, 2\sqrt{a})$ (respectively $(0, -2\sqrt{a})$ connects in the upper (respectively lower) half phase-space the equilibria $(-\pi, 0)$ and $(\pi, 0)$ (*heteroclinic orbits*) (or *homoclinic orbits* connecting $(\pi, 0)$ to itself in the phase-cylinder). Thus $(\pi, 0)$ is a *saddle point*. Those orbits correspond to positive (respectively negative) *asymptotic solutions* y of (5) such that

$$\lim_{t\to-\infty} y(t) = -\pi, \qquad \lim_{t\to+\infty} y(t) = \pi$$

(respectively

$$\lim_{t\to-\infty} y(t) = \pi, \qquad \lim_{t\to+\infty} y(t) = -\pi),$$

and

$$\lim_{t\to\pm\infty} y'(t) = 0.$$

Physically, those solutions correspond to motions which are *asymptotic* (in the past or in the future) to the unstable vertical equilibrium.

Notice that, if $T \leqslant \frac{2\pi}{\sqrt{a}}$, the only T-periodic solutions of Equation (5) are the equilibria $y \equiv 0$ and $y \equiv \pi$. If $T > \frac{2\pi}{\sqrt{a}}$, and if $m \geqslant 1$ is the largest integer such that $\frac{T}{m} > \frac{2\pi}{\sqrt{a}}$, in other words if

$$m = \left[\frac{T\sqrt{a}}{2\pi}\right],$$

where $[s]$ denotes the integer part of s, then Equation (5) has, besides the two equilibria, nontrivial T-periodic solutions corresponding to the closed orbits of amplitude A_k such that

$$T(A_k) = \frac{T}{k} \quad (k = 1, 2, \ldots, m).$$

2.2. *The free damped pendulum*

The study of the *free damped pendulum*

$$y'' + cy' + a \sin y = 0, \tag{10}$$

where, without loss of generality, we assume $c > 0$ and $a > 0$, is more delicate, as no energy integral exists. Equation (10) still admits the constant solutions

$$y \equiv 0, \qquad y \equiv \pi,$$

but has no nonconstant T-periodic solution of any period T. Indeed, if y is a T-periodic solution of (10), integrating the identity

$$y''y' + cy'^2 + (a \sin y)y' = 0,$$

over $[0, T]$, and using the periodicity gives

$$c \int_0^T \big(y'(t)\big)^2 \, \mathrm{d}t = 0.$$

Thus $y' = 0$, and $y = 0 \pmod{\pi}$.

Now, the variational equation around the zero equilibrium

$$z'' + cz' + az = 0$$

has characteristic roots which are both real and negative if $c \geqslant 2\sqrt{a}$ and complex conjugate with negative real part if $c < 2\sqrt{a}$. Thus the zero equilibrium is *asymptotically stable.* The variational equation around the equilibrium $y \equiv \pi$

$$z'' + cz' - az = 0$$

has a positive and a negative characteristic root. It is unstable of *saddle point* type.

Using the Lyapunov function

$$V(y, z) = z^2 + (cy + z)^2 + 4a(1 - \cos y),$$

for the associated first order system obtained in letting $y' = z$ in (10), one can show that, in the phase-cylinder, with the exception of the two orbits which constitute the stable manifold of the saddle point $(\pi, 0)$, all other orbits correspond to solutions tending to $(0, 0)$ when $t \to +\infty$. The stable equilibrium $(0, 0)$ is a focus or a node according to $c < 2\sqrt{a}$ or $c \geqslant 2\sqrt{a}$.

Thus Equation (10) has no nonconstant periodic solution, no periodic solution of the second kind, no homoclinic orbit. It has two heteroclinic orbits connecting $(\pi, 0)$ to $(0, 0)$ and corresponding to the unstable manifolds of the saddle point $(\pi, 0)$. We refer to [6] for the corresponding phase plane or phase cylinder portraits.

2.3. *The pendulum with constant torque*

The differential equation of the *conservative pendulum with constant torque* is

$$y''(t) + a \sin y(t) = b, \tag{11}$$

where $a > 0,\ b > 0$, and its energy integral

$$\frac{y'^2(t)}{2} - a\cos y(t) - by = C \quad (C \in \mathbb{R}), \tag{12}$$

gives the equations of the family of the orbits of (11) in the phase space (y, y').

Equation (11) has zero, one unstable $(y(t) \equiv \pi/2)$, or one stable $(y(t) \equiv \arcsin(b/a))$ and one unstable $(y(t) \equiv \pi - \arcsin(b/a))$ equilibria, according to $b > a$, $b = a$ or $b < a$. If $b < a$, Equation (11) has T-periodic solutions for each $T \in]2\pi/\sqrt{a^2 - b^2}, +\infty[$, and, if $b \geqslant a$, Equation (11) has no (nonconstant) T-periodic solution. If $b < a$, Equation (11) has one homoclinic orbit to the saddle point $(\pi, 0)$ and no homoclinic orbit if $b \geqslant a$. Finally, Equation (11) has no heteroclinic orbit and no periodic solution of the second kind. See [6] for more details and corresponding phase portraits.

The differential equation of the *damped pendulum with constant torque*

$$y'' + cy' + a\sin y = b, \tag{13}$$

where $a > 0, b > 0, c > 0$, is also the equation of synchronous electrical motors considered by Tricomi in [125,126]. Because of the absence of a first integral, the discussion of its qualitative behavior in the phase space or phase cylinder is much more delicate, and, in contrast to the undamped case, periodic solutions of the second kind may exist.

Equation (13) has no, one unstable or one stable and one unstable (saddle point) equilibrium, according to $b > a$, $b = a$ or $b < a$. It has no nonconstant T-periodic solution and no homoclinic orbit. For $b > a$, it has no heteroclinic orbit, and one periodic solution of the second kind, with $y'(t) > 0$ for all $t \in R$. For $b \leqslant a$, there exists $c_0 > 0$ such that Equation (13) has no periodic solution of the second kind and two heteroclinic orbits if $c > c_0$, and one periodic solution of the second kind and one heteroclinic orbit if $c \leqslant c_0$. The effective determination of c_0 is an important and delicate problem.

3. Periodic solutions of the forced pendulum

3.1. *The problems*

We now consider the (possibly dissipative) *periodically forced pendulum equation*

$$y'' + cy' + a\sin y = h(t) \quad \bigl(= \bar{h} + \tilde{h}(t)\bigr), \tag{14}$$

where we can assume, without loss of generality, that $c \geqslant 0$, $a > 0$, and where $h = \bar{h} + \tilde{h}$ is T-periodic, for some period $T > 0$, and corresponding frequency $\omega := \frac{2\pi}{T}$. For the simplicity of exposition, we assume that h is continuous. Most results hold under weaker regularity conditions.

We use the following notations:

$$L^p_T = \bigl\{h \in L^p_{\mathrm{loc}}(R)\colon\ h(t+T) = h(t) \text{ for a.e. } t \in R\bigr\},$$

$$H^1([0,T]) = \{h \in AC([0,T]):\ h' \in L^2([0,T])\},$$
$$C_T = \{h \in C(R):\ h(t+T) = h(t) \text{ for all } t \in R\},$$
$$H_T^1 = \{h \in AC_{\mathrm{loc}}(R) \cap C_T:\ h' \in L^2([0,T])\},$$
$$\|h\|_p = \left(\int_0^T |h(t)|^p \,\mathrm{d}t\right)^{1/p}, \qquad \|h\|_\infty = \max_{t\in[0,T]} |h(t)|,$$
$$\|h\|_{H^1([0,T])} = \|h\|_{H_T^1} = \left(\|h\|_2^2 + \|h'\|_2^2\right)^{1/2},$$
$$\bar{h} = \frac{1}{T}\int_0^T h(t)\,\mathrm{d}t, \qquad \tilde{h}(t) = h(t) - \bar{h},$$
$$\widetilde{L}_T^p = \{h \in L_T^p:\ \bar{h} = 0\}, \qquad \widetilde{C}_T = \{h \in C_T:\ \bar{h} = 0\}.$$

Consequently,

$$L_T^p = R \oplus \widetilde{L}_T^p, \qquad C_T = R \oplus \widetilde{C}_T,$$

with the corresponding decomposition $y = \bar{y} + \tilde{y}$.

DEFINITION 1. A *T-periodic solution* of Equation (14) is a solution $y:\mathbb{R} \to \mathbb{R}$ such that $y(t+T) = y(t)$ for all $t \in \mathbb{R}$.

We will sometimes use an interesting *equivalent formulation* of the problem of T-periodic solutions for Equation (14).

LEMMA 1. *If $\widetilde{H}(t)$ denotes the unique T-periodic solution in $\widetilde{C}_T$ of equation*

$$y'' + cy' = \tilde{h}, \tag{15}$$

then $y(t)$ is a T-periodic solution of Equation (14) *if and only if $x(t) = y(t) - \widetilde{H}(t)$ is a T-periodic solution of equation*

$$x'' + cx' + a\sin\left(x + \widetilde{H}(t)\right) = \bar{h}. \tag{16}$$

When h is not constant, Equation (14) has no constant solution, and one expects that the equilibria will by replaced by T-periodic solutions. By integrating Equation (14) over $[0,T]$, we immediately obtain a *necessary condition* for the existence of a T-periodic solution to Equation (14).

PROPOSITION 1. *If Equation* (14) *has a T-periodic solution, then*

$$-a \leqslant \bar{h} \leqslant a. \tag{17}$$

The main questions one can raise about the T-periodic solutions of Equation (14) are the following ones:

1. Determine the *nature* and the *properties* of the set

$$\mathcal{R} = \mathcal{R}(c, a, T) \subset [-a, a] \oplus \widetilde{C}_T$$

of T-periodic forcings h such that Equation (14) has at least one T-periodic solution, i.e., the *range* of the nonlinear operator

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} + c\frac{\mathrm{d}}{\mathrm{d}t} + a\sin(\cdot)$$

over the space of C^2 T-periodic functions.
2. For $h \in \mathcal{R}$, discuss the *multiplicity* of the T-periodic solutions.
3. For $h \in \mathcal{R}$, discuss the *stability* of the T-periodic solutions.
4. Discuss the existence of *subharmonic solutions*, i.e., solutions with minimum period kT for some integer $k \geqslant 2$.

Concerning the multiplicity, it is clear that if y is a T-periodic solution of Equation (14), the same is true for $y + 2k\pi$, $k \in \mathbb{Z}$. Consequently, we say that y_1 and y_2 are *geometrically distinct* T-periodic solutions of (14) if they do not differ by a multiple of 2π.

3.2. *The possibly dissipative case $c \geqslant 0$*

The *Lyapunov–Schmidt's decomposition* (see, e.g., [50]) consists in the following elementary fact.

LEMMA 2. *$y = \bar{y} + \tilde{y}$ is a T-periodic solution of Equation* (14) *if and only if it is a solution of the system*

$$\tilde{y}'' + c\tilde{y}' + a\sin(\bar{y} + \tilde{y}) - \overline{a\sin(\bar{y} + \tilde{y})} = \tilde{h}(t), \tag{18}$$

$$\overline{a\sin(\bar{y} + \tilde{y})} = \bar{h}. \tag{19}$$

Of course, instead of $y = \bar{y} + \tilde{y}$, one can also decompose y as $y = y(\tau) + [y - y(\tau)]$, for some $\tau \in \mathbb{R}$.

In the classical Liapunov–Schmidt's method, the first equation in (18) is solved with respect to $\tilde{y}$ for fixed $\bar{y}$ (using a fixed point or implicit function theorem, or critical point theory) and this solution is introduced in the second equation, which then becomes the (one-dimensional) *bifurcation equation*. One can also study directly the equivalent system (18)–(19) by *degree* or *critical point theory*.

The *method of upper and lower solutions* for the periodic solutions of Equation (14) (see, e.g., [76]) consists in the following statement.

LEMMA 3. *If α and β are of class C^2, T-periodic and such that, for all $t \in R$,*

$$\begin{aligned} &\alpha(t) \leqslant \beta(t), \\ &\alpha''(t) + c\alpha'(t) + a\sin\alpha(t) \geqslant h(t) \geqslant \beta''(t) + c\beta'(t) + a\sin\beta(t), \end{aligned} \tag{20}$$

then (14) *has at least one T-periodic solution y such that $\alpha(t) \leqslant y(t) \leqslant \beta(t)$.*

Furthermore, if the inequalities are strict in (20)*, the coincidence degree of the T-periodic boundary value problem for* (14) *with respect to the open bounded set*

$$\Omega = \{y \in C_T\colon \forall t \in \mathbb{R},\ \alpha(t) < y(t) < \beta(t)\}$$

is defined and equal to one.

The coincidence degree is the Leray–Schauder degree of a suitable associated nonlinear operator in C_T, whose fixed points give the T-periodic solutions of (14). See, e.g., [50,74].

The following results are now classical and can be found in [85,47,76]. Some of them were first proved in [36] and some have been reobtained in [61]. We recall first the main ideas and results of the used methods.

THEOREM 1. *For each $\tilde{h} \in \widetilde{C}_T$, there exists*

$$m_{\tilde{h}} = m_{\tilde{h}}(c,a,T) \leqslant M_{\tilde{h}} = M_{\tilde{h}}(c,a,T)$$

such that the following hold.

1. *Equation* (14) *has at least one T-periodic solution if and only if $\bar{h} \in [m_{\tilde{h}}, M_{\tilde{h}}]$.*
2. $-a \leqslant m_{\tilde{h}} \leqslant M_{\tilde{h}} \leqslant a$, $m_0 = -a$, $M_0 = a$.
3. $\mathcal{R}(c,a,T) = \bigcup_{\tilde{h} \in \widetilde{C}_T} [m_{\tilde{h}}, M_{\tilde{h}}] \times \{\tilde{h}\} \subset [-a,a] \oplus \widetilde{C}_T$ *is closed.*
4. *Equation* (14) *has at least two distinct T-periodic solutions if $\bar{h} \in\,]m_{\tilde{h}}, M_{\tilde{h}}[$.*

To prove this theorem, one first uses the Lyapunov–Schmidt decomposition. Leray–Schauder's fixed point theorem applied to an integral formulation of (18) in the space $\widetilde{C}_T$ with parameter $\bar{y} \in S^1$ implies the existence of a connected closed set $\mathcal{C}(c,a,T,\tilde{h}) \subset S^1 \times \widetilde{C}_T$ of solutions $(\bar{y},\tilde{y})$ of (18), whose projection on S^1 is equal to S^1. Hence, Equation (14) has at least one T-periodic solution if $\bar{h}$ belongs to the nonempty set

$$\mathcal{I}(c,a,T,\tilde{h}) = \left\{\overline{a\sin(\bar{y}+\tilde{y})}\colon (\bar{y},\tilde{y}) \in \mathcal{C}(c,a,T,\tilde{h})\right\}.$$

The fact that $\mathcal{I}(c,a,T,\tilde{h})$ is an interval follows from the fact that if $\bar{h}_1 < \bar{h}_2$ belong to $\mathcal{I}(c,a,T,\tilde{h})$, and if $\bar{h} \in\,]\bar{h}_1,\bar{h}_2[$, then a corresponding solution $y_i = \bar{y}_i + \tilde{y}_i$ of equation

$$y'' + cy' + a\sin y = \bar{h}_i + \tilde{h} \quad (i = 1,2)$$

satisfies the equations

$$\tilde{y}_i'' + c\tilde{y}_i' + a\sin(\bar{y}_i + \tilde{y}_i) - \overline{a\sin(\bar{y}_i+\tilde{y}_i)} = \tilde{h}(t),$$
$$\overline{a\sin(\bar{y}_i+\tilde{y}_i)} = \bar{h}_i \quad (i = 1,2),$$

and hence $(\bar{y}_i,\tilde{y}_i) \in \mathcal{C}(c,a,T,\tilde{h})$ $(i = 1,2)$ and are such that

$$\overline{a\sin(\bar{y}_1+\tilde{y}_1)} < \bar{h} < \overline{a\sin(\bar{y}_2+\tilde{y}_2)}.$$

As the real map

$$(\bar{y}, \tilde{y}) \mapsto \overline{a \sin(\bar{y} + \tilde{y})} \tag{21}$$

is continuous, there exists, by connexity, some $(\bar{y}, \tilde{y}) \in \mathcal{C}(c, a, T, \tilde{h})$ such that

$$\overline{a \sin(\bar{y} + \tilde{y})} = \bar{h},$$

and $y = \bar{y} + \tilde{y}$ is a T-periodic solution of Equation (14). We set

$$m_{\tilde{h}} = \inf \mathcal{I}(c, a, T, \tilde{h}), \qquad M_{\tilde{h}} = \sup \mathcal{I}(c, a, T, \tilde{h}).$$

It follows from (17) that $-a \leqslant m_{\tilde{h}} \leqslant M_{\tilde{h}} \leqslant a$, and it is trivial that $m_0 = -a$ and $M_0 = a$.

To prove that $\mathcal{R}(c, a, T)$ is closed, one considers a sequence (h_i) in $\mathcal{R}(c, a, T)$ converging uniformly to $h \in C_T$, and $(\bar{y}_i, \tilde{y}_i)$ a corresponding sequence in $S^1 \times \widetilde{C}_T$ of T-periodic solutions. It is easy to show from (18) that $(\tilde{y}_i)$ is bounded in C^2_T, and Ascoli–Arzelá theorem implies that, up to a subsequence, $(\bar{y}_i + \tilde{y}_i)$ converges to some T-periodic solution y of Equation (14). This implies that $\mathcal{I}(c, a, T, \tilde{h})$ is closed.

To prove the multiplicity result when $m_{\tilde{h}} < \bar{h} < M_{\tilde{h}}$, we observe that in this case, if y_m (respectively y_M) is a T-periodic solution of Equation (14) with $h = m_{\tilde{h}} + \tilde{h}$ (respectively $h = M_{\tilde{h}} + \tilde{h}$), then y_m (respectively y_M) is an strict upper (respectively lower) solution for Equation (14) with periodic boundary conditions. As Equation (14) is invariant under the substitution $y \to y + 2k\pi$ $(k \in \mathbb{Z})$, we can assume without loss of generality that $y_M(t) \leqslant y_m(t)$ for all $t \in \mathbb{R}$ and that $y_m(\tau) - y_M(\tau) < 2\pi$ for some $\tau \in \mathbb{R}$. An easy maximum principle type reasoning then shows that $y_M(t) < y_m(t)$ for all $t \in \mathbb{R}$. Hence Lemma 3 implies that (14) has at least one T-periodic solution $\hat{y}_1$ such that

$$y_M(t) < \hat{y}_1(t) < y_m(t) \tag{22}$$

for all $t \in \mathbb{R}$, and such that the associated coincidence degree with respect to the open set

$$\Omega_1 = \left\{ y \in C_T \colon \forall t \in \mathbb{R},\ y_M(t) < y(t) < y_m(t) \right\},$$

is equal to one. Now, $(y_M + 2\pi, y_m + 2\pi)$ and $(y_M, y_m + 2\pi)$ are also two strictly ordered couples of lower and upper solution for Equation (14) with T-periodic boundary conditions. Consequently, the coincidence degree respectively associated to the T-periodic boundary value problem for Equation (14) with respect to the open sets

$$\Omega_2 = \left\{ y \in C_T \colon \forall t \in \mathbb{R},\ y_M(t) + 2\pi < y(t) < y_m(t) + 2\pi \right\}$$

and

$$\Omega_3 = \left\{ y \in C_T \colon \forall t \in \mathbb{R},\ y_M < y(t) < y_m + 2\pi \right\}$$

are equal to one. As

$$\Omega_1 \subset \Omega_3, \qquad \Omega_2 \subset \Omega_3, \qquad \overline{\Omega}_1 \cap \overline{\Omega}_2 = \emptyset,$$

it follows from its additivity property that the coincidence degree associated to the T-periodic boundary value problem for Equation (14) with respect to the open set $\Omega_3 \setminus (\overline{\Omega}_1 \cup \overline{\Omega}_2)$ is equal to -1, and Equation (14) has at least one T-periodic solution $\hat{y}_2 \in \Omega_3 \setminus (\overline{\Omega}_1 \cup \overline{\Omega}_2)$. By construction, $\hat{y}_1$ and $\hat{y}_2$ do not differ by a multiple of 2π.

3.3. *The conservative case* $c = 0$

In the conservative case

$$y'' + a \sin y = h(t) \quad \bigl(= \bar{h} + \tilde{h}(t)\bigr), \tag{23}$$

the *direct method of the calculus of variations* allows to prove that $0 \in [m_{\tilde{h}}, M_{\tilde{h}}]$. The starting point of this approach or of any application of *critical point theory* to the periodic solutions of the forced pendulum equation without dissipation is the following classical observation.

LEMMA 4. *y is a T-periodic solution of Equation* (23) *if and only if y is a critical point of the* action functional

$$A_h : H_T^1 \to R, \qquad y \mapsto \int_0^T \left(\frac{y'^2(t)}{2} + a \cos y(t) + h(t) y(t) \right) \mathrm{d}t. \tag{24}$$

If convenient, we also write $A_h(\bar{y}, \tilde{y})$ instead of $A_h(y)$. We recall a few definitions and results. Let E be a Banach space and $\varphi \in C^1(E, \mathbb{R})$. We denote its Fréchet differential at y by $\varphi'(y)$, and its value at $v \in E$ by $\langle \varphi'(y), v \rangle$.

DEFINITION 2. $y \in E$ is a *critical point* of φ if $\varphi'(y) = 0$. $c \in \mathbb{R}$ is a *critical value* of φ if $c = \varphi(y)$ for some critical point y of φ.

We recall a classical and easy to prove sufficient condition for the existence of a minimum to φ.

PROPOSITION 2. *Let E be a reflexive Banach space and $\varphi \in C^1(E, \mathbb{R})$ be bounded from below and sequentially weakly lower semi-continuous. Then φ has a minimum on E.*

A slight extension of Hamel's result [55,129,36] follows from Proposition 2.

THEOREM 2. *For each $\tilde{h} \in \widetilde{C}_T$, Equation* (23) *has at least one T-periodic solution which minimizes A_h over H_T^1. In other words, $0 \in [m_{\tilde{h}}, M_{\tilde{h}}]$.*

First observe that if $\bar{h} = 0$, A_h is 2π-periodic and, using the Sobolev inequality for $y = \bar{y} + \tilde{y} \in H^1_T$,

$$\max_{[0,T]} |\tilde{y}| \leqslant \frac{T^{1/2}}{2\sqrt{3}} \|y'\|_{L^2}, \tag{25}$$

we get

$$\begin{aligned} \varphi(y) &= \int_I \left[\frac{y'^2(t)}{2} + a \cos y(t) + \tilde{h}(t)\tilde{y}(t) \right] \mathrm{d}t, \\ &\geqslant \frac{1}{2} \int_I \left| y'(t) \right|^2 \mathrm{d}t - aT - \left(\int_I \left| \tilde{h}(t) \right| \mathrm{d}t \right) \max_I |\tilde{y}| \\ &\geqslant \frac{1}{2} \int_I \left| y'(t) \right|^2 \mathrm{d}t - aT - \left(\int_0^T \left| \tilde{h}(t) \right| \mathrm{d}t \right) \frac{T^{1/2}}{2\sqrt{3}} \left(\int_I \left| y'(t) \right|^2 \mathrm{d}t \right)^{1/2}. \end{aligned} \tag{26}$$

This shows that A_h is bounded from below. From the sequential weak lower semicontinuity of the map $y \mapsto \int_0^T |y'(t)|^2 \, \mathrm{d}t$, and the compact embedding of H^1_T into C_T, it is easy to show that A_h is sequentially weakly lower semi-continuous.

The existence of a *second solution* for the conservative forced pendulum when $\bar{h} = 0$ can be proved by a *generalized mountain pass lemma*, which extends both the geometrical and the compactness assumptions of the classical *Ambrosetti–Rabinowitz mountain pass lemma* [4]. This abstract extension, due to Ghoussoub–Preiss [51] and Yihong Du [42], was motivated by the first proof of a second T-periodic solution to Equation (23) given in [85], as it was the case for the slightly less general version given earlier by Pucci–Serrin [113]. We first introduce two definitions.

DEFINITION 3. A *Palais–Smale sequence at level $c \in \mathbb{R}$* for $\varphi \in C^1(E, \mathbb{R})$ is a sequence (y_n) in E such that

$$\varphi(y_n) \to c, \quad \varphi'(y_n) \to 0, \quad \text{if } n \to \infty.$$

The function φ satisfies the $(\mathrm{PS})_c$-*condition* if the existence of a Palais–Smale sequence at level c for φ implies that c is a critical value for φ.

DEFINITION 4. The function $\varphi \in C^1(E, \mathbb{R})$ satisfies the *Palais–Smale condition* (PS) (respectively the *bounded Palais–Smale condition* (BPS)) if any sequence (respectively bounded sequence) (y_n) such that $(\varphi(y_n))$ is bounded and $\varphi'(y_n) \to 0$ has a convergent subsequence.

We now state the *generalized mountain pass lemma*.

LEMMA 5. *Let $d, e \in E$, $0 < r < \|e - d\|$, be such that*

$$a := \max\bigl\{\varphi(d), \varphi(e)\bigr\} \leqslant b := \inf_{r \leqslant \|y - d\| \leqslant R} \varphi(y). \tag{27}$$

Let

$$\Gamma = \{\gamma \in C([0,1], E) \colon \gamma(0) = d,\ \gamma(1) = e\},$$
$$c := \inf_{\gamma \in \Gamma} \sup_{s \in [0,1]} \varphi(\gamma(s)).$$

Then $c \geqslant b$, *and if* φ *satisfies the* $(\mathrm{PS})_c$ *and* (BPS) *conditions,* c *is a critical value for* φ. *Moreover, if* $c = a$, *there is a critical point* z *such that* $\varphi(z) = c$ *and* $\|z - d\| = r$.

The original mountain pass lemma requires that $a < b$ and that the classical Palais–Smale compactness condition holds.

To apply this Lemma 5 to A_h when $h \in \widetilde{C}_T$, the most technical part is to check that A_h verifies the Palais–Smale-type conditions.

LEMMA 6. *If* $h \in \widetilde{C}_T$, A_h *satisfies the* (BPS)*-condition and the* $(\mathrm{PS})_c$*-condition for each* $c \in \mathbb{R}$.

To show the (BPS)-condition, assume that $(A_h'(y_n))$ converges to 0 and (y_n) is bounded in H_T^1. Then, up to a subsequence, (y_n) converges weakly in H_T^1 and uniformly on I to some $y \in H_T^1$. Consequently

$$\langle A_h'(y_n) - A_h'(y), y - y_n \rangle \to 0,$$

as $n \to \infty$. But

$$\begin{aligned}&\langle A_h'(y_n) - A_h'(y), y - y_n \rangle \\ &\quad = \int_I |y_n'(t) - y'(t)|^2 \,\mathrm{d}t - a \int_I [\sin y_n(t) - \sin y(t)][y_n(t) - y(t)] \,\mathrm{d}t,\end{aligned}$$

and

$$\int_I [\sin y_n(t) - \sin y(t)][y_n(t) - y(t)] \,\mathrm{d}t \to 0$$

as $n \to \infty$. Consequently,

$$\int_I |y_n'(t) - y'(t)|^2 \,\mathrm{d}t \to 0,$$

as $n \to \infty$, and hence $y_n \to y$ in H_T^1. Thus A_h satisfies the (BPS)-condition. Finally, let (y_n) be such that $A_h(y_n) \to c$ and $A_h'(y_n) \to 0$ as $n \to \infty$. Thus $(A_h(y_n))$ is bounded and, using (26) we see that $(\|y_n'\|_{L^2})$ is bounded. Now, by 2π-periodicity of A_h, there exists $z_n \in [0, 2\pi[$ such that $z_n = \bar{y}_n \pmod{2\pi}$. Letting $w_n(t) = z_n + \tilde{y}_n(t)$, we have $A_h(y_n) = A_h(w_n)$, $A_h'(y_n) = A_h'(w_n)$ and (w_n) is bounded in H_T^1. By the reasoning for

(BPS) applied to this sequence, we see that, up to a subsequence, $w_n \to w \in H_T^1$ and this easily implies that c is a critical value of A_h.

We can now prove the existence of a second T-periodic solution.

THEOREM 3. *For each $h \in \widetilde{C}_T$, Equation* (23) *has at least two geometrically distinct T-periodic solutions.*

A first solution y^* has been obtained which minimizes A_h over H_T^1. To obtain a second one, it remains to check the geometry of the generalized mountain pass lemma. As

$$A_h\big(y^*\big) = A_h\big(y^* + 2\pi\big) = \min_{H_T^1} A_h,$$

let us take $0 < r < 2\pi$, $d = y^*, e = y^* + 2\pi$. Then, all conditions of Proposition 5 are satisfied. If the critical value $c > \min_{H_T^1} A_h$, then the corresponding critical point is geometrically distinct from y^*. If $c = \min_{H_T^1} A_h$, then A_h has a critical point z such that $0 < \|z - y^*\| = r < 2\pi$, so that y^* and z are two geometrically distinct T-periodic solutions of Equation (23).

The periodicity property of A_h when $\bar{h} = 0$ allows also the use of a *Lusternik–Schnirelman type* argument to prove directly that A_h has two distinct critical points (see [78,34,114]). Let G be a discrete subgroup of a Banach space E and $\pi : E \to E/G$ be the canonical surjection.

DEFINITION 5. $A \subset X$ is called *G-invariant* if $\pi^{-1}(\pi(A)) = A$. $\varphi : E \to \mathbb{R}$ is called *G-invariant* if $\varphi(y + g) = \varphi(y)$ for every $y \in E$ and every $g \in G$.

When φ is differentiable, the same G-invariance holds for φ', so that if y is a critical point of φ, then $\pi^{-1}(\pi(y))$ is a set of critical point of φ, called a *critical orbit* of φ. A suitable Palais–Smale condition for G-invariant functions is the following one.

DEFINITION 6. A G-invariant differentiable functional $\varphi : E \to \mathbb{R}$ satisfies the (PS)$_G$-*condition* if, for every sequence (y_k) in E such that $(\varphi(y_k))$ is bounded and $\varphi'(y_k) \to 0$, the sequence $(\pi(y_k))$ contains a convergent subsequence.

The following multiplicity theorem holds for G-invariant functionals [114]. Its proof is based upon Lusternik–Schnirelman category [86].

PROPOSITION 3. *Let $\varphi \in C^1(E, \mathbb{R})$ be a G-invariant functional satisfying the* (PS)$_G$-*condition. If φ is bounded from below and G generates a subspace of finite dimension N, then φ has at least $N + 1$ critical orbits.*

If we consider A_h with $h \in \widetilde{C}_T$, so that $A_h(y + 2\pi) = A_h(y)$ for all $y \in H_T^1$, we can take the discrete subgroup $G = \{2k\pi\colon\ k \in \mathbb{Z}\}$ of H_T^1. If $y = \bar{y} + \tilde{y} \in H_T^1$, there exists a unique $k \in \mathbb{Z}$ such that $\bar{y}_0 := \bar{y} - 2k\pi \in [0, 2\pi[$, and $\bar{y}_0 + \tilde{y}$ is a representative of $[y] \in H_T^1/G$.

So the (PS)$_G$-condition for A_h essentially reduces to the (BPS)-condition checked above. Proposition 3 gives another proof of Theorem 3.

In [119], Serra and Tarallo have introduced a new *reduction method of Liapunov–Schmidt's type*, which provides equivalent formulations for some of problems for the forced conservative pendulum equation. One first observes that, for each $\xi \in \mathbb{R}$, the functional A_h is bounded below on the hyperplane $H^1_{T,\xi} = \{y \in H^1_T\colon\ \bar{y} = \xi\}$. Consequently, by the argument of Theorem 2, it reaches its minimum on $H^1_{T,\xi}$ and we let

$$\varphi_h(\xi) := \min_{\bar{y}=\xi} A_h(y) = \min_{\tilde{y}\in\widetilde{H}^1_T} A_h(\xi, \tilde{y}), \tag{28}$$

defining in this way the real function $\varphi_h : \mathbb{R} \to \mathbb{R}$, $\xi \mapsto \varphi_h(\xi)$. Define also

$$\begin{aligned} M_h(\xi) &= \bigl\{y \in H^1_T\colon\ \bar{y} = \xi,\ A_h(y) = \varphi_h(\bar{y})\bigr\} \\ &= \Bigl\{y \in H^1_T\colon\ \bar{y} = \xi,\ A_h(\xi, \tilde{y}) = \min_{\tilde{z}\in\widetilde{H}^1_T} A_h(\xi, \tilde{z})\Bigr\} \end{aligned}$$

(the set of y in the hyperplane $\bar{y} = \xi$ where A_h reaches its minimum on this hyperplane),

$$\begin{aligned} M_h &= \bigcup_{\xi\in R} M_h(\xi) = \bigl\{y \in H^1_T\colon\ A_h(y) = \varphi_h(\bar{y})\bigr\} \\ &= \Bigl\{y \in H^1_T\colon\ A_h(\bar{y}, \tilde{y}) = \min_{\tilde{z}\in\widetilde{H}^1_T} A_h(\bar{y}, \tilde{z})\Bigr\}. \end{aligned}$$

THEOREM 4. *The following properties hold for φ_h.*

1. *φ_h is locally Lipschitz continuous on R.*
2. *$M_h(\xi) \neq \emptyset$ and compact for each $\xi \in \mathbb{R}$, M_h is weakly closed and $M_h : \mathbb{R} \to 2^{H^1_T}$ is upper semi-continuous.*
3. *If $y \in M_h$ and $\bar{y}$ gives a local minimum to φ_h, y gives a local minimum to A_h.*
4. *φ_h is differentiable at ξ if and only if $y \mapsto \int_0^T (a \sin y(t) - h(t))\,\mathrm{d}t$ is constant on $M_h(\xi)$.*
5. *If φ_h has a critical point, A_h has a critical point.*
6. *If φ_h is not strictly monotone, A_h has a critical point.*

Properties 1, 2 and 3 are easy to prove, and Property 5 is a consequence of Property 4. To prove this last one, define

$$\lambda(\xi, \tilde{y}) = \frac{\partial}{\partial \xi} A_h(\xi, \tilde{y}) = \int_0^T \bigl\{a \cos\bigl[\xi + \tilde{y}(s)\bigr] + h(s)\bigr\}\,\mathrm{d}s,$$

and show, using the following inequalities, valid when $A_h(\xi, \tilde{y}) = \varphi_h(\xi)$,

$$\varphi_h(\xi + \delta) - \varphi_h(\xi) \leqslant A_h(\xi + \delta, \tilde{y}) - A_h(\xi, \tilde{y})$$

and the mean value theorem, that if φ_h is differentiable at ξ, then

$$\varphi_h'(\xi) \leqslant \lambda(\xi + \tilde{y}) \leqslant \varphi_h'(\xi).$$

The proof of the converse is a little more elaborate and uses the compactness property in Property 2 and the uniform coercivity of $A_h(\xi, \cdot)$ on bounded sets. Finally, the proof of Property 6 uses, at a local maximum of φ_h, an argument on upper and lower solutions similar to the one used in Theorem 1.

It is interesting to compare this approach to the classical method of Lyapunov–Schmidt. In this case, one proves that, for each $\xi \in \mathbb{R}$, the set

$$K_h(\xi) = \{y \in C_T\colon\ \bar{y} = \xi,\ \tilde{y} \text{ solves Equation (18)}\}$$

is not empty, and then the problem is reduced to find the elements of the set $K_h = \bigcup_{\xi \in \mathbb{R}} K_h(\xi)$ such that

$$\overline{a \sin(\bar{y} + \tilde{y})} = \bar{h}.$$

In the Serra–Tarallo's approach, on each *slice* $\xi + \widetilde{H}_T^1$ of H_T^1, one considers only the elements of $K_h(\xi)$ which minimize the restriction of A_h on this slice, which provides the subset $M_h(\xi) \subset K_h(\xi)$, and then, instead of trying to solve Equation (19) on this set, one concentrates on the reduced functional φ_h and relates its critical points to those of A_h. Hence the spirit is more variational than in earlier approaches combining a Lyapunov–Schmidt argument with some variational method (like Castro–Lazer's one used in [32]), in that the emphasis, at each step, remains on the functional instead of on its gradient. One of the main features of this approach is that, in contrast to most other ones, it still works when $a \sin y$ is replaced by a more general *almost periodic* function.

3.4. *The dissipative case $c > 0$*

We just have shown that $0 \in [m_{\tilde{h}}, M_{\tilde{h}}]$ when $c = 0$. A natural question is to find an explicit element in $[m_{\tilde{h}}, M_{\tilde{h}}]$ when $c > 0$. The following result, proved in [85] by topological degree arguments, shows that $m_{\tilde{h}} < M_{\tilde{h}}$ and $0 \in]m_{\tilde{h}}, M_{\tilde{h}}[$ when c is sufficiently large.

THEOREM 5. *If $\frac{c}{T} > \frac{1}{\pi\sqrt{3}}\|\tilde{h}\|_2$, then $m_{\tilde{h}} < 0 < M_{\tilde{h}}$.*

The question was then raised to know if $0 \in [m_{\tilde{h}}, M_{\tilde{h}}]$ for each $c > 0$. A *negative answer* was first given by a counterexample of Ortega [99], and this result was improved by Alonso [2] in the following form.

THEOREM 6. *For each $c > 0$, there exists $\tilde{h} \in \widetilde{C}_T$ and $T_0 = T_0(a, c)$ such that for each $T > T_0$, $0 \notin [m_{\tilde{h}}, M_{\tilde{h}}]$.*

The idea of Alonso's result consists in constructing a forcing term close to a piecewise constant function $h(t)$ taking a large positive value p in the interval $[0, \tau]$ and a small negative value $-q$ in the interval $[\tau, T]$, where $p\tau - q(T - \tau) = 0$.

Further information has been given by Ortega, Serra and Tarallo [107] who have proved the following interesting result. Let $\widetilde{H}$ be defined in Lemma 1 and $\mathcal{B}_T \in L^2_T$ be the T-periodic function defined by

$$\mathcal{B}_T(t) = 2\pi\left(\frac{t}{T} - \left[\frac{t}{T}\right]\right) - \pi. \tag{29}$$

THEOREM 7. *Given $a > 0$, $c > 0$ and $T > 0$, there exists $\varepsilon > 0$ such that Equation* (14) *has no T-periodic solution if $h \in \widetilde{L}^1_T$ and*

$$\|\widetilde{H} - \mathcal{B}_T\|_2 < \varepsilon.$$

Moreover, ε can be explicitly computed in terms of a, c and T.

To prove this result, the authors first consider the equivalent formulation (16) of Equation (14) which, for $\widetilde{H} = \mathcal{B}_T$ is itself equivalent to equation

$$y'' + cy' - a\sin(y + \omega t) = 0. \tag{30}$$

Multiplying both members of Equation (30) by $y' + \omega$ and integrating over $[0, T]$ shows immediately that Equation (30) has no T-periodic solution. This is in contrast with the case where $c = 0$, for which Bates [9] has shown the existence of a continuum of T-periodic solutions (see also [86]). The following lemma provides estimates for the possible T-periodic solutions of equation

$$y'' + cy' - a\sin\bigl(y + \omega t + P(t)\bigr) = 0, \tag{31}$$

where $P \in L^2_T$.

LEMMA 7. *Let y be a possible T-periodic solution of Equation* (31). *Then*

$$\|y'\|_2^2 \leqslant \varphi\left(\frac{a}{c}\|P\|_2\right),$$

where

$$\varphi(s) = \frac{1}{2}\left(s^2 + \sqrt{s^4 + \frac{16\pi^2}{T}s^2}\right),$$

and

$$\|y'\|_2^2 \geqslant \frac{Ta^2}{2(\omega^2 + c^2)} - 2\frac{\sqrt{T}a^2}{\omega^2 + c^2}\left(\frac{1}{\omega^2}\varphi\left(\frac{a}{c}\|P\|_2\right) + \|P\|_2^2\right)^{1/2}.$$

The first inequality is obtained by subtracting and adding $a\sin(y+\omega t)$ to Equation (31), multiplying both members by $y'+\omega$, integrating over $[0,T]$ and using T-periodicity and Cauchy–Schwarz inequality. For the second inequality, one first observes that the right-hand side of the equivalent form of Equation (31)

$$y''+cy'-a\sin(\bar{y}+\omega t)=-a\sin(\bar{y}+\omega t)+a\sin\bigl(y+\omega t+P(t)\bigr):=f(t)$$

has mean value zero if y solves (31) and we call $\widetilde{F}$ the solution of Equation (15) with right-hand side $f(t)$. Setting $z=y-\widetilde{F}$, we obtain the equivalent equation

$$z''+cz'-a\sin(\bar{y}+\omega t)=0,$$

with solution

$$z(t)=-\frac{a}{\omega^2+c^2}\sin(\bar{y}+\omega t)-\frac{ac}{\omega(\omega^2+c^2)}\cos(\bar{y}+\omega t)$$

such that

$$\|z'\|_2^2=\frac{Ta^2}{2(\omega^2+c^2)}.$$

Then

$$\|y'\|_2^2\geqslant\frac{Ta^2}{2(\omega^2+c^2)}-\sqrt{\frac{2Ta^2}{\omega^2+c^2}}\|F'\|_2\geqslant\frac{Ta^2}{2(\omega^2+c^2)}-\sqrt{\frac{2Ta^2}{\omega^2+c^2}}\frac{\|f\|_2}{\omega^2+c^2}.$$

The result then follows from

$$\bigl|f(t)\bigr|\leqslant a\bigl(\bigl|\tilde{y}(t)\bigr|+\bigl|P(t)\bigr|\bigr)$$

and the use of Wirtinger's inequality.

To deduce Theorem 7 from Lemma 7, it suffices to observe that if

$$\begin{aligned}&\frac{Ta^2}{2(\omega^2+c^2)}\varphi\Bigl(\frac{a}{c}\|P\|_2\Bigr)+2\frac{\sqrt{T}a^2}{\omega^2+c^2}\Bigl(\frac{1}{\omega^2}\varphi\Bigl(\frac{a}{c}\|P\|_2\Bigr)+\|P\|_2^2\Bigr)^{1/2}\\&\quad<\frac{Ta^2}{2(\omega^2+c^2)},\end{aligned}$$

then Equation (31) has no T-periodic solution, and, since $\varphi(s)\to 0$ as $s\to 0$, this happens when $\|P\|_2$ is sufficiently small.

A consequence of Theorem 7 is that, under its conditions, every solution y of Equation (14) is such that $\lim_{t\to+\infty}|y(t)|=\infty$. This follows from the fact that the *Poincaré mapping*

$$\mathcal{P}:\bigl(y(0),y'(0)\bigr)\to\bigl(y(T),y'(T)\bigr) \tag{32}$$

associated to an arbitrary solution y of Equation (14) is an orientation preserving homeomorphism of the plane, whose fixed points are the T-periodic solutions of Equation (14). As $\mathcal{P}$ has no fixed point under the conditions of Theorem 7, the theory of Brouwer for fixed point free homeomorphisms [53] is applicable, and the orbits of $\mathcal{P}$ must go to infinity. The result then follows from the easily verified fact that the derivatives of the solutions of Equation (14) are bounded.

3.5. *The degeneracy problem*

The *degeneracy problem* for the T-periodic solutions of Equation (14), which is still open, consists in proving or disproving the existence of some h such that $m_{\tilde{h}} = M_{\tilde{h}}$. Here is the known partial information.

We start with some results valid for $c \geqslant 0$.

THEOREM 8. *For the T-periodic problem for Equation* (14)*, the set* $\{\tilde{h} \in \widetilde{C}_T\colon\ m_{\tilde{h}} < M_{\tilde{h}}\}$ *is open and dense.*

This has been proved using various arguments [85,76,68], and in particular a *generalized Sard–Smale's theorem.* Thus, *generically,* $[m_{\tilde{h}}, M_{\tilde{h}}]$ is a nondegenerate interval.

We now describe some contributions of Ortega and Tarallo [108], which generalize in various directions earlier results of Donati [41] and of Serra, Tarallo and Terracini [121].

DEFINITION 7. Equation (14) is said to be *degenerate* if the set of $\bar{h}$ such that Equation (14) has a T-periodic solution is a singleton $\{\bar{h}_{\tilde{h}}\}$, i.e., if $m_{\tilde{h}} = M_{\tilde{h}}$.

In the (excluded) case where $a = 0$, Equation (14) is degenerate. It only admits T-periodic solutions when $\bar{h} = 0$, in which case it has an unbounded path of T-periodic solutions. This fact will be extended to the case where $a > 0$. Let $\mathcal{T}$ be the set of T-periodic solutions of Equation (14). Using the Lyapunov–Schmidt decomposition and the Leray–Schauder argument of the proof of Theorem 1, we obtain the existence of a connected and closed subset $\mathcal{C} \subset \mathcal{T}$ such that $\{y(0)\colon\ y \in \mathcal{C}\} = \mathbb{R}$. By the same argument, one indeed proves that given a closed interval $I \subset \mathbb{R}$ and $\tau \in \mathbb{R}$, a closed connected set $\mathcal{C}_{\tau,I} \subset \mathcal{T}$ exists such that $\{y(\tau)\colon\ y \in \mathcal{C}_{\tau,I}\} = I$.

THEOREM 9. *The following statements for Equation* (14) *are equivalent.*

(i) *The equation is degenerate.*

(ii) *For any $\xi \in \mathbb{R}$, there exists a unique $y_\xi \in \mathcal{T}$ which satisfies $y_\xi(0) = \xi$.*

(iii) *There exists a continuous path $\xi \mapsto y_\xi$ in $\mathcal{T}$ which satisfies*

$$\lim_{\xi \to \pm\infty} y_\xi(t) = \pm\infty$$

uniformly in $t \in \mathbb{R}$.

Moreover, if one of those conditions holds, the map $\xi \mapsto y_\xi$ is continuous, monotone ($y_\xi(t) < y_\eta(t)$ for all $t \in \mathbb{R}$ when $\xi < \eta$), and such that $y_{\xi+2\pi} = y_\xi + 2\pi$.

To show that condition (i) implies condition (ii), one first observes that if two elements $y, z \in \mathcal{T}$ intersect at some t_0 then they coincide or intersect transversally ($y'(t_0) \neq z'(t_0)$). This implies the following useful intermediate proposition.

PROPOSITION 4. *If $\mathcal{A} \subset \mathcal{T}$ be connected and $z \in \mathcal{T} \setminus \mathcal{A}$, and if, for some $y^* \in \mathcal{A}$, z and y^* do not intersect, then z does not intersect any $y \in \mathcal{A}$.*

Indeed, $\{y \in \mathcal{A}\colon\ y \text{ does not intersect } v\}$ is trivially open, and $\{y \in \mathcal{A}\colon\ y \text{ intersects } v\}$ is open by the transversality property.

Coming back to the proof of Theorem 9, one then shows that, for any ξ and τ, letting $\mathcal{C}_{\tau,\xi} = \mathcal{C}_{\tau,]-\infty,\xi]}$, one has $\mathcal{C}_{\tau,\xi} = \mathcal{T}_{\tau,\xi} := \{y \in \mathcal{T}\colon\ y(\tau) \leqslant \xi\}$. This depends in particular on the property

$$\|y\|_\infty \leqslant \left|y(\tau)\right| + C, \tag{33}$$

satisfied by any T-periodic solution of Equation (14). One next proves using again Proposition 4 that given $y, z \in \mathcal{T}$ and $\tau \in \mathbb{R}$ such that $y(\tau) < z(\tau)$, one has $y(t) < z(t)$ for all $t \in \mathbb{R}$. This follows from the fact that, for $\xi = y(\tau)$, $z \notin \mathcal{T}_{\tau,\xi}$ and cannot intersect the very large functions in $\mathcal{T}_{\tau,\xi}$. Then the existence of y_ξ follows from the existence of $\mathcal{C}$ and its uniqueness from the above monotonicity property.

To show that condition (ii) implies condition (iii), one uses (33), the uniqueness property and the local compactness of $\mathcal{T}$. By uniqueness also $y_{\xi+2\pi} = y_\xi + 2\pi$ and this implies the limiting properties of y_ξ for $\xi \to \pm\infty$. Finally, to prove that (iii) implies (i), assume that y is a T-periodic solution of Equation (14) and define

$$A = \bigl\{\xi \in \mathbb{R}\colon\ \forall t \in \mathbb{R},\ y_\xi(t) \leqslant y(t)\bigr\}, \qquad B = \bigl\{\xi \in \mathbb{R}\colon\ \forall t \in \mathbb{R},\ y_\xi \geqslant y(t)\bigr\},$$

and $\xi_A = \max A$, $\xi_B = \min B$. One has $y_{\xi_A} \leqslant y$, $y_{\xi_B} \geqslant y$ and they must touch somewhere, namely at t_A and t_B, so that

$$y'_{\xi_A}(t_A) = y'(t_A), \qquad y''_{\xi_A}(t_A) \leqslant y''(t_A),$$

and the reverse inequality for y_{ξ_B} at t_B. Then,

$$\begin{aligned}
\bar{h}_{\tilde{h}} &= y''_{\xi_A}(t_A) + c y'_{\xi_A}(t_A) + a \sin y_{\xi_A}(t_A) - \tilde{h}(t_A) \\
&\leqslant y''(t_A) + c y'(t_A) + a \sin y(t_A) - \tilde{h}(t_A) = \bar{h} \\
&= y''_{\xi_B}(t_B) + c y'_{\xi_B}(t_B) + a \sin y_{\xi_B}(t_B) - \tilde{h}(t_B) \\
&\leqslant y''_{\xi_B}(t_B) + c y'_{\xi_B}(t_B) + a \sin y_{\xi_B}(t_B) - \tilde{h}(t_B) = \bar{h}_{\tilde{h}},
\end{aligned}$$

so that $\bar{h} = \bar{h}_{\tilde{h}}$.

Consequences of Theorem 9 are that the solutions bounded over $\mathbb{R}$ of a degenerate equation are either T-periodic solutions, or heteroclinic connections between different T-periodic solutions, and some information about the stability of the T-periodic solutions.

We now describe some results valid in the conservative case $c = 0$, with, first, an improvement of Theorem 8 proved in [68] using a generalized Sard–Smale lemma.

DEFINITION 8. A *regular value* for a mapping f of class C^1 between two smooth Banach manifolds is the image by f of a point c such that f'_c is onto.

THEOREM 10. *The set $\mathcal{G}$ of regular values for $y \to y'' + a \sin y$ on C_T^2 is open and dense in $\widetilde{C}_T$, and, for every $g \in \mathcal{G}$, there exists $\varepsilon > 0$ such that, if $\|h - g\|_\infty \leqslant \varepsilon$, Equation* (23) *has a T-periodic solution.*

Serra, Tarallo and Terracini [121] have found another characterization of degeneracy in the conservative case. If $h \in \widetilde{C}_T$, let

$$c_0 = \min_{H_T^1} A_h, \tag{34}$$

and for $\xi \in \mathbb{R}$, let

$$\mathcal{K}_{c_0}(\xi) := \{ y \in H_T^1 \colon A_h(y) = c_0,\ \bar{y} = \xi \}.$$

PROPOSITION 5. *If $\mathcal{K}_{c_0}(\xi) \neq \emptyset$ for each $\xi \in \mathbb{R}$, then*

(i) *For each $\xi \in \mathbb{R}$, $\mathcal{K}_{c_0}(\xi) = \{y_\xi\}$.*
(ii) *The map $\xi \mapsto y_\xi$ is continuous and $y_\xi(t) < y_\eta(t)$ for all t if $\xi < \eta$.*
(iii) *There are no other periodic solutions* (*of any period*) *to* (23).

To prove (i) one supposes that $\mathcal{K}_{c_0}$ contains two points v and w. As $\overline{v - w} = 0$, one has $v(\tau) = w(\tau)$ for some τ and we show that $v'(\tau) = w'(\tau)$. Indeed, $y = \max(v, w) \in H_T^1$ and it is easy to show that $A_h(y) = c_0$, so that y solves Equation (23), thus is differentiable at τ, and the result follows. The second part of (ii) is proved in an analogous way and the first one easily follows from the fact that if $\xi_n \to \xi$, then (y_{ξ_n}) is a bounded Palais–Smale sequence, and hence is relatively compact. To show (iii) one first assume that u is a periodic solution of Equation (23) with period rational with T, and one defines

$$B = \{ \xi \in \mathbb{R} \colon \forall t \in \mathbb{R},\ y_\xi(t) \geqslant u(t) \}.$$

Like in Ortega–Tarallo reasoning above, B is nonempty, bounded below and closed, and one shows that, if $\xi_0 = \inf B$, then $u = y_{\xi_0}$. If $y_{\xi_0}(t) > u(t)$ for all $t \in \mathbb{R}$, then the rationality of the period of u with T implies that $\varepsilon := \inf_{t \in \mathbb{R}} [y_{\xi_0}(t) - u(t)] > 0$. For $\delta > 0$ such that $|\xi - \xi_0| \leqslant \delta$ implies $\|y_\xi - y_{\xi_0}\|_\infty \leqslant \varepsilon$, we see that $\xi_0 - \delta \in B$, a contradiction with the definition of ξ_0. Thus $y_\xi(t_0) = u(t_0)$ for some t_0 and then also $y'_\xi(t_0) = u'(t_0)$, so that $y_\xi = u$. If the period of u is irrational with T, then $h \equiv 0$ is constant and the assumption is not satisfied.

THEOREM 11. *Let $h \in C_T$ and $c_0 = \min_{H_T^1} A_{\tilde{h}}$. Then $m_{\tilde{h}} < M_{\tilde{h}}$ if and only if*

$$\mathcal{K}_{c_0}(\xi) = \emptyset \quad \textit{for some } \xi \in R. \tag{35}$$

In this case, 0 *is an interior point of* $[m_{\tilde{h}}, M_{\tilde{h}}]$.

If $\mathcal{K}_{c_0}(\xi) = \emptyset$, then one can find $\rho > 0$ and $\delta > 0$ such that if $\|\bar{y} - \xi\| \leqslant \rho$, and $|A_{\tilde{h}}(y) - c_0| \leqslant \rho$, then $\|A'_{\tilde{h}}(y)\| \geqslant \delta$. If not, for each positive integer n, there exist y_n such that

$$\|\bar{y}_n - \xi\| \leqslant \frac{1}{n}, \qquad \left|A_{\tilde{h}}(y_n) - c_0\right| \leqslant \frac{1}{n}, \qquad \left\|A'_{\tilde{h}}(y_n)\right\| \leqslant \frac{1}{n}.$$

The (BPS)-condition leads to a contradiction. Now let $z \in H^1_T$ be such that $A_{\tilde{h}}(z) < c_0 + \delta$ and m, k be two integers such that $\xi + 2k\pi < \bar{z} < \xi + 2m\pi$. Let $B = \{y \in H^1_T\colon \xi + 2k\pi \leqslant \bar{y} \leqslant \xi + 2m\pi\}$. Since by construction and its 2π-periodicity $A_{\tilde{h}}(z) < c_0 + \delta \leqslant \min_{\partial B} A_{\tilde{h}}$, we have $A_h(z) < \min_{\partial B} A_h$ provided $\bar{h} \in [-\mu, \mu]$, for some sufficiently small $\mu > 0$. Taking a minimizing sequence (y_n) for A_h restricted to B, which, up to a subsequence, weakly converges to y, we obtain $A_h(y) = \min_B A_h$, and, from $A_h(y) \leqslant A_h(z) < \min_{\partial B} A_h$, y is a local minimum for A_h on H^1_T and hence a solution to Equation (23). The proof of the converse result is essentially the same as the proof of the implication (iii) ⇒ (i) in Theorem 9.

Serra and Tarallo [119] have used their reduction method to obtain more precise information.

THEOREM 12. *Let* $h \in C_T$.

1. *If* φ_h *is constant, then* $M_h(\xi) = \{y_\xi\}$, *and if* y *is a periodic solution of Equation* (23), *then* $\bar{h} = 0$ *and* $y = y_\xi$ *for some* $\xi \in \mathbb{R}$.
2. φ_h *is not constant if and only if there exists* $\varepsilon_0 > 0$ *such that Equation* (23) *has at least one* T*-periodic solution for each* $|\bar{h}| < \varepsilon_0$.
3. $[m_{\tilde{h}}, M_{\tilde{h}}] = \{0\}$ *if and only if* φ_h *is constant.*
4. $\{\tilde{h} \in \widetilde{C}_T\colon \varphi_{\tilde{h}}$ *is not constant*$\}$ *is open and dense in* $\widetilde{C}_T$.
5. *If* φ_h *is constant and* $\bar{h} \neq 0$, *then Equation* (23) *has no bounded solution.*

From the study of some dynamical properties of diffeomorphisms of the plane having a continuum of fixed points, Campos, Dancer and Ortega [27] have refined conclusion (iii) of Proposition 5.

THEOREM 13. *If* (23) *is degenerate, then every solution which is bounded in the future is a* T*-periodic solution.*

The corresponding diffeomorphism of the plane is of course Poincaré's map.

Finally, Kannan and Ortega [61] have proved the following asymptotic result, showing that $M_{\tilde{h}} - m_{\tilde{h}}$ can be small, and have given an example showing that the involved set is not open. The proof makes use of some Riemann–Lebesgue lemma and asymptotic analysis techniques.

THEOREM 14. *For* $c = 0$,

$$\left\{\tilde{h} \in \widetilde{C}_T\colon \lim_{|\lambda| \to \infty} m_{\lambda\tilde{h}} = \lim_{|\lambda| \to \infty} M_{\lambda\tilde{h}} = 0\right\}$$

contains an open and dense subset of $\widetilde{C}_T$.

3.6. *Multiplicity and stability results under restrictions of the coefficients*

More precise results in the conservative case $c = 0$ can be obtained if one assumes that condition

$$a < \omega^2 \tag{36}$$

holds. Donati [41] has proved the following result about the structure of the solution set.

THEOREM 15. *If condition* (36) *holds and* $\bar{h} \in [m_{\tilde{h}}, M_{\tilde{h}}]$, *then Equation* (23) *has at most finitely many distinct* T*-periodic solutions when* $[m_{\tilde{h}}, M_{\tilde{h}}] \neq \{0\}$. *Otherwise, Equation* (23) *has an analytic unbounded curve of solutions.*

Starting from the Lyapunov–Schmidt's reduction, one first proves that for each $\bar{y} \in \mathbb{R}$ and $\tilde{h} \in \widetilde{C}_T$, Equation (18) has a unique solution $\tilde{y} = \tilde{y}(\bar{y}, \tilde{h}) \in \widetilde{C}^2_T$, which depends analytically upon $\bar{y}$ and $\tilde{h}$, and is such that

$$\|\tilde{y}\|_{C^2} \leqslant C\big(|\bar{y}| + \|\tilde{h}\|_\infty\big)$$

for a suitable constant $C > 0$. The problem is reduced to the study of the real valued function $\mathcal{F}$ defined over $\mathbb{R} \times \widetilde{C}_T$ by

$$\mathcal{F}(\bar{y}, \tilde{h}) := \frac{1}{T} \int_0^T a \sin\big[\bar{y} + \tilde{y}(\bar{y}, \tilde{h})\big]\, \mathrm{d}t.$$

One shows that, for each fixed $\tilde{h} \in \widetilde{C}_T$, $\mathcal{F}(\cdot, \tilde{h})$ is a real analytic 2π-periodic function and that

$$m_{\tilde{h}} = \min_{\bar{y} \in [0,2\pi]} \mathcal{F}(\bar{y}, \tilde{h}) \leqslant 0 \leqslant M_{\tilde{h}} = \max_{\bar{y} \in [0,2\pi]} \mathcal{F}(\bar{y}, \tilde{h}).$$

The result follows from properties of analytical functions.

By imposing further restrictions upon c, a, and T, it is possible to obtain on one hand *exact multiplicity results* for the T-periodic solutions, and, on the other hand, informations upon their *Lyapunov stability*. The pioneering work in the first direction is due to Tarantello [123] (using a Lyapunov–Schmidt approach) and, in the second direction, to Ortega [100–102] (using relations between stability and the Brouwer degree of Poincaré's operator).

If we first assume that $c > 0$, a recent paper of Čepička, Drábek and Jenšiková [33] provides the sharpest known conditions.

THEOREM 16. *If*

$$c > 0, \qquad a < \max\left\{\frac{c^2}{4} + \omega^2, \omega\sqrt{c^2 + \omega^2}\right\},$$

then $m_{\tilde{h}} < M_{\tilde{h}}$ *and Equation* (14) *has*:

(1) *exactly one T-periodic solution if either $\bar{h} = m_{\tilde{h}}$ or $\bar{h} = M_{\tilde{h}}$;*

(2) *exactly two T-periodic solutions if $\bar{h} \in]m_{\tilde{h}}, M_{\tilde{h}}[$.*

If

$$c > 0, \qquad a < \max\left\{\frac{c^2+\omega^2}{4}, \frac{\omega}{2}\sqrt{c^2+\frac{\omega^2}{4}}\right\},$$

then the conclusions 1 *and* 2 *remain true and the periodic solution obtained in case* 1 *is unstable while one solution obtained in case* 2 *is asymptotically stable and the other one unstable.*

The proof of the exact multiplicity results in Theorem 16 is based upon the Lyapunov–Schmidt reduction method together with the real analytic version of the implicit function theorem, to analyze the bifurcation equation. The uniqueness in the solution of Equation (18) is deduced from some preliminary study of the T-periodic solutions of linear equations of the type

$$y'' + cy' + g(t)y = 0,$$

with g T-periodic. The stability conclusion is obtained in the same way as in Ortega's papers.

Assume now that $c = 0$. The difficulty in analyzing the stability in the conservative case is that asymptotic stability can no more be expected. In a recent paper, Dancer and Ortega [37] have proved the following proposition.

LEMMA 8. *A stable isolated fixed point of an orientation preserving local homeomorphism on $\mathbb{R}^2$ has fixed point index equal to one.*

The proof of this result depends upon a variant of Brouwer's lemma on translation arcs. One of the given applications is the following result. Let $V : \mathbb{R}^2 \to \mathbb{R}$, $(t, y) \mapsto V(t, y)$ be continuous together with $\frac{\partial V}{\partial y}$, and T-periodic in t.

LEMMA 9. *If y is an isolated T-periodic solution of equation*

$$y'' = \frac{\partial V}{\partial y}(t, y), \tag{37}$$

and y reaches a local minimum on H_T^1 of the action functional

$$\varphi(y) = \int_0^T \left(\frac{y'^2(t)}{2} + V\big(t, y(t)\big)\right) \mathrm{d}t,$$

then y is unstable.

This result is proved by showing first, through a result of Amann on the computation of degree of gradient mappings and a relatedness principle of Krasnosel'skii–Zabreiko, that the index of y is equal to minus one. The result then follows from the previous one.

An immediate consequence for the pendulum equation is the following one.

THEOREM 17. *If $h \in \widetilde{C}_T$, any isolated T-periodic solution of Equation* (23) *minimizing A_h is unstable.*

One can then ask if the above result still holds without the isolatedness assumption. Ortega [105] has proved the following interesting property of fixed points in the plane.

LEMMA 10. *If $D \subset \mathbb{R}$ is a domain and $F : D \subset \mathbb{R}^2 \to \mathbb{R}^2$ is real analytical and not the identity on D, its Jacobian is equal to* 1 *on D, and if p is a stable fixed point of F, then p is isolated in the fixed points set of F.*

The delicate proof of this result uses Brouwer's plane translation theorem.

As an application, the following unstability result is proved in [105].

LEMMA 11. *If V is real analytic, and y is a T-periodic solution of Equation* (37) *on which φ has a local minimum on H^1_T, then y is unstable.*

An immediate consequence for the forced pendulum equation is the following one.

THEOREM 18. *If $h \in \widetilde{C}_T$ is analytical, the number of T-periodic solutions of Equation* (23) *that are stable and geometrically distinct is finite.*

Calanchi and Tarallo [26] have used the Serra–Tarallo reduction method to show the following result.

THEOREM 19. *Assume that $a < \omega^2$. Then there exists $K = K(a, T) > 0$ such that if $\|h\|_2 < K$, there exists an increasing and continuous map $\eta : \mathbb{R} \to \mathbb{R}$ such that, for the reduction function φ_h, one has*

$$\#\varphi_h^{-1}(\lambda) = \#\big[aT\cos(\cdot)\big]^{-1}\big(\eta(\lambda)\big).$$

Furthermore, each critical point of A_h over H^1_T is a local minimum or a point of mountain pass type.

3.7. *Many T-periodic solutions for special forcings*

We now raise the question of the possibility of having more than two geometrically distinct T-periodic solutions for the forced pendulum. In [40], Donati proved the following result.

THEOREM 20. *Given $a > 0$ and $T > 0$, there exists $h^* \in \widetilde{C}_T$ and a neighborhood V of h^* in $\widetilde{C}_T$ such that for each $h \in V$, Equation* (23) *has at least four geometrically distinct T-periodic solutions.*

The proof is based upon a *classification of singularities* of the nonlinear Fredholm operator $\frac{d^2}{dt^2} + a\sin(\cdot)$ over the space of C^2 T-periodic functions.

Applying to Equation (16) a classical *perturbation method* as used for example by Loud for Duffing's equation, Ortega [104] has improved this result by replacing 4 by any even number.

THEOREM 21. *Given $a > 0$ and an integer $N \geqslant 1$, there exists $h^* \in \widetilde{C}_T$ and $\delta > 0$ such that if $h \in \widetilde{C}_T$ is such that $\|h - h^*\|_{L^1} < \delta$, then Equation* (23) *has at least $2N$ geometrically distinct T-periodic solutions.*

The idea of the proof consists in considering the equation

$$y'' + a\sin\bigl(y + \mathcal{B}_T(t)\bigr) = 0, \tag{38}$$

where $\mathcal{B}_T$ is defined in (29), which has a continuum $(y_c)_{c\in\mathbb{R}}$ of T-periodic solutions, and in considering a perturbation of Equation (38)

$$y'' + a\sin\bigl(y + \mathcal{B}_T(t) + \Psi(t,\varepsilon)\bigr) = 0, \tag{39}$$

with conditions upon Ψ insuring that $\mathcal{B}_T(t) + \Psi(t,\varepsilon)$ is smooth and that one has at least $2N$ periodic simultaneous bifurcations for $\varepsilon = 0$.

We finally describe some recent results obtained by Ureña [127], which cover the possibly dissipative case.

THEOREM 22. *For any $N \in \mathbb{N}$, the set*

$$\mathcal{S}_N := \bigl\{h \in L^1_T\colon (14) \textit{ has at least } N \textit{ geometrically distinct solutions}\bigr\}$$

has nonempty interior in L^1_T. More precisely,

(1) $\operatorname{int}\mathcal{S}_N \cap \widetilde{L}^1_T \neq \emptyset$ *if $c = 0$;*

(2) $\operatorname{int}\mathcal{S}_N \cap \{h \in L^1_T\colon -\varepsilon < \bar{h} < 0\} \neq \emptyset \neq \operatorname{int}\mathcal{S}_N \cap \{h \in L^1_T\colon 0 < \bar{h} < \varepsilon\}$ *if $c \neq 0$.*

The proof uses the equivalent formulation (16). One first observes that there exists a necessarily constant $\bar{h}$ such that the problem

$$y'' + cy' + a\sin y = \bar{h}, \quad y(0) = 0,\ y(t + T) = y(t) + 2\pi$$

has a unique solution, denoted by $y_{c,T}$. This is proved by considering the equivalent problem of finding $w \in C_T$ solution of equation

$$v'' + cv' + a\sin(v + \omega t) = \bar{h} - \omega c,$$

and using Schauder's fixed point theorem to obtain $\bar{h} = \bar{h}_{c,T}$ for which a solution $v_{c,T}$ exists. Notice that such a problem is degenerate, and hence has a nontrivial curve of solutions $\gamma : \mathbb{R} \to C^1_T$, $a \mapsto v_{c,T}(\cdot + a) + \omega a$. Some bifurcation technique based upon Lyapunov–Schmidt decomposition is then used to obtain many solutions bifurcating from such a closed loop at a constant external force. See [127] for the quite involved details, as well as for the proof of the following result, when $c = 0$.

THEOREM 23. *Let $N \in \mathbb{N}$ be given. If $c = 0$ and*

$$T \geqslant 12 \log\left(\frac{\sqrt{3}+1}{\sqrt{2}}\right)\frac{N}{\sqrt{a}},$$

there exists an open set $\mathcal{O}_N \subset L^1_T$ with $\mathcal{O}_N \cap \widetilde{L}^1_T \neq \emptyset$, such that, for any $h \in \mathcal{O}_N$, Equation (23) *has exactly $2N$ geometrically distinct T-periodic solutions.*

3.8. *Many T-periodic solutions for small length pendulum*

To motivate a further multiplicity result of perturbation type proved by Fonda and Zanolin [46] in the forced case, let us recall that for the conservative free pendulum equation (23), relation (7) implies that, given any positive integer N, Equation (5) has a closed orbit with least period $\frac{T}{k}$ for each $k = 1, 2, \ldots, N$, if $a > \frac{4\pi^2 N^2}{T^2}$. To deal with the forced case, Fonda and Zanolin use Weiyue Ding's generalization of the Poincaré–Birkhoff fixed point theorem [39] for the area-preserving twist homeomorphism given by Poincaré's operator $\mathcal{P}$, that we recall now.

Let A be the annulus $S^1 \times [a, b]$ in $\mathbb{R}^2$. If $\phi : A \to \mathbb{R}^2 \setminus \{0\}$, we denote by

$$F(\theta, r) : \mathbb{R} \times [a, b] \to \mathbb{R} \times [0, +\infty[,\ (\theta, r) \mapsto \big(\theta + f(\theta, r), g(\theta, r)\big) \tag{40}$$

its *lift* on the polar coordinates covering space, where f and g are 2π-periodic in θ.

The eldest fixed-point theorem for maps in an annulus is the *Poincaré–Birkhoff's theorem* [112,15].

LEMMA 12. *Every area-preserving homeomorphism $\phi : A \to A$ which rotates the two boundaries in opposite directions, i.e., is such that $f(\theta, a) f(\theta, b) < 0$ for all $\theta \in \mathbb{R}$, has at least two fixed points in the interior of A.*

A more effective version for applications has been given by Weiyue Ding [39].

LEMMA 13. *Let $A \subset \mathbb{R}^2 \setminus \{0\}$ denote an annular region whose inner boundary Γ_1 and outer boundary Γ_2 are closed simple curves around the origin* 0 *and denote by D_i the open region bounded by Γ_i, $i = 1, 2$. Let $\varphi : A \to \mathbb{R}^2 \setminus \{0\}$ be an area-preserving homeomorphism. Suppose that*

(i) *the inner boundary curve* Γ_1 *is star-shaped around the origin,*
(ii) φ *admits a lift* F *onto the polar coordinate covering space* (40), *such that* $f(\theta, r) > 0$ *on the lift of* Γ_1 *and* $f(\theta, r) < 0$ *on the lift of* Γ_2,
(iii) *there exists an area-preserving homeomorphism* $\varphi_0 : \overline{D}_2 \to \mathbb{R}^2$ *which satisfies* $\varphi_0|_A \equiv \varphi$ *and* $0 \in \varphi_0(D_1)$.

Then φ *has at least two fixed points in the interior of* A.

The result of Fonda and Zanolin can be stated as follows.

THEOREM 24. *Given any positive integer* N, *there exists* $a_0 > 0$ *such that, for any* $a \geqslant a_0$, *Equation* (23) *has at least* N *periodic solutions with minimal period* T, *which can be chosen to have exactly* $2j$ *simple crossings with* 0 *in the interval* $[0, T[$ $(j = 1, 2, \ldots, N)$.

To prove this result, the authors notice that a simple change of variable transforms Equation (23) into equation

$$y'' + \sin y = a^{-1} h\big(a^{-1/2} t\big) \tag{41}$$

and the T-periodic solutions of Equation (23) correspond to the $T\sqrt{a}$-periodic solutions of Equation (41). They combine then a perturbation argument together with Ding's twist theorem.

3.9. *Subharmonic solutions when* $c = 0$

Subharmonic solutions of a T-periodic equation are solutions whose minimal period is a proper multiple of T.

DEFINITION 9. If $k \geqslant 2$ is an integer, a *subharmonic solution* of order k of Equation (14) is a periodic solution of Equation (14) with minimal period kT.

The first existence theorem for the subharmonic solutions of Equation (23) with $h \in \widetilde{C}_T$ have been obtained by Fonda and Willem [45] using Morse theory (see, e.g., [86]). Offin [96] has proved a close result using an index theory for periodic extremals and a variant of the mountain pass lemma. The subharmonics of order k are associated to the critical points of the functional $A_{h,k}$ defined by

$$A_{h,k}(y) = \int_0^{kT} \left(\frac{y'^2(t)}{2} + a \cos y(t) + h(t) y(t) \right) \mathrm{d}t,$$

over the Sobolev space H^1_{kT}.

DEFINITION 10. The *Morse index* of an isolated T-periodic solution y of Equation (23) is the supremum of the dimensions of the subspaces of H^1_T on which the quadratic form

$$\chi_y(v) := \int_0^T (1/2)\{v'^2(t) - [a\cos y(t)]v^2(t)\}\,\mathrm{d}t$$

is negative definite.

If, for $\lambda \in \mathbb{R}$, we denote by $\sigma'_{\lambda,T}$ and $\sigma''_{\lambda,T}$ the characteristic multipliers of the T-periodic differential equation

$$v'' + [a\cos y(t)]v + \lambda v = 0,$$

then y is said to be *nondegenerate* if $1 \notin \{\sigma'_{\lambda,T}, \sigma''_{\lambda,T}\}$. Given $\sigma \in S^1$, one defines $J(y, T, \sigma)$ to be the number of negative λ's for which $\sigma \in \{\sigma'_{\lambda,T}, \sigma''_{\lambda,T}\}$. Then $J(y, T, 1)$ is equal to the Morse index of the T-periodic solution y of (23). *Bott's iteration formula* [24] ensures that

$$J(y, kT, 1) = \sum_{\sigma^k=1} J(y, T, \sigma).$$

THEOREM 25. *Suppose that the T-periodic solutions of Equation* (23) *are isolated and that every T-periodic solution of Equation* (23) *having Morse index equal to zero is nondegenerate. Then there exists $k_0 \geqslant 2$ such that, for every prime $k \geqslant k_0$, there is a periodic solution of Equation* (23) *with minimal period kT. If moreover the kT-periodic solutions of Equation* (23) *are nondegenerate for $k = 1$ and every prime k, there exists $k_0 \geqslant 3$ such that, for every prime $k \geqslant k_0$, Equation* (23) *has at least two periodic solutions with minimal period kT.*

By assumption and an easy reasoning, $A_{h,1} = A_h$ has a finite number of critical points $y_0, y_1, \ldots, y_n$, which, of course, are also critical points of $A_{h,k}$ for any $k \geqslant 2$. The first ingredient of the proof consists in showing the existence of some integer k_0 such that, for $k \geqslant k_0$ and $0 \leqslant i \leqslant n$, either the Morse index $J(y_i, kT, 1)$ of y_i is equal to 0 and y_i is nondegenerate, or $J(y_i, kT, 1) \geqslant 2$. This is done using Bott's iteration formula. Now, let $k \geqslant k_0$ be prime, so that the critical points of $A_{h,k}$ have minimal period T or kT. Assume now by contradiction that $y_0, \ldots, y_n$ are the only critical points of $A_{h,k}$. The *Poincaré polynomial* of the space $S^1 \times \widetilde{H}^1_T$ (whose coefficient of t^n is the nth Betti number of $S^1 \times \widetilde{H}^1_T$) is equal to $1 + t$, and hence, by the *Morse inequalities* [86], one has

$$\sum_{j=0}^n M_k(t, y_j) = (1+t)\big(1 + Q(t)\big) \tag{42}$$

for some polynomial $Q(t)$ with nonnegative integer coefficients. In this formula,

$$M_k(t, y_j) = \sum_i \dim C_i(A_{h,k}, y_j), \quad C_i(A_{h,k}, y_j)$$

is the ith *critical group* of y_j [86], so that $\dim C_i(A_{h,k}, y_j) = \delta_{i,J(y_j,kT,1)}$ when y_j is nondegenerate. By the claim, if $J(y_j, kT, 1) = 0$, then $M_k(t, y_j) = 1$, and otherwise, if $J(y_j, kT, 1) \geqslant 0$, then $\dim C_i(A_{h,k}, y_j) = 0$ for $i = 0, 1$, and $M_k(t, y_j)$ starts with terms in t^2 at least. So, the left-hand member of (42) contains no term in t, a contradiction. The proof of the second part of Theorem 25 is similar.

Combining the Fonda–Willem's theorem with the results of [68], one obtains the generic existence of subharmonic solutions.

THEOREM 26. *There exists an open dense subset $\mathcal{G}$ of $\widetilde{C}_T$ such that for every $h \in \mathcal{G}$, there exists $k_0 \geqslant 2$ such that, for every prime $k \geqslant k_0$, Equation* (23) *has a periodic solution with minimal period kT.*

As shown in [121], the modified Lyapunov–Schmidt reduction method also provides some information about subharmonic solutions, by relating their existence to the properties of φ_h.

THEOREM 27. *Equation* (23) *with $h \in \widetilde{C}_T$ has subharmonics of infinitely many distinct levels if and only if φ_h is not constant. If $\min_{H^1_T} A_h$ is isolated in the set of critical levels of A_h, then Equation* (23) *with $h \in \widetilde{C}_T$ admits subharmonics of arbitrary large minimal period if and only if φ_h is not constant. Finally, the isolatedness assumption in the previous statement can be dropped if $a < \omega^2$.*

Fonda–Zanolin's multiplicity result [46] has a counterpart for subharmonic solutions, proved using the same fixed point technique.

THEOREM 28. *Given any two positive integers M, N, there exists $a_0 > 0$ such that, for any $a \geqslant a_0$, Equation* (23) *has, for each $k = 1, 2, \ldots, M$, at least N periodic solutions with minimal period kT.*

4. Rotating solutions and Mather sets

4.1. *Periodic solutions of the second kind*

Besides periodic solutions, we have seen the free pendulum has also periodic solutions of the second kind, which are the sum of a linear function of t and of a periodic term. We shall study the existence of such solutions for the forced conservative pendulum (23).

DEFINITION 11. y is a *periodic solution of the second kind* of (23) if there exists $p \in \mathbb{Z} \setminus \{0\}$ and $q \in \mathbb{N} \setminus \{0\}$ such that, for each $t \in \mathbb{R}$, one has

$$y(t + qT) = y(t) + 2p\pi. \tag{43}$$

Notice that such a solution is such that

$$y'(t + qT) = y'(t)$$

for all $t \in \mathbb{R}$, and hence, integrating and using (43), we find that

$$y(t) = \frac{p}{q}\omega t + v(t) \tag{44}$$

with $\omega = 2\pi/T$ and v some qT-periodic function. Conversely, every solution of (23) of the type (44) is a periodic solution of second kind. Such a solution has the *rotation number*

$$\lim_{t \to \pm\infty} \frac{y(t)}{\omega t} = \frac{p}{q}. \tag{45}$$

Under some conditions, the conservative forced pendulum (23) can also admit periodic solutions of the second kind. First, the change of unknown $y(t) \mapsto v(t)$ defined by (44) and the use of direct methods of the calculus of variations to the transformed equation allows a very simple proof of the following result [79].

THEOREM 29. *For each $a > 0$, $T > 0$, $q \in \mathbb{N} \setminus \{0\}$, $p \in \mathbb{Z} \setminus \{0\}$, and each $h \in \widetilde{C}_T$, Equation* (23) *has at least one solution y satisfying* (43).

Such a solution, also called a *rotating solution of Equation* (23) *with rotation number $\frac{p}{q}$*, is entirely determined by its values on $[0, qT]$, and satisfies the boundary conditions

$$y(qT) = y(0) + 2p\pi, \qquad y'(qT) = y'(0). \tag{46}$$

A second geometrically distinct solution also follows from the mountain pass argument.

In the case of an analytic h, Ortega's approach described in Section 3.7 provides information about the number and stability of those rotating solutions [105].

THEOREM 30. *If $h \in \widetilde{C}_T$ is analytic, given $p \in \mathbb{Z} \setminus \{0\}$ and $q \in \mathbb{N} \setminus \{0\}$, the number of stable and distinct rotating solutions with rotation number $\frac{p}{q}$ of Equation* (23) *is finite.*

4.2. *Solutions with an arbitrary rotation number*

We shall now show that Equation (23) with $h \in \widetilde{C}_T$ also admits solutions having an arbitrary rotation number. This follows from some results of Mather [69] and Moser [89,88], that we sketch in the form given by Denzler [38], which is closer in spirit to the calculus of variations (see also [95]). Let $I = [t_1, t_2]$ be given and consider the action functional A_h^I defined on $H^1(I)$ by

$$A_h^I(y) = \int_I \left(\frac{y'^2(t)}{2} + a\cos y(t) + h(t)y(t) \right) \mathrm{d}t. \tag{47}$$

DEFINITION 12. A function $y \in H^1(I)$ is called *minimal with respect to fixed boundary conditions on $I = [t_1, t_2]$*, if $A_h^I(z) \geqslant A_h^I(y)$ for all $z \in H^1(I)$ equal to y at t_1 and t_2.

A function $y \in H^1_{\mathrm{loc}}(\mathbb{R})$ is called a *minimal* if it is minimal with respect to fixed boundary conditions on every interval. Finally, a function $y \in H^1([0, qT])$ is *minimal with respect to condition* (43) if, for $I = [0, qT]$, $A^I_h(z) \geqslant A^I_h(y)$ for all $z \in H^1([0, qT])$ satisfying condition (46).

Notice that the direct method of the calculus of variations easily implies that, for $h \in \widetilde{C}_T$, minimal solutions with respect to fixed boundary conditions on a given interval I always exist, are of class C^2, and satisfy Equation (23). From Theorem 29, the same is true for the functions which are minimal with respect to condition (43). In particular those minimals have no corners, a fact which is useful in several proofs.

The existence of minimals is not clear from the direct methods of the calculus of variations, and will follow from a delicate sequence of arguments. We first give some properties of the minimals [38]. The first one deals with the possible intersections of two different minimals.

LEMMA 14. *Two different minimals can intersect at most once, and are not tangent to each other anywhere. Two different minimals which are C^1 asymptotic as $t \to \infty$ cannot intersect at all.*

The non-tangency is a consequence of the uniqueness of the Cauchy problem. The first intersection property is proved by contradiction, assuming that two minimals y and z intersect twice in the interior of some interval I, and noticing that if $y^* = \min(y, z)$, $z^* = \max(y, z)$, then

$$\begin{aligned} &A^I_h(y^*) \geqslant A^I_h(y), \qquad A^I_h(z^*) \geqslant A^I_h(z), \\ &A^I_h(y^*) + A^I_h(z^*) = A^I_h(y) + A^I_h(z), \end{aligned}$$

so that $A^I_h(y^*) = A^I_h(y)$, y^* is minimal and hence y and z must intersect tangentially and be identical. The asymptotic property follows from the fact that up to an arbitrarily small error, asymptoticity counts like an intersection.

LEMMA 15. *The following statements are equivalent for $y \in H^1_{\mathrm{loc}}(\mathbb{R})$.*

(a) *y is minimal and satisfies* (43);
(b) *y is minimal with respect to condition* (43);
(c) *y is minimal with respect to condition*

$$y(t + NqT) = y(t) + 2Np\pi$$

for some positive integer N.

To prove that (b) $\Leftrightarrow$ (c), let z be minimal in the sense of (c) and y minimal in the sense of (b). The translate $v(t) = z(t + qT) - 2\pi p$ must intersect z at least once because, if not, one gets inductively a contradiction with the condition in (c). By the periodicity conditions

v and z intersect infinitely many times and hence coincide. Thus $z(t+qT)=z(t)+2\pi p$, and

$$NA_h^{[0,qT]}(z) \geqslant NA_h^{[0,qT]}(y) = A_h^{[0,NqT]}(y) \geqslant A_h^{[0,NqT]}(z) = NA_h^{[0,qT]}(z),$$

implying (b) $\Leftrightarrow$ (c). It is easy to show that (c) $\Leftrightarrow$ (a). That (a) $\Leftrightarrow$ (b) is equivalent to showing that all minimals y with fixed boundary conditions verifying (43) yield the same action over one period. If not the difference between the two actions increases indefinitely with the number of periods but remains bounded by minimality on the corresponding large interval.

If y is minimal, the same is true for any of its translates $y(t-qT)+2p\pi$. They correspond to the same orbit on the torus $\mathbb{R}/2\pi\mathbb{Z} \times T\mathbb{Z}$ parametrized by the (y,t) coordinates. The following lemma shows that there are no self-intersections on this torus.

LEMMA 16. *A minimal $y(t)$ does not intersect any of its translates $y(t-qT)+2p\pi$.*

There can be at most one intersetion, say at t_0 and w.l.g. we can assume that $y(t+qT) < y(t)+2p\pi$ for $t > t_0$ and $y(t+qT) > y(t)+2p\pi$ for $t < t_0$. Let $\xi(t) = y(t) - \frac{p}{q}\omega t$ and $\xi_n(t) = \xi(t+qnT)$. It is easy to show that for $t > t_0$ (respectively $t < t_0$) the sequence $(\xi_n(t))$ (respectively $(\xi_{-n}(t))$) is decreasing and that one of the two sequences is bounded from below, say $(\xi_n(t))$. Its limit $\eta(t)$ as $n \to \infty$ satisfies, by construction, $\eta(t+qT) = \eta(t)$. Now, from the differential equation satisfied by $\xi(t)$, one gets that $|\xi''(t)|$ is bounded on the real line and the boundedness of $|\xi(t)|$ on $[t_0, +\infty[$ follows from the above reasoning. Then the same is true for $|\xi'(t)|$ and Arzela–Ascoli theorem applied to (ξ_n) implies its convergence in the C^1-topology. As ξ_n and ξ_{n-1} have the same limit η, y and its translate are C^1-asymptotic, a contradiction to the existence of the intersection t_0.

THEOREM 31. *Every minimal y has a rotation number*

$$\alpha = \lim_{t\to\pm\infty} \frac{y(t)}{\omega t}.$$

Moreover,

$$\bigl|y(t) - y(0) - \alpha\omega t\bigr| \leqslant c\bigl(1 + 2\pi|\alpha|\bigr), \tag{48}$$

$$\bigl|y'(t)\bigr| \leqslant c, \tag{49}$$

where the constant c does not depend upon y.

Using Lemma 16 and Denjoy theory for mappings on the circle applied to the conjugate of the Poincaré mapping over $[0,T]$ with respect to the projection on the first component of the phase space, restricted to its translates, one obtains the existence of some α such that

$$\bigl|y(jT) - y(0) - 2\pi\alpha j\bigr| \leqslant 1 \quad (j \in \mathbb{Z}). \tag{50}$$

To obtain the inequalities, notice that, as $h \in \widetilde{C}_T$, we can find, using Sobolev inequality, $\delta > 0$ independent upon y such that, for each $z \in H^1([0,T])$,

$$-aT + \frac{\delta}{4}\int_0^T \left|z'(t)\right|^2 \mathrm{d}t \leqslant A_h^{[0,T]}(z) \leqslant aT + \delta^{-1}\int_0^T \left|z'(t)\right|^2 \mathrm{d}t.$$

Now if x is the function interpolating linearly y with respect to its values at 0 and T, we get

$$\begin{aligned}-aT + \frac{\delta}{4}\int_0^T \left|y'(t)\right|^2 \mathrm{d}t &\leqslant A_h^{[0,T]}(y) \leqslant A_h^{[0,T]}(x)\\ &\leqslant aT + \delta^{-1}\int_0^T \left|x'(t)\right|^2 \mathrm{d}t\\ &= aT + (T\delta)^{-1}\left|y(T) - y(0)\right|^2.\end{aligned}$$

Consequently, using (50),

$$\int_0^T \left|y'(t)\right|^2 \mathrm{d}t \leqslant \frac{8aT}{\delta} + \frac{4}{T\delta^2}\left(1 + |\alpha|\omega\right)^2,$$

which gives, for $t \in [0,T]$,

$$\left|y(t) - y(0)\right|^2 = \left(\int_0^t y'(s)\,\mathrm{d}s\right)^2 \leqslant t\int_0^t \left|y'(s)\right|^2 \mathrm{d}s \leqslant \frac{8a}{\delta} + \frac{4}{\delta^2}\left(1 + 2\pi|\alpha|\right)^2.$$

Using inequality (50), inequalities (48) and (49) easily follow.

We can now prove the existence of minimals.

THEOREM 32. *For every $\alpha \in \mathbb{R}$, there exists a minimal with rotation number α. Furthermore, C^1-limits* (*in the sense of uniform convergence of the function and its derivative on compact subsets of* $\mathbb{R}$) *of minimals are minimals, and the rotation number is continuous* (*i.e., if a sequence of minimals (y_n) with corresponding rotation numbers (α_n) converge to a minimal y with rotation number α in the C^1-topology, then $\alpha = \lim_{n\to\infty}\alpha_n$*).

The result follows from Theorem 29 and Lemma 15 when α is rational. If $\alpha = \lim_{n\to\infty}\frac{p_n}{q_n}$ is irrational, let y_n be a minimal for the boundary conditions

$$y_n(t + q_nT) = y_n(t) + 2\pi p_n$$

such that (translation invariance) $|y_n(0)| \leqslant 2\pi$. For $t \in [-R,R]$, Theorem 31 yields estimates

$$\left|y_n(t)\right| \leqslant MR, \qquad \left|y_n'(t)\right| \leqslant M, \qquad \left|y_n''(t)\right| \leqslant M$$

with M independent of t and n, which gives, by Ascoli–Arzela theorem some $C^1([-R, R])$-converging subsequence and by a diagonal procedure a subsequence which converges C^1 on every compact interval to some limit y. Such a y is minimal with rotation number α.

4.3. *Mather set*

We now restrict to the case where α is irrational.

THEOREM 33. *The translates $y(t + jT) - 2k\pi$ of a minimal y are totally ordered. The ordering is the same as the ordering of $j\alpha - k$, and is independent of y.*

The mapping $(j, k) \mapsto j\alpha - k$ is one-to-one and we may write

$$y(t + jT) - 2k\pi := u(t|\beta), \quad \beta = j\alpha - k,$$

with $u(t|\cdot)$ so defined for a dense set of reals β and strictly monotone in β. It can therefore be easily extended to all real β by

$$u^+(t|\beta) := \lim_{\gamma\to\beta+} u(t|\gamma), \qquad u^-(t|\beta) := \lim_{\gamma\to\beta-} u(t|\gamma),$$

DEFINITION 13. The *Mather sets* M_α and $M_\alpha(t)$ of Equation (23) are defined by

$$M_\alpha^\pm = \{u^\pm(\cdot|\beta)\colon \beta \in \mathbb{R}\}, \qquad M_\alpha^\pm(t) = \{u^\pm(t|\beta)\colon \beta \in \mathbb{R}\},$$
$$M_\alpha = M_\alpha^+ \cup M_\alpha^-, \qquad M_\alpha(t) = M_\alpha^+(t) \cup M_\alpha^-(t).$$

It can be shown that they do not depend upon the minimal y chosen for their construction and that they have the following properties.

THEOREM 34. *The Mather sets have the following properties.*

(1) *$M_\alpha(t)$ is invariant under the Poincaré map $y(t) \mapsto y(t + T)$ and the translation $y \mapsto y + 2\pi$.*
(2) *$M_\alpha(t)$ is closed without isolated points.*
(3) *$M_\alpha(t)$ does not contain any nonempty closed strict subset invariant under Poincaré map and translation.*
(4) *$M_\alpha(t)$ is either $\mathbb{R}$ or is nowhere dense.*
(5) *M_α consists of trajectories of some vector field $y' = \psi(y, t)$ defined either on a Cantor-like part or on the whole of the (y, t)-plane.*

Hai Huang has proved in [57], for pendulum-type equations, the generic existence of *invariant cantori* (minimal orbits defining a Cantor set on the torus).

5. KAM theory and Lagrange stability

5.1. *Twist mappings*

The existence of quasi-periodic solutions to the forced pendulum can be obtained by applying to the Poincaré's operator some theorem for twist mappings, that we recall now.

Let $A = S^1 \times [a, b]$ be an annulus in $\mathbb{R}^2$, and $\phi : A \to \mathbb{R}^2$ a mapping with lift (40). *Moser's twist theorem* [87] can be stated as follows.

LEMMA 17. *Let* $l \geqslant 5$, $\beta \in C^5(\mathbb{R})$ *be such that* $|\beta'(r)| \geqslant \nu > 0$ *for all* $r \in [a, b]$, *let* $\varepsilon > 0$, *and let* α *be an irrational number satisfying the Diophantine conditions*

$$\left|\alpha - \frac{m}{n}\right| \geqslant \gamma n^{-\tau},$$

for some positive γ, τ, *and all integers* $n > 0$, m. *Then there exists* $\delta = \delta(\varepsilon, l, \alpha) > 0$ *such that each area-preserving mapping* $\phi : A \to \mathbb{R}^2$ *whose lift* (40)*, with* $f, g \in C^l$, *is such that*

$$|f - \beta|_{C^l} + |g - r|_{C^l} < \nu\delta,$$

possesses an invariant curve of the form

$$\theta = \tau + w(\tau), \qquad r = c + z(\tau),$$

in A, *where* w, z *are of class* C^1, 1*-periodic,* $|w|_{C^1} + |z|_{C^1} < \varepsilon$, $c \in]a, b[$, *on which* ϕ *takes the form* $\tau \to \tau + \alpha$.

The following result has been proved independently by Levi [66] and by Moser [89,94], using Moser twist theorem.

THEOREM 35. *Assume that* $h \in \widetilde{C}_T$. *For any* $\alpha \in]0, 1[$ *satisfying the set of Diophantine inequalities*

$$\left|\alpha - \frac{m}{n}\right| > \frac{\gamma}{n^{2+\mu}},$$

for some $\gamma > 0$ *and* $\mu > 0$, *and all integers* m, n *with* $n > 0$, *there exists an integer* $k_0 = k_0(\gamma, \mu)$ *such that the Poincaré mapping associated to* (23) *possesses, for all integers* k *with* $|k| \geqslant k_0$, *a countable set of invariant curves* $r = f_{\alpha+k}(\theta) \equiv f_{\alpha+k}(\theta + 2\pi)$.

Physically, the solutions correspond to quasiperiodic rotations with average angular velocity $\omega + k$. The basic idea of the proof is that, for large velocities y', the forced pendulum equation has solutions which are close to those of the integrable system $y'' = 0$. One first shows that, because of condition $\bar{h} = 0$, the Poincaré map $\mathcal{P}$ is exact on the phase cylinder $\{(y, z) \in S^1 \times \mathbb{R}\}$ (i.e., $\int_{C_0} z\,\mathrm{d}y = \int_{C_1} z\,\mathrm{d}y$ if C_0 is an arbitrary noncontractible circle going once around the cylinder and C_1 its image under $\mathcal{P}$). Then one shows that $\mathcal{P}$ is C^4-close to a twist mapping for large y', namely

$$\mathcal{P}_1(\theta, r) = \theta + r + \overline{H} + r_1(\theta, r), \qquad \mathcal{P}_2(\theta, r) = r + r_2(\theta, r),$$

where $\overline{H} = \int_0^T \int_0^t h(\tau)\,\mathrm{d}\tau\,\mathrm{d}t$ and $\|r_1\|_{C^4} + \|r_2\|_{C^4} < C|z|^{-1}$.

Independently, and by the same approach, Jiangong You [130] has proved a similar result, and has completed it as follows.

Notice that an invariant curve with rotation number α of the Poincaré mapping of Equation (23) give rise to an invariant torus with rotation numbers $(1, \alpha)$ in the extended phase space $S^1 \times \mathbb{R} \times \mathbb{R}$ of the corresponding (t, y, y').

THEOREM 36. *Equation* (23) *has an invariant torus if and only if* $h \in \widetilde{C}_T$. *If it is the case, there exists* $\alpha^* > 0$ *such that, for every irrational number* $\alpha > \alpha^*$ *satisfying the Diophantine condition*

$$\left| \alpha - \frac{m}{n} \right| > \frac{\gamma}{n^{5/2}},$$

for all integers m and n with $n > 0$, *and some* $\gamma > 0$, *Equation* (23) *has an invariant torus with rotation number* $(1, \alpha)$.

The necessity is proved by observing that the *Calabi invariant* of $\mathcal{P}$, defined by $C(\mathcal{P}) = \int_{\mathcal{P}(\gamma)} r\,\mathrm{d}\theta - \int_\gamma r\,\mathrm{d}\theta$, where γ is a circle in $\mathbb{R} \times S^1$ homotopic to a circle $\{r\} \times S^1$, is equal to $\bar{h}$. Thus, if $\mathcal{P}$ possesses an invariant closed curve which is homotopic to $\{r\} \times S^1$, one has $C(\mathcal{P}) = 0$, and hence $\bar{h} = 0$.

We now show that those results imply a property called Lagrange stability, answering a question raised by Moser in the Introduction of [87].

DEFINITION 14. Equation (23) is called *Lagrange stable* if any solution of (14) is bounded over $\mathbb{R}$ in the phase cylinder $\{(y \bmod 2\pi, y')\}$.

Physically, this means that any solution of (14) has angular velocity bounded over $\mathbb{R}$. The Lagrange stability is a consequence of the results of Levi and Moser described above. See also [67].

THEOREM 37. *If* $h \in \widetilde{C}_T$, *then for any sufficiently large* $N > 0$, *there exists* $M = M(N)$ *such that any solution* $y(t)$ *of Equation* (23) *with* $|y'(0)| \leqslant M$ *satisfies* $|y'(t)| \leqslant N$ *for all* $t \in R$.

Independently, You [130] also proved the Lagrange stability of (23) and completed the result as follows.

THEOREM 38. *Equation* (23) *is Lagrange stable if and only if* $h \in \widetilde{C}_T$.

Indeed, if $\bar{h} = 0$, then, from Theorem 36, Equation (23) has an invariant torus with rotation number $(1, \alpha)$ for infinitely many sufficiently large irrational numbers α. Letting $u = -y$, $v = -z$, Equation (23) becomes

$$u'' + a \sin u = -h(t),$$

and has an invariant torus with rotation number $(1, \alpha)$ for infinitely many sufficiently large irrational numbers α. Thus Equation (23) has an invariant torus with rotation number $(1, -\alpha)$. One uses two such invariant tori to confine any solutions in their interior. To prove the necessity, one shows, using the estimates on $\mathcal{P}$, that Equation (23) has a solution with y' unbounded if $\bar{h} \neq 0$.

The destruction of invariant tori with rotation number $(1, \omega)$ and ω a Liouville irrational in pendulum-type equations has been considered by Hai Huang [56].

Let us notice the Mather's theory of previous section can also be expressed in the frame of mappings on an annulus, which enlightens its relation with KAM theory. We keep the notations of Section 3.8.

DEFINITION 15. $\phi : A \to A$ is a *monotone twist homeomorphism* if it preserves orientation, preserves boundary components of A and if its lift (40) is such that $f(\theta, \cdot)$ is strictly monotone for each θ. For definiteness, we assume it to be increasing.

Let $F^j(\theta, r) = (\theta_j, r_j)$, and

$$\alpha_r(\phi) = \lim_{j \to \infty} \frac{\theta_j}{j}$$

be its *rotation number*. The *twist interval* of ϕ, $[\alpha_a(\phi), \alpha_b(\phi)]$, is defined up to an integral translation.

If $\phi^q(z) = z$, then $F^q(\theta, r) = (\theta + 2p\pi, r)$, for some integer p determined up to a multiple of q.

$\frac{p}{q}$ is the rotation number of z. One calls a point $z = (\theta, r)$ a *Birkhoff point of type* (p, q) if there exists a sequence $(\theta_n, r_n)_{n \in \mathbb{Z}}$ such that $(\theta_0, r_0) = (\theta, r)$, $\theta_{n+1} > \theta_n$, $(\theta_{n+q}, r_{n+q}) = (\theta_n + 2\pi, r_{n+q})$, $(\theta_{n+q}, r_{n+q}) = F(\theta_n, r_n)$ $(n \in \mathbb{N})$. One has *Birkhoff's twist theorem* [16].

LEMMA 18. *Let $\phi : A \to A$ be an area-preserving monotone twist homeomorphism and*

$$\frac{p}{q} \in \left[\alpha_a(\phi), \alpha_b(\phi)\right],$$

with p, q relatively prime. Then ϕ has at least two Birkhoff points of type (p, q).

The situation is different for α irrational.

DEFINITION 16. A *Mather set of rotation number* α for F is a closed invariant set for F with representation $\theta = u(\tau), r = v(\tau)$, where u is monotone increasing, $u - Id$ and v are 1-periodic (not necessarily continuous !), and

$$u(\tau + \alpha) = u(\tau) + f\big(u(\tau), v(\tau)\big), \qquad v(\tau + \alpha) = g\big(u(\tau), v(\tau)\big).$$

One has the *Aubry–Mather's twist theorem* [7,69].

LEMMA 19. *Let $\phi : A \to A$ be an area-preserving monotone twist homeomorphism and let $\alpha \in [\alpha_a(\phi), \alpha_b(\phi)]$ be irrational. Then there exists an invariant Mather set Γ_α with rotation number α. Furthermore, Γ_α is a subset of a closed curve $r = \gamma(\theta)$ where γ is 2π-periodic and Lipschitz continuous, i.e., $v(\tau) = \gamma(u(\tau))$.*

When u and v are continuous, Γ_α defines a Lipschitz continuous invariant curve. When u and v have countably many discontinuities, Γ_α can be seen as a Cantor set on the curve $r = \gamma(\theta)$.

5.2. *Chaotic dynamics*

We first recall the definition of chaotic dynamics.

DEFINITION 17. Equation (23) displays *chaotic dynamics* if
- (i) the solutions of Equation (23) depend sensitively on the initial conditions;
- (ii) Equation (23) has infinitely many periodic solutions with diverging periods;
- (iii) Equation (23) has an uncountable number of bounded, nonperiodic solutions;
- (iv) the Poincaré map associated to Equation (23) has positive topological entropy.

The following definition is due to Serra, Tarallo and Terraccini [121]. Without loss of generality we take $T = 1$. Let $n \geqslant 1$ be an integer, $h \in \widetilde{C}_1$, and

$$A_{h,n} : H_n^1 \to \mathbb{R}, \qquad y \mapsto \int_0^n \left[\frac{y'^2(t)}{2} + a \cos y(t) + h(t) y(t) \right] \mathrm{d}t.$$

DEFINITION 18. y_0 and y_1 are *consecutive minimizers* of A_h on H_n^1 if
- (1) $A_{h,n}(y_0) = A_{h,n}(y_1) = \min_{H_n^1} A_{h,n} := c_{h,n}$.
- (2) $y_0(t) < y_1(t)$ for all $t \in [0, n]$.
- (3) $y \in H_n^1 \cap [y_0, y_1]$ and $A_{h,n}(y) = c_{h,n}$ imply $y \in \{y_0, y_1\}$.

This notion is weaker than the isolatedness or nondegeneracy of minimizers in the variational sense.

Write $y(\pm\infty) = u$ if $\lim_{t\to\pm\infty}(y(t) - u(t)) = 0$, and, following an idea of Rabinowitz [115], define the functional J by

$$J(y) = \sum_{j\in\mathbb{Z}} \left(\int_{jn}^{(j+1)n} \left[\frac{y'^2(t)}{2} + a \cos y(t) + h(t) y(t) \right] \mathrm{d}t - c_{h,n} \right) \tag{51}$$

over the classes of functions

$$\Gamma(y_0, y_1) = \{y \in H^1_{\mathrm{loc}}\colon\ y(-\infty) = y_0,\ y(+\infty) = y_1\},$$
$$\Gamma(y_1, y_0) = \{y \in H^1_{\mathrm{loc}}\colon\ y(-\infty) = y_1,\ y(+\infty) = y_0\}. \tag{52}$$

Define

$$\mathcal{I} =]y_0(0), y_1(0)[,$$
$$\mathcal{S}(y_0, y_1) = \Big\{y(0) \in \mathcal{I}\colon : y \in \Gamma(y_0, y_1),\ J(y) = \min_{\Gamma(y_0,y_1)} J\Big\},$$
$$\mathcal{S}(y_1, y_0) = \Big\{y(0) \in \mathcal{I}\colon\ y \in \Gamma(y_1, y_0),\ J(y) = \min_{\Gamma(y_1,y_0)} J\Big\}.$$

The following result has been proved in [21].

THEOREM 39. *Let* $h \in \widetilde{C}_1$. *If* $A_{h,n}$ *has two consecutive minimizers* y_0 *and* y_1 *and if*

$$\mathcal{S}(y_0, y_1) \neq \mathcal{I} \quad \text{and} \quad \mathcal{S}(y_1, y_0) \neq \mathcal{I}, \tag{53}$$

then Equation (23) *displays chaotic dynamics.*

A first important step in proving the above theorem, which is interesting in itself, is the following existence result for *heteroclinic* (or *one-bump*) solutions between two consecutive minimizers y_0 and y_1 on H^1_1.

PROPOSITION 6. *Let* $c_0 = \inf_{\Gamma(y_0,y_1)} J,\ c_1 = \inf_{\Gamma(y_1,y_0)} J$. *There exists* $q_0 \in \Gamma(y_0, y_1)$ *and* $q_1 \in \Gamma(y_1, y_0)$ *such that* $J(q_0) = c_0,\ J(q_1) = c_1$. *The functions* q_0 *and* q_1 *solve Equation* (23).

A second step is the existence of multi-bump solutions. One makes the convention that the indices in $y, c, q, \ldots$ have to be taken mod 2.

PROPOSITION 7. *Assume that* y_0 *and* y_1 *are two consecutive minimizers of* A_h *over* H^1_1 *and that condition* (53) *holds. Then, for every sufficiently small* $\delta > 0$, *there exists* $m = m(\delta) \in \mathbb{N}$ *such that for every sequence* $(p_i)_{i\in\mathbb{Z}}$ *such that* $p_{i+1} - p_i \geqslant 4m$, *and for every* $j, k \in \mathbb{Z}$ *with* $j < k$, *there exists a classical solution* q *of Equation* (23) *satisfying*

$$y_0(t) < q(t) < y_1(t) \quad (t \in \mathbb{R}),$$
$$q(-\infty) = y_j, \qquad q(+\infty) = y_{k+1},$$

and, for all $i = j, \ldots, k$,

$$\big|q(p_i - m) - y_i(p_i - m)\big| \leqslant \delta, \qquad \big|q(p_i + m) - y_{i+1}(p_i + m)\big| \leqslant \delta.$$

From this result, one deduces the existence of solutions with infinitely many bumps.

PROPOSITION 8. *Assume that* y_0 *and* y_1 *are two consecutive minimizers of* A_h *over* H_1^1 *and that condition* (53) *holds. Then, for every sufficiently small* $\delta > 0$, *there exists* $m = m(\delta) \in \mathbb{N}$ *such that for every sequence* $(p_i)_{i \in \mathbb{Z}}$ *such that* $p_{i+1} - p_i \geqslant 4m$, *there exists a classical solution* q *of Equation* (23) *satisfying*

$$y_0(t) < q(t) < y_1(t) \quad (t \in \mathbb{R}),$$

and

$$\left|q(p_i - m) - y_i(p_i - m)\right| \leqslant \delta, \quad \left|q(p_i + m) - y_{i+1}(p_i + m)\right| \leqslant \delta \quad (i \in \mathbb{Z}).$$

Notice that, by choosing the points p_i in a periodic way, a slight modification of the argument gives the existence of infinitely many periodic solutions of Equation (23) with arbitrary large period. Notice also that is has been shown in [21] that condition (53) is weaker than the standard *transversality* assumption.

A natural question is to see for which h is assumption (53) satisfied. A result in this direction has been given in [22,23]. Let

$$\widetilde{C} := \bigcup_{n \in \mathbb{N}} \widetilde{C}_n,$$

endowed with the L^∞-topology.

THEOREM 40. *There exists a dense subset* $\mathcal{H}$ *of* $\widetilde{C}$ *such that for every* $h \in \mathcal{H}$, *Equation* (23) *displays chaotic dynamics. Moreover, for all* $n \in \mathbb{N}$, $\mathcal{H} \cap \widetilde{H}_n^1$ *is open in* $\widetilde{H}_n^1$.

To prove this theorem, one first refers to Theorem 10 to prove the result in a simplified setting. If $h \in \widetilde{C}_n$, $y \in H_n^1$ is a *true minimizer* for $A_{h,n}$ over H_n^1 if $A_{h,n}(y) = c_{h,n}$ and $\bar{y} \in [0, 2\pi[$.

PROPOSITION 9. *For each positive integer* n, *the set* $\mathcal{H}_n$ *of forcing terms* $h \in \widetilde{C}_n$, *such that* $A_{h,n}$ *has only one nondegenerate true minimizer over* H_n^1 *is dense in* H_n^1, *and open in* H_{kn}^1 *for any fixed* $k \in \mathbb{N}$.

Taking then $h \in \mathcal{H}_n$, fixing $k \in \mathbb{N}$ and $h_1 \in \widetilde{H}_{kn}^1$, one considers the perturbed equation

$$y'' + a \sin y = h(t) + \varepsilon h_1(t) \tag{54}$$

for ε small, y_0^ε its true minimizer and $y_1^\varepsilon = y_0^\varepsilon + 2\pi$. Regularity estimates are obtained on some associated heteroclinic orbits, using in particular a Lyapunov–Schmidt argument, and this regularity is used to compute some analogue of a primitive of the Melnikov function, in the line of [3].

Similar results have been obtained by Offin and Yu Hongfan [97].

6. Bounded forcing

6.1. *Bounded functions and their averages*

T-periodic functions are special cases of bounded functions.

DEFINITION 19. The function $p:\mathbb{R} \to \mathbb{R}$ is called *bounded* if there exists $M > 0$ such that, for all $t \in \mathbb{R}$, $|p(t)| \leqslant M$.

The concept of mean value of a T-periodic function has been extended by Tineo [124] to some classes of functions containing the bounded ones.

DEFINITION 20. Let $p = p^* + p^{**}:\mathbb{R} \to \mathbb{R}$ be continuous,with p^{**} bounded and p^* having a bounded primitive. The *lower* (respectively *upper*) *average* of p is defined by

$$\bar{p}_L = \lim_{r\to+\infty} \inf_{t-s\geqslant r} \frac{1}{t-s}\int_s^t p(u)\,\mathrm{d}u,$$

(respectively

$$\bar{p}_U = \lim_{r\to+\infty} \sup_{t-s\geqslant r} \frac{1}{t-s}\int_s^t p(u)\,\mathrm{d}u).$$

It is easy to verify that

$$-\infty < \bar{p}_L \leqslant \bar{p}_U < +\infty,$$
$$\bar{p}_L = \bar{p}_L^{**}, \qquad \bar{p}_U = \bar{p}_U^{**},$$

and that, if p is continuous and T-periodic,

$$\bar{p}_L = \bar{p}_U = \bar{p}.$$

Notice also that if p is continuous and T-periodic, then $\bar{p} = 0$ if and only if p has a T-periodic primitive.

6.2. *A necessary condition for the existence of bounded solutions*

Let $g:\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be bounded over $\mathbb{R}\times[-r, r]$ for each $r > 0$ and continuous, and consider the differential equation

$$y'' + cy' = g(t, y). \tag{55}$$

DEFINITION 21. A solution $y:\mathbb{R} \to \mathbb{R}$ of Equation (55) is said to be *bounded* if y and y' are bounded.

Consider the dissipative forced pendulum-type equation

$$y'' + cy' + a\sin y = h(t), \tag{56}$$

where $a > 0$, $c \geqslant 0$, and $h : \mathbb{R} \to \mathbb{R}$ is continuous and bounded. It is easy to extend Proposition 1 to bounded forcings and bounded solutions.

PROPOSITION 10. *If Equation* (56) *has a bounded solution, then*

$$-a \leqslant \bar{h}_L \leqslant \bar{h}_U \leqslant a. \tag{57}$$

To show this result, one notices that, if y is a bounded solution of (56), one has, for each $t > s$,

$$\frac{y'(t) - y'(s)}{t - s} + c\frac{y(t) - y(s)}{t - s} + \frac{a}{t - s}\int_s^t \sin y(\tau)\,\mathrm{d}\tau = \frac{1}{t - s}\int_s^t h(\tau)\,\mathrm{d}\tau,$$

and, given $\varepsilon > 0$, one can find T_0 such that, for all $T \geqslant T_0$ and all t, s such that $t - s \geqslant T$, one has

$$-\varepsilon \leqslant \frac{y(t) - y(s)}{t - s} \leqslant \varepsilon, \qquad -\varepsilon \leqslant \frac{y'(t) - y'(s)}{t - s} \leqslant \varepsilon.$$

Consequently, for $t - s \geqslant T$, we get

$$-(1 + c)\varepsilon - a \leqslant \frac{1}{t - s}\int_s^t h(\tau)\,\mathrm{d}\tau \leqslant (1 + c)\varepsilon + a,$$

and (57) follows easily.

6.3. *Sufficient conditions for the existence of bounded solutions*

The following existence result for bounded solutions of Equation (55) goes back Opial [98] (see also [80]).

LEMMA 20. *If there exists $r_- < r_+$ such that*

$$g(t, r_-) \leqslant 0 \leqslant g(t, r_+)$$

for all $t \in \mathbb{R}$, then Equation (55) *has at least one bounded solution such that*

$$r_- \leqslant y(t) \leqslant r_+$$

for all $t \in \mathbb{R}$.

The first easy consequence for Equation (56) is the one-dimensional case of a result for elliptic partial differential equations due to Fournier, Szulkin and Willem [48].

THEOREM 41. *If* $c \geqslant 0$ *and if*

$$-a \leqslant h(t) \leqslant a, \tag{58}$$

for all $t \in \mathbb{R}$, *Equation* (56) *has at least one bounded solution* y *such that*

$$\frac{\pi}{2} \leqslant y(t) \leqslant \frac{3\pi}{2}$$

for all $t \in \mathbb{R}$.

REMARK 1. *It is clear that* condition (58) implies condition (57).

Another existence result, first proved in [81], follows from applying Lemma 20 to an equivalent formulation for the forced pendulum problem together with the following result of Ortega on bounded solutions of second order linear equations [103] (see also [80]). We define as usual the *oscillation* $\operatorname{osc}_{\mathbb{R}} z$ of the function $z : \mathbb{R} \to \mathbb{R}$ by

$$\operatorname{osc}_{\mathbb{R}} z = \sup_{\mathbb{R}} z - \inf_{\mathbb{R}} z.$$

LEMMA 21. *If* $c > 0$ *and* $h : \mathbb{R} \to \mathbb{R}$ *is continuous, then equation*

$$y'' + cy' = h(t) \tag{59}$$

has a bounded solution if and only if h *has a bounded primitive. If it is the case, any bounded solution* y *of Equation* (59) *verifies the inequality*

$$\operatorname{osc}_{\mathbb{R}} y \leqslant \frac{1}{c} \operatorname{osc}_{\mathbb{R}} H, \tag{60}$$

where

$$H(t) = \int_0^t h(s)\, \mathrm{d}s.$$

Furthermore, for each h *with bounded primitive, Equation* (59) *has a unique* (*symmetrized*) *bounded solution* H_c *such that*

$$\sup_{\mathbb{R}} H_c = -\inf_{\mathbb{R}} H_c = \frac{1}{2} \operatorname{osc}_{\mathbb{R}} H_c. \tag{61}$$

THEOREM 42. *If* $c > 0$, $h = h^* + h^{**}$ *where* h^{**} *is bounded and* h^* *has a bounded primitive over* $\mathbb{R}$, *and if inequalities*

$$\operatorname{osc}_{\mathbb{R}} H_c^* \leqslant \pi, \tag{62}$$

and

$$\|h^{**}\|_\infty \leqslant a \cos\left(\frac{\operatorname{osc}_{\mathbb{R}} H_c^*}{2}\right) \tag{63}$$

hold, where H_c^* *is the symmetrized bounded solution of* $y'' + cy' = h^*(t)$, *then Equation* (56) *has at least one solution* y *such that*

$$\frac{\pi}{2} + H_c^*(t) \leqslant y(t) \leqslant \frac{3\pi}{2} + H_c^*(t),$$

for all $t \in \mathbb{R}$. *When* $c = 0$, *the above result holds if* $h^{**} = 0$, $h = h^*$ *has a second primitive* H^1 *bounded over* $\mathbb{R}$ *and* H_c^* *is replaced by* H^1 *in* (67).

To prove this theorem, one sees that y is a bounded solution of Equation (56), if and only if z defined by

$$y(t) = z(t) + H_c^*(t),$$

is a bounded solution of equation

$$z'' + cz' + a \sin\bigl(z + H_c^*(t)\bigr) = h^{**}(t). \tag{64}$$

One checks easily that the conditions of Lemma 20 hold for $r_- = \pi/2$ and $r_+ = 3\pi/2$.

REMARK 2. If follows from inequality (60) that, given $h = h^* + h^{**}$, condition (62) holds as soon as

$$c \geqslant \frac{1}{\pi} \operatorname{osc}_{\mathbb{R}} H^*.$$

Consequently, when h has a bounded primitive, Equation (56) has a bounded solution for all sufficiently large c. The question remains open for other values of c.

6.4. *Local uniqueness of bounded solutions*

By straightening the assumptions of Theorems 41 and 42, one obtains the local uniqueness of the obtained bounded solutions by using a maximum principle proved in [19].

LEMMA 22. *Let* $\alpha > 0$, $\beta \in \mathbb{R}$ *and* $\gamma : \mathbb{R} \to \mathbb{R}$ *bounded and continuous. If the function* $r \in C^2(\mathbb{R}, \mathbb{R})$ *is bounded together with its first two derivatives and satisfies the differential inequality*

$$r''(t) \geqslant \gamma(t) r'(t) + \alpha r(t) - \beta,$$

then

$$\sup_{\mathbb{R}} r \leqslant \frac{\beta}{\alpha}.$$

LEMMA 23. *Let* $c \in \mathbb{R}$ *and* $f : \mathbb{R} \to \mathbb{R}$ *be bounded, continuous and such that*

$$f(t) \leqslant -\delta < 0$$

for all $t \in \mathbb{R}$. *Then the unique bounded solution of equation*

$$y'' + cy' + f(t) y = 0$$

is the trivial one.

THEOREM 43. *If* $c \geqslant 0$ *and if*

$$\|h\|_\infty < a, \tag{65}$$

holds, there exists $\varepsilon > 0$ *such that Equation* (56) *has a unique solution* y *such that*

$$\frac{\pi}{2} + \varepsilon \leqslant y(t) \leqslant \frac{3\pi}{2} - \varepsilon \tag{66}$$

for all $t \in \mathbb{R}$.

This result follows from the fact that for some sufficiently small $\varepsilon > 0$, $r_- = \frac{\pi}{2} + \varepsilon$ and $r_+ = \frac{3\pi}{2} - \varepsilon$ satisfy the condition of Lemma 20 and that $\cos u \leqslant -\delta$ for some $\delta > 0$ when $u \in [r_-, r_+]$. The uniqueness then follows from Lemma 23.

THEOREM 44. *If* $c > 0$, $h = h^* + h^{**}$ *where* h^{**} *is bounded and* h^* *has a bounded primitive over* $\mathbb{R}$, *and if inequalities*

$$\operatorname{osc}_{\mathbb{R}} H_c^* < \frac{\pi}{2}, \tag{67}$$

$$\|h^{**}\|_\infty \leqslant \frac{a\sqrt{2}}{2}\left[\sin\left(\frac{\operatorname{osc}_{\mathbb{R}} H_c^*}{2}\right) + \cos\left(\frac{\operatorname{osc}_{\mathbb{R}} H_c^*}{2}\right)\right], \tag{68}$$

hold, then there exists $\varepsilon > 0$ *such that Equation* (56) *has a unique solution* y *satisfying the inequality*

$$\frac{\pi}{2} + \varepsilon \leqslant y(t) \leqslant \frac{3\pi}{2} - \varepsilon, \tag{69}$$

for all $t \in \mathbb{R}$. *When* $c = 0$, *the above result holds if* $h^{**} = 0$, $h = h^*$ *has a second primitive* H^1 *bounded over* $\mathbb{R}$ *and* H_c^* *is replaced by* H^1 *in* (67).

The proof follows lines similar to that of Theorem 43.

7. Almost periodic forcings

7.1. *Almost periodic functions*

An interesting intermediate class between the bounded and the periodic functions is the class of almost periodic solutions in the sense of H. Bohr [44].

DEFINITION 22. $f : \mathbb{R} \to \mathbb{R}$ is (Bohr)-*almost periodic* if, for each $\varepsilon > 0$, there exists $L > 0$ such that any interval of length L contains at least some τ such that $\forall t \in \mathbb{R}$, $|f(t+\tau) - f(t)| < \varepsilon$.

Such a function is necessarily bounded and uniformly continuous over $\mathbb{R}$. The following result gives an alternative definition of the space $AP(\mathbb{R})$ of (Bohr)-almost periodic functions. Let us denote by $TP(\mathbb{R})$ the space of real trigonometric polynomials

$$p_N(t) = \sum_{k=-N}^{N} p_k \exp(i\lambda_k t),$$

where $p_k = \bar{p}_{-k}$ and $\lambda_k \in \mathbb{R}$ $(-N \leqslant k \leqslant N)$.

LEMMA 24. *The space* $AP(\mathbb{R})$ *is the closure of* $TP(\mathbb{R})$ *for the uniform norm over* $\mathbb{R}$.

If $f \in AP(\mathbb{R})$ and $\lambda \in \mathbb{R}$, the limit

$$f_\lambda := \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t) \exp(-i\lambda t)\, \mathrm{d}t$$

exists, $\bar{f_\lambda} = f_{-\lambda}$, and the set Λ of $\lambda \in \mathbb{R}$ for which $f_\lambda \neq 0$ is at most countable. The series

$$\sum_{\lambda : f_\lambda \neq 0} f_\lambda \exp(i\lambda t) \tag{70}$$

is called the *Fourier series* associated to f. The Fourier coefficient

$$f_0 = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(t)\,\mathrm{d}t$$

is called the *mean value* of f. If f has an almost-periodic primitive, then $f_0 = 0$, but, in contrast to the periodic case, an almost periodic function with mean value zero needs not to have an almost-periodic primitive. It is only the case if its primitive is itself bounded.

More general classes of almost periodic functions have been defined by Besicovitch [14]. If $p \in TP(\mathbb{R})$ define $\|p\|_{B^2}$ by

$$\|p\|_{B^2} = \left[\limsup_{T\to\infty}\left\{\frac{1}{T}\int_0^T \left|p(t)\right|^2 \mathrm{d}t\right\}\right]^{1/2},$$

and define $B^2(\mathbb{R})$ to be the space of equivalence classes of functions $f : \mathbb{R} \to \mathbb{R}$ such that $\lim_{n\to\infty} \|f - p_n\|_{B^2} = 0$ for some sequence (p_n) in $TP(\mathbb{R})$, under the equivalence relation

$$f \sim g \quad \Leftrightarrow \quad \|f - g\|_{B^2} = 0. \tag{71}$$

If f and g belong to $B^2(\mathbb{R})$, and $\lambda \in \mathbb{R}$, one can show that the limits

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \left|f(t)\right|^2 \mathrm{d}t,$$

$$\langle f, g\rangle_{B^2} := \lim_{T\to\infty} \frac{1}{T}\int_0^T f(t)g(t)\,\mathrm{d}t,$$

$$f_\lambda := \lim_{T\to\infty} \frac{1}{T}\int_0^T f(t)\exp(-i\lambda t)\,\mathrm{d}t$$

exist and that $f_\lambda \neq 0$ on an at most countable set, so that the *Fourier series* (70) of f is well defined. Because of the fact that two functions in the same equivalence class for (71) can differ on a set of positive and even of infinite measure, the use of $B^2(\mathbb{R})$ in the study of differential equations gives rather weak existence assertions. Work in this direction for Equation (56) has been done by Belley, Fournier and Saadi Drissi [11,12], who have also considered some subspaces of $B^2(\mathbb{R})$ whose functions behave more like periodic ones.

7.2. *Almost periodic solutions*

Combining some results on the existence and uniqueness of bounded solutions over $\mathbb{R}$ with Amerio's criterion on the existence of almost periodic solutions (see, e.g., [44]), Fink [43] has given in 1968 some partial extension of the *method of upper and lower solutions* to almost periodic solutions. A special case of his results is the following proposition.

LEMMA 25. *Let* $c \in \mathbb{R}$, $g \in C^1(\mathbb{R}, \mathbb{R})$ *and* h *be continuous and almost periodic. Assume that there exist* $r_- < r_+$ *such that* $g'(x) > 0$ *for all* $x \in [r_-, r_+]$, *and*

$$g(r_-) + h(t) \leqslant 0 \leqslant g(r_+) + h(t)$$

for all $t \in \mathbb{R}$. *Then equation*

$$y'' + cy' = g(y) + h(t)$$

has a unique almost periodic solution y *such that* $a \leqslant y(t) \leqslant b$ *for all* $t \in \mathbb{R}$.

Notice that, as shown by Ortega and Tarallo [109], the monotonicity condition on g cannot be dropped in Lemma 25.

Lemma 25 implies the following existence theorem [81], also proved independently by Fournier, Szulkin and Willem [48] as a special case of a more general result for elliptic partial differential equations. When $c = 0$, Theorem 45 generalizes an earlier approximate solvability result of Blot [17] for Equation (23), based upon variational techniques and convex analysis, which only provides the existence for a dense subset of forcing functions h.

THEOREM 45. *For each* $c \geqslant 0$ *and each* $h \in AP(\mathbb{R})$ *such that* $\|h\|_\infty < a$, *Equation* (56) *has a unique solution* $y \in AP(\mathbb{R})$ *such that* $\pi/2 < y(t) < 3\pi/2$ *for all* $t \in \mathbb{R}$.

Indeed, the condition upon $\|h\|_\infty$ implies the existence of $\varepsilon > 0$ such that $a = \frac{\pi}{2} + \varepsilon$ and $b = \frac{3\pi}{2} - \varepsilon$ satisfy the conditions of Lemma 13.

Similar arguments applied to the equivalent formulation of the forced pendulum equation lead to the following existence theorem, first proved in [81]. When $c = 0$, Theorem 46 generalizes an earlier approximate solvability result of Blot [18] for Equation (23), based upon variational techniques and convex analysis, which gives existence for a dense subset of forcing functions h only. Subsequently, Blot and Pennequin [20] have extended Blot's result to the case of quasi-periodic forcing terms.

THEOREM 46. *If* $c > 0$, $h = h^* + h^{**}$ *where* h^{**} *is almost periodic and* h^* *has an almost periodic primitive, and if conditions* (67) *and* (68) *are satisfied, then there exists* $\varepsilon > 0$ *such that Equation* (56) *has a unique almost periodic solution verifying inequality* (69). *If* $c = 0$, *and* $h \in C$ *has an almost periodic second primitive* H^1 *satisfying* (67) *with* H_c^* *replaced by* H^1, *then the same conclusion holds.*

The following result of Lagrange unstability-type for almost periodic forcings has been proved by Hai Huang [58].

THEOREM 47. *Equation* (23) *possesses infinitely many unbounded solutions on a cylinder* $S^1 \times \mathbb{R}$ *for any almost periodic* h *with nonvanishing mean value.*

References

[1] F. Alessio, M. Calanchi and E. Serra, *Complex dynamics in a class of reversible equations,* Nonlinear Analysis and its Applications to Differential Equations (Lisbon, 1998), Birkhäuser, Boston (2001), 147–159.

[2] J.M. Alonso, *Nonexistence of periodic solutions for a damped pendulum equation,* Differential Integral Equations **10** (1997), 1141–1148.

[3] A. Ambrosetti and M. Badiale, *Homoclinics: Poincaré–Melnikov type results via a variational approach*, Ann. Inst. H. Poincaré, Anal. Non Linéaire **15** (1998), 233–252.

[4] A. Ambrosetti and P. Rabinowitz, *Dual variational methods in critical point theory and applications,* J. Functional Anal. **14** (1973), 349–381.

[5] Z. Amine and R. Ortega, *Existence of asymptotically stable periodic solutions of a forced equation of Liénard type,* Nonlinear Anal. **22** (1994), 993–1003.

[6] A.A. Andronov, A.A. Vitt and S.E. Khaikin, *Theory of Oscillators,* Moscow (1959) (in Russian). English translation: Pergamon, Oxford (1966) and Dover, New York (1987).

[7] S. Aubry and P.Y. Le Daeron, *The discrete Fraenkel–Kontorova model and its extensions I,* Physica D **8** (1983), 381–422.

[8] V. Bangert, *Mather sets for twist maps and geodesics on tori,* Dynam. Report. **1** (1988), 1–56.

[9] P.W. Bates, *Reduction theorems for a class of semilinear equations at resonance,* Proc. Amer. Math. Soc. **24** (1982) 73–78.

[10] J.M. Belley, G. Fournier and J. Hayes, *Existence of almost periodic weak type solutions for the conservative forced pendulum equation*, J. Differential Equations **124** (1996), 205–224.

[11] J.M. Belley, G. Fournier and K. Saadi Drissi, *Almost periodic weak solutions to forced pendulum type equations without friction*, Aequationes Math. **44** (1992), 100–108.

[12] J.M. Belley, G. Fournier and K. Saadi Drissi, *Solutions faibles presque périodiques d'équation différentielle du type du pendule forcé*, Bull. Cl. Sci. Acad. R. Belgique **3** (1992), 173–186.

[13] J.M. Belley and F. Vaillancourt, *Variational methods and the nonconservative pendulum equation with periodic forcing*, to appear.

[14] A.S. Besicovitch, *Almost Periodic Functions*, Cambridge Univ. Press, Cambridge (1932). Reprinted: Dover, New York (1954).

[15] G.D. Birkhoff, *Proof of Poincaré's last geometric theorem*, Trans. Amer. Math. Soc. **14** (1913), 14–22.

[16] G.D. Birkhoff, *On the periodic motions of dynamical systems*, Acta Math. **50** (1927), 359–379.

[17] J. Blot, *Une méthode hilbertienne pour les trajectoires presque périodiques*, C. R. Acad. Sci. Paris, Ser. I Math. **313** (1991), 487–490.

[18] J. Blot, *Almost periodically forced pendulum*, Funkcialaj Ekvacioj **36** (1993), 235–250.

[19] J. Blot, P. Cieutat and J. Mawhin, *Almost periodic oscillations of monotone second order systems*, Adv. Differential Equations **2** (1997), 693–714.

[20] J. Blot and D. Pennequin, *Spaces of quasi-periodic functions and oscillations in differential equations*, Acta Appl. Math. **65** (2001), 83–113.

[21] E. Bosetto and E. Serra, *A variational approach to chaotic dynamics in periodically forced nonlinear oscillators*, Ann. Inst. H. Poincaré, Anal. Non Linéaire **17** (2000), 673–709.

[22] E. Bosetto, E. Serra and S. Terracini, *Density of chaotic dynamics in periodically forced pendulum-type equations*, Atti Accad. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Mat. Appl. (9) **12** (2001), 107–113.

[23] E. Bosetto, E. Serra and S. Terracini, *Generic-type results for chaotic dynamics in equations with periodic forcing terms*, J. Differential Equations **180** (2002), 99–124.

[24] R. Bott, *On the intersection of a closed geodesics and Sturm intersection theory*, Comm. Pure Appl. Math. **9** (1956), 176–206.

[25] M. Calanchi and E. Serra, *Homoclinic solutions to periodic motions in a class of reversible equations*, Calculus Variations Partial Differential Equations **9** (1999), 157–184.

[26] M. Calanchi and M. Tarallo, *On the count and the classification of periodic solutions to forced pendulum type equations*, Differential Integral Equations **11** (1998), 503–520.

[27] J. Campos, E.N. Dancer and R. Ortega, *Dynamics in the neighbourhood of a continuum of fixed points*, Ann. Mat. Pura Appl. (4) **180** (2002), 483–492.

[28] A. Cañada and D. Ruiz, *Resonant problems with multidimensional kernel and periodic nonlinearities*, Differential Integral Equations **16** (2003), 499–512.

[29] A. Cañada and D. Ruiz, *Resonant nonlinear boundary value problems with almost periodic nonlinearity*, Bull. Belgian Math. Soc. **9** (2002), 193–204.

[30] A. Cañada and A. Ureña, *An asymptotic result on the solvability of pendulum-type equations with linear damping and homogeneous Dirichlet conditions*, Proc. USA–Chile Workshop on Nonlinear Analysis (Viña del Mar-Valparaiso, 2000), Electron. J. Differential Equations Conf. **6** (2001), 55–64.

[31] A. Cañada and A. Ureña, *Solvability of some pendulum-type equations with linear damping and homogeneous Dirichlet conditions*, Nonlinear Anal. **49** (2002), 1149–1165.

[32] A. Castro, *Periodic solutions of the forced pendulum equation*, Differential Equations, Ahmad, Keener, Lazer, eds., Academic Press, New York (1980), 149–160.

[33] J. Čepička, P. Drábek and J. Jenšiková, *On the stability of periodic solutions of the damped pendulum equation*, J. Math. Anal. Appl. **209** (1997), 712–723.

[34] K.C. Chang, *On the periodic nonlinearity and the multiplicity of solutions*, Nonlinear Anal. **13** (1989), 527–537.

[35] A. Chenciner, *La dynamique au voisinage d'un point fixe elliptique conservatif : de Poincaré et Birkhoff à Aubry et Mather*, Séminaire Bourbaki No. 622, Astérisque, Vols. 121–122 (1985), 147–170.

[36] E.N. Dancer, *On the use of asymptotics in nonlinear boundary value problems*, Ann. Mat. Pura Appl. (4) **131** (1982), 167–185.

[37] E.N. Dancer and R. Ortega, *The index of Lyapunov stable fixed points in two dimensions*, J. Dynamics Differential Equations **6** (1994), 631–637.

[38] J. Denzler, *Mather sets for plane Hamiltonian systems*, J. Appl. Math. Phys. (ZAMP) **38** (1987), 791–812.

[39] Weiyue Ding, *A generalization of the Poincaré–Birkhoff theorem*, Proc. Amer. Math. Soc. **88** (1983), 341–346.

[40] F. Donati, *Sur l'existence de quatre solutions périodiques pour l'équation du pendule forcé*, C. R. Acad. Sci. Paris I **317** (1993), 667–672.

[41] F. Donati, *Some remarks about periodic solutions to the forced pendulum equation*, Differential Integral Equations **8** (1995), 141–149.

[42] Yihong Du, *A deformation lemma and some critical point theorems*, Bull. Austral. Math. Soc. **43** (1991), 161–168.

[43] A.M. Fink, *Uniqueness theorems and almost periodic solutions to second order differential equations*, J. Differential Equations **4** (1968), 543–548.

[44] A.M. Fink, *Almost Periodic Differential Equations*, Lecture Notes in Math., Vol. 377, Springer, Berlin (1974).

[45] A. Fonda and M. Willem, *Subharmonic oscillations of forced pendulum-type equations*, J. Differential Equations **81** (1989), 215–220.

[46] A. Fonda and F. Zanolin, *Periodic oscillations of forced pendulums with very small length*, Proc. Roy. Soc. Edinburgh A **127** (1997), 67–76.

[47] G. Fournier and J. Mawhin, *On periodic solutions of forced pendulum-like equations*, J. Differential Equations **60** (1985), 381–395.

[48] G. Fournier, A. Szulkin and M. Willem, *Semilinear elliptic equations in R^N with almost periodic or unbounded forcing term*, SIAM J. Math. Anal. **27** (1996), 1653–1660.

[49] S. Fučík, *Solvability of Nonlinear Equations and Boundary value Problems*, Reidel, Dordrecht (1980).

[50] R.E. Gaines and J. Mawhin, *Coincidence Degree and Nonlinear Differential Equations*, Lecture Notes in Math., Vol. 568, Springer, Berlin (1977).

[51] N. Ghoussoub and D. Preiss, *A general mountain pass principle for locating and classifying critical points*, Ann. Inst. H. Poincaré Anal. Non Linéaire **6** (1989), 321–330.

[52] P. Girg and F. Roca, *On the range of certain pendulum-type equation*, J. Math. Anal. Appl. **249** (2000), 127–149.

[53] L. Guillou, *Théorème de translation plane de Brouwer et généralisations du théorème de Poincaré–Birkhoff*, Topology **33** (1994), 331–351.

[54] P. Habets and P.J. Torres, *Some multiplicity results for periodic solutions of a Rayleigh differential equation*, Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal. **8** (2001), 335–351.

[55] G. Hamel, *Ueber erzwungene Schingungen bei endlischen Amplituden*, Math. Ann. **86** (1922), 1–13.

[56] Hai Huang, *Destruction of invariant tori in pendulum-type equations*, J. Differential Equations **146** (1998), 67–89.

[57] Hai Huang, *Generic existence of invariant Cantori in pendulum-type equations*, J. Math. Anal. Appl. **218** (1998), 379–394.

[58] Hai Huang, *Unbounded solutions of almost-periodically forced pendulum-type equations*, Acta Math. Sinica (Engl. Ser.) **17** (2001), 391–396.

[59] Hai Huang and Rong Yuan, *Boundedness of solutions and existence of invariant tori for generalized pendulum type equations*, Chinese Sci. Bull. **42** (1997), 1673–1675.

[60] R. Kannan and R. Ortega, *Periodic solutions of pendulum-type equations*, J. Differential Equations **59** (1985), 123–144.

[61] R. Kannan and R. Ortega, *An asymptotic result in forced oscillations of pendulum-type equations*, Appl. Anal. **22** (1986), 45–54.

[62] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge Univ. Press, Cambridge (1995).

[63] G. Katriel, *Many periodic solutions for pendulum-type equations*, Nonlinear Anal. **34** (1998), 687–699.

[64] G. Katriel, *Periodic solutions of the forced pendulum: exchange of stability and bifurcation*, J. Differential Equations **182** (2002), 1–50.

[65] M. Levi, *KAM theory for particles in periodic potential*, Ergodic Theory Dynamical Systems **10** (1990), 777–785.

[66] M. Levi, *Invariant foliations in forced oscillations*, Geometry and Analysis in Nonlinear Dynamics, Broer-Takens, ed., Longman, London (1992), 34–40.

[67] M. Levi and E. Zehnder, *Boundedness of solutions for quasiperiodic potentials*, SIAM J. Math. Anal. **26** (1995), 1233–1256.

[68] P. Martinez-Amores, J. Mawhin, R. Ortega and M. Willem, *Generic results for the existence of nondegenerate periodic solutions of some differential systems with periodic nonlinearities*, J. Differential Equations **91** (1991), 138–148.

[69] J.N. Mather, *Existence of quasi-periodic orbits for twist homeomorphisms of the annulus*, Topology **21** (1982), 457–467.

[70] J.N. Mather, *Modulus of continuity for Peierl's barrier*, Periodic Solutions of Hamiltonian Systems and Related Topics, P.H. Rabinowitz, et al., eds., NATO ASI Series C, Vol. 209, Reidel, Dordrecht (1987), 177–202.

[71] J.N. Mather, *Destruction of invariant circles*, Ergodic Theory Dynamical Systems **8*** (1988), 199–214.

[72] J.N. Mather, *Variational construction of orbits for twist diffeomorphisms*, J. Amer. Math. Soc. **4** (1991), 203–267.

[73] J.N. Mather and G. Forni, *Action minimizing orbits in Hamiltonian systems*, Transition to Chaos in Classical and Quantum Mechanics, Lecture Notes in Math., Vol. 1589, Springer, Berlin (1994), 92–186.

[74] J. Mawhin, *Topological Degree Methods in Nonlinear Boundary Value Problems*, Amer. Math. Soc., Providence, RI (1979).

[75] J. Mawhin, *Periodic oscillations of forced pendulum-like equations*, Lecture Notes in Math., Vol. 964, Springer, Berlin (1982), 458–476.

[76] J. Mawhin, *Points fixes, points critiques et problèmes aux limites*, Presses Univ. Montréal, Montréal (1985).

[77] J. Mawhin, *The forced pendulum: a paradigm for nonlinear analysis and dynamical systems*, Exposition Math. **6** (1988), 271–287.

[78] J. Mawhin, *Forced second order conservative systems with periodic nonlinearity*, Ann. Inst. H. Poincaré Anal. Non Linéaire Suppl. **6** (Special issue dedicated to J.J. Moreau) (1989), 415–434.

[79] J. Mawhin, *Nonlinear oscillations: one hundred years after Liapunov and Poincaré*, Z. Angew. Math. Mech. **73** (1993), T54–T62.

[80] J. Mawhin, *Bounded solutions of nonlinear ordinary differential equations*, Nonlinear Analysis and Boundary Value Problems for Ordinary Differential Equations, CISM Courses and Lectures, Vol. 371, Udine 1995, Zanolin, ed., Springer, Wien (1996), 121–147.

[81] J. Mawhin, *Remarques sur les solutions bornées ou presque périodiques de l'équation du pendule forcé*, Ann. Sci. Math. Québec **22** (1998), 213–224.

[82] J. Mawhin, *Seventy-five years of global analysis around the forced pendulum equation*, Equadiff 9 Proceedings, Masaryk Univ. Brno and Electronic Publ. House, Stony Brooks (1998), 115–145.

[83] J. Mawhin, *From Tricomi's equation for synchronous motors to the periodically forced pendulum*, Tricomi's Ideas and Contemporary Applied Mathematics (Roma–Torino, 1997), Atti dei Convegni Lincei **147**, Accad. Naz. Lincei, Roma (1998), 251–269.

[84] J. Mawhin, *Bounded and almost periodic solutions of nonlinear differential equations: variational vs. nonvariational approach*, Calculus of Variations and Differential Equations (Technion, 1998), A. Ioffe, S. Reich and I. Shafrir, eds., Res. Notes Math., Vol. 410, Chapman and Hall/CRC, Boca Raton (1999), 167–184.

[85] J. Mawhin and M. Willem, *Multiple solutions of the periodic boundary value problem for some forced pendulum-type equations*, J. Differential Equations **52** (1984), 264–287.

[86] J. Mawhin and M. Willem, *Critical Point Theory and Hamiltonian Systems*, Springer, New York (1989).

[87] J. Moser, *Stable and Random Motions in Dynamical Systems*, Ann. Math. Studies, Vol. 77, Princeton Univ. Press, Princeton (1973).

[88] J. Moser, *Monotone twist mappings and the calculus of variations*, Ergodic Theory Dynamical Systems **6** (1986), 401–413.

[89] J. Moser, *Minimal solutions of variational problems on the torus*, Ann. Inst. H. Poincaré, Anal. Non Linéaire **3** (1986), 229–272.

[90] J. Moser, *Recent developments in the theory of Hamiltonian systems*, SIAM Rev. **28** (1986), 459–485.

[91] J. Moser, *Breakdown of stability*, Lecture Notes in Phys., Vol. 247, Springer, Berlin (1986), 492–518.

[92] J. Moser, *Minimal foliations on a torus*, Topics in Calculus of Variations, CIME Montecatini 1987, Lecture Notes in Math., Vol. 1365, Springer (1989), 62–99.

[93] J. Moser, *A stability theory for minimal foliations on a torus*, Ergodic Theory Dynamical Systems **8*** (1988), 251–288.

[94] J. Moser, *Quasi-periodic solutions of nonlinear elliptic partial differential equations*, Bol. Soc. Brasil. Mat. **20** (1989), 29–45.

[95] J. Moser, *Selected Chapters in the Calculus of Variations*, Lectures in Math. ETH Zürich, Birkhäuser, Basel (2003).

[96] D. Offin, *Subharmonic oscillations for forced pendulum type equations*, Differential Integral Equations **3** (1990), 965–972.

[97] D. Offin and Yu Hongan, *Homoclinic orbits in the forced pendulum system*, Conservative Systems and Quantum Chaos (Waterloo, ON, 1992), Fields Inst. Commun., Vol. 8, Amer. Math. Soc., Providence, RI (1996) 1–18.

[98] Z. Opial, *Sur les intégrales bornées de l'équation* $u'' = f(t, u, u')$, Ann. Polon. Math. **4** (1958), 314–324.

[99] R. Ortega, *A counterexample for the damped pendulum equation*, Bull. Cl. Sci. Acad. R. Belgique (5) **73** (1987), 405–409.

[100] R. Ortega, *Stability and index of periodic solutions of an equation of Duffing type*, Boll. Un. Mat. Ital. (7) **3B** (1989), 533–546.

[101] R. Ortega, *Topological degree and stability of periodic solutions for certain differential equations*, J. London Math. Soc. (2) **42** (1990) 505–516.

[102] R. Ortega, *Some applications of the topological degree to stability theory*, Topological Methods in DEs and Inclusions, Kluwer, Dordrecht (1995), 377–409.

[103] R. Ortega, *A boundedness result of Landesman–Lazer type*, Differential Integral Equations **8** (1995), 729–734.

[104] R. Ortega, *A forced pendulum equation with many periodic solutions*, Rocky Mountain J. Math. **27** (1997), 861–876.

[105] R. Ortega, *The number of stable periodic solutions of time-dependent Hamiltonian systems with one degree of freedom*, Ergodic Theory Dynamical Systems **18** (1998), 1007–1018.

[106] R. Ortega, *Counting periodic solutions of the forced pendulum equation*, Nonlinear Anal. **42** (2000), 1055–1062.

[107] R. Ortega, E. Serra and M. Tarallo, *Non-continuation of the periodic oscillations of a forced pendulum in the presence of friction*, Proc. Amer. Math. Soc. **128** (2000), 2659–2665.

[108] R. Ortega and M. Tarallo, *Degenerate equations of pendulum-type*, Comm. Contemporary Math. **2** (2000), 127–149.

[109] R. Ortega and M. Tarallo, *Almost periodic upper and lower solutions*, to appear.

[110] I.C. Percival, *Variational principles for the invariant toroids of classical dynamics*, J. Phys. A **7** (1974), 794–802.

[111] I.C. Percival, *A variational principle for invariant tori of fixed frequency*, J. Phys. A **12** (1979), L57–L60.

[112] H. Poincaré, *Sur un théorème de géométrie*, Rend. Circ. Mat. Univ. Palermo **33** (1912), 375–407.

[113] P. Pucci and J. Serrin, *Extensions of the mountain pass theorem*, J. Functional Anal. **59** (1984), 185–210.

[114] P. Rabinowitz, *On a class of functionals invariant under a* Z_n *action*, Trans. Amer. Math. Soc. **310** (1988), 303–311.

[115] P. Rabinowitz, *Heteroclinics for a reversible Hamiltonian system*, Ergodic Theory Dynamical Systems **14** (1994), 817–829.

[116] P. Roselli, *A multiplicity result for the periodically forced N-pendulum with nonzero mean valued forcings*, Nonlinear Anal. **43** (2001), 1019–1041.

[117] R.Z. Sagdeev, P.A. Usikov and G.M. Zaslavsky, *Nonlinear Physics: from the Pendulum to Turbulence and Chaos*, Harwood Acad. Publ., Chur (1988).

[118] R. Schaaf and K. Schmitt, *A class of nonlinear Sturm–Liouville problems with infinitely many solutions*, Trans. Amer. Math. Soc. **306** (1988), 853–859.

[119] E. Serra and M. Tarallo, *A reduction method for periodic solutions of second order subquadratic equations*, Adv. Differential Equations **3** (1998), 199–226.

[120] E. Serra and M. Tarallo, *A unified approach to boundary value problems at resonance with the first eigenvalue*, Proc. Int. Conf. Dynamical Systems Differential Equations (Sprinfield, MO, 1996), Discrete Contin. Dynam. Systems, Added Vol. II (1998), 182–195.

[121] E. Serra, M. Tarallo and S. Terracini, *On the structure of the solution set of forced pendulum-type equations*, J. Differential Equations **131** (1996), 189–208.

[122] E. Serra, M. Tarallo and S. Terracini, *Subharmonic solutions to second order differential equations with periodic nonlinearities*, Nonlinear Anal. **41** (2000), 649–667.

[123] G. Tarantello, *On the number of solutions for the forced pendulum equations*, J. Differential Equations **80** (1989), 79–93.

[124] A. Tineo, *An iterative scheme for the N-competing species problem*, J. Differential Equations **116** (1995), 1–15.

[125] F. Tricomi, *Sur une équation différentielle de l'électrotechnique*, C. R. Acad. Sci. Paris **193** (1931), 635–636.

[126] F. Tricomi, *Integrazione di un equazione differenziale presentatasi in elettrotecnica*, Ann. R. Scuola Norm. Sup. Pisa **2** (1933), 1–20.

[127] A.J. Ureña, *On the multiplicity of—up to constants—periodic solutions for dissipative pendulum-type equations*, to appear.

[128] J.R. Ward Jr., *A boundary value problem with a periodic nonlinearity*, Nonlinear Anal. **10** (1986), 207–213.

[129] M. Willem, *Oscillations forcées de l'équation du pendule*, Pub. IRMA Lille **3** (1981), V1–V3.

[130] Jiangong You, *Invariant tori and Lagrange stability of pendulum-type equations*, J. Differential Equations **85** (1990), 54–65.

[131] Jiangong You, *Periodic solutions to pendulum type equations*, Nanjing Daxie Xuebao Shuxue Bannian Kan **7** (1990), 218–222 (in Chinese).

[132] Jiangong You, *Boundedness of solutions and existence of quasiperiodic solutions for a non-conservative pendulum-type equation*, Chinese Science Bull. **21** (1991), 1606–1609.

[133] Jiangong You, *Global attraction of forced pendulum equations with positive damping*, Acta Math. Sinica **34** (1991), 690–695 (in Chinese).

[134] Jiangong You, *Boundedness for solutions of non-conservative pendulum-type equations*, Preprint.

[135] S.F. Zakharin and I.O. Parasyuk, *Forced almost periodic oscillations of a pendulum and their extremal properties* (in Ukrainian), Dopov. Nats. Akad. Nauk Ukr. Mat. Prirodozn. Tekh. Nauki **6** (1998), 19–21

This Page Intentionally Left Blank

CHAPTER 7

Ważewski Method and Conley Index

Roman Srzednicki*

Institute of Mathematics, Jagiellonian University, ul. Reymonta 4, 30-059 Kraków, Poland
E-mail: srzednic@im.uj.edu.pl

Contents

1. Introduction . . . 593
   1.1. Historical background of topological methods . . . 593
   1.2. An outline of the Ważewski method in a historical context . . . 594
   1.3. On the origin of the Conley index theory . . . 600
   1.4. Review of the current exposition . . . 601
2. Ważewski method for local semiflows and flows . . . 602
   2.1. Local semiflows and Ważewski sets . . . 602
   2.2. Properties of Ważewski sets . . . 605
   2.3. Versions of the Ważewski theorem . . . 606
3. Ważewski method in ordinary differential equations and inclusions . . . 610
   3.1. Polyfacial sets . . . 610
   3.2. Equations without the uniqueness property and differential inclusions . . . 614
   3.3. Weak inequalities in polyfacial sets . . . 616
   3.4. Generalized polyfacial sets . . . 617
4. Ważewski method for retarded functional differential equations . . . 620
   4.1. Local semiflows generated by RFDEs . . . 620
   4.2. Ważewski-type theorems for RFDEs . . . 621
5. Some topological concepts . . . 626
   5.1. Topological pairs, quotient spaces, and pointed spaces . . . 626
   5.2. Homotopy types . . . 627
   5.3. Absolute neighborhood retracts . . . 629
   5.4. Lusternik–Schnirelmann category . . . 630
   5.5. Cup-length . . . 632
   5.6. Lefschetz theorem and the fixed point index . . . 634
6. Properties of the asymptotic parts of Ważewski sets . . . 636
   6.1. Estimates on the category of the asymptotic part . . . 636
   6.2. Stationary points of gradient-like local semiflows . . . 637
   6.3. Stationary points of local semiflows in the general case . . . 637
7. Isolating blocks and segments . . . 638

---

[*] Supported by KBN, grant 2 P03A 041 24.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 1
Edited by A. Cañada, P. Drábek and A. Fonda
© 2004 Elsevier B.V. All rights reserved

- 7.1. Isolating blocks and their structure . . . 638
- 7.2. Estimates on the cohomology of the invariant part . . . 642
- 7.3. On the number of stationary points of gradient-like flows . . . 648
- 7.4. Fixed point index and the index of stationary points . . . 651
- 7.5. Isolating segments . . . 652

8. Selected applications of the Ważewski method . . . 658
   - 8.1. Asymptotic solutions . . . 659
   - 8.2. Two-point boundary value problems . . . 662
   - 8.3. Detection of chaotic dynamics . . . 664
9. Conley index . . . 667
   - 9.1. Isolated invariant sets and the Conley index . . . 667
   - 9.2. Properties of the Conley index . . . 671
   - 9.3. Examples of Conley indices and an application . . . 675
   - 9.4. Concluding remarks . . . 680

References . . . 680

1. Introduction

The main subject of this note is the *retract method* introduced by Tadeusz Ważewski. It is a method of proving the existence of solutions which remain in a given set and refers to differential equations describing some evolution in time. The sets under consideration should satisfy the condition "all egress points are strict" or its less restrictive variant. Now they are called *Ważewski sets*. The method is based on theorems which roughly assert that there is a solution contained a Ważewski set for all positive values of time if the subset of egress points is not a retract of the whole set (which explains the name of the method). If, moreover, the set is compact then its *invariant part* (i.e., the set of full solutions contained in it) is nonempty. For *isolating blocks*, i.e., compact Ważewski sets which do not contain any full solutions intersecting their boundaries, Charles Conley discovered a homotopical invariant which provides a quantitative information on their invariant parts. It is called the *Conley index*. The Conley index theory, both from a point of view of continuous and discrete-time dynamical systems, was presented by Mischaikow and Mrozek in Chapter 9 of *Handbook of Dynamical Systems*, vol. 2, within the current series of handbooks edited by Elsevier Science B.V. (see [50]).

In this note we describe the Ważewski method in details and provide an information on foundations of the Conley index theory which directly relates to the method and essentially does not overlap with the exposition of Mischaikow and Mrozek. We begin with a short introduction to the method and to the theory related to the index in a historical context.

1.1. *Historical background of topological methods*

The modern understanding of mechanics begins since the fundamental research of Newton in the XVII century. In order to formulate the rules of motion, Newton introduced the concept of differential equation. The equations which appear in mechanics belong mainly to the class of ordinary differential equations. They are characterized by a distinguished parameter interpreted as time and, if some regularity conditions are imposed, the initial data determine the whole solution. It was clear from the beginning that in general one cannot find analytical formulas for such solutions. Instead of rigorous description, numerical methods usually provide satisfactory approximations of solutions in bounded time-intervals. Possibility of calculation of such approximations contributed mainly to the successful development of applied mechanics.

However, numerical methods fail in general if one is interested in the global behavior of the solution, i.e., in the whole domain, up to the left and right range of time of its existence. For a given initial data, usually it is impossible to determine whether the solution represents a periodic motion and to predict the asymptotic behavior at the limits of its existence. As it was observed by Henri Poincaré at the end of the XIX century, the behavior of solutions can be very complicated and have a chaotic character, hence one cannot expect to determine the full information on the dynamics generated by a given equation. Even a proof of the existence of a solution with prescribed behavior (periodic, almost-periodic, bounded, chaotic, satisfying some boundary conditions, etc.) is usually a difficult task.

Poincaré initiated systematic studies on the global understanding of solutions and dynamics, which are known under the name of qualitative theory of differential equations, or, in a more general context, theory of dynamical systems. Besides analytical arguments he used methods of topology and algebra. Research on limit cycles and indexes of zeroes of planar vector fields, periodic solutions of Hamiltonian systems (in particular, based on his "twist theorem"), and stability of the Solar system belong to his most known results on that field. He also contributed essentially in the development of variational methods applied to boundary value problems.

One of the greatest achievements of Poincaré was the introduction of algebraic topology (called by him "analysis situs"). Among the results which were proved by its application, there were the fixed point theorems of L.E.J. Brouwer (1912) and Salomon Lefschetz (1923) which found natural applications in the qualitative theory of differential equations. Brouwer introduced also the concept of the degree of a vector-field. A related notion of the fixed point index was invented later by Jean Leray. The Brouwer fixed point theorem and the degree were generalized to the infinite-dimensional case by Juliusz Schauder in 1930 and, respectively, by Leray and Schauder in 1934. Together with variational methods, they form the main tools in proofs of results on existence of solutions of initial and boundary value problems in nonlinear differential equations. The idea of their application consists in suitable choice of an operator acting on a function space in which the original problem reduces to the existence of its fixed point or its zero. In variational methods, the existence of a solution of the problem is equivalent to the existence of a critical point of a suitable functional. An essential progress in development of these methods were done by Marston Morse in 1925 who invented the theory of critical points which now bears his name and by Lusternik and Schnirelmann by their research on the concept of category in 1930. It seems that the next major achievement in topological approach to differential equations was done by Tadeusz Ważewski (1896–1972) by introducing the retract method in the middle of the twentieth century.

1.2. *An outline of the Ważewski method in a historical context*

The following definition provides the notion of a retract introduced by Karol Borsuk in the paper [13] from 1931. Let X be a topological space and let A be its subset.

DEFINITION 1.1. A continuous map $r : X \to A$ such that $r(a) = a$ for every $a \in A$ is called a *retraction*. A is called a *retract of* X if there exists a retraction $X \to A$. A continuous map $r : X \to A$ is called a *strong deformation retraction* if it is equal to the map $x \mapsto h(x, 1)$, where $h : X \times [0, 1] \to X$ is a continuous map such that $h(x, 0) = x$ and $h(x, 1) \in A$ for every $x \in X$, and $h(a, t) = a$ for every $a \in A$ and $t \in [0, 1]$. A is called a *strong deformation retract* of X if there exists a strong deformation retraction $X \to A$.

Obviously, a strong deformation retract is also a retract and a retract of a connected set is also connected. Moreover, it is well known that the Brouwer fixed point theorem is equivalent to the fact that the unit sphere S^{n-1} in $\mathbb{R}^n$ is not a retract of the closed unit ball B^n.

In 1947, Ważewski presented the paper [85] in which he proved a theorem on the existence of solutions which are contained in given set for all positive values of time. In the contemporary used mathematical notation the theorem can be described as follows. Assume that v is a continuous vector-field on M, where M is an open subset of $\mathbb{R}^n$ (it can be a smooth manifold without boundary as well). We refer to M as to the *phase-space*. It is assumed that through each point x_0 of M passes a unique saturated solution $t \mapsto \phi(x_0, t)$ of the Cauchy problem

$$\dot{x} = v(x), \tag{1.1}$$

$$x(0) = x_0. \tag{1.2}$$

The map $\phi : (x, t) \mapsto \phi(x, t)$ is a called *local flow*; its purely topological description is provided in Definition 2.1. Let us mention here that usually $\phi(x, t)$ is not defined for all t and by the *trajectory* (respectively, the *positive semitrajectory*, the *negative semitrajectory*) of x we mean the set of all points $\phi(x, t)$ (respectively, the set of all points $\phi(x, t)$ with $t \geqslant 0$, all points $\phi(x, t)$ with $t \leqslant 0$) whenever they are defined.

Let V be an open subset of the phase space. Let $x \in \partial V$. Some types of behavior of the trajectory of x with respect to V are described in the following

DEFINITION 1.2. x is called an *egress point* of V if there exists an $\varepsilon > 0$ such that $\phi(x, t) \in V$ for $-\varepsilon < t < 0$. If, moreover, $\phi(x, t) \in \backslash\overline{V}$ for $0 < t < \varepsilon$ then x is called a *strict egress point*. Symmetrically, by reversing t to $-t$, one defines an *ingress point* and a *strict ingress point*. Finally, x is called an *outward tangency point* of V if there exists an $\varepsilon > 0$ such that $\phi(x, t) \in \backslash\overline{V}$ for $t \in (-\varepsilon, \varepsilon)$ and $t \neq 0$.

The points distinguished in the above definition can be visualized on Figure 1, where there are shown fragments of trajectories of some planar equation near the boundary of a square. The open vertical sides of the square consist of strict egress points, the open horizontal sides consists of strict ingress points, and the four vertices form the set of outward tangency points.

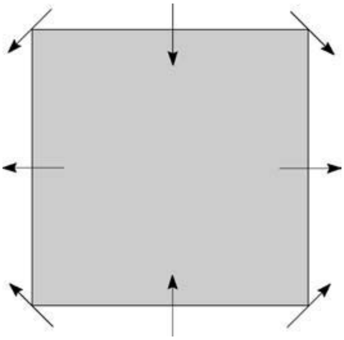


Fig. 1.

Denote by E the set of egress points. There is a distinguished subset V^* of V consisting of points which escape from V in a positive time, i.e., points $x \in V$ for which there exists a $t > 0$ such that $\phi(x,t) \notin V$. For every $x \in V^*$ one can assign the first time in which it gets the boundary of V. That time is called the *escape-time* of x and is denoted by $\sigma(x)$; by definition, it is a positive number such that $\phi(x,[0,\sigma(x))) \subset V$ and $\phi(x,\sigma(x)) \in E$. The latter point is called the *consequent* of x. Both notions are extended to E: for $x \in E$ put $\sigma(x) = 0$ and define the consequent of x as x itself.

Let $Y \subset E$ and let $Z \subset V \cup Y$. The main theorem [85, Theorem 1] is the following:

WAŻEWSKI THEOREM. *If all egress points are strict egress points and*

- $Z \cap Y$ *is a retract of* Y,
- $Z \cap Y$ *is not a retract of* Z

then there exists an $x_0 \in Z$ *such that*

- *either the positive semitrajectory of* x_0 *is contained in* V
- *or* $x_0 \in V^*$ *and the consequent of* x_0 *belongs to* $E \setminus Y$.

In particular, the second possibility in the conclusion is trivially excluded if $Y = E$ (compare [85, Theorem 2]). If we exchange egress points by ingress points in the statement of the theorem, we obtain a result on the existence of a negative semitrajectory.

Originally, the *Ważewski method* (or the *retract method*) consisted in applications of the theorem in proofs of the existence of solutions representing some special properties, usually referred to its asymptotic behavior. As it was indicated in [85], a suitable choice of the sets Z and Y leads also to proofs of the existence of solutions of some two-point boundary value problems. In [85], Ważewski pointed out also how to drop the assumption on the uniqueness of the Cauchy problem if the considered set Z is compact and V is a *polyfacial set*, which roughly means that it is described by some strong inequalities (compare Sections 3.1 and 3.2).

In the following examples we illustrate the usability of the Ważewski method.

EXAMPLE 1.1. Consider a planar equation given by

$$\begin{aligned} \dot{x} &= f(t,x,y), \\ \dot{y} &= g(t,x,y). \end{aligned} \tag{1.3}$$

in the (x,y)-coordinates. Assume that

$$xf(t,x,y) > 0, \quad \text{if } t \text{ is arbitrary, } |x| = 1,\ |y| \leqslant 1, \tag{1.4}$$

$$yg(t,x,y) < 0, \quad \text{if } t \text{ is arbitrary, } |x| \leqslant 1,\ |y| = 1. \tag{1.5}$$

After extending the plane by the time variable, (1.3) provides an autonomous equation represented by

$$\dot{t} = 1, \qquad \dot{x} = f(t,x,y), \qquad \dot{y} = g(t,x,y) \tag{1.6}$$

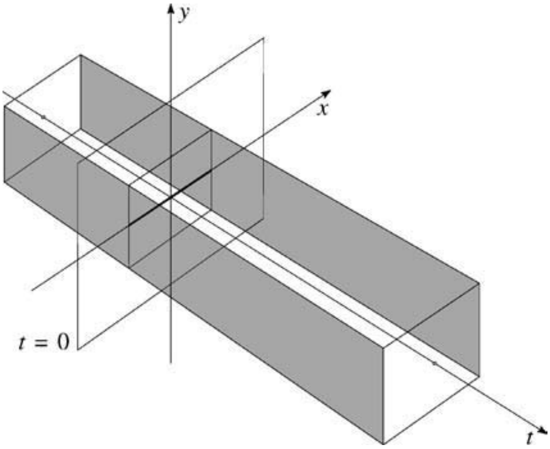


Fig. 2.

in the coordinates of $\mathbb{R}^3$. Assume that the uniqueness of the Cauchy problem is guaranteed. Put

$$V =: \left\{(t, x, y) \in \mathbb{R}^3 \colon |x| < 1,\ |y| < 1\right\}. \tag{1.7}$$

It is an infinite rectangular tube which part is shown on Figure 2. It follows by the inequalities (1.4) and (1.5) that the set of egress points

$$E := \left\{(t, x, y) \in \mathbb{R}^3 \colon |x| = 1,\ |y| < 1\right\} \tag{1.8}$$

consists of strict egress points. It forms two shaded sides of the tube on the figure. Let

$$Z := \left\{(0, x, 0) \in \mathbb{R}^3 \colon |x| \leqslant 1\right\}, \tag{1.9}$$

i.e., the black bar inside the plane $\{t = 0\}$. It follows that $Z \cap E$ consists of two ends of Z, hence it is not a retract of Z. On the other hand, each end is a retract of the corresponding side, hence there is a retraction $E \to Z \cap E$. It follows by the Ważewski theorem that there exists an x_0, $|x_0| < 1$ such that the solution (ϕ, ψ) of (1.3) with $\phi(0) = x_0$ and $\psi(0) = 0$, satisfies

$$\left|\phi(t)\right| < 1, \qquad \left|\psi(t)\right| < 1$$

for each $t \geqslant 0$.

EXAMPLE 1.2. Consider the boundary-value problem

$$y(0) = x(1) = 0 \tag{1.10}$$

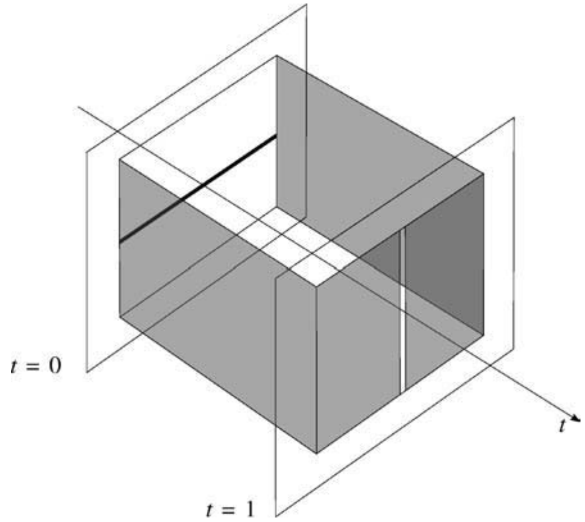


Fig. 3.

associated to (1.3). It means that we are looking for a solution of (1.3) such that its second coordinate at time 0 and the first coordinate at time 1 are equal to zero. In order to solve the problem we consider again the local flow in $\mathbb{R}^3$ generated by (1.6). Now we put

$$V := \{(t, x, y) \in \mathbb{R}^3 \colon\ t < 1,\ |x| < 1,\ |y| < 1\}. \tag{1.11}$$

The set of egress points of V is equal to

$$E := \{(t, x, y) \colon\ t \leqslant 1,\ |x| = 1,\ |y| < 1\} \cup \{(1, x, y) \colon\ |x| \leqslant 1,\ |y| < 1\}. \tag{1.12}$$

A fragment of V over the interval $[0, 1]$ is shown on Figure 3. It is the rectangular prism and the corresponding fragment of E consists of its three sides. Let Z be the same as in Example 1.1, i.e., given by (1.9). Put

$$Y := E \setminus \{(1, 0, y) \colon\ |y| < N\}. \tag{1.13}$$

On the figure, Z is shown as the black bar and Y is represented by the shaded surface. It is easy to see that $Z \cap Y$ is not a retract of Z but it is a retract of Y. It follows by the Ważewski theorem that there exists $x_0 \in (-1, 1)$ such that the solution of (1.6) starting at $(0, x_0, 0)$ either remains in V or intersects $E \setminus Y$. Since $\dot{t} = 1$, the first possibility is ruled out and, moreover, the time in which the solution gets $E \setminus Y$ is equal to 1. Thus it determines the required solution of the boundary value problem (1.3), (1.10).

In the above examples we assumed the uniqueness of solutions of the Cauchy problem associated to (1.3). Actually, that assumption can be dropped and still we can get the existence of solutions which remain in $\overline{V}$ in Example 1.3 and solve the boundary value problem in Example 1.2. Indeed, if (1.3) does not satisfy the uniqueness property, one can find approximating sequences of smooth functions $\{f_n\}$ and $\{g_n\}$ for f and g, respectively, such that they satisfy the inequalities (1.4) and (1.5). The corresponding solutions obtained for them by the Ważewski theorem are determined by initial points $z_n \in Z$. Their accumulation point z_0 is the initial point of a required solution.

The Ważewski theorem was announced also in [86,87], and later in the proceedings of the Congress of Mathematicians in Amsterdam [89]. Its proof follows easily from a lemma which asserts the continuity of the *consequent map* $V^* \cup E \to E$ provided the assumption "all egress points are strict" is satisfied (compare [85, Lemma 3]). Actually, the lemma is an immediate consequence of the following fact implicitly stated in the paper:

WAŻEWSKI LEMMA. *If all egress points are strict then the escape-time function*

$$V^* \cup E \ni x \to \sigma(x) \in [0, \infty)$$

is continuous.

The lemma has important consequences for the homotopy theory methods in differential equations. At present, the name "Ważewski method" refers also to arguments applying Ważewski lemma. Complete proofs of the Ważewski theorem and related results will be presented in Section 2.3.

In the years after publication of [85], the Ważewski method was intensively developed and applied by many authors. Both the assumptions on retraction and egress points in the Ważewski theorem were modified. Improved versions of the theorem using variants of the notion of strong deformation retract were given, among others, in the papers [1,58,62]; the proofs of those versions were based on the Ważewski lemma. In Section 2.3, we provide examples of such results. On the other hand, in [11] Bielecki observed that all results concerning the Ważewski method which are based on the hypothesis "all egress points are strict" can be directly strengthen if strict egress points are replaced by *strong egress points*. (In the above notation, an egress point $x \in E$ is called a strong egress point if for every $t > 0$ there exists $0 < s < t$ such that $\phi(x, s) \notin \overline{V}$.) Later, in [20] (and also in [21]) Charles Conley (1933–1984) presented a version of the Ważewski theorem with a more general and convenient form of that assumption. He introduced the notion of the *exit set* W^- of an arbitrary set W (in a similar way as we defined above the set of strong egress points), and with its use, he defined the concept of *Ważewski set*. We provide its definition later in Section 2.1. At this moment we mention that it is particularly easy to check in one case: if W and W^- are closed then W satisfies the definition of a Ważewski set. Moreover, referring to our previous consideration, the set $V \cup E$ is a Ważewski set provided each egress point is a strong egress point; in that case E is its exit set. The main feature of [21] was the observation that the Ważewski lemma extends to Ważewski sets: the escape-time function $\sigma : x \mapsto \sigma(x)$ defined for points of W such that their positive semitrajectory leaves W, is continuous. (Here, similarly as before, $\sigma(x)$ denotes the first time in which the semitrajectory of x hits W^-.) Our presentation of the results related to the Ważewski theorem in Section 2.3 is inspired by the ideas of Conley.

1.3. *On the origin of the Conley index theory*

The Ważewski theorem, like the Schauder fixed point theorem is an existence result. Complementary to the Schauder theorem, the Leray–Schauder degree and the fixed point index provide a quantitative information on the number of solutions. In some sense, similar complementary quantitative information on solutions given by the Ważewski method is provided by the *Conley index*, a topological invariant related to Ważewski sets of a special form, called *isolating blocks*. By an isolating block we mean a compact subset B of the phase space of a local flow ϕ such that $B = \overline{\operatorname{int} B}$ and every $x \in \partial B$ is either a strict egress point, or a strict ingress point, or a point of outward tangency of $\operatorname{int} B$ (its definition is stated in Section 7.1 in a slightly more general form). An example of isolating block is shown on Figure 1. For an isolating block B, its exit set B^- consists of all strict egress and outward tangency points. Because of the compactness of B, if some positive semitrajectory is contained in B then its limit set is nonempty and also contained in B. Since the limit set is invariant, it is contained in the invariant part of B

$$\operatorname{Inv}(B) := \{x \in B\colon\ \phi(x,t) \in B\ \forall t\}$$

which is nonempty in this case. The set $\operatorname{Inv}(B)$ is an example of an *isolated invariant set*, i.e., a compact set which is maximal invariant in some of its neighborhood. (Obviously, the block itself is such a neighborhood for $\operatorname{Inv}(B)$.) In the above sense, isolating blocks have several other nice properties due to symmetry with respect to reversing the time-axis. (Sometimes they are defined only as compact Ważewski sets which are simultaneously isolating neighborhoods.)

The first extensive research on isolating blocks for smooth flows appeared in 1971 in the paper [22] by Conley and Richard Easton. In that paper an important theorem was proved:

FIRST CONLEY THEOREM. *Each isolated invariant set is an invariant part of some isolating block.*

Also in 1971, Conley announced another important theorem (compare [19]):

SECOND CONLEY THEOREM. *If S is an isolated invariant set then the homotopy type $[B/B^-, *]$ of the pointed space obtained from B by collapsing B^- to a point $*$, does not depend on the choice of an isolating block B such that $S = \operatorname{Inv}(B)$.*

(We refer to Section 5.2 for rigorous definitions of the notions involved in that statement.) A proof of the second Conley theorem appeared later in [18] (on the Čech cohomology level only) and in [21] in the full generality. Both the theorems ensure the correctness of the following

DEFINITION 1.3. The *Conley index* of an isolated invariant set S, denoted by $h(\phi, S)$, is defined as the homotopy type $[B/B^-, *]$ for an arbitrary isolating block B having S as the invariant part.

In the Conley's publications the index appeared under the names *homotopy index* and *generalized Morse index*. The origin of the later name is the Morse theory. If $f : M \to \mathbb{R}$ is a Morse function and ϕ is the (local) flow generated by the gradient equation

$$\dot{x} = -\nabla f(x)$$

then each critical point x_0 of f is also an isolated invariant set for ϕ and its Conley index $h(\phi, \{x_0\})$ is equal to the homotopy type of the pointed k-dimensional sphere S^k, where k is the Morse index of x_0. (Recall that the Morse index is defined as the number of negative eigenvalues of the Hessian of f at x_0.) Thanks to the above relation between the indexes, some results of the Morse theory were extended to the theory of isolated invariant sets.

A simple application of the Ważewski lemma leads to the conclusion that if $h(\phi, S)$ is nontrivial (i.e., it is not the homotopy type of a one-point space) then S is nonempty. The converse is not true in general (see Example 9.5) and in this respect the Conley index is not a better tool then the Ważewski method. The Conley index has the additivity and multiplicativity properties, which explain in which sense it "counts" solutions in S. The main advantage of the index is its continuation property: two isolated invariant sets (with respect to possibly different flows) which can be linked by a kind of homotopy, have the same Conley indices. This property is in an analogy to the homotopy properties of the Brouwer and Leray–Schauder degrees, and the fixed point index. It enables calculation of the index in a complicated system by its continuation to a simpler one.

1.4. *Review of the current exposition*

Our exposition is rather self-contained; we present almost all necessary definitions and we prove a majority of the stated results, although in a few cases we provide sketches of proofs only.

In Section 2, we define the Ważewski set (Definition 2.4), state its main properties (Lemma 2.1) and state strong versions of the Ważewski theorem: Theorem 2.1 valid for local semiflows and Theorem 2.2, based on the notion of quasi-isotopic deformation retract (Definition 2.7), valid for local flows. Section 3 presents methods of construction of Ważewski sets for ordinary differential equations. In particular, so-called polyfacial sets and generalized polyfacial sets are defined (Definitions 3.1 and 3.4). Moreover, a version of the Ważewski theorem for equations without the uniqueness property and differential inclusions is stated (Theorem 3.1). The Ważewski method for local semiflows generated by retarded functional differential equations is described in Section 4, where the corresponding main results are Theorems 4.1 and 4.2. Some topological concepts are recalled in Section 5. They include the notions of the quotient space (Definition 5.1), the absolute and pointed homotopy types (Definition 5.2), absolute neighborhood retract (Definition 5.3), relative Lusternik–Schnirelmann category (Definition 5.6), and cup length (Definition 5.9). Moreover, we recall the Lefschetz fixed point theorem (Proposition 5.6) and properties of the fixed point index (Proposition 5.7). In Section 6, we provide some general results on properties of the sets of solutions given by the Ważewski method. In particular, Theorem 6.1 estimates their category and Theorem 6.2 is a result on the existence of stationary points.

Section 7 presents the notion of isolating block (Definition 7.1) and its topological properties. The related results on the category (Theorem 7.1), exact sequences (Theorem 7.3), cup-length (Theorem 7.4), stationary points of gradient-like local flows (Theorem 7.5), and fixed point index (Theorem 7.6) are given. Moreover, we define the notion of an isolating segment (Definition 7.4) and we prove Theorem 7.7, a related to that notion result on the fixed point index of the evolutionary map. In Section 8, we present three applications of results stated in previous sections. The first of them, Theorem 8.1, generalizes a result of Perron on linear equations, the second one is a result on the existence of solutions of two-point boundary value problems under the classical Bernstein–Nagumo condition (Theorem 8.2), and the last one is a result on the existence of chaotic dynamics (Theorem 8.3). An introduction to the Conley index theory is contained in Section 9. Definition 9.1 provides the notion of an isolated invariant set, Theorems 9.1 and 9.2 of Conley lead to the definition of the index, and properties of the index are stated in Theorems 9.3 and 9.4. Some calculations of the index are given in Propositions 9.1 and 9.2, and an example of its applications is given in the proof of Proposition 9.3. Finally, in Section 9.4 we indicate extensions and improvements of the Conley index.

The bibliography contains positions which usually directly refer to topics presented in this note and by far it does not pretend to be the complete list of books and articles on the Ważewski method and the Conley index. We point out that biographical information on Tadeusz Ważewski and Charles Conley can be found in the articles [54] and, respectively, [47].

2. Ważewski method for local semiflows and flows

In this section we introduce the notion of a Ważewski set and we prove Ważewski type theorems in an abstract topological setting. For this reason we begin with the definitions of the basic concepts of the theory of continuous-time dynamical systems.

2.1. *Local semiflows and Ważewski sets*

Let X be a topological space.

DEFINITION 2.1. A *local semiflow* on X is a continuous map $\phi : D \to X$, where D is an open subset of $X \times [0, \infty)$ such that for every $x \in X$ the set $\{t \in [0, \infty)\colon (x, t) \in D\}$ is equal to an interval $[0, \omega_x)$ for some $0 < \omega_x \leqslant \infty$, if $t \in [0, \omega_x)$ then $\omega_{\phi(x,t)} = \omega_x - t$ and the following equations hold:

$$\phi(x, 0) = x, \tag{2.1}$$

$$\phi(x, s + t) = \phi\big(\phi(s, x), t\big). \tag{2.2}$$

If $D = X \times [0, \infty)$ then ϕ is called a *semiflow*.

A *local flow* is a again a continuous map $\phi : D \to X$, but now D is an open subset of $X \times \mathbb{R}$, for every $x \in X$ the set $\{t\colon\ (x,t) \in D\}$ is equal to an open interval (α_x, ω_x), where

$$-\infty \leqslant \alpha_x < 0 < \omega_x \leqslant \infty,$$

if $t \in (\alpha_x, \omega_x)$ then $\alpha_{\phi(x,t)} = \alpha_x - t$ and $\omega_{\phi(x,t)} = \omega_x - t$, and Equations (2.1) and (2.2) hold. Finally, if $D = X \times \mathbb{R}$ then a local flow is called a *flow*.

Obviously, a restriction of a local flow $D \to X$ to $D \cap (X \times [0,\infty))$ is a local semiflow. Ordinary differential equations generate local flows in a well-known way provided they are autonomous and satisfy the uniqueness condition for solutions of the Cauchy problem. In a similar way retarded functional differential equations (compare Section 4.1) and semilinear parabolic equations generate local semiflows.

In the sequel we frequently use the following notation: we write $\phi_t(x)$ instead of $\phi(x,t)$ and if $A \subset X$ and $J \subset \mathbb{R}$ then we write $\phi(A, J)$ instead of $\phi(A \times J)$.

In the following definitions we assume that ϕ is a local semiflow on X. Let $x \in X$.

DEFINITION 2.2. The set

$$\phi^+(x) := \phi\bigl(x, [0, \omega_x)\bigr)$$

is called the *positive semitrajectory* of x. If, moreover, ϕ is a local flow then the sets

$$\phi(x) := \phi\bigl(x, (\alpha_x, \omega_x)\bigr),$$
$$\phi^-(x) := \phi\bigl(x, (\alpha_x, 0]\bigr)$$

are called, respectively, the *trajectory* and the *negative semitrajectory* of x.

A map $\sigma : \mathbb{R} \to X$ such that

$$\sigma(0) = x, \qquad \phi_t\bigl(\sigma(\tau)\bigr) = \sigma(t + \tau).$$

is called a *full solution through* x. A set $A \subset X$ is called *invariant* if for each $x \in A$ there exists a full solution σ through x such that $\sigma(\mathbb{R}) \subset A$. The *$\omega$-limit set of* x is defined as

$$\omega(x) := \bigcap_{s \in [0,\omega_x)} \overline{\phi\bigl(x, [s, \omega_x)\bigr)}.$$

Obviously, if σ is a full solution through x then $\sigma(\tau) = \phi_\tau(x)$ for $\tau \geqslant 0$ in the case of a local semiflow and for $\tau \in \mathbb{R}$ in the case of a local flow. If x is a *stationary point* (i.e., $\phi^+(x) = \{x\}$) then the constant map $t \mapsto x$ is a full solution through x. More general, if x is a *periodic point* (i.e., $\phi_T(x) = x$ for some $T > 0$) then the periodic extension of $t \mapsto \phi_t(x)$ from $[0, T]$ to the whole $\mathbb{R}$ is also a full solution. It is easy to prove that if $\omega(x) \neq \emptyset$ then necessarily $\omega_x = \infty$ and $\omega(x)$ is invariant.

Let ϕ be a local semiflow on X and let $W \subset X$. Two subsets of W are distinguished:

$$\begin{aligned} W^- &:= \bigl\{x \in W\colon\ \phi(x, [0, t]) \not\subset W\ \forall t > 0\bigr\}, \\ W^* &:= \bigl\{x \in W\colon\ \exists t > 0\colon\ \phi_t(x) \notin W\bigr\}. \end{aligned}$$

DEFINITION 2.3. W^- is called the *exit set of* W. The complement of W^* in W, i.e., the set $W \setminus W^*$ is called the *asymptotic part of* W.

Clearly, $W^- \subset W^*$ and $W \setminus W^*$ consists of all points $x \in W$ such that the whole positive semitrajectory $\phi(x, [0, \omega_x))$ is contained in W, which justifies its name. Note that if W is given, usually it is difficult to determine the set W^* (hence also the asymptotic part), since one should follow the whole positive semitrajectory of every point $x \in W$ unless it reaches the complement of W. It is even difficult to check whether $W \setminus W^*$ is nonempty. The Ważewski method provides answers to that question in some reasonable situations. On the other hand, the set W^- can be determined easier, because in order to verify whether $x \in W^-$ one needs to know the position of $\phi_t(x)$ for $t \in (0, \varepsilon)$, where $\varepsilon > 0$ is arbitrarily small.

DEFINITION 2.4. We call W a *Ważewski set* provided

(a) if $x \in W$, $t > 0$, and $\phi(x, [0, t]) \subset \overline{W}$ then $\phi(x, [0, t]) \subset W$,

(b) W^- is closed relative to W^*.

That definition seems to be difficult to verify in practice, at least because it is difficult to determine the set W^*. However, in the following result we indicate that in a reasonable case we do not need any information on W^*:

PROPOSITION 2.1. *If both W and W^- are closed subsets of X then W is a Ważewski set.*

In Section 1.2, we defined the notions of egress and strict egress points for an open set in the phase-space (Definition 1.2) and with their help we stated the Ważewski theorem. Using the notion of a Ważewski set we can formulate more general results. In fact, it is easy to observe that if V is an open set and every egress point is a strict egress point then $W := V \cup E$ is a Ważewski set, $W^- = E$, and $W^* = V^* \cup E$. In the following example we show that the above implication cannot be reversed.

EXAMPLE 2.1. Consider the flow generated by the equation $\dot{x} = 1$ on $\mathbb{R}$. Put

$$V := (-\infty, 0) \cup \bigcup_{n=1}^{\infty} \left(\frac{1}{2^n - 2^{n+2}}, \frac{1}{2^n} \right).$$

Here the set E consists of 0 and the points $1/2^n$ for $n \geqslant 1$. It follows that 0 is not a strict egress point although $E = W^-$ and $V \cup E$ is a Ważewski set.

In [11], there is an example illustrating a similar conclusion in which V is homeomorphic to a ball in $\mathbb{R}^3$.

2.2. *Properties of Ważewski sets*

Let W be an arbitrary subset of X.

DEFINITION 2.5. The function $\sigma : W^* \to [0, \infty)$,

$$\sigma(x) := \sup\bigl\{t \in [0, \infty)\colon \phi(x, [0, t]) \subset W\bigr\}$$

is called the *escape-time function of* W.

It follows by (2.1) and (2.2) that

$$\sigma(\phi_t(x)) = \sigma(x) - t, \quad \text{if } 0 \leqslant t \leqslant \sigma(x). \tag{2.3}$$

The most important properties of Ważewski sets (including the Ważewski lemma from Section 1.2) are summarized in the following

LEMMA 2.1. *If W is a Ważewski set and σ is its escape-time function then*

(i) *if $x \in W^*$ then $\phi_{\sigma(x)}(x) \in W^-$,*

(ii) *$x \in W^-$ if and only if $x \in W^*$ and $\sigma(x) = 0$,*

(iii) *σ is continuous,*

(iv) *W^* is open relative to W.*

PROOF. Ad (i). It is clear that if $x \in W^*$ then $\phi(x, [0, \sigma(x)]) \subset \overline{W}$, hence $\phi_{\sigma(x)}(x) \in W$ by (a). Moreover $\phi(x, [0, \sigma(x) + t]) \not\subset W$ for every $t > 0$, hence, by (2.2), $\phi_{\sigma(x)}(x) \in W^-$. Thus (i) is proved.

Ad (ii). If $x \in W^-$ then, clearly, $\sigma(x) = 0$. On the other hand, if $\sigma(x) = 0$, then $x \in W^-$ by (2.1) and (i), hence (ii) follows.

Ad (iii). For the proof of (iii) assume that $x \in W^*$ and $\varepsilon > 0$. At first we prove that there exists a neighborhood U of x such that $\sigma(y) < \sigma(x) + \varepsilon$ for every $y \in U \cap W^*$. Indeed, $\phi(x, [0, \sigma(x) + \varepsilon]) \not\subset W$, hence, by (a), there exists a $t \in [\sigma(x), \sigma(x) + \varepsilon]$ such that $\phi_t(x) \notin \overline{W}$. By continuity of ϕ, there exists a neighborhood U of x such that $\phi_t(U) \in X \setminus \overline{W}$. It is clear, that $\sigma(y) < t$ for every $y \in U \cap W^*$, hence the assertion follows.

Now we prove that there exists a neighborhood V of x such that $\sigma(x) - \varepsilon < \sigma(y)$ if $y \in V \cap W^*$. There is nothing to prove in the case $\sigma(x) = 0$, hence we can assume without loss of generality that $0 < \varepsilon < \sigma(x)$. Let $t \in [\sigma(x) - \varepsilon, \sigma(x))$. Then $\phi(x, [0, t]) \subset W^* \setminus W^-$. By (b), there is an open set Z such that $Z \cap W^* = W^* \setminus W^-$. By a standard argument, the continuity of ϕ and the compactness of $[0, t]$ imply the existence of a neighborhood V of x such that $\phi(V, [0, t]) \subset Z$. Let $y \in V \cap W^*$. By (i), $\phi_{\sigma(y)}(y) \in W^-$. Since W^- is disjoint from Z, $\sigma(y) > t$ and the proof of (iii) is finished.

Ad (iv). Finally, in order to prove (iv) assume that $x \in W^*$. By (i) and (a) there exists a $t > \sigma(x)$ such that $\phi_t(x) \notin \overline{W}$. Thus, if y belongs to some neighborhood of x then $\phi_t(y) \notin W$ and the proof of the lemma is finished. □

We consider also the *extended escape-time function* given by

$$\sigma^+ : W \ni x \to \sup\bigl\{t \geqslant 0\colon\ \phi\bigl(x, [0, t]\bigr) \subset W\bigr\} \in [0, \infty].$$

Actually, that function is not continuous even if W is a compact Ważewski set as an obvious example, in which a stationary point belongs to the closure of W^-, shows. However, the following statement can be easily proved:

LEMMA 2.2. *If both W and W^- are compact then σ^+ is continuous.*

PROOF. Since W is compact, $\sigma^+(x) = \infty$ if and only if $x \in W \setminus W^*$, hence, by Lemma 2.1(iii), it suffices to prove that σ^+ is continuous at each x_0 such that $\sigma^+(x_0) = \infty$. The required continuity follows from the fact that $\mathrm{Inv}^+(B)$ is compact (see Lemma 2.1(iv) and disjoint from W^-. Indeed, if $r > 0$ then there exists an open neighborhood U of x_0 in W such that $\phi(y, [0, r]) \cap W^- = \emptyset$, hence $\sigma^+(y) \geqslant r$ for every $y \in U$. □

2.3. *Versions of the Ważewski theorem*

The Ważewski theorem appears in the literature in various forms and each of them is a direct application of Lemma 2.1 stated above, especially its part (iii) (i.e., the Ważewski lemma). Our purpose is to formulate the theorem in a form convenient for applications. In fact, we present two theorems. The first one is a direct generalization of the main theorems of [85] and [20] obtained in a possibly most "economic" way. It is sufficient for a great majority of applications existing in the literature. However, as we show in a relatively simple example, it is sometimes convenient to have its stronger version and such a version will be presented at the end of the current section. In any case we should extend the notion of a strong deformation retract given in Definition 1.1. To this purpose we follow the ideas from the papers [58,62].

Assume that X is a topological space and A and B are subsets of X such that $A \subset B$.

DEFINITION 2.6. A is called a *strong deformation retract of B in X* if there exists a continuous map $h : B \times [0, 1] \to X$ (for which we also write $h_t(x)$ instead of $h(x, t)$) such that

(a) $h(x, 0) = x$ for every $x \in B$,
(b) $h(x, t) = x$ for every $x \in A$ and $t \in [0, 1]$,
(c) $h_1(B) \subset A$.

Obviously, in the case $B = X$, A is a strong deformation retract of B in the usual sense (see Definition 1.1). Moreover, if A is a strong deformation retract of B in X then A is a retract of B, if $X' \supset X$ then A is also a strong deformation retract of B in X', and if A is a strong deformation retract of X then A is also a strong deformation retract of B in X for every B such that $A \subset B \subset X$. A comparison among various types of retracts can be visualized on Figure 4. Here X is the grey rectangle with a hole, A is its bottom edge, and L and R are the circles—only R surrounds the hole. It is readily seen that A is a strong

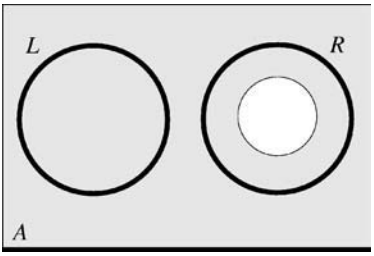


Fig. 4.

deformation retract of $A \cup L$ in X, A is a retract of $A \cup R$, and A is not a strong deformation retract of $A \cup R$ in X.

In the following sequence of results we assume that

- W is a Ważewski set,
- σ is the escape-time function of W,
- Z is a subset of W,
- Y is a subset of W^- such that $Z \cap W^- \subset Y$.

THEOREM 2.1 (Ważewski theorem—an improved statement). *If Y is not a strong deformation retract of $Z \cup Y$ in $(W \setminus W^-) \cup Y$ then there exists an $x_0 \in Z$ such that*

- *either $\phi^+(x_0) \subset W \setminus W^-$,*
- *or else $x_0 \in W^*$ and $\phi_{\sigma(x_0)}(x_0) \in W^- \setminus Y$.*

PROOF. Assume that the conclusion is false. It follows that $Z \subset W^*$ and for every $x \in Z$ its *consequent* $\phi_{\sigma(x)}(x)$ is contained in Y. For every $x \in Z \cup Y$ and every $s \in [0, \sigma(x)]$ the point $\phi_s(x) \notin W^- \setminus Y$, hence, by Lemma 2.1, the map

$$(Z \cup Y) \times [0, 1] \ni (x, t) \to \phi_{t\sigma(x)}(x) \in (W \setminus W^-) \cup Y$$

is continuous and satisfies (a)–(c) in Definition 2.6 contradictory to the assumption, hence the result follows. □

As an immediate consequence we get

COROLLARY 2.1. *If Y is not a retract of $Z \cup Y$ then the conclusion of Theorem* 2.1 *holds.*

Theorem 2.1 does not follow from Corollary 2.1 as an example based on Figure 4 shows. The original Ważewski theorem stated in Section 1.2 is a corollary of the above one. We repeat it below in a more general setting (as we already mentioned in Section 2.1, Ważewski sets naturally generalize the sets considered in [85]).

COROLLARY 2.2. *If $Z \cap Y$ is a retract of Y and $Z \cap Y$ is not a retract of Z then the conclusion of Theorem* 2.1 *holds.*

PROOF. It suffices to prove that under the imposed hypotheses there is no retraction $Z \cup Y \to Y$. Indeed, if there exists a retraction of $Z \cup Y$ onto Y, its composition with a retraction $Y \to Z \cap Y$ provides a retraction $Z \to Z \cap Y$ contrary to the assumption. □

In the case $Y = W^-$, there is an essential simplification in the statement of Theorem 2.1:

COROLLARY 2.3. *If W^- is not a strong deformation retract of $Z \cup W^-$ in W then there exists an $x_0 \in Z$ such that $\phi^+(x_0) \subset W$.*

In particular, the version of the Ważewski theorem given in [20] is the following

COROLLARY 2.4 (Ważewski theorem—the Conley's version). *If W^- is not a strong deformation retract of W then there exists an $x_0 \in W$ such that $\phi^+(x_0) \subset W$.*

In the case of compact W one can get the full solution contained in W.

COROLLARY 2.5. *Let W be a compact set. If W^- is not a strong deformation retract of W then there exists an $x_0 \in W$ such that $\emptyset \neq \omega(x_0) \subset W$. In particular, W contains a nonempty invariant set.*

PROOF. By Corollary 2.4, there exists an x such that $\phi^+(x) \subset W$. Since W is compact, the ω-limit set of x is nonempty. Since the latter set is invariant, the result follows. □

Throughout reminder of this section we will consider local flows. The last version of the Ważewski theorem presented in this section is due to Andrzej Pliś in [62]. In order to state it, we again modify the notion of a strong deformation retract. Let X be a topological space and let A be a subset of $B \subset X$.

DEFINITION 2.7. A is called a *quasi-isotopic deformation retract of B in X* if there exists a continuous map $h : B \times [0,1] \to X$ such that the conditions (a), (b), and (c) in Definition 2.6 are satisfied, and

(d) $B \ni x \to h_t(x) \in h_t(B)$ is a homeomorphism if $0 \leqslant t < 1$.

In the case $B = X$ we simply call A a *quasi-isotopic deformation retract of X*. Obviously, a quasi-isotopic deformation retract of B in X is a strong deformation retract of B in X. The following version of the Ważewski theorem essentially appeared in [62].

THEOREM 2.2 (Ważewski theorem—the version of Pliś). *Let ϕ be a local flow. If Y is not a quasi-isotopic deformation retract of $Z \cup Y$ in $(W \setminus W^-) \cup Y$ then the conclusion of Theorem* 2.1 *holds.*

PROOF. We assume that the conclusion is false. It follows that $Z \subset W^*$ and $\phi_{\sigma(x)}(x) \in Y$ for $x \in Z$. As in the proof of Theorem 2.1, for $x \in Z \cup Y$ and $t \in [0,1]$ we put

$$h(x,t) := \phi\bigl(x, t\sigma(x)\bigr)$$

and using Lemma 2.1 we conclude that $h(x,t) \in (W \setminus W^-) \cup Y$ and h is continuous and satisfies (a)–(c) in Definition 2.6. Thus, in order get a contradiction by proving that Y is a quasi-isotopic deformation retract of $Z \cup Y$ in $(W \setminus W^-) \cup Y$ it remains to prove that if $t < 1$ then h_t is a homeomorphism as a map $Z \cup Y \to h_t(Z \cup Y)$. Fix $t < 1$ and for $y \in h_t(Z \cup Y)$ put

$$g(y) := \phi\bigl(y, -(1-t)^{-1} t\sigma(y)\bigr).$$

Let $y = h_t(x)$ and $x \in Z \cup Y$. Since $\sigma(y) = (1-t)\sigma(x)$ by (2.3) and $y = \phi(x, t\sigma(x))$,

$$g(y) = \phi\bigl(y, -t\sigma(x)\bigr) = x.$$

It follows that $g : h_t(Z \cup Y) \to Z \cup Y$ and $g \circ h_t$ is equal to the identity. On the other hand,

$$(h_t \circ g)\bigl(h_t(x)\bigr) = h_t(x),$$

hence $h_t \circ g$ is equal to the identity on $h_t(Z \cup Y)$. Thus h_t is a homeomorphism (and g is its inverse), which concludes the proof. ☐

Obviously, Theorem 2.2 implies Theorem 2.1 as a corollary in the case ϕ is a local flow. Using an example based on Figure 5, below we indicate that it is a stronger result.

EXAMPLE 2.2. Let the phase-space of a local flow be equal to $\mathbb{R}^3$ and let W be equal to a cube as shown on Figure 5. Assume that the shaded rear and front sides of the cube form the exit set W^-. Let the set Z consists of two linked disjoint arcs, each of them is attached by its ends to one of the components of W^-. We claim that there exists a positive semitrajectory starting in Z and contained in W.

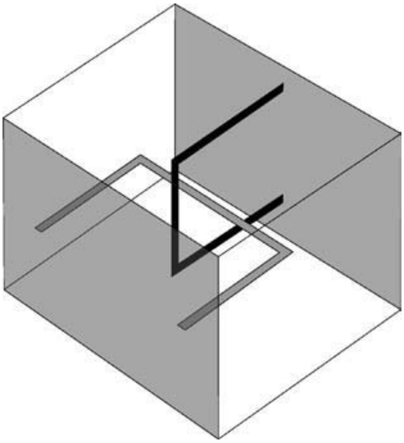


Fig. 5.

Corollary 2.3 is useless in attempts to prove the claim because W^- is a strong deformation retract of $Z \cup W^-$ in W. Indeed, it suffices to push each of the components of Z to the corresponding side of the cube. Nevertheless, we assert that W^- it is not a quasi-isotopic deformation retract of $Z \cup W^-$ in W, hence, by Theorem 2.2, there is a required positive semitrajectory.

In order to verify the assertion, assume on the contrary that W^- is a quasi-isotopic deformation retract of $Z \cup W^-$ in W and denote by h an associated homotopy which satisfies the conditions (a)–(d). Let Γ_1 and Γ_2 be two sets homeomorphic to the circle S^1, each of them is contained in a different components of $Z \cup W^-$ and surrounds the corresponding hole. Thus the *linking number* (compare [17, II.15] in the smooth case or [28, VII.4.17] in the purely topological setting) of Γ_1 and Γ_2 is equal to ± 1. On the other hand, since both the sets are compact, h is continuous, and (c) is valid, there exists an $\varepsilon > 0$ such that if $1 - \varepsilon < t < 1$ then each of the sets $h_t(\Gamma_i)$ is so close to the side to which it is attached that there is no link between $h_t(\Gamma_1)$ and $h_t(\Gamma_2)$. Thus their linking number is equal to 0. This is a contradiction, since $h_s(\Gamma_1) \cap h_s(\Gamma_2) = \emptyset$ for every $s \in [0, t]$ by (d), which implies that the linking number is preserved under the continuation over the interval $[0, t]$.

REMARK 2.1. In Example 2.2, we concluded by Theorem 2.2 that for every local flow there is a positive semitrajectory starting at Z and contained in W. Actually, in this example the same conclusion holds for local semiflows. Indeed, this is a consequence of the homotopy property of the linking number and the modified Theorem 2.2 valid for local semiflows, for which the notion of quasi-isotopic deformation retract is changed by substituting the condition

(e) if $x, y \in B$, $t < 1$, and $h_t(x) = h_t(y)$ then $h_s(x) = h_s(y)$ for every $s \in (t, 1]$, instead of (d) in Definition 2.7.

3. Ważewski method in ordinary differential equations and inclusions

Abstract local flows and semiflows, for which we describe the Ważewski method in Section 2, arise in the theory of differential equations. Here we present how to construct Ważewski sets in ordinary differential equations and give some extensions of the method to equations which do not generate local semiflows.

3.1. *Polyfacial sets*

A usual construction of Ważewski sets is based on existence of functions which behave similarly to Liapunov functions on some parts of their zero-levels. In the sequel by M we denote a smooth (i.e., of C^∞ class) Riemannian manifold without boundary, by TM its tangent bundle, by the dot $\cdot$ the induced scalar product in each tangent space T_xM, and by $\nabla f(x)$ the corresponding gradient of a function $f : M \to \mathbb{R}$ differentiable at x. Let v be a vector-field on M, i.e., a continuous map $v : M \to TM$ such that $v(x) \in T_xM$ for every $x \in M$. Let p and q be nonnegative integers, $p + q \geqslant 1$, and for $i = 1, \ldots, p$

and $j = 1, \dots, q$ let ℓ^i and m^j be continuous functions $M \to \mathbb{R}$ of C^1-class in some neighborhoods of their zero levels. Put

$$V := \{x \in M\colon\ \ell^i(x) < 0\ \forall i = 1, \dots, p,\ m^j(x) < 0\ \forall j = 1, \dots, q\}$$

and for the above i and j define

$$L^i := \{x \in \partial V\colon\ \ell^i(x) = 0\}, \tag{3.1}$$

$$M^j := \{x \in \partial V\colon\ m^j(x) = 0\}. \tag{3.2}$$

(We do not exclude the possibility $p = 0$ or $q = 0$; in such a case one of the corresponding families of functions and sets is empty.)

DEFINITION 3.1. The set V is called a *polyfacial set determined by the set of functions* $\{\ell^i, m^j\}$. If, moreover,

$$v(x) \cdot \nabla \ell^i(x) > 0, \quad \text{if } x \in L^i, \tag{3.3}$$

$$v(x) \cdot \nabla m^j(x) < 0, \quad \text{if } x \in M^j \tag{3.4}$$

for every $i = 1, \dots, p$ and every $j = 1, \dots, q$ then it is called a *polyfacial set for the vector-field* v (or, alternatively, for the differential equation (1.1)).

The concept of a polyfacial set comes from the original paper [85]. In that paper the subsets L^i were called its *positive faces*, while M^j were its *negative faces*, which terminology is justified by (3.3) and (3.4). (Here we try to avoid that terminology because, as we will see soon, L^i is contained in the exit set $\overline{V}^-$.)

Throughout reminder of this section we assume that V is the polyfacial set for v determined by $\{\ell^i, m^j\}$. Assume that Equation (1.1) generates a local flow ϕ. It follows by (3.3) that if $x \in L^i$ then there exists an $\varepsilon > 0$ such that $\ell^i(\phi_t(x)) > 0$ (hence $\phi_t(x) \notin \overline{V}$) for each $t \in (0, \varepsilon)$ and $\ell^i(\phi_t(x)) < 0$ for each $t \in (-\varepsilon, 0)$. If $x \in M^j$ then (3.4) implies the reversed inequalities for $m^j(\phi_t(x))$. Recall, that in Definition 1.2 we provided the notions of strict egress, strict ingress, and outer tangency points on the boundary of a given open set.

PROPOSITION 3.1. *If $x \in \partial V$ then*

(i) *x is a strict egress point if and only if $x \in L^i$ for some $i = 1, \dots, p$ and $x \notin M^j$ for every $j = 1, \dots, q$,*

(ii) *x is a strict ingress point if and only if $x \in M^j$ for some $j = 1, \dots, q$ and $x \notin L^i$ for every $i = 1, \dots, p$,*

(iii) *x is an outward tangency point if and only if $x \in L^i \cap M^j$ for some $i = 1, \dots, p$ and $j = 1, \dots, q$,*

(iv) *x is a strict egress or strict ingress, or outward tangency point* (*hence no other possibility holds*).

PROOF. Let $x \in \partial V$.

Ad (i). Assume that x is a strict egress point. Then $x \notin M^j$ for every j. Since

$$\partial V \subset \bigcup_{i=1}^{p} L^i \cup \bigcup_{j=1}^{q} M^j, \tag{3.5}$$

the implication from the left to the right follow. On the other hand, assume that $x \in L^{i_0}$ for some i_0 and $x \notin M^j$ for every j. One has $\ell^i(x) \leqslant 0$ for every i and $m^j(x) \leqslant 0$ for every j since $x \in \overline{V}$. It follows by the assumption that $m^j(x) < 0$ for every j, hence also $m^j(\phi_t(x)) < 0$ for $t < 0$ such that $|t|$ is small. If $\ell^i(x) < 0$ then, as before, $\ell^i(\phi_t(x)) < 0$ for $t < 0$ with $|t|$ small. The other possibility $\ell^i(x) = 0$ implies that $\phi(x, (0, \varepsilon_i)) \subset \backslash\overline{V}$ and $\ell^i(\phi_t(x)) < 0$ if $-\varepsilon_i < t < 0$ for some $\varepsilon_i > 0$. Thus, there exists an $\varepsilon > 0$ such that $\phi_t(x) \in V$ if $-\varepsilon < t < 0$. Moreover, $\phi(x, (0, \varepsilon_{i_0})) \subset \backslash\overline{V}$, hence the result follows.

Ad (ii). After reversing the time-direction it follows from (i).

Ad (iii). Assume that x is an outward tangency point. It follows by (3.5) and (i) and (ii) that $x \in L^i$ for some i and $x \in M^j$ for some j. The other implication is obvious.

Ad (iv). It follows by (3.5) and (i), (ii), and (iii). □

Denote by E the set of strict egress points and by T the set of outward tangency points. As an immediate consequence of Proposition 3.1 we get the following equations

$$E = \bigcup_{i=1}^{p} L^i \setminus \bigcup_{j=1}^{q} M^j, \tag{3.6}$$

$$E \cup T = \bigcup_{i=1}^{p} L^i. \tag{3.7}$$

Proposition 3.1 indicates also which subsets of $\overline{V}$ are Ważewski sets. Usually there are infinitely many of them. In the following straightforward result two of them are distinguished:

PROPOSITION 3.2.

(i) *The set* $V \cup E$ *is a Ważewski set and its exit set is equal to* E.

(ii) *The set* $\overline{V}$ *is a Ważewski set and its exit set is equal to* $E \cup T$.

The sets considered in [85] assumed the form (i). Ważewski sets obtained from polyfacial sets have an advantage in comparison to some other constructions: they are stable with respect to small perturbations of the vector-field (since the inequalities (3.3) and (3.4) are strict). In the following definition we distinguish a special class of polyfacial sets.

DEFINITION 3.2. Let V be a polyfacial set determined by $\{\ell^i, m^j\}$. It is called *regular* if for every $x \in \partial V$ the set $\{\nabla\ell^{i_1}(x), \ldots, \nabla\ell^{i_r}(x), \nabla m^{j_1}(x), \ldots, \nabla^{j_s}(x)\}$ is linearly independent provided $1 \leqslant i_1 < \cdots < i_r \leqslant p$ and $1 \leqslant j_1 < \cdots < j_s \leqslant q$, and

$$\{i_1, \ldots, i_r\} = \{i = 1, \ldots, p\colon\ \ell^i(x) = 0\},$$
$$\{j_1, \ldots, j_s\} = \{j = 1, \ldots, q\colon\ m^j(x) = 0\}.$$

REMARK 3.1. If V is a regular polyfacial set then $\overline{V}$ is a topological submanifold with boundary in M and the exit set $\overline{V}^- = E \cup T$ is a topological submanifold of ∂V. In particular, all homology and cohomology functors for the pair $(\overline{V}, \overline{V}^-)$ coincide. One can prove that for every such functor H,

$$H(\overline{V}, \overline{V}^-) = H(V \cup E, E).$$

If, moreover, $p = q = 1$, i.e., $V = \{\ell < 0,\ m < 0\}$ then the set $\overline{V}$ is a C^1-class submanifold with corners and the exit set is a C^1-class submanifold with boundary.

In practice, the Ważewski method is most frequently applied to non-autonomous equations. Let U be an open subset of $\mathbb{R} \times M$ and let $w : U \to TM$ be time-dependent vector-field, i.e., a continuous map such that $w(t, x) \in T_x M$ for every $(t, x) \in U$. Assume that it generates the equation

$$\dot{x} = w(t, x) \tag{3.8}$$

for which the Cauchy problem $x(t_0) = x_0$ has the unique solution. We consider the induced system

$$\begin{aligned} \dot{t} &= 1, \\ \dot{x} &= w(t, x) \end{aligned} \tag{3.9}$$

in the extended phase space U and the vector-field $v := (1, w)$ on its right-hand side. All the previous results can be applied to v since U is also a manifold without boundary. In the case $M = \mathbb{R}^n$ the inequalities (3.3) and (3.4) assume the forms

$$\frac{\partial \ell^i}{\partial t}(t, x) + \sum_{k=1}^{n} \frac{\partial \ell^i}{\partial x_k}(t, x) w_k(t, x) > 0, \quad \text{if } (t, x) \in L^i, \tag{3.10}$$

$$\frac{\partial m^j}{\partial t}(t, x) + \sum_{k=1}^{n} \frac{\partial m^j}{\partial x_k}(t, x) w_k(t, x) < 0, \quad \text{if } (t, x) \in M^j. \tag{3.11}$$

In particular, the set $\{(t, x, y) \in \mathbb{R}^3\colon\ |x| < 1,\ |y| < 1\}$ considered in Example 1.1 and shown on Figure 2 is a polyfacial set determined by the functions

$$\ell : (t, x, y) \mapsto x^2 - 1, \qquad m : (t, x, y) \mapsto y^2 - 1$$

and the set $\{(t,x,y) \in \mathbb{R}^3\colon\ t < 1,\ |x| < 1,\ |y| < 1\}$ in Example 1.2 (see Figure 3) is a polyfacial set determined by the functions

$$\ell^1 : (t,x,y) \mapsto x^2 - 1, \qquad \ell^2 : (t,x,y) \mapsto t - 1, \qquad m : (t,x,y) \mapsto y^2 - 1.$$

3.2. *Equations without the uniqueness property and differential inclusions*

Using polyfacial sets, in this section we extend the Ważewski method to the case in which the vector-field $v : M \to TM$ does not satisfy the uniqueness property of the initial value problem (1.1), (1.2). Since it does not involve any essential difficulty, we will extend the Ważewski method even further: we will consider differential inclusions. The idea of an extension to equations without uniqueness comes from the original paper [85] of Ważewski, while the extension to differential inclusions was initiated by Bielecki in [12] (he used the term *paratingent equations*). Our presentation is based on the Bielecki's approach.

As in Section 3.1, we assume that M is Riemannian manifold with a given metric. Let F be a multivalued vector-field on M, i.e., for every $x \in M$ we assign a nonempty compact and convex set $F(x) \subset T_x M$. In what follows we assume that F is *upper semi-continuous*, i.e., for every $x \in M$ and every ε there exists a neighborhood U of x such that if $y \in U$ then $F(y)$ is contained in the ε-neighborhood $N(F(x), \varepsilon)$ of $F(x)$ in TM. (In particular, every single-valued continuous vector-field has that property.) By a solution of a *differential inclusion*

$$\dot{x} \in F(x) \tag{3.12}$$

we mean an absolutely continuous mapping $t \mapsto x(t)$ defined in some interval such that $\dot{x}(t) \in F(x(t))$ for almost every t. We assume also that the domain of x is maximal, i.e., the solution is saturated.

Below we assume that V be a polyfacial set determined by $\{\ell^i, m^j\}$, $i = 1, \ldots, p$, $j = 1, \ldots, q$, $p + q \geqslant 1$, and the sets L^i, M^j, and let E are given by, respectively, (3.1), (3.2), and (3.6).

DEFINITION 3.3. V is called a *polyfacial set for (3.12)* if for every i and j,

$$y \cdot \nabla \ell^i(x) > 0, \quad \text{for all } y \in F(x) \text{ and } x \in L^i,$$
$$y \cdot \nabla m^j(x) < 0, \quad \text{for all } y \in F(x) \text{ and } x \in M^j.$$

In the sequel we assume that V is a polyfacial set for (3.12). Our aim is to prove the following extension of the Ważewski theorem, which essentially were given in [12]:

THEOREM 3.1. *Let $Y \subset E$ and $Z \subset V \cup Y$. Assume that Z and $E \setminus Y$ are compact and Y is not a quasi-isotopic deformation retract of $Z \cup Y$ in $V \cup E$. Then there exists a solution x of* (3.12) *such that* $x(0) \in Z$ *and*

- *either* $x(t) \in \overline{V}$ *for all* $t \geqslant 0$
- *or there exists a* $t_0 > 0$ *such that* $x(t) \in \overline{V}$ *if* $t \in [0, t_0]$ *and* $x(t_0) \in E \setminus Y$.

PROOF. We will base on the results stated in [4], although they concern the case $M = \mathbb{R}^n$ only. In the case of an arbitrary manifold M they can be stated in a similar way. Let v_n be a sequence of locally Lipschitzean vector-fields such that their graphs in TM approximate the graph of F and preserve the strong inequalities (3.3) and (3.4) (compare [4, Theorem 1.12.1]). By Theorem 2.2, for each n either there exists a point $x_n \in Z$ such that the solution of the equation

$$\dot{x} = v_n(x)$$

starting at x_n remains in V or there exists an y_n and $t_n > 0$ such that the solution starting at y_n remains in V for $t \in [0, t_n)$ and $y_n(t_n) \in E \setminus Y$. At least one of the sequences $\{x_n\}$ and $\{y_n\}$ must be infinite. An accumulation point of that infinite sequence has the required properties (compare First Proof of Theorem 2.1.3 in [4]; observe moreover that the time t_0 in the conclusion is positive since the distance between Z and $E \setminus Y$ is nonzero). ☐

We get the following corollary as an immediate consequence of the theorem in the case $Y = E$:

COROLLARY 3.1. *If* $H(V \cup E, E) \neq 0$, *where* H *is the singular homology functor with coefficients in a given ring, then there exists a solution* x *of* (3.12) *such that* $x(t) \in \overline{V}$ *for all* $t \geqslant 0$.

PROOF. Indeed, by the assumption there exists a singular chain c in $V \cup E$ which is a homologically nontrivial cycle with respect to the pair $(V \cup E, E)$, hence its support $Z := |c|$ cannot be continuously deformed to E. Thus Z is compact and satisfies the assumption of Theorem 3.1, hence the result follows. ☐

The questions concerning the existence of solutions contained in a given set are important in the theory of differential inclusions; the area of research related to them bears the name *viability theory* (compare [3,4]). In that theory, a set W is called *viable at* x_0 if $x_0 \in W$ and there exists a solution x with $x(0) = x_0$ and $x(t) \in W$ for all $t \geqslant 0$.

Apart from the Bielecki's paper, Ważewski-type theorems for equations without uniqueness and differential inclusions appeared in several publications, compare for example [6–9,40,43]. Opposite to the approximation argument in the above proof, a purely topological approach based on set-valued retractions was frequently presented and the obtained results were limited in practice to questions of lack of connectedness. Recently, in the paper [35], Gabor and Quincampoix applied homology theory methods for *admissible* multivalued mappings (see [33]) to obtain stronger results in that topological setting. In particular, [35] contains a theorem similar to Corollary 3.1 valid for closed sets which resemble the Ważewski ones. We state it below.

Assume that $M = \mathbb{R}^n$. F is a called a *Marchaud map* if there exists a $C > 0$ such that

$$\sup\bigl\{|y|\colon\ y \in F(y)\bigr\} \leqslant C\bigl(1 + |x|\bigr)$$

for every $x \in \mathbb{R}^n$. For $x_0 \in \mathbb{R}^n$ denote by $S_F(x_0)$ the set of solutions of the initial value problem (3.12), (1.2). We extend the notion of exit set and escape-time function to the case of differential inclusions; for $W \subset \mathbb{R}^n$ put

$$W^- := \bigl\{x_0 \in \partial W\colon \exists x \in S_F(x_0)\ \forall t > 0\colon\ x\bigl([0,t]\bigr) \not\subset W\bigr\}$$

and for each solution x of (3.12) satisfying $x(0) \in W$ define

$$\Sigma(x) := \sup\bigl\{t \geqslant 0\colon\ x\bigl([0,t]\bigr) \in W\bigr\}.$$

THEOREM 3.2 (compare Theorem A and Remark 2.1 in [35]). *Let F be a Marchaud map and let W and W^- be closed subsets of $\mathbb{R}^n$. Assume that for every $x_0 \in W^-$ and every $x \in S_F(x_0)$,*

$$x\bigl(\bigl[0, \Sigma(x)\bigr]\bigr) \subset W^-$$

provided $\Sigma(x) < \infty$. If $\check{H}(W, W^-) \neq 0$ then there exists a solution x of (3.12) *such that $x(t) \in W$ for all $t \geqslant 0$, where $\check{H}$ denotes the Čech homology functor with compact supports and coefficients in $\mathbb{Q}$.*

In the case of Marchaud maps, Theorem 3.2 generalizes Corollary 3.1 if the considered polyfacial set V is regular (see Definition 3.2) and the ring of coefficients of H is equal to $\mathbb{Q}$. Indeed, in that case the Čech and singular homologies coincide by Remark 3.1, but Theorem 3.2 allows empty interior of W and sliding along W^-. In the next section we indicate that for regular polyfacial sets there is a stronger version of Corollary 3.1 which also allows sliding along the boundary.

3.3. *Weak inequalities in polyfacial sets*

The proof of Theorem 3.1 was essentially based on the strong inequalities (3.13) and (3.13). One can ask whether the replacement of those strong inequalities by the weak ones still provides results similar to that theorem. That question was considered in the paper [37]; in particular, for regular polyfacial sets. The following version of the Ważewski theorem is based on [37, Theorem 3].

THEOREM 3.3. *Let V be a regular polyfacial set determined by $\{\ell^i, m^j\}$, and let L^j, M^i, and E be given by, respectively,* (3.1), (3.2), *and* (3.6). *Let Y and Z satisfy the assumptions of Theorem* 3.1. *Assume moreover that F is a multivalued vector-field such that the weak inequalities*

$$y \cdot \nabla \ell^i(x) \geqslant 0, \quad \textit{for all } y \in F(x) \textit{ and } x \in L^i,$$
$$y \cdot \nabla m^j(x) \leqslant 0, \quad \textit{for all } y \in F(x) \textit{ and } x \in M^j.$$

are satisfied for every i and j. Then the conclusion of Theorem 3.1 *holds.*

PROOF. The idea of a proof is to add to the right-hand side of (3.12) a sequence of perturbation terms of the form $\phi_n h$, where h is a bounded vector-field satisfying

$$h(x) \cdot \nabla \ell^i(x) > 0, \quad \text{if } x \in L^i,$$
$$h(x) \cdot \nabla m^j(x) < 0, \quad \text{if } x \in M^j,$$

for every i and j, and ϕ_n is a sequence of continuous, positive-valued functions such that ϕ_n tends to zero uniformly on compact sets. It follows that one can apply Theorem 3.1 to the inclusions

$$\dot{x} \in F(x) + \phi_n(x)h(x).$$

As in the proof of Theorem 3.1, the required point x_0 is obtained as an accumulation point of the obtained sequence of points. The vector-field h can be constructed as follows: for every x_0 in the boundary of V find a vector w such that

$$w \cdot \nabla \ell^i(x_0) > 0, \qquad w \cdot \nabla m^j(x_0) < 0$$

if $\ell^i(x_0) = 0$ and $m^j(x_0) = 0$. Such a vector exists thanks to the linear independence of the gradients (in order to verify that assertion one can adapt the argument in the proof of Farkas Lemma in linear programming). The vector w can be extended to a vector-field in a neighborhood of x_0 such that the strong inequalities are preserved. Those vector-fields defined locally can by glued using a partition of unity to the global vector-field satisfying the required properties. Details of the construction in the case $M = \mathbb{R}^n$ can be found in [37]. ☐

3.4. *Generalized polyfacial sets*

Let us return to the case in which $v : M \to TM$ is a vector-field on the Riemannian manifold M. If v generates a local flow ϕ, one can still obtain a Ważewski set on the similar way to the one described in Section 3.1 if a more general type of inequalities then (3.3) and (3.4) occurs. Let $n : M \to \mathbb{R}$ be a C^1-function, let N be equal to the set $\{x \in M\colon\ n^k(x) \leqslant 0\}$, and let $x_0 \in N$ be such that

$$v(x_0) \cdot \nabla n(x_0) = 0. \tag{3.13}$$

In order to get the outward tangency property of the trajectory of x_0 with respect to N, the value of the function $t \mapsto n(\phi_t(x_0))$ should assume a sharp local minimum at $t = 0$. This is guaranteed if its second derivative at $t = 0$ exists and is positive (or, more generally, the lowest-order nonzero derivative is even and positive). By a simple calculation we conclude that the second-order criterium is satisfied if the function

$$v \cdot \nabla n : M \ni x \to \nabla n(x) \cdot v(x) \in \mathbb{R}$$

is differentiable at x_0 and the condition

$$v(x_0) \cdot \nabla(v \cdot \nabla n)(x_0) > 0 \tag{3.14}$$

holds. That observation leads to the following generalization of the notion of polyfacial set (compare [24,25]). Let $p+q+r \geqslant 1$ and let ℓ^i for $i=1,\dots,p$, m^j for $j=1,\dots,q$, and n^k for $k=1,\dots,r$ be continuous functions $M \to \mathbb{R}$ of C^1-class in some neighborhoods of their zero levels. Similarly as in Section 3.1, put

$$\begin{aligned}
V &:= \{x \in M\colon\ \ell^i(x) < 0\ \forall i,\ m^j(x) < 0\ \forall j,\ n^k(x) < 0\ \forall k\},\\
L^i &:= \{x \in \partial V\colon\ \ell^i(x) = 0\},\\
M^j &:= \{x \in \partial V\colon\ m^j(x) = 0\},\\
N^k &:= \{x \in \partial V\colon\ n^k(x) = 0\}.
\end{aligned}$$

DEFINITION 3.4. The set V is called a *generalized polyfacial set determined by* $\{\ell^i, m^j, n^k\}$. It is called a *generalized polyfacial set for* v (or, alternatively, for (1.1)) provided the inequalities (3.3) and (3.4) hold for every $i=1,\dots,p$ and $j=1,\dots,q$, and, moreover, for every $k=1,\dots,r$ the function $v \cdot \nabla n^k$ is of C^1-class in a neighborhood of its zero level, and for every $x \in N^k$ one of the following conditions is satisfied:

$$v(x) \cdot \nabla n^k(x) > 0, \tag{3.15}$$

$$v(x) \cdot \nabla n^k(x) < 0, \tag{3.16}$$

$$v(x) \cdot \nabla n^k(x) = 0, \qquad v(x) \cdot \nabla\bigl(v \cdot \nabla n^k\bigr)(x) > 0. \tag{3.17}$$

In the case $M = \mathbb{R}^n$, the expressions which appear on the left sides of the inequalities assume the forms

$$v(x) \cdot \nabla n^k(x) = \sum_{i=1}^{n} \frac{\partial n^k}{\partial x_i}(x) v_i(x), \tag{3.18}$$

$$v(x) \cdot \nabla\bigl(\nabla n^k \cdot v\bigr)(x) = \sum_{i,j=1}^{n} \left(\frac{\partial^2 n^k}{\partial x_i \partial x_j}(x) v_i(x) v_j(x) + \frac{\partial n^k}{\partial x_i}(x) \frac{\partial v_i}{\partial x_j}(x) v_j(x)\right) \tag{3.19}$$

whenever their right-hand sides are defined. They can be directly translated to the case of Equation (3.9) in $\mathbb{R}^{n+1}$ in order to get analogous inequalities to (3.11) and (3.10). In particular, for a scalar second order equation

$$x'' = g(t, x, x'), \tag{3.20}$$

where $g\colon \mathbb{R}^3 \to \mathbb{R}$ is continuous, the corresponding equation of the form (3.9) is given by

$$\dot t = 1, \qquad \dot x = y, \qquad \dot y = g(t, x, x'), \tag{3.21}$$

i.e., $v(t,x,y) = (1, y, g(t,x,x'))$, and the counterparts of (3.18) and (3.19) are

$$\begin{aligned} &v(t,x,y)\cdot\nabla n^k(t,x,y)\\ &\quad = \frac{\partial n^k}{\partial t}(t,x,y) + \frac{\partial n^k}{\partial x}(t,x,y)y + \frac{\partial n^k}{\partial y}(t,x,y)g(t,x,y), \end{aligned} \tag{3.22}$$

$$\begin{aligned} &v(t,x,y)\cdot\nabla\bigl(\nabla n^k\cdot v\bigr)(t,x,y)\\ &\quad = \begin{bmatrix} 1 & y \end{bmatrix} \begin{bmatrix} \frac{\partial^2 n^k}{\partial t^2}(t,x) & \frac{\partial^2 n^k}{\partial t\partial x}(t,x) \\ \frac{\partial^2 n^k}{\partial x\partial t}(t,x) & \frac{\partial^2 n^k}{\partial x^2}(t,x) \end{bmatrix} \begin{bmatrix} 1 \\ y \end{bmatrix} + \frac{\partial n^k}{\partial x}(t,x)g(t,x,y). \end{aligned} \tag{3.23}$$

Let us return to the general case and for $k = 1,\dots,r$ define

$$\begin{aligned} N^k_- &:= \bigl\{x \in N^k\colon\ x \text{ satisfies (3.15)}\bigr\},\\ N^k_+ &:= \bigl\{x \in N^k\colon\ x \text{ satisfies (3.16)}\bigr\},\\ N^k_0 &:= \bigl\{x \in N^k\colon\ x \text{ satisfies (3.17)}\bigr\}. \end{aligned}$$

Assume again that v generates a local flow ϕ, i.e., (1.1) satisfies the uniqueness property of the corresponding Cauchy problem. Taking into account remarks on the beginning of the current section, by an argument similar to the one in the proof of Proposition 3.1 one can observe that the sets of strict egress points E and the set of outward tangency points T of V are determined by

$$E = \left(\bigcup_{i=1}^{p} L^i \cup \bigcup_{k=1}^{r} N^k_-\right) \Bigm\backslash \left(\bigcup_{j=1}^{q} M^j \cup \bigcup_{k=1}^{r}\bigl(N^k_+ \cup N^k_0\bigr)\right), \tag{3.24}$$

$$E \cup T = \bigcup_{i=1}^{p} L^i \cup \bigcup_{k=1}^{r}\bigl(N^k_- \cup N^k_0\bigr) \tag{3.25}$$

and the same assertion as in Proposition 3.2 holds:

PROPOSITION 3.3. *If V is a generalized polyfacial set for v then*:

(i) *the set $V \cup E$ is a Ważewski set and its exit set is equal to E,*

(ii) *the set $\overline{V}$ is a Ważewski set and its exit set is equal to $E \cup T$.*

REMARK 3.2. If V is a generalized polyfacial set is determined by only one function n, i.e., $V = \{n < 0\}$, and $\nabla n(x) \neq 0$ for every x such that $n(x) = 0$ then $\overline{V}$ is a C^1-class submanifold with boundary in the phase-space M and the set $\overline{V}^{-}$ is a C^1-class submanifold with boundary in ∂V. Moreover, the same properties on homology or cohomology as stated in Remark 3.1 hold.

REMARK 3.3. Let V be a generalized polyfacial set for (1.1). If (1.1) satisfies the uniqueness property then Theorem 2.2 holds for the sets described in Proposition 3.3. On the other hand, if it does not satisfy the uniqueness property, a theorem similar to Theorem 3.1 can be formulated for those sets provided there is a sequence of locally Lipschitzean vector-fields v_n approximating v such that (3.15), (3.16), and (3.17) with v replaced by v_n hold on, respectively, N_-^k, N_+^k, and N_0^k. This can be justified by analyzing the argument in the proof of Theorem 3.1.

4. Ważewski method for retarded functional differential equations

Now we consider local semiflows generated by retarded functional differential equations (RFDEs). We provide results concerning the Ważewski method for those equations which go beyond direct substitutions of the abstract Ważewski-type theorems from Section 2 to those semiflows.

4.1. *Local semiflows generated by RFDEs*

Let M be a Riemannian manifold. For each continuous function $x : J \to M$, where $J \subset \mathbb{R}$, and for each $t \in \mathbb{R}$ put $x_t(\tau) := x(t+\tau)$ whenever it is defined. For some $r > 0$ let $\mathbf{C} := C([-r,0], M)$ be the set of continuous mappings $u : [-r,0] \to M$ endowed with the topology of uniform convergence. Let $F : \mathbf{C} \to TM$ be a continuous map such that $F(u) \in T_{u(0)}M$ for every $u \in \mathbf{C}$. By a solution of the retarded functional differential equation

$$\dot{x}(t) = F(x_t) \tag{4.1}$$

we mean a continuous function $x : [-r, \omega) \to M$ such that (4.1) is satisfied for $t \in [0, \omega)$ and $\omega > 0$ is maximal. (Here and in the sequel $\dot{x}(t)$ denotes the right-hand side derivative in the case t is equal to the left end of the interval in which x is defined.) Moreover, if for every $u \in \mathbf{C}$ there is the unique solution x of (4.1) satisfying the initial condition

$$x|_{[-r,0]} = u \tag{4.2}$$

then the right end of the domain interval of x is denoted by ω_u (i.e., $x : [-r, \omega_u) \to M$) and for $t \in [0, \omega_u)$ we define $\phi(u,t) := x_t$ (hence, in particular, $\phi(u,t)(0) = x(t)$). The mapping ϕ defined in this way is a local semiflow on the space $\mathbf{C}$.

Now let us consider a nonautonomous equation, i.e., an equation of the form

$$\dot{x}(t) = G(t, x_t), \tag{4.3}$$

where $G : U \to TM$ is continuous on an open subset U of $\mathbb{R} \times \mathbf{C}$ and $G(t,u) \in T_{u(0)}M$ for every $(t,u) \in U$. In this case the description of the corresponding local semiflow is slightly more complicated. Let $(\sigma, u) \in U$ and let $x : [\sigma - r, \omega) \to M$ be a solution of (4.3) (i.e., x satisfies the equation on $[\sigma, \omega)$ and $\omega > \sigma$ is maximal) which satisfies

$$x_\sigma|_{[-r,0]} = u. \tag{4.4}$$

In this case we call x a solution of (4.3) with the initial condition (σ, u). If for every (σ, u) there exists the unique such solution x, we define $\omega_{(\sigma,u)}$ in such a way that

$$\sigma + \omega_{(\sigma,u)} = \omega$$

(recall that ω is the right end of the interval of the existence of x) and for $t \in [0, \omega_{(\sigma,u)})$ we put

$$\psi\big((\sigma, u), t\big) := (\sigma + t, x_{\sigma+t}).$$

It is easy to check that ψ obtained in this way is a local semiflow on U.

4.2. *Ważewski-type theorems for RFDEs*

Our presentation of the basic results concerning the Ważewski method in Section 2.3 was purely topological, hence it also directly applies to local semiflows described in the previous section. However, since their phase-spaces are infinite-dimensional, usually the theorems given in Section 2.3 have limited applications, at least because the sphere in such a space is a retract of the ball. Therefore one should provide separate Ważewski-type theorems more suitable for retarded functional differential equations. To this aim we will follow an idea in the paper [67] by Rybakowski. Actually, it was preceded by [57], in which considered conditions were too restrictive, and [49], where the idea later used in [67] appeared in a special case. We will not provide Rybakowski's approach in the full generality, because it would require the introduction of several new concepts, but we rather concentrate on results which can be directly used in applications.

In the current section we assume the uniqueness of the Cauchy problem for the considered equations, hence they generated local semiflows as it was described above. At first we will deal with the autonomous case, i.e., with the local flow ϕ generated by Equation (4.1) on $\mathbf{C}$. As in Section 3.1, let V be a polyfacial set determined by $\{\ell^i, m^j\}$, i.e.,

$$V := \big\{x \in M\colon\ \ell^i(x) < 0\ \forall i,\ m^j(x) < 0\ \forall j\big\}.$$

Recall that L^i and M^j denote the sets of zeros of ℓ^i and, respectively, m^j on the boundary of V (see (3.1) and (3.2)).

DEFINITION 4.1. The set V is called a *polyfacial set for* (4.1) if

$$F(u) \cdot \nabla\ell^i\big(u(0)\big) > 0, \quad \text{if } u(0) \in L^i \text{ and } u\big([-r, 0)\big) \subset V, \tag{4.5}$$

$$F(u) \cdot \nabla m^j\big(u(0)\big) < 0, \quad \text{if } u(0) \in M^j \text{ and } u\big([-r, 0)\big) \subset V \tag{4.6}$$

for $u \in \mathbf{C}$, $i = 1, \ldots, p$, and $j = 1, \ldots, q$.

EXAMPLE 4.1 (*compare Example* 3.1 *in* [67]). Consider the scalar equation

$$\dot{x} = -ax(t) - bx(t-r), \tag{4.7}$$

where $a, b \in \mathbb{R}$. If $|a| \geqslant |b|$ then for an arbitrary $w > 0$, the interval $(-w, w)$ is a polyfacial set for (4.7) determined by the functions $t \mapsto \pm t - w$.

In the sequel we assume that V is a polyfacial set for (4.1); in particular we assume that (4.5) and (4.6) are satisfied. Let the set E be given by the formula (3.6). Put

$$\widetilde{V} := \{v \in \mathbf{C}\colon\ v(\tau) \in V\ \forall \tau \in [-r, 0]\},$$

$$\widetilde{E} := \{u \in \mathbf{C}\colon\ u(0) \in E,\ u(\tau) \in V\ \forall \tau \in [-r, 0)\}.$$

PROPOSITION 4.1. *The set* $W := \widetilde{V} \cup \widetilde{E}$ *is a Ważewski set for the local semiflow* ϕ *generated by* (4.1) *on* $\mathbf{C}$ *and its exit set* W^- *is equal to* $\widetilde{E}$.

PROOF. The set W consists of two parts. It is obvious that there are no elements of W^- contained in $\widetilde{V}$. Moreover, by (4.5), each element of $\widetilde{E}$ is in W^-, hence $W^- = \widetilde{E}$. The latter set is equal to $W \cap \partial \widetilde{V}$, hence it is closed in W, so (b) in Definition 2.4 holds. If the positive semitrajectory of a point in $\widetilde{V}$ reaches the boundary of $\widetilde{V}$, the first point of intersection belongs to $\widetilde{E}$ by (4.5) and (4.6), hence also (a) is satisfied and the result follows. □

Thus all Ważewski-type theorems for semiflows stated in Section 2.3 can be applied W. Here we do not repeat them, but instead we indicate other, more useful results. In the following theorems we assume that

- $Y \subset E$,
- $Z \subset V \cup Y$,
- $\pi : \overline{Z} \cap (Z \cup Y) \to \mathbf{C}$ is a continuous map.

The first result is a counterpart of Theorem 2.1.

THEOREM 4.1. *Assume that for every* $z \in \overline{Z} \cap (Z \cup Y)$,

$$\pi(z)(0) = z, \qquad \pi(z)\big([-r, 0)\big) \subset V.$$

If Y *is not a strong deformation retract of* $Z \cup Y$ *in* $V \cup Y$ *then there exists a* $z_0 \in Z$ *such that the solution* $x : [-r, \omega) \to M$ *of Equation* (4.1) *with the initial condition* $x|_{[-r,0]} = \pi(z_0)$ *satisfies*

- *either* $x([0, \omega)) \subset V$
- *or else there exists* $t_0 \in (0, \omega)$ *such that*

$$x\big([0, t_0)\big) \subset V, \qquad x(t_0) \in E \setminus Y.$$

PROOF. Assume on the contrary that the conclusion does not hold. In our notation it means that for every $z \in \overline{Z} \cap (Z \cup Y)$,

$$\pi(z) \in W^*, \qquad \sigma\big(\pi(z)\big) \in Y,$$

where $\sigma : W^* \to [0, \infty)$ is the escape-time function. By Lemma 2.1 and Proposition 4.1, σ is continuous. For $y \in Z \cup Y$ and $t \in [0, 1]$ define

$$h(y,t) := \begin{cases} \phi\big(\pi(y), t\sigma\big(\pi(y)\big)\big)(0), & \text{if } y \in Z, \\ y, & \text{if } y \in Y. \end{cases}$$

It is easy to verify that h satisfies the conditions required in Definition 2.6. □

The presented below counterpart of Corollary 2.1 cannot by directly derived from Theorem 4.1 because of a weaker hypothesis imposed on π. Nevertheless, its proof based on already presented arguments is straightforward.

THEOREM 4.2. *Assume that for every $z \in Z \setminus Y$,*

$$\pi(z)\big([-r, 0]\big) \subset V$$

and for every $y \in \overline{Z} \cap Y$,

$$\pi(y)(0) = y, \qquad \pi(y)\big([-r, 0)\big) \subset V.$$

If Y is not a retract of $Z \cup Y$ then the conclusion of Theorem 4.1 *holds.*

PROOF. If the conclusion is false then the map $r : Z \cup Y \to Y$ given by the formula

$$r(y) := \begin{cases} \phi\big(\pi(y), \sigma\big(\pi(y)\big)\big)(0), & \text{if } y \in Z, \\ y, & \text{if } y \in Y, \end{cases}$$

is a retraction which contradicts to the assumption. □

Essentially, the above theorems were stated in [67, Section 3]. In order to match to our presentation in Section 2.3, we strengthened them slightly (in particular, by considering the set Y). However, a direct strengthening of Theorem 4.1 in order to get a counterpart of Theorem 2.2 is not possible. A counterexample with a suitable choice of V, Z, and the maps F and π can be based on Figure 5. (Such a counterexample does not contradict to Remark 2.1.) The theorems can be extended to equations without the uniqueness property in a similar way as described in Section 3.2 (compare [67, Section 4]).

In the theorems we assumed the existence of a map π having some special properties. One can ask whether such a mapping can be constructed for a given polyfacial set V and a set Z. In practice this is always possible; if the polyfacial set is regular (see Definition 3.2) then, by Remark 3.1, $\overline{V}$ is a topological manifold with boundary and thus it has a *collar*, i.e., there is an embedding of $\partial V \times [0,1]$ into $\overline{V}$ for which $\partial V \times \{0\}$ corresponds to ∂V (compare [17, Section VI.9]). Now it is easy to construct π such that $\pi(z)$ is equal to the constant map $\tau \mapsto z$ if z is outside of the image of the embedding.

One can formulate similar theorems for the nonautonomous equation (4.3). Formally, contrary to the case of ordinary differential equations, they are not direct consequences of the results stated for the autonomous ones by attaching the equation $\dot{t} = 1$, at least because the phase-space of the local flow generated by (4.3) does not coincide with $C([-r,0], \mathbb{R} \times M)$. However, in spite of slightly greater complexity, there are no essential problems in their formulating and proving—that can be done in the same way as above for the autonomous case.

We skip the statements of those general theorems and for simplicity concentrate on the case $M = \mathbb{R}^n$ where we assume that the domain U of the function G on the right-hand side of the equation (4.3) is equal to $\mathbb{R} \times \mathbf{C}$ and $\mathbf{C} = C([-r,0], \mathbb{R}^n)$. (For a more general situation the reader is referred to [67].) Let V be a polyfacial set in $\mathbb{R}^{n+1}$ determined by $\{\ell^i, m^j\}$ and let E be given by (3.6). The counterparts of (4.5) and (4.6) are

$$\frac{\partial \ell^i}{\partial t}\big(\sigma, x(\sigma)\big) + \sum_{k=1}^{n} \frac{\partial \ell^i}{\partial x_k}\big(\sigma, x(\sigma)\big) G_k(\sigma, x_\sigma) > 0,$$

$$\text{if } x_\sigma \in \mathbf{C},\ \big(\sigma, x(\sigma)\big) \in L^i,\ \big(\tau, x(\tau)\big) \subset V\ \forall \tau \in [\sigma - r, \sigma), \tag{4.8}$$

$$\frac{\partial m^j}{\partial t}\big(\sigma, x(\sigma)\big) + \sum_{k=1}^{n} \frac{\partial m^j}{\partial x_k}\big(\sigma, x(\sigma)\big) G_k(\sigma, x_\sigma) < 0,$$

$$\text{if } x_\sigma \in \mathbf{C},\ \big(\sigma, x(\sigma)\big) \in M^j,\ \big(\tau, x(\tau)\big) \subset V\ \forall \tau \in [\sigma - r, \sigma), \tag{4.9}$$

The following result resembles Theorems 4.1 and 4.2.

THEOREM 4.3. *Assume that* (4.8) *and* (4.9) *are satisfied and one of the following conditions hold*:

(i) *for every* $(\sigma, z) \in \overline{Z} \cap (Z \cup Y)$,

$$\pi\big((\sigma, z)\big)(0) = z, \quad \big(\sigma + \tau, \pi(\sigma, z)(\tau)\big) \in V\ \forall \tau \in [-r, 0),$$

and Y is not a strong deformation retract of $Z \cup Y$ in $V \cup Y$,

(ii) *for every* $(\sigma, z) \in Z \setminus Y$,

$$\pi(\sigma, z)\big([-r, 0]\big) \subset V$$

and for every $(\sigma, y) \in \overline{Z} \cap Y$,

$$\pi\big((\sigma, y)\big)(0) = z, \quad \big(\sigma + \tau, \pi(\sigma, y)(\tau)\big) \in V\ \forall \tau \in [-r, 0),$$

and, moreover, Y is not a retract of $Z \cup Y$.

Then there exists a $(\sigma_0, z_0) \in Z$ *such that the solution* $x : [-r + \sigma_0, \omega) \to \mathbb{R}^n$ *of the equation* (4.1) *with the initial condition* $x_{\sigma_0}|_{[-r,0]} = \pi(\sigma_0, z_0)$ *satisfies*

- *either* $(t, x(t)) \subset V$ *for every* $t \in [\sigma_0, \omega)$
- *or else there exists* $t_0 \in (\sigma_0, \omega)$ *such that*

$$\bigl(t, x(t)\bigr) \in V \ \forall t \in [\sigma_0, t_0), \qquad \bigl(t_0, x(t_0)\bigr) \in E \setminus Y.$$

PROOF. One can follow the ideas of the proofs of Theorems 4.1 and 4.2. Define

$$\begin{aligned} \widetilde{V} &:= \bigl\{(\sigma, u) \in \mathbb{R} \times \mathbf{C}\colon\ \bigl(\sigma + \tau, u(\tau)\bigr) \in V \ \forall \tau \in [-r, 0]\bigr\}, \\ \widetilde{E} &:= \bigl\{(\sigma, u) \in \mathbb{R} \times \mathbf{C}\colon\ \bigl(\sigma, u(0)\bigr) \in E,\ \bigl(\sigma + \tau, u(\tau)\bigr) \in V \ \forall \tau \in [-r, 0)\bigr\}. \end{aligned}$$

Their union is a Ważewski set for the local semiflow ψ on $\mathbb{R} \times \mathbf{C}$ described in the previous section and the continuity of its escape-time function is used in definitions of the suitable mappings. Details are straightforward and are omitted here. □

The equation considered in the following two examples of application of Theorem 4.3 is taken from [67, Example 3.2]. Similar applications for ordinary differential equation were given in Examples 1.1 and 1.2.

EXAMPLE 4.2. In the planar equation

$$\begin{aligned} \dot{x}(t) &= ax(t) + bx(t - r) + f(t, x_t, y_t), \\ \dot{y}(t) &= -cy(t) + dy(t - r) + g(t, x_t, y_t) \end{aligned} \tag{4.10}$$

let $a, c > 0$ and let

$$(f, g) : \mathbb{R} \times \mathbf{C} \times \mathbf{C} \to \mathbb{R}^2,$$

where $\mathbf{C} = C([-r, 0], \mathbb{R})$, be a continuous map. Assume that (4.10) satisfies the uniqueness of the Cauchy problem,

$$\bigl|f(t, u, v)\bigr| < |a| - |b|$$

for all $(t, u, v) \in \mathbb{R} \times \mathbf{C} \times \mathbf{C}$ for which

$$\begin{aligned} &\bigl|u(0)\bigr| = 1, \qquad \bigl|v(0)\bigr| \leqslant 1, \\ &\bigl|u(\tau)\bigr| < 1, \qquad \bigl|v(\tau)\bigr| < 1 \quad \forall \tau \in [-r, 0), \end{aligned}$$

and

$$\bigl|g(t, u, v)\bigr| < |c| - |d|$$

for all $(t, u, v) \in \mathbb{R} \times \mathbf{C} \times \mathbf{C}$ for which

$$\begin{aligned} &\bigl|u(0)\bigr| \leqslant 1, \quad \bigl|v(0)\bigr| = 1, \\ &\bigl|u(\tau)\bigr| < 1, \quad \bigl|v(\tau)\bigr| < 1 \quad \forall \tau \in [-r, 0). \end{aligned}$$

Let the sets V, E, and Z be the same as in Example 1.1, i.e., they are given by, respectively, (1.7), (1.8), and (1.9) (see also Figure 2). V is a polyfacial set for (4.10). Define $\pi : Z \cup E \to \mathbf{C} \times \mathbf{C}$ by

$$\pi(t, x, y)(\tau) := \left(\frac{r + \tau}{r} x, y\right)$$

for $(t, x, y) \in Z \cup E$ and $\tau \in [-r, 0]$. By Theorem 4.3 we conclude that there is an $x_0 \in (-1, 1)$ such that the graph of the solution of (4.10) starting at $(0, \pi(0, x_0, 0))$ is contained in V.

EXAMPLE 4.3. We associate the boundary value problem (1.10) to Equation (4.10) in Example 4.2. As in Example 1.2, we define V, E, Z, and Y by, respectively (1.11), (1.12), (1.9), and (1.13). Recall, that those sets are shown on Figure 3. Let π be the mapping defined in Example 4.2. By Theorem 4.3 we conclude the existence of a solution of the problem (4.10), (1.10) which starts at $(0, \pi(0, x_0, 0))$ for some $x_0 \in (-1, 1)$.

5. Some topological concepts

In order to state variants of the Ważewski theorem, we have already defined the notion of retract and its modifications. In this section we recall several other topological notions and results which will be used in description of properties of the sets of solutions obtained by the Ważewski method.

5.1. *Topological pairs, quotient spaces, and pointed spaces*

By a *topological pair* we mean a pair (X, A) consisting of a topological space X and its subset A. If A is closed in X then (X, A) is a called a *closed pair*. If $x_0 \in X$, we frequently write (X, x_0) for the topological pair $(X, \{x_0\})$. Let (X, A) and (Y, B) be topological pairs. We write $f : (X, A) \to (Y, B)$ if f is continuous map $X \to Y$ and $f(A) \subset B$.

Now we recall the notion of quotient space. We impose a special attention to the case in which the empty space is involved. Let (X, A) be a topological pair.

DEFINITION 5.1. The *quotient space* X/A is the set whose elements are all points of $X \setminus A$ and the set A (denoted by $[A]$ in order to avoid confusions), i.e.,

$$X/A := (X \setminus A) \cup \bigl\{[A]\bigr\},$$

endowed with the topology for which $U \subset X/A$ is open if and only if the set $\{x \in X\colon\ [x] \in U\}$ is open in X, where

$$[x] := \begin{cases} x, & \text{if } x \in X \setminus A, \\ [A], & \text{if } x \in A. \end{cases}$$

It follows that the *quotient map*

$$q : X \ni x \to [x] \in X/A$$

is continuous. By definition, $X/\emptyset$ is equal to $X \cup \{[\emptyset]\}$ and its topology is equal to the direct sum topology of X and the one-point space $\{[\emptyset]\}$. In particular, $\emptyset/\emptyset$ is equal to $\{[\emptyset]\}$. Since $[A]$ is distinguished in X/A in a canonical way, in order to simplify the notation, for an arbitrary pair (X, A) we write shortly $(X/A, *)$ instead of $(X/A, [A])$.

For presentation of results involving quotient spaces we use the Čech cohomology functor $\check{H}^* = \{\check{H}^q\}_{q \in \mathbb{Z}}$ having coefficients in a fixed commutative ring-with-unit R. We recall that it is isomorphic to the Alexander–Spanier cohomology functor (compare [71]).

REMARK 5.1. By the strong excision of the Čech cohomology, if (X, A) is a closed pair then the quotient map induces an isomorphism

$$\check{H}^*(q) : \check{H}^*(X/A, *) \xrightarrow{\cong} \check{H}^*(X, A).$$

By a *pointed space* we mean a topological pair (X, x_0), where $x_0 \in X$. In this case x_0 is called the *base point*. For example, $(X/A, *)$ is a pointed space with the base point $*$ equal to $[A]$. Let (X, x_0) and (Y, y_0) be two pointed spaces. In this case $f : (X, x_0) \to (Y, y_0)$ means $f : X \to Y$ continuous such that $f(x_0) = y_0$. We put

$$X \vee Y := X \times \{y_0\} \cup \{x_0\} \times Y$$

and define the *wedge sum of* (X, x_0) *and* (Y, y_0) by

$$(X, x_0) \vee (Y, y_0) := \bigl(X \vee Y, (x_0, y_0)\bigr).$$

Similarly, we put

$$X \wedge Y := X \times Y/(X \vee Y),$$

$$(X, x_0) \wedge (Y, y_0) := (X \wedge Y, *)$$

and call $(X, x_0) \wedge (Y, y_0)$ the *smash product of* (X, x_0) *and* (Y, y_0).

5.2. *Homotopy types*

Recall that maps $f, g : X \to Y$ are *homotopic* (denoted $f \simeq g$) if for some continuous map $h : X \times [0, 1] \to Y$, f is equal to $x \mapsto h(x, 0)$ and g is equal to $x \mapsto h(x, 1)$. If, moreover,

$f, g : (X, x_0) \to (Y, y_0)$ then they are called *pointed-homotopic* (denoted $f \simeq_* g$) if the map h additionally satisfies $h(x_0, t) = y_0$ for every $t \in [0, 1]$.

DEFINITION 5.2. We say that topological spaces X and Y have the same *homotopy type* (and we write $X \simeq Y$) if for some maps $f : X \to Y$ and $g : Y \to X$ (called *homotopy equivalences*),

$$g \circ f \simeq \mathrm{id}_X, \qquad f \circ g \simeq \mathrm{id}_Y.$$

We say that the pointed spaces (X, x_0) and (Y, y_0) have the same *pointed homotopy type* (we write $(X, x_0) \simeq_* (Y, y_0)$) if there exist $f : (X, x_0) \to (Y, y_0)$ and $g : (Y, y_0) \to (X, x_0)$ such that

$$g \circ f \simeq_* \mathrm{id}_X, \qquad f \circ g \simeq_* \mathrm{id}_Y.$$

The *absolute homotopy type* (shortly: *homotopy type*) of X is the equivalence class of X in the relation $\simeq$. It is denoted by $[X]$. Similarly, the *pointed homotopy type* of (X, x_0) is the equivalence class of (X, x_0) in the relation $\simeq_*$ and is denoted by $[X, x_0]$.

Obviously, $[X, x_0]$ is a more restrictive class of spaces then $[X]$. Since a homology or cohomology functor H is homotopy invariant, we define

$$H[X] := H(X), \qquad H[X, x_0] := H(X, x_0).$$

One can prove that the following definitions of the *wedge sum* and *smash product* of homotopy types are correctly stated:

$$[X, x_0] \vee [Y, y_0] := \bigl[(X, x_0) \vee (Y, y_0)\bigr],$$
$$[X, x_0] \wedge [Y, y_0] := \bigl[(X, x_0) \wedge (Y, y_0)\bigr].$$

We distinguish the *trivial pointed homotopy type*:

$$\overline{0} := \bigl[\{[\emptyset]\}, [\emptyset]\bigr],$$

i.e., the homotopy type of a pointed one-point space, which is the neutral element with respect to the wedge sum.

REMARK 5.2. There is no inverse to a nontrivial pointed homotopy type with respect to the wedge sum, i.e.,

$$[X, x_0] \vee [Y, y_0] = \overline{0}$$

implies

$$[X, x_0] = \overline{0}, \qquad [Y, y_0] = \overline{0}.$$

Indeed, if there is a continuous map $F : X \vee Y \times [0, 1] \to X \vee Y$ such that

$$F\big((x, y), 0\big) = (x, y), \qquad F\big((x, y), 1\big) = F\big((x_0, x_0), t\big) = (x_0, x_0)$$

for every t, x, y then the map $G : X \times [0, 1] \to X$ defined by

$$G(x) := \begin{cases} \pi F\big((x, y_0), t\big), & \text{if } F\big((x, y_0), t\big) \in X \times \{y_0\}, \\ x_0, & \text{otherwise,} \end{cases}$$

where $\pi : X \times Y \to X$ denotes the projection, provides

$$(X, x_0) \simeq_* \big(\{x_0\}, x_0\big).$$

The following simple lemma is useful in the Ważewski method.

LEMMA 5.1. *If A is a strong deformation retract of X then*

$$[X/A, *] = \overline{0}.$$

In the following example we illustrate the difference between the absolute and pointed homotopy types.

EXAMPLE 5.1. The unit zero- and one-dimensional spheres S^0 and S^1 are, respectively, the set $\{-1, 1\}$ and the circle $\{x\colon |x| = 1\}$ in $\mathbb{C}$. They naturally provide pointed spaces $(S^0, 1)$ and $(S^1, 1)$. By definition, $S^0 \vee S^1$ is homeomorphic to the disjoint union of S^1 and a one-point space, hence it is homeomorphic to $S^1/\emptyset$ and

$$\big[S^0 \vee S^1\big] = \big[S^1/\emptyset\big].$$

On the other hand, each homotopy equivalence $S^0 \vee S^1 \to S^1/\emptyset$ transforms the base point $(1, 1)$ of $(S^0, 1) \vee (S^1, 1)$ to a point in the other component of $S^1/\emptyset$ then $[\emptyset]$, hence

$$\big[S^0, 1\big] \vee \big[S^1, 1\big] = \big[\big(S^0, 1\big) \vee \big(S^1, 1\big)\big] \neq \big[S^1/\emptyset, *\big].$$

5.3. *Absolute neighborhood retracts*

We recall the topological notions of absolute and Euclidean neighborhood retracts.

DEFINITION 5.3 (*compare* [14,33]). A metrizable space X is called an *absolute neighborhood retract* (shortly: ANR) if there is an open set U in a normed space E and a map $h : X \to E$ which is a homeomorphism onto its image $h(X)$ such that $h(X)$ is a retract of U.

REMARK 5.3. There is an equivalent definition: A metrizable space X is an ANR if and only if for each closed subset A of a metrizable space W and every continuous map $f : A \to X$ there exists a neighborhood U of A in W and a continuous map $F : U \to X$ such that $F|_A = f$ (compare [14, Theorem IV(4.2)(ii)]).

We refer to [14] for other properties of ANRs.

DEFINITION 5.4 (*compare* [28]). A metrizable space X is called an *Euclidean neighborhood retract* (shortly: ENR) if there is an $n \in \mathbb{N}$, an open set U in $\mathbb{R}^n$, and a map $h : X \to \mathbb{R}^n$ which is a homeomorphism onto its image $h(X)$ such that $h(X)$ is a retract of U.

The class of ENRs coincides with the class of locally compact, separable, and finite-dimensional ANRs. Examples of ENRs include all compact polyhedra and manifolds with boundary. The Hilbert cube is a compact ANR which is not an ENR.

5.4. *Lusternik–Schnirelmann category*

We provide the concept of the Lusternik–Schnirelmann category for topological pairs, which appears to be very convenient in connection to the Ważewski method. At first we recall the notion of contractibility.

DEFINITION 5.5. A set $A \subset X$ is called *contractible in X* if there exists a continuous map $h : A \times [0, 1] \to X$ such that $h(x, 0) = x$ for every $x \in A$ and $h(x, 1) = h(y, 1)$ for every $x, y \in A$.

Notice that the above definition is also valid also for $A = \emptyset$. Let (X, Y) be a closed pair. Motivated by the definition posed by Reeken, Fournier, and Willem (compare [31]) we put

DEFINITION 5.6. The *relative category* $\mathrm{cat}_{X,Y}(A)$ of a closed set $A \subset X$ is an element of $\mathbb{N} \cup \{\infty\}$ such that

$$\begin{array}{ll} \mathrm{cat}_{X,Y}(A) := 0, & \text{if } Y \text{ is a strong deformation retract of } A \cup Y \text{ in } X. \\ \mathrm{cat}_{X,Y}(A) \leqslant n, & \text{if there exist closed sets } A_i \subset A,\ i = 0, \dots, n, \text{ such that} \\ & A = \displaystyle\bigcup_{i=0}^{n} A_i, \\ & A_i \text{ is contractible in } X \text{ for } i = 1, \dots, n, \text{ and} \\ & Y \text{ is a strong deformation retract of } A_0 \cup Y \text{ in } X. \end{array}$$

(Compare Definition 2.6 for the notion of a strong deformation retract in a space.)

Formally, we did not exclude any possibility in which the empty set appears. In particular, for closed $A \subset X$ we define

$$\mathrm{cat}_X(A) := \mathrm{cat}_{X,\emptyset}(A).$$

It follows that if $X \neq \emptyset$ then $\mathrm{cat}_X(A)$ is the minimal number of closed contractible in X subsets of A which cover A (since $A_0 = \emptyset$) or, if such a number does not exist, it is equal to ∞.

DEFINITION 5.7. Define

$$\mathrm{cat}(X, Y) := \mathrm{cat}_{X,Y}(X),$$
$$\mathrm{cat}(X) := \mathrm{cat}_X(X)$$

and call them the *category of* (X, Y) and, respectively the *category of* X.

It is easy to check that for a closed $A \subset X$,

$$\mathrm{cat}_X(A) \leqslant \mathrm{cat}(A), \tag{5.1}$$
$$\mathrm{cat}_{X,Y}(A) \leqslant \mathrm{cat}_X(A) \leqslant \mathrm{cat}_{X,Y}(A) + \mathrm{cat}_X(Y). \tag{5.2}$$

In particular,

$$\mathrm{cat}(X, Y) \geqslant \mathrm{cat}(X) - \mathrm{cat}_X(Y) \geqslant \mathrm{cat}(X) - \mathrm{cat}(Y) \tag{5.3}$$

and if x_0 is a point in X then

$$\mathrm{cat}_{X,x_0}(A) \leqslant \mathrm{cat}_X(A) \leqslant \mathrm{cat}_{X,x_0}(A) + 1. \tag{5.4}$$

EXAMPLE 5.2. In the notation used in Example 5.1,

$$\mathrm{cat}\big(S^1\big) = \mathrm{cat}\big(S^1/\emptyset, *\big) = \mathrm{cat}\big(\big(S^0, 1\big) \vee \big(S^1, 1\big)\big) = 2, \qquad \mathrm{cat}\big(S^1/\emptyset\big) = 3.$$

It follows immediately by definition that the category is an invariant of the absolute and pointed homotopy types. Hence we can put the following definitions:

DEFINITION 5.8. The categories of the homotopy types are defined as

$$\mathrm{cat}[X] := \mathrm{cat}(X),$$
$$\mathrm{cat}[X, x_0] := \mathrm{cat}(X, x_0).$$

We list the basic properties of the category.

PROPOSITION 5.1 (compare [31]). *Let A and B be closed subsets of X.*

(i) *If $A \subset B$ then* $\mathrm{cat}_{X,Y}(A) \leqslant \mathrm{cat}_{X,Y}(B)$.

(ii) $\mathrm{cat}_{X,Y}(A \cup B) \leqslant \mathrm{cat}_{X,Y}(A) + \mathrm{cat}_X(B)$.

(iii) *If there exists a continuous map $h : A \times [0, 1] \to X$ such that $h(x, 0) = x$ for $x \in A$, $h(x, t) = x$ for $x \in A \cap Y$ and $t \in [0, 1]$ then*

$$\mathrm{cat}_{X,Y}\big(h_1(A)\big) \leqslant \mathrm{cat}_{X,Y}(A).$$

(iv) *If both X and Y are ANRs then there exists a closed neighborhood U of A such that*

$$\mathrm{cat}_{X,Y}(U) = \mathrm{cat}_{X,Y}(A).$$

PROOF. Proofs of (i), (ii), and (iii) are immediate. As for (iv), we refer to [31, Proposition 2.9]. □

PROPOSITION 5.2. *If (X, A) is a closed pair of ANRs then*

$$\mathrm{cat}(X, A) \geqslant \mathrm{cat}(X/A, *) \geqslant \mathrm{cat}(X/A) - 1.$$

PROOF. We can assume that $\mathrm{cat}(X, A) = n < \infty$. Let A_i, $i = 0, \ldots, n$, be a closed covering of X such that $A \subset A_0$, A is a strong deformation retract of A_0, and A_i are contractible in X for $i \geqslant 1$. Since X and A are ANRs, as in the proof of [31, Proposition 2.9] one can find a closed neighborhood U_0 of A_0 for which A is a strong deformation retract. For $i = 1, \ldots, n$ define A_i' as $A_i \setminus \mathrm{int}\, U_0$, hence $A_i' \cap A = \emptyset$. Using the subsets U_0/A, $A_1', \ldots, A_n'$ of X/A one gets

$$\mathrm{cat}(X/A, *) \leqslant n,$$

hence the left inequality follows. The right-hand inequality is given in (5.4). □

5.5. *Cup-length*

Here we use the cup-product (denoted by $\smile$) in the Čech cohomologies (compare [28] for results which concern that notion and which we apply in the sequel). We provide the notion of cup-length in the context of pair of spaces, compare [31, Definition 3.2]. Notice that in our definition, the cup-length of a pair is shifted by one with respect to the cup-length considered in that paper.

DEFINITION 5.9. The *cup-length* of a topological pair (X, Y) is an element $\ell(X, Y)$ of $\mathbb{N} \cup \{\infty\}$ which satisfies

$$\begin{aligned} &\ell(X, Y) = 0, && \text{if } \check{H}^*(X, Y) = 0, \\ &\ell(X, Y) = 1, && \text{if } \check{H}^0(X, Y) \neq 0,\ \check{H}^q(X) = 0 \text{ for } q \geqslant 1, \\ &\ell(X, Y) \geqslant n, && \text{if there exist } q \geqslant 0,\ v \in \check{H}^q(X, Y),\ q_i \geqslant 1, \text{ and } u_i \in \check{H}^{q_i}(X) \\ & && \text{for } i = 1, \ldots, n-1 \text{ such that } v \smile u_1 \smile \cdots \smile u_{n-1} \neq 0. \end{aligned}$$

We write $\ell(X) := \ell(X, \emptyset)$. It follows $\ell(\emptyset) = 0$, $\ell(X) = 1$ if $X \neq \emptyset$ and $\check{H}^q(X) = 0$ for $q \geqslant 1$, and $\ell(X) \geqslant n$ if there exists $q_i \geqslant 1$ and $u_i \in \check{H}^{q_i}(X)$, $i = 1, \ldots, n-1$, such that

$u_1 \smile \cdots \smile u_{n-1} \neq 0$. Indeed, it suffices to put $v = 1 \in \check{H}^0(X, \emptyset)$. Moreover, if $x_0 \in X$ then

$$\ell\bigl(X, \{x_0\}\bigr) \leqslant \ell(X) \leqslant \ell\bigl(X, \{x_0\}\bigr) + 1, \tag{5.5}$$

since $\check{H}^q(X, \{x_0\}) = \check{H}^q(X)$ for $q \geqslant 1$. Moreover, it follows that the cup-length is an invariant of the Čech cohomology ring of (X, x_0), hence also of the pointed homotopy type $[X, x_0]$. Thus we can put

DEFINITION 5.10. The cup-lengths of the homotopy types are defined as

$$\ell[X] := \ell(X),$$
$$\ell[X, x_0] := \ell(X, x_0).$$

PROPOSITION 5.3. *If (X, A) is a closed pair then*

$$\ell(X, A) \geqslant \ell(X/A, *) \geqslant \ell(X/A) - 1.$$

PROOF. Assume $\ell(X/A, *) \geqslant n$. The diagram

$$\begin{tikzcd}
\check{H}^*(X/A, *) \otimes \check{H}^*(X/A) \otimes \cdots \otimes \check{H}^*(X/A) \arrow[r, "\smile"] \arrow[d, "\cong"'] & \check{H}^*(X/A, *) \arrow[d, "\cong"] \\
\check{H}^*(X, A) \otimes \check{H}^*(X) \otimes \cdots \otimes \check{H}^*(X) \arrow[r, "\smile"] & \check{H}^*(X, A)
\end{tikzcd}$$

with the vertical lines induced by the quotient map, is commutative. By Remark 5.1 and the commutativity, if

$$v \smile u_1 \smile \cdots \smile u_{n-1} \neq 0$$

in $\check{H}^*(X/A, *)$, then the corresponding element in $\check{H}^*(X, A)$ is also nonzero, hence $\ell(X, A) \geqslant n$ and the left inequality follows. The right-hand inequality is given in (5.5). $\square$

The most important result concerning the notion of cup-length relates it to the Lusternik–Schnirelmann category.

PROPOSITION 5.4 (compare Theorem 3.6 in [31]). *If (X, Y) is a closed pair then*

$$\operatorname{cat}(X, Y) \geqslant \ell(X, Y).$$

5.6. *Lefschetz theorem and the fixed point index*

Recall that if $f:(X,A)\to(X,A)$ is a continuous map, H is a homology or cohomology functor over $\mathbb{Q}$, and the vector-spaces $H_q(X,A)$ are finitely dimensional and equal to zero for almost all q then the *Lefschetz number* of f (with respect to H) is defined as

$$\Lambda(f):=\sum_{q=0}^{\infty}(-1)^q\operatorname{tr}H_q(f),$$

where $\operatorname{tr}H_q(f)$ denotes the trace of the endomorphism $H_q(f)$ of $H_q(X,A)$. If f is equal to the identity on (X,A) then its Lefschetz number is denoted by $\chi(X,A)$ and is called the Euler characteristic of (X,A) (with respect to H), i.e.,

$$\chi(X,A)=\sum_{q=0}^{\infty}(-1)^q\dim H_q(X,A).$$

If (X,A) is a pair of compact ANRs then the Lefschetz number is always defined and independent of the choice of H, and

$$\Lambda(f)=\Lambda(f|_X)-\Lambda(f|_A),$$

where $f|_X:X\to X$ and $f|_A:A\to A$ are the restrictions of f. In particular,

$$\chi(X,A)=\chi(X)-\chi(A).$$

Since homologies and cohomologies are homotopy invariants, we define

$$\begin{aligned}&\chi[X]:=\chi(X),\\&\chi[X,x_0]:=\chi(X,x_0)=\chi[X]-1.\end{aligned}$$

PROPOSITION 5.5. *If (X,A) is a closed pair and the Euler characteristic with respect to the Čech cohomologies $\check{\chi}(X,A)$ is defined then*

$$\check{\chi}(X,A)=\check{\chi}(X/A,*)=\check{\chi}(X/A)-1.$$

Moreover, if X and A are compact ENRs then X/A is also a compact ENR and

$$\chi(X,A)=\chi(X/A,*)=\chi(X/A)-1$$

does not depend on the choice of homology or cohomology.

PROOF. This is a consequence of Remark 5.1 and [28, Chapter IV.8]. ☐

The Lefschetz number is used in formulation of the following well-known result:

PROPOSITION 5.6 (Lefschetz Fixed Point Theorem). *If X is a compact ANR, $f : X \to X$ is a continuous map, and $\Lambda(f) \neq 0$ then f has a fixed point.*

Assume that X is an ENR. For a continuous map $f : D \to X$, where $D \subset X$, define its set of fixed points as

$$\operatorname{Fix}(f) := \{x \in D\colon\ f(x) = x\}.$$

DEFINITION 5.11. A set P is called an *isolated set of fixed points* of f if it is compact and there exists an open set U in X such that $U \subset D$ and

$$P = U \cap \operatorname{Fix}(f).$$

If Q is an isolated set of fixed points of $g : D \to X$, we write $(f, P) \simeq (g, Q)$ (and call (f, P) *homotopic to* (g, Q)) if for every $t \in [0, 1]$ there exist a map $f_t : D \to X$ such that

$$F : D \times [0, 1] \ni (x, t) \to \big(f_t(x), t\big) \in X \times [0, 1]$$

is continuous and there exists an isolated set P^* of fixed points of F such that $f_0 = f$, $f_1 = g$, $P = \{x\colon\ (x, 0) \in P^*\}$, and $Q = \{x\colon\ (x, 1) \in P^*\}$.

To such an isolated set P one can associate an integer number $\operatorname{ind}(f, P)$, called the *fixed point index*, see the book [28]. We use a slightly different notation with respect to that in the book of Dold; our $\operatorname{ind}(f, P)$ means the same as the index of $f|_U$ in [28]. Below we list its properties:

PROPOSITION 5.7 (compare [28]).

(i) (Solvability) *If* $\operatorname{ind}(f, P) \neq 0$ *then* $P \neq \emptyset$.

(ii) (Additivity) *If P and Q are isolated sets of fixed points of f and $P \cap Q = \emptyset$ then*

$$\operatorname{ind}(f, P \cup Q) = \operatorname{ind}(f, P) + \operatorname{ind}(f, Q).$$

(iii) (Multiplicativity) *If $f : X \to X$ and $g : Y \to Y$, P and Q are isolated sets of fixed points of f and g, respectively, then*

$$\operatorname{ind}(f \times g, P \times Q) = \operatorname{ind}(f, P)\operatorname{ind}(g, Q).$$

(iv) (Homotopy Invariance) *If* $(f, P) \simeq (g, Q)$ *then* $\operatorname{ind}(f, P) = \operatorname{ind}(g, Q)$.

(v) (Commutativity) *If $f : D \to X'$ and $f' : D' \to X$, where $D \subset X$ and $D' \subset X'$, and P is an isolated set of fixed points of $f' \circ f$ then $f(P)$ is an isolated set of fixed points of $f \circ f'$ and*

$$\operatorname{ind}(f' \circ f, P) = \operatorname{ind}\big(f \circ f', f(P)\big).$$

(vi) (Lefschetz Property) *If X is compact and $f : X \to X$ then*

$$\operatorname{ind}(f, \operatorname{Fix} f) = \Lambda(f).$$

6. Properties of the asymptotic parts of Ważewski sets

Roughly speaking, abstract Ważewski-type theorems provide sufficient conditions for

$$W \setminus W^* \neq \emptyset$$

if W is a Ważewski set. Here we consider the question what additional information on that asymptotic part can be obtained from the properties of W and W^-; in particular whether it contains stationary or periodic points and what is its topology. In this section we concentrate on qualitative results valid for local semiflows. For some more accurate results which are easier to handle if local flows are considered we refer to the next section.

6.1. *Estimates on the category of the asymptotic part*

We assume that ϕ be a local semiflow on X. It seems that the paper [63] of Poźniak was the first one in which the notion of category appeared in the context of the Ważewski method and Conley index. We base on some ideas from that paper.

THEOREM 6.1. *Let $W \subset X$ be a Ważewski set, Assume that both W and W^- are ANRs closed in X and Z is a closed subset of W. Then*

$$\operatorname{cat}_W\bigl(Z \setminus W^*\bigr) \geqslant \operatorname{cat}_{W,W^-}(Z).$$

PROOF. $Z \setminus W^*$ is closed by Lemma 2.1. By Proposition 5.1, there exists U, a closed neighborhood of $Z \setminus W^*$ in W such that $\operatorname{cat}_W(W \setminus W^*) = \operatorname{cat}_W(U)$. Put $V := Z \setminus \operatorname{int} U$ (where int denotes the interior relative to W). It follows by Proposition 5.1

$$\operatorname{cat}_W\bigl(Z \setminus W^*\bigr) = \operatorname{cat}_W(U) = \operatorname{cat}_W(U) + \operatorname{cat}_{W,W^-}(V) \geqslant \operatorname{cat}_{W,W^-}(Z)$$

since $Z = U \cup V$ and $V \subset W^*$, hence W^- is a strong deformation retract of $V \cup W^-$ in W. ☐

Under the imposed conditions on W and W^-, the above theorem can be regarded as an extension of Corollary 2.3, since if W^- is not a strong deformation retract of $Z \cup W^-$ in W then, by definition, $\operatorname{cat}_{W,W^-}(Z) \geqslant 1$, hence $\operatorname{cat}(Z \setminus W^*) \geqslant 1$, hence $Z \setminus W^*$ is nonempty.

COROLLARY 6.1. *If W is a Ważewski set and W and W^- are closed ANRs then*

$$\operatorname{cat}_W\bigl(W \setminus W^*\bigr) \geqslant \operatorname{cat}\bigl(W, W^-\bigr).$$

Recall that $\operatorname{cat}(W, W^-)$ can be estimated using (5.3) or Proposition 5.2. Another estimate using the notion of cup-length is given by Proposition 5.4.

6.2. *Stationary points of gradient-like local semiflows*

Let ϕ be a local semiflow on X.

DEFINITION 6.1. $g : X \to \mathbb{R}$ is called a *Liapunov function* for ϕ provided it is decreasing along nonconstant trajectories, i.e.,

$$g\big(\phi_s(x)\big) > g\big(\phi_t(x)\big)$$

if $s < t$ whenever x is not a stationary point. ϕ is called a *gradient-like* local semiflow if there exists a Liapunov function for ϕ.

An example which justifies the name is a gradient local flow on a Riemannian manifold, i.e., the local flow generated by

$$\dot{x} = -\nabla g(x) \tag{6.1}$$

for a C^1-class function g. In this case stationary points are critical points of g.

It is not difficult to prove the following result:

PROPOSITION 6.1. *If ϕ is gradient-like then each nonempty ω-limit set consists of stationary points.*

By Corollary 2.5, it immediately follows:

COROLLARY 6.2. *If a local semiflow ϕ on X is gradient-like, $W \subset X$ is a compact Ważewski set and W^- is not a strong deformation retract of W then W contains a stationary point of ϕ.*

6.3. *Stationary points of local semiflows in the general case*

Results on the existence of stationary points of semiflows without any gradient-like structure can be obtained using fixed points theorems. The most useful for our purposes is the Lefschetz theorem. It is a applied in a proof of the following result:

THEOREM 6.2 (compare Corollary 1 in [73]). *If ϕ is a local semiflow on X, W and W^- are compact ANRs in X, and*

$$\chi\big(W, W^-\big) = \chi(W) - \chi\big(W^-\big) \neq 0$$

then W contains a stationary point of ϕ.

PROOF. *Step* 1. At first we prove that for every $T > 0$ there exists a point $x \in W$ such that $\phi_T(x) = x$. Let S^1 be the unit circle in the complex plane $\mathbb{C}$. In the topological direct sum

$$W \amalg \big(W^- \times S^1\big),$$

i.e., a disjoint union of W and $W^- \times S^1$ endowed with the inductive topology, identify each point of $W^- \subset W$ with a point of $W^- \times S^1$ via the map $j : w \mapsto (w, 1)$. The resulting quotient space is the adjunction

$$\widetilde{W} := W \cup_j \big(W^- \times S^1\big).$$

Define a semiflow ψ on $\widetilde{W}$ by

$$\psi(z,t) := \begin{cases} \phi(z,t), & \text{if } z \in W \text{ and } \sigma^+(z) \geqslant t, \\ \big(\phi\big(z, \sigma^+(z)\big), \exp\big(\pi i \big(t - \sigma^+(z)\big)/T\big)\big), & \text{if } z \in W \text{ and } \sigma^+(z) \leqslant t, \\ \big(x, \alpha \exp(\pi i t/T)\big), & \text{if } z = (x,\alpha) \in W^- \times S^1. \end{cases}$$

The set $\widetilde{W}$ is a compact ANR since W, W^-, and $W^- \times S^1$ are ANRs, and

$$\chi(\widetilde{W}) = \chi(W) + \chi\big(W^- \times S^1\big) - \chi\big(W^-\big) = \chi(W) - \chi\big(W^-\big) \neq 0.$$

Since $\phi_T \simeq \mathrm{id}$, there exists $z \in \widetilde{W}$ such that $\psi_T(z) = z$ by the Lefschetz fixed point theorem (Proposition 5.6). It follows by the definition of ψ that $\psi_T(z) = z$ if and only if

$$z \in W, \qquad \phi_T(z) = z, \qquad \phi^+(z) = \phi\big(z, [0,T]\big) \subset W,$$

hence the required conclusion follows.

Step 2. Now we can use a standard argument to prove that ϕ has a stationary point in W (compare [32, Proposition 7(4.8)]). By Step 1, there exists $x_n \in \widetilde{W}$ such that $\phi(x_n, 2^{-n}) = x_n$. An accumulation point of the sequence $\{x_n\}$ is a stationary point of ϕ, hence the result follows. □

7. Isolating blocks and segments

We continue the series of results on properties of the asymptotic part of a Ważewski set. Here we additionally assume that the considered local semiflow is a local flow and the Ważewski set is an isolating block. These assumptions enable us to provide various quantitative properties of the invariant part of the set in an easier way.

7.1. *Isolating blocks and their structure*

Let ϕ be a local flow on a metrizable space X. At first we should fix terminology concerning behavior of ϕ with respect to a given subset of X. Recall, that we have already defined the exit set and the asymptotic part in Section 2.1, the extended escape-time function in Section 2.2, and we mentioned on an invariant part in Section 1.3. Let $Z \subset X$. We provide a list of notions related to Z which will be of our interest (for the full account we repeat the definitions of Z^-, $\mathrm{Inv}(Z)$, and σ^+). Define the sets

$$\begin{aligned}
&Z^- := \{x \in Z\colon\ \phi(x, [0,t]) \not\subset Z\ \forall t > 0\},\\
&Z^+ := \{x \in Z\colon\ \phi(x, [-t,0]) \not\subset Z\ \forall t > 0\},\\
&\mathrm{Inv}^{\pm}(Z) := \{x \in Z\colon\ \phi^{\pm}(x) \subset Z\},\\
&\mathrm{Inv}(Z) := \mathrm{Inv}^-(Z) \cap \mathrm{Inv}^+(Z),
\end{aligned}$$

and functions $\sigma^{\pm}: Z \to [0, \infty]$ by

$$\begin{aligned}
&\sigma^+(x) := \sup\{t \geqslant 0\colon\ \phi(x, [0,t]) \subset Z\},\\
&\sigma^-(x) := \sup\{t \geqslant 0\colon\ \phi(x, [-t,0]) \subset Z\}.
\end{aligned}$$

Obviously, $\mathrm{Inv}^+(Z)$ is the asymptotic part of Z; it is also called the *positive invariant part*, and $\mathrm{Inv}(Z)$ is the *invariant part of* Z. Moreover, $\mathrm{Inv}^-(Z)$ is called the *negative invariant part* and Z^+ is called the *entrance* set of Z. (It should be pointed out that all the above notions are defined with respect to a fixed ϕ and Z. In fact, a more proper notation should indicate on ϕ, for example $\mathrm{Inv}_\phi(Z)$ instead of $\mathrm{Inv}(B)$, $\sigma^{\pm}_{\phi,Z}$ instead of $\sigma^{\pm}$, etc. Nevertheless, for simplicity we avoid that notation whenever it does not lead to confusions.)

In Section 1.3, it was given a rough definition of the concept of isolating block. Here we state it in a slightly more general context.

DEFINITION 7.1. A set $B \subset X$ is called an *isolating block* if

$$B = \overline{\mathrm{int}\, B},$$

B, B^-, and B^+ are compact, and for every $x \in \partial B \setminus (B^- \cup B^+)$,

$$\begin{aligned}
&\sigma^-(x) < \infty, \qquad \sigma^+(x) < \infty,\\
&\phi\big(x, [-\sigma^-(x), \sigma^+(x)]\big) \subset \partial B.
\end{aligned}$$

(In practice, one usually considers isolating blocks without sliding on the boundary, i.e., $\partial B = B^+ \cup B^-$). By Proposition 2.1 and Lemma 2.2 we get immediately the following properties of isolating blocks:

PROPOSITION 7.1. *If B is an isolating block then*

(i) *B is a Ważewski set for both ϕ and its reverse obtained by the transformation $t \mapsto -t$ of time,*
(ii) *the functions $\sigma^{\pm}$ are continuous.*

In the sequel we assume that B is an isolating block. In particular, it follows by Corollary 2.5 that if B^- or B^+ is not a strong deformation retract of B then its invariant part

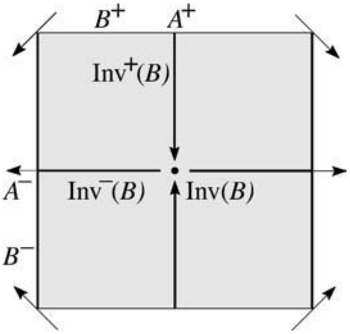


Fig. 6.

$\mathrm{Inv}(B)$ is nonempty. Our aim is to provide a more accurate information on its properties that it was given for $W \setminus W^*$ in Section 6. For this reason we introduce some other concepts related to B. Denote by $A(B)$ (or shortly A, if it does not lead to confusions) the subset of B given by

$$A := A(B) := \mathrm{Inv}^-(B) \cup \mathrm{Inv}^+(B).$$

It follows, in particular, that its exit and entrance sets are given as

$$A^\pm = A \cap B^\pm.$$

Below we frequently use abbreviated notation for the invariant part of the isolating block B; we put:

$$I := \mathrm{Inv}(B).$$

The position of the mentioned sets in the block B can be seen on Figure 6.

Let B be an isolating block (with respect to the flow ϕ).

DEFINITION 7.2. A set $Z \subset B$ is called *B-invariant* if for every $x \in Z$ the connected component of $\phi(x) \cap B$ containing x is a subset of Z.

The connected component which appears in the above definition is equal to $\phi(x, \Delta)$, where Δ is the closed interval having $\sigma^-(x)$ and $\sigma^+(x)$ as the ends. In particular, it is equal to $\phi(x)$ if $x \in I$. The sets A and $\mathrm{Inv}^\pm(B)$ are examples of B-invariant sets.

For $n \in \mathbb{N}$ define

$$A_n := \{x \in A\colon \sigma^-(x) \geqslant n,\ \sigma^+(x) \geqslant n\},$$
$$B_n := \{x \in B\colon \sigma^-(x) + \sigma^+(x) \geqslant n\} \supset A.$$

In particular, B_n is B-invariant, $A_0 = A$, $B_0 = B$, and $B_n^\pm = B^\pm \cap B_n$. We list some properties of those sets:

LEMMA 7.1.

(i) *B_n is a neighborhood of A in B and for every neighborhood U of A there exists n such that $B_n \subset U$.*

(ii) *For every n there exists a compact set $D_n \subset B$,*

$$B_{n+1} \cup B^- \subset D_n \subset B_n \cup B^-$$

such that D_n is a strong deformation retract of $Z \cup B^-$ for every B-invariant set Z such that $B_n \subset Z$. In particular, D_n is a strong deformation retract of B as well as of $B_n \cup B^-$.

(iii) *For every neighborhood V of I there exists n such that $A_n \subset V$.*

(iv) *A_{n+1} is a strong deformation retract of A_n for every n.*

(v) $\mathrm{Inv}^\pm(B) \cap A_{n+1}$ *is a strong deformation retract of* $\mathrm{Inv}^\pm(B) \cap A_n$ *for every n.*

PROOF. In the proof we frequently refer to the continuity of $\sigma^\pm$ (see Proposition 7.1).

Ad (i) and (iii). They immediately follow by that continuity.

Ad (ii). Let $f : B^- \to [0, 1]$ be a continuous function such that $f(x) = 0$ if $\sigma^-(x) \leqslant n$ and $f(x) = 1$ if $\sigma^-(x) \geqslant n + 1$. Define $g : B^- \to [0, n + 1]$ by

$$g(x) = \begin{cases} f(x)\sigma^-(x), & \text{if } \sigma^-(x) \leqslant n + 1, \\ n + 1, & \text{if } \sigma^-(x) \geqslant n + 1. \end{cases}$$

In particular, $g(x) \leqslant \sigma^-(x)$ for $x \in B^-$. The set

$$D_n := B_{n+1} \cup \bigl\{\phi(x, -t) \in B\colon\ x \in B^-,\ t \leqslant g(x)\bigr\}$$

has the required properties. Indeed, the map r defined by

$$r(x) := \begin{cases} \phi\bigl(x, \sigma^+(x)\bigr), & \text{if } \sigma^-(x) + \sigma^+(x) \leqslant n, \\ \phi\bigl(x, \sigma^+(x) - g\bigl(\phi\bigl(x, \sigma^+(x)\bigr)\bigr)\bigr), & \text{if } n \leqslant \sigma^-(x) + \sigma^+(x) \leqslant n + 1 \\ & \text{and } g\bigl(\phi\bigl(x, \sigma^+(x)\bigr)\bigr) \leqslant \sigma^+(x), \\ x, & \text{if } n \leqslant \sigma^-(x) + \sigma^+(x) \leqslant n + 1 \\ & \text{and } g\bigl(\phi\bigl(x, \sigma^+(x)\bigr)\bigr) \geqslant \sigma^+(x) \\ & \text{or if } x \in B_{n+1} \end{cases}$$

is a strong deformation retraction $B \to D_n$ and each its restriction to $Z \cup B^- \to D_n$ is a strong deformation retraction as well.

Ad (iv) and (v). The map $s : A_n \to A_{n+1}$,

$$s(x) := \begin{cases} \phi\bigl(x, n + 1 - \sigma^-(x)\bigr), & \text{if } \sigma^-(x) \leqslant n + 1, \\ \phi\bigl(x, \sigma^+(x) - n - 1\bigr), & \text{if } \sigma^+(x) \leqslant n + 1, \\ x, & \text{otherwise} \end{cases}$$

is a strong deformation retraction required in (iv); its restrictions apply to (v). ☐

As a consequence we get the following estimate on the category of I:

THEOREM 7.1 (compare Theorem 3.1 in [63]). *If B and B^- are ANRs then*

$$\operatorname{cat}(I) \geqslant \operatorname{cat}\bigl(B, B^-\bigr).$$

PROOF. Lemma 7.1(iii) implies

$$I = \bigcap_{n\in\mathbb{N}} A_n \cap \operatorname{Inv}^+(B),$$

hence the result follows by Proposition 5.1(iv), Corollary 6.1, and Lemma 7.1(v). □

7.2. *Estimates on the cohomology of the invariant part*

We begin with the case B^- or B^+ empty. It follows I is an *attractor* (or, in other terminology, an *asymptotically stable set*) for the flow ϕ or, respectively, its reverse. In that case we have a satisfactory information on the topology of I given by the following result:

THEOREM 7.2. *If $B^- = \emptyset$ or $B^+ = \emptyset$ then the inclusion $I \hookrightarrow B$ induces an isomorphism $\check{H}^*(B) \to \check{H}^*(I)$.*

PROOF. Assume that $B^- = \emptyset$ (for a proof in the other case it suffices to reverse the direction of time). For every $t \geqslant 0$ the inclusion $\phi_t(B) \hookrightarrow B$ induces an isomorphism in the cohomologies,

$$\bigcap_{t\geqslant 0} \phi_t(B) = I,$$

hence the result is an immediate consequence of the continuity property of the Čech cohomologies (compare [71]). □

REMARK 7.1. Actually, it is not difficult to prove a stronger result: $I \hookrightarrow B$ is a *shape equivalence* (compare [38]; more information on the notion of shape can be found in [15]). In [38], there is an example in which B is an annular region of the plane and I is the *Warsaw circle*, hence, in general, the shape equivalence cannot be further improved to the homotopy equivalence.

On the other hand, if neither B^- nor B^+ is empty, usually one cannot get enough information from B to determine the topology of I, as it is indicated on the following example:

EXAMPLE 7.1. Let ϕ be a local flow generated by a vector-field v in $\mathbb{R}^n$. Assume that B is an isolating block for ϕ and there exists an $x_0 \in \operatorname{int} B$ such that $0 < \sigma^{\pm}(x_0) < \infty$. In particular, $x_0 \notin I$, hence $v(x_0) \neq 0$. Let D be an $(n-1)$-dimensional disc centered at

x_0 and contained in the affine hyperplane perpendicular to $v(x_0)$ such that v is transversal to D. By the continuity of $\sigma^{\pm}$, we can assume that

$$0 < \sigma^{\pm}(y) < \infty, \quad \text{for every } y \in D. \tag{7.1}$$

Let K be an arbitrary compact set contained in D and let $g : \mathbb{R}^n \to [0, \infty)$ be a continuous function such that

$$K = \{x \in \mathbb{R}^n \colon g(x) = 0\}.$$

It follows by (7.1) that B is an isolating block for the local flow generated by

$$\dot{x} = g(x)v(x),$$

its invariant part is equal to $I \cup K$, and $I \cap K = \emptyset$. The latter implies that $\check{H}^*(K)$ is a direct factor of $\check{H}^*(I \cup K)$. Except of some limitations which we do not precise here, $\check{H}^q(K)$ is an arbitrary countably generated Abelian group if $0 \leqslant q \leqslant n - 2$ and is equal to zero if $q \geqslant n - 1$ since K is homeomorphic to a subset of $\mathbb{R}^{n-1}$, hence $\check{H}^q(I \cup K)$ is at least as complicated.

From the above example we conclude that the Čech cohomology of the invariant part of B can be almost arbitrarily complex. Thus, in general one can at most expect to estimate the "lower bound" of the cohomology of I. We provide several results in this direction. A majority of them can be duplicated by substituting B^+ instead of B^-, etc.

Our first aim is to present an exact sequence which connects the cohomologies of I, (B, B^-), and A^- due to Churchill in [18].

PROPOSITION 7.2 (compare Lemma 4.3 in [18]). *The inclusion* $(A, A^-) \hookrightarrow (B, B^-)$ *induces an isomorphism* $\check{H}^*(B, B^-) \to \check{H}^*(A, A^-)$.

PROOF. By Lemma 7.1, the homotopy and excision properties of the cohomology imply that the inclusions induce a sequence of isomorphisms

$$\check{H}^*(B, B^-) \cong \check{H}^*(D_n, B^-) \cong H^*(B_n \cup B^-, B^-) \cong H^*(B_n, B_n^-)$$

for every n. It follows that the inclusion $(B_{n+1}, B_{n+1}^-) \hookrightarrow (B_n, B_n^-)$ induces an isomorphism, hence there is also an isomorphism

$$\check{H}^*(B, B^-) \cong \operatorname{dir}\lim \check{H}^*(B_n, B_n^-)$$

induced by the inclusions. Since

$$\bigcap_{n \in \mathbb{N}} (B_n, B_n^-) = (A, A^-),$$

the result follows by the continuity of the Čech cohomology. □

PROPOSITION 7.3 (compare Proposition 4.6 in [18]). *The inclusion $I \hookrightarrow A$ induces an isomorphism $\check{H}^*(A) \to \check{H}^*(I)$.*

PROOF. It follows by Lemma 7.1 that the inclusion $A_{n+1} \hookrightarrow A_n$ induces an isomorphism in cohomologies and

$$\bigcap_{n\in\mathbb{N}} A_n = I,$$

hence the result follows by the properties of homotopy and continuity of the Čech cohomology. □

THEOREM 7.3 (compare Theorem 4.7 in [18]). *There exists an exact sequence*

$$\cdots \xrightarrow{\delta^{q-1}} \check{H}^q(B, B^-) \xrightarrow{\beta^q} \check{H}^q(I) \xrightarrow{\gamma^q} \check{H}^q(A^-) \xrightarrow{\delta^q} \check{H}^{q+1}(B, B^-) \xrightarrow{\beta^{q+1}} \cdots$$

in which β^q is induced by the inclusion for every $q \in \mathbb{Z}$.

PROOF. By Propositions 7.2 and 7.3, it suffices to substitute $\check{H}^q(A, A^-) \to \check{H}^q(A)$ by β^q in the exact sequence of the pair (A, A^-). □

The above result is not completely satisfactory since the term $\check{H}^*(A^-)$ in the exact sequence cannot be determined from B and B^- alone. Nevertheless, some estimates can be obtained by its application. In the following corollary and example, χ and $\check{\chi}$ denote the Euler characteristic with respect to singular cohomologies and, respectively, Čech cohomologies. A standard fact from homological algebra on exact sequences and the corresponding Euler characteristics leads to the following

COROLLARY 7.1. *If two Euler characteristics among $\check{\chi}(I)$, $\check{\chi}(B, B^-)$, and $\check{\chi}(A^-)$ are defined, then the third one is also defined and*

$$\check{\chi}(I) = \check{\chi}\big(B, B^-\big) + \check{\chi}\big(A^-\big).$$

EXAMPLE 7.2. Let a 2-dimensional manifold be the phase space of the considered local flow. Assume that an isolating block B is a manifold with boundary and B^- is a submanifold with boundary of ∂B (for example, given as a generalized polyfacial set like in Remark 3.2). In that case B and B^- are compact ENRs, hence the Čech cohomologies of (B, B^-) are isomorphic to the corresponding singular cohomologies. Moreover, $\check{H}^q(B, B^-)$ is finite-dimensional for every q and equal to zero if $q \geqslant 3$ and B^- is the union of components each of which is homeomorphic to the circle S^1 or to the interval $[0, 1]$. Thus the set A^- consists of the union of some number of the components of B^- homeomorphic to the circle and a set K homeomorphic to a compact subset of $\mathbb{R}$. If K has infinitely many components then from the exact sequence in Theorem 7.3 we conclude that $\check{H}^0(I)$ is infinitely generated (hence I has infinitely many components) because

$\operatorname{Im}\gamma^0 = \ker\delta^0$ and $\ker\delta^0$ is infinite-dimensional since $\dim \check{H}^1(B, B^-) < \infty$. In the opposite case, the subset K consists of a finite number of isolated points and sets homeomorphic to $[0, 1]$. Since the Euler characteristic of the circle is equal to zero,

$$\check{\chi}(A^-) = \dim \check{H}^0(K) \geqslant 0.$$

In particular, as a consequence of Corollary 7.1 we get

$$\check{\chi}(I) \geqslant \chi(B, B^-). \tag{7.2}$$

Denote by 1 the unit element in the ring of coefficients R of the Čech cohomology. The following notation will also be used: if $(Y, Y') \subset (X, X')$ and $u \in \check{H}^*(X, X')$ then $u|_{(Y,Y')}$ denotes the image of u under the homomorphism induced by the inclusion $(Y, Y') \hookrightarrow (X, X')$. The next theorem is an equivalent version of a result of Floer (compare [30, Proposition 2]—it is written there that another version of that result was stated by Benci in an unpublished report). The proof given here is taken form [77].

THEOREM 7.4. *If for $u \in \check{H}^*(B)$ there exists $v \in \check{H}^*(B, B^-)$ such that $u \smile v \neq 0$ then $u|_I \neq 0$ in $\check{H}^*(I)$.*

PROOF. In the commutative diagram

$$\begin{tikzcd}
\check{H}^*(B) \otimes \check{H}^*(B, B^-) \arrow[r, "\smile"] \arrow[d, "\cong"] & \check{H}^*(B, B^-) \arrow[d, "\cong"] \\
\check{H}^*(A) \otimes \check{H}^*(A, A^-) \arrow[r, "\smile"] & \check{H}^*(A, A^-)
\end{tikzcd}$$

the vertical arrows are generated by the inclusions and those induced by $(A, A^-) \hookrightarrow (B, B^-)$ is an isomorphism by Proposition 7.2, hence

$$u|_A \smile v|_{(A,A^-)} = (u \smile v)|_{(A,A^-)} \neq 0.$$

Thus $u|_A \neq 0$. Since $I \hookrightarrow A$ induces an isomorphism by Proposition 7.3,

$$u|_I = (u|_A)|_I \neq 0,$$

hence the result follows. □

By Theorem 7.4, if $v \in \check{H}^*(B, B^-)$ is nonzero then $1 \smile v = v \neq 0$ (where 1 is treated as the element of $\check{H}^0(B)$), hence $1|_I \neq 0$, which means $I \neq \emptyset$. Thus, in the case of isolating block, Theorem 7.4 is a generalization of a weaker variant of Corollary 2.5, in which the condition $\check{H}^*(B, B^-) \neq 0$ replaces the condition B^- is not a strong deformation retract of B.

In the following simple example we explain the meaning of Theorem 7.4.

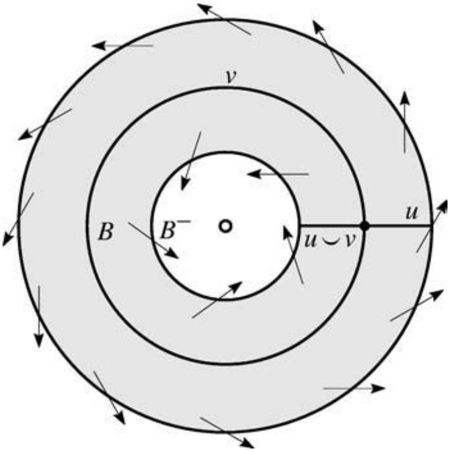


Fig. 7.

EXAMPLE 7.3. Consider the flow on the punctured plane $\mathbb{R}^n \setminus 0$ such the annulus

$$B = \{(r, \theta)\colon\ 1 \leqslant r \leqslant 2\}$$

(written in the polar coordinates) is an isolating block for the flow and its exit set B^- is equal the boundary of B, see Figure 7. Assume for simplicity that $R = \mathbb{Z}_2$, hence there is one generator u of $\check{H}^1(B)$ and one generator v of $\check{H}^1(B, B^-)$. On the figure, they are represented by their supports: u as the vertical segment and v as the circle inside B. Their cup-product $u \smile v$, represented as the intersection of the supports of u and v, is the generator of $\check{H}^2(B, B^-)$. It follows, by Theorem 7.4 that $u|_{\mathrm{Inv}(B)} \neq 0$.

Actually, in that particular example (which was chosen because of its visualization on the plane) the same conclusion is given by Theorem 7.2. This is not the case in the next example which is a direct generalization of the above one.

EXAMPLE 7.4. Now we use different letters to denote the sets considered in the previous example: put

$$G := \{(r, \theta) \in \mathbb{R}^2\colon\ 1 \leqslant r \leqslant 2\},$$
$$H := \{(r, \theta) \in \mathbb{R}^2\colon\ r \in \{1, 2\}\}.$$

As usual, D^q and S^{q-1} denote the unit ball and, respectively, the unit sphere centered at zero in $\mathbb{R}^q$. Consider a local flow in the phase space

$$X = (\mathbb{R}^2 \setminus 0) \times \mathbb{R}^{m+n}$$

and assume that the isolating block and its exit set are given by

$$B := G \times D^m \times D^n,$$

$$B^- := H \times D^m \times D^n \cup G \times S^{m-1} \times D^n.$$

We assume again that $R = \mathbb{Z}_2$ and let $u \in \check{H}^1(G)$, $v \in \check{H}^1(G, H)$, and $z \in \check{H}^m(D^m \times D^n, S^{m-1} \times D^n)$ be the generators. It follows that

$$(u \times 1) \smile (v \times z) = (u \smile v) \times z \neq 0,$$

where × denotes the cohomological cross-product (compare [28]) and $1 \in \check{H}^0(D^m \times D^n)$ is the unit element, hence Theorem 7.4 implies

$$(u \times 1)|_I \neq 0.$$

From that conclusion we can get slightly more information than just $\check{H}^1(I) \neq 0$. Let

$$\pi : G \times D^m \times D^n \to G$$

be the projection. For an arbitrarily chosen $\theta_0 \in [0, 2\pi)$ put

$$R := \{(r, \theta) \in G \colon \theta = \theta_0\}.$$

We assert that

$$\pi(I) \cap R \neq \emptyset.$$

Indeed, since $\check{H}^1(\pi)(u) = u \times 1$,

$$u|_{\pi(I)} \neq 0.$$

Since the inclusion $G \hookrightarrow (G, G \setminus R)$ induces an isomorphism in $\check{H}^1$, there exists $w \in \check{H}^1(G, G \setminus R)$ such that $u = w|_G$. It follows that

$$w|_{(\pi(I), \pi(I) \setminus R)} \neq 0,$$

hence $\pi(I) \setminus R \subsetneq \pi(I)$ which verifies the assertion.

The conclusions in the above example do not follow from Theorem 7.2 if $n \geqslant 1$. It should also be noted that the usage of the Čech cohomologies is essential; in general the above results are non valid for the singular cohomologies by the example mentioned in Remark 7.1. In the paper [30], the version of Theorem 7.4 was formulated using the notion of cup-length. The following immediate consequence of Theorem 7.4 estimates the cup-length of the invariant part of an isolating block.

COROLLARY 7.2. $\ell(I) \geqslant \ell(B, B^-)$.

7.3. *On the number of stationary points of gradient-like flows*

Let $g : X \to \mathbb{R}$ be a Liapunov function for a local flow ϕ on X. For $a \in \mathbb{R}$ we write

$$g_a := g^{-1}(a).$$

DEFINITION 7.3. The set g_a is called a *level of g*. It is called a *critical level* if it contains a stationary point; in this case a is called a *critical value*. In the other case it is called *noncritical.*

For $a < b$ put

$$g_a^b := g^{-1}\bigl([a, b]\bigr).$$

Now we recall the most fundamental result concerning the notion of category. Usually, it is stated for gradient flows. Here we provide its topological version.

PROPOSITION 7.4. *Let g be a Liapunov function on X. If g_a^b is a compact ANR, g_a and g_b are noncritical levels then*

$$\#\bigl\{x \in X\colon\ x \text{ stationary},\ a < g(x) < b\bigr\} \geqslant \operatorname{cat}\bigl(g_a^b, g_a\bigr).$$

PROOF. Put $k =: \operatorname{cat}(g_a^b, g_a)$. First of all observe that $k \leqslant \infty$ since the category of a compact ANR is finite (which is easy to prove; see [16]) and (5.2) holds. We can assume that the number of stationary points is finite and $k \geqslant 1$; otherwise there is nothing to prove. Thus there is a finite number of critical levels between a and b. We adapt the standard mini-max argument. For $i = 1, \ldots, k$ define

$$\mathcal{C}_i := \bigl\{C \subset g_a^b\colon\ C \text{ compact},\ \operatorname{cat}_{g_a^b, g_a}(C) \geqslant i\bigr\},$$
$$c_i := \inf_{C \in \mathcal{C}_i} \max_{x \in C} g(x).$$

We claim that c_i is a critical value. Indeed, in the other case there exists an $\varepsilon > 0$ such that $a < c_i - \varepsilon < c_i + \varepsilon < b$ and there are no critical values in the interval $[c_i - \varepsilon, c_i + \varepsilon]$ since the set of critical values is finite. It follows that $g_{c_i-\varepsilon}^{c_i+\varepsilon}$ is an isolating block and $g_{c_i-\varepsilon}$ is its exit set. Since that set has the invariant part empty by Remark 6.1, there is a strong deformation retraction $r : g_{c_i-\varepsilon}^{c_i+\varepsilon} \to g_{c_i-\varepsilon}$. It follows that if $C \in \mathcal{C}_i$ and $\max_{x \in C} g(x) < c_i + \varepsilon$ then $r(C) \in \mathcal{C}_i$ and

$$\max_{x \in r(C)} g(x) = c_i - \varepsilon$$

which is a contradiction. The critical values c_i are ordered

$$a < c_1 \leqslant \cdots \leqslant c_k < b$$

since $\mathcal{C}_j \subset \mathcal{C}_i$ if $i < j$. Denote by K_{c_i} the set of stationary points contained in g_{c_i}. It suffices to prove that if $c_i = c_{i+q}$ then

$$\operatorname{cat}_{g_a^b}(K_{c_i}) = q + 1,$$

since in that case K_{c_i} contain at least $q + 1$ different points.

Let $\varepsilon > 0$ be such that $a < c_i - \varepsilon < c_i + \varepsilon < b$ and $c_i = c_{i+q}$ is the only critical value in $[c_i - \varepsilon, c_i + \varepsilon]$. Since there exists $C \in \mathcal{C}_{i+q}$ such that $C \subset g_a^{c_i+\varepsilon}$, by Proposition 5.1(i),

$$\operatorname{cat}_{g_a^b, g_a}\bigl(g_a^{c_i+\varepsilon}\bigr) \geqslant i + q. \tag{7.3}$$

It follows also by the choice of c_i that

$$\operatorname{cat}_{g_a^b, g_a}\bigl(g_a^{c_i-\varepsilon}\bigr) \leqslant i - 1. \tag{7.4}$$

Denote $B := g_{c_i-\varepsilon}^{c_i+\varepsilon}$. It is an isolating block, $B^- = g_{c_i-\varepsilon}$ and $\operatorname{Inv}(B) = K_{c_i}$ by Remark 6.1. We can apply Proposition 5.1(iv) since g_a is an ANR. Indeed, let $\delta > 0$ be such that there are no critical points in $[a, a + \delta]$. It follows that g_a is a strong deformation retract of $\{x\colon\ a \leqslant g(x) < a + \delta\}$. The latter set is open in g_a^b, hence it is an ANR by Theorem of Hanner (compare [13]) and g_a is an ANR as well. Thus, by Lemma 7.1 and by Proposition 5.1(iii) and (iv), for n and m sufficiently large we have the following sequence of equations:

$$\operatorname{cat}_{g_a^b}(K_{c_i}) = \operatorname{cat}_{g_a^b}(A_m) = \operatorname{cat}_{g_a^b}(A) = \operatorname{cat}_{g_a^b}(B_n).$$

Thus, again by Lemma 7.1 and Proposition 5.1,

$$\begin{aligned} i + q &\leqslant \operatorname{cat}_{g_a^b, g_a}\bigl(g_a^{c_i+\varepsilon}\bigr) = \operatorname{cat}_{g_a^b, g_a}\bigl(g_a^{c_i-\varepsilon} \cup D_n\bigr) \leqslant \operatorname{cat}_{g_a^b, g_a}\bigl(g_a^{c_i-\varepsilon}\bigr) + \operatorname{cat}_{g_a^b}(B_n) \\ &\leqslant i - 1 + \operatorname{cat}_{g_a^b}(K_{c_i}), \end{aligned}$$

which finishes the proof. □

THEOREM 7.5. *If ϕ is gradient-like and B and B^- are ANR-s then*

$$\#\{x \in I\colon\ x \text{ stationary}\} \geqslant \operatorname{cat}\bigl(B, B^-\bigr).$$

PROOF. Let $g : X \to \mathbb{R}$ be a Liapunov function. In the topological direct sum

$$B \amalg \bigl(B^- \times (-\infty, 0]\bigr) \amalg \bigl(B^+ \times [0, \infty)\bigr)$$

identify each element of $B^- \subset B$ with an element of $B^- \times (-\infty, 0]$ via the map $j^- : z \to (z, 0)$ and, similarly, each element of $B^+ \subset B$ with an element of $B^+ \times [0, \infty)$ via the map $j^+ : z \mapsto (z, 0)$. Denote the obtained adjunction by $\widetilde{B}$, i.e.,

$$\widetilde{B} := B \cup_{j^-} \bigl(B^- \times (-\infty, 0]\bigr) \cup_{j^+} \bigl(B^+ \times [0, \infty)\bigr).$$

In a natural way we define a flow $\psi : \widetilde{B} \times \mathbb{R} \to \widetilde{B}$ as an extension of ϕ restricted to B. If $z \in B$ put

$$\psi(z,t) := \begin{cases} \phi(z,t) \in B, & \text{if } -\sigma^-(x) \leqslant t \leqslant \sigma^+(x), \\ \big(\phi\big(x, \sigma^+(x)\big), \sigma^+(x) - t\big) \in B^- \times (-\infty, 0], \\ \quad \text{if } t \geqslant \sigma^+(x), \\ \big(\phi\big(x, -\sigma^-(x)\big), \sigma^-(x) - t\big) \in B^+ \times [0, \infty), \\ \quad \text{if } t \leqslant -\sigma^-(x), \end{cases}$$

if $(x, \tau) \in B^- \times (-\infty, 0]$ put

$$\psi\big((x,\tau), t\big) := \begin{cases} (x, \tau - t), & \text{if } t \geqslant \tau, \\ \psi(x, t - \tau), & \text{if } t \leqslant \tau, \end{cases}$$

and similarly, if $(x, \tau) \in B^+ \times [0, \infty)$ put

$$\psi\big((x,\tau), t\big) := \begin{cases} \psi(x, t - \tau), & \text{if } t \geqslant \tau, \\ (x, \tau - t), & \text{if } t \leqslant \tau. \end{cases}$$

Define also a function $G : \widetilde{B} \to \mathbb{R}$,

$$G(z,t) := \begin{cases} g(z), & \text{if } z \in B, \\ g(x) + \tau, & \text{if } z = (x, \tau) \in B^- \times (-\infty, 0] \text{ or } B^+ \times [0, \infty). \end{cases}$$

G is a Liapunov function for ψ and the stationary points of ϕ restricted to B and ψ coincide. Let $a < b$ be such that

$$a < \min g(B) \leqslant \max g(B) < b$$

It follows that a and b are noncritical values of G. Put

$$M := \big\{(x, \tau) \in B^- \times (-\infty, 0],\ \tau \geqslant b - g(x)\big\},$$
$$N := \big\{(x, \tau) \in B^+ \times [0, \infty),\ \tau \leqslant a - g(x)\big\},$$

hence

$$G_a^b = B \cup M \cup N$$

and G_a^b is a compact ANR since M and N are compact ANRs as retracts of ANRs $B^- \times (\infty, 0]$ and $B^+ \times [0, \infty)$ and the intersections of B with M and N coincide with B^- and B^+, respectively. Using the definition of the relative category it is easy to prove that

$$\operatorname{cat}\big(G_a^b, G_a\big) = \operatorname{cat}\big(B, B^-\big).$$

The result is a consequence of Proposition 7.4. □

EXAMPLE 7.5. The following fact was essential in the proof of an Arnold's conjecture by Conley and Zehnder (compare [23]): if

$$B = T^q \times D^m \times D^n,$$

is an isolating block for a gradient-like flow, where T^q is the q-dimensional torus, and

$$B^- = T^q \times S^{m-1} \times D^n$$

then there are at least $q + 1$ stationary points contained in B. Indeed, in this case

$$\ell\bigl(B, B^-\bigr) = q + 1$$

hence the result follows by Proposition 5.4 and Theorem 7.5.

REMARK 7.2. Results on the existence and the number of critical points of functionals are essential in variational methods. Since the spaces considered in those methods are infinite-dimensional, usually direct applications of results based on algebraic topology fail there. For some functionals it is possible, however, to pass to a finite dimension by the saddle-point reduction due to Amann and Zehnder (compare [2]). In particular, it was applied in the paper [23] to reduce the initial problem to calculation of the number of critical points inside the isolating block mentioned in Example 7.5.

7.4. *Fixed point index and the index of stationary points*

Using Theorem 6.2 one can check whether the set of stationary points inside of an isolating block is nonempty. Below we provide quantitative information on properties of that set using the notion of fixed point index.

Assume that X is a ENR, ϕ is a local flow on X, and B is an isolating block.

LEMMA 7.2. *For every* $T > 0$*,*

$$I_T := \bigl\{x \in I\colon\ \phi_T(x) = x\bigr\}$$

is an isolated set of fixed points of ϕ_T*.*

Let U be an open subset of X such that $\overline{U}$ is compact and there are no stationary points of ϕ in ∂U. It follows that there is an $\varepsilon > 0$ such that the set

$$K_t(U) := \bigl\{x \in U\colon\ \phi_t(x) = x\bigr\}$$

is an isolated set of fixed points of ϕ_t for every t such that $0 < t < \varepsilon$. By Proposition 5.7(iv), the indexes of $K_s(U)$ and $K_t(U)$ coincide if s and t are nonzero and small, hence we can define an integer number

$$i(\phi, U) := \lim_{0<t\to 0} \operatorname{ind}\bigl(\phi_t, K_t(U)\bigr).$$

We call the number $i(\phi, U)$ the *index of stationary points of* ϕ *in* U (compare [72, Section 4]). It inherits properties of the fixed point index. It can be proved that if $X = \mathbb{R}^n$ and ϕ is generated by a C^1-class vector-field v then

$$i(\phi, U) = (-1)^n \deg(0, v, U),$$

where deg denotes the Brouwer degree.

THEOREM 7.6 (compare Theorem 4.4 in [72]). *If* B *is an isolating block on* X *and* B *and* B^- *are ENRs, and* $T > 0$ *then*

$$\operatorname{ind}(\phi_T, I_T) = \chi\bigl(B, B^-\bigr),$$
$$i(\phi, \operatorname{int} B) = \chi\bigl(B, B^-\bigr).$$

PROOF. The second equation is a consequence of the first one and the definition of the index. Since the first equation is a particular case of the more general equation in Theorem 7.7 below, we do not provide its rigorous proof here. We mention only that it can be proved by a careful examination of Step 1 in the proof of Theorem 6.2. Indeed, the corresponding indexes of ϕ_T in B and ψ_T at I_T coincide by Proposition 5.7(v) and the index of fixed points of ψ_T is equal to $\chi(\widetilde{B})$ by (vi) in that proposition. By those facts it is easy to conclude the required property; for a detailed argument we refer to the proof of Theorem 7.7 (or to [72]). ☐

It should be pointed out that all proper T-periodic points do not contribute to $\operatorname{ind}(\phi_T, I_T)$ since both indexes are equal to the same number $\chi(B, B^-)$. In the case of a local flow generated by a smooth vector-field on a Riemannian manifold, B being a manifold with boundary and B^- its submanifold (like in Remark 3.2) the above result is a particular case of the formula of Pugh generalizing the Poincaré index formula (compare [34,64]).

EXAMPLE 7.6. The index of stationary points contained in the isolating block in Figure 1 or Figure 6 is equal to -1, while in the isolating block in Figure 7 it is equal to 0.

EXAMPLE 7.7. As in Example 7.2, we assume that X is a 2-dimensional manifold, the isolating block B is a manifold with boundary, and B^- is a submanifold with boundary of ∂B. Assume that the invariant part I of B consists only of a single stationary point. It follows by (7.2) and Theorem 7.6 that

$$i(\phi, \operatorname{int} B) \leqslant 1.$$

7.5. *Isolating segments*

It is convenient to use special kinds of isolating blocks if the considered local semiflow is generated by a nonautonomous equation in its extended phase space. As in Section 3.1, let

M be a smooth Riemannian manifold, U open subset of $\mathbb{R} \times M$, and let $w : U \to TM$ be a continuous mapping such that $w(t,x) \in T_x M$. We assume that the system (3.9) satisfies the uniqueness property of the Cauchy problem and let ϕ be the induced local flow on U. It follows that ϕ is of a special form: there is a map $\Phi : (\sigma, \tau, x) \mapsto \Phi_{\sigma,\tau}(x)$, called the *evolutionary operator for* (3.8), such that

$$\phi_t(\sigma, x) = \bigl(\sigma + t, \Phi_{\sigma,\sigma+t}(x)\bigr)$$

for $(\sigma, x) \in U$. In particular,

$$\Phi_{s,s} = \mathrm{id}, \qquad \Phi_{s,u} = \Phi_{t,u} \circ \Phi_{s,t}.$$

Actually, basing on that properties, the evolutionary operator can be defined in a purely topological way and the results presented below are valid in that more general setting; here we do not develop that approach.

At first we fix the notation concerning the extended phase space. For a subset Z of $\mathbb{R} \times M$ we put

$$Z_t := \bigl\{x \in M\colon\ (t,x) \in Z\bigr\}$$

and by π_1 and π_2 we denote the projections $\mathbb{R} \times M \to \mathbb{R}$ and, respectively, $\mathbb{R} \times M \to M$. Let a and b be real numbers, $a < b$.

DEFINITION 7.4. A compact set $W \subset U$ is called an *isolating segment over* $[a,b]$ *for* (3.8) if the exit and entrance sets $W^{\pm}$ with respect to the local flow ϕ are also compact and

$$\partial W = W^{+} \cup W^{-}$$

(hence W is an isolating block for ϕ), there exist compact subsets W^{--} and W^{++} of ∂W (called, respectively, the *proper exit set* and the *proper entrance set*) such that

$$W^{-} = \bigl(\{b\} \times W_b\bigr) \cup W^{--}, \qquad W^{+} = \bigl(\{a\} \times W_a\bigr) \cup W^{++},$$

and there exists a homeomorphism

$$h : [a,b] \times W_a \to W$$

satisfying $\pi_1 \circ h = \pi_1$ such that

$$h\bigl([a,b] \times W_a^{--}\bigr) = W^{--}, \qquad h\bigl([a,b] \times W_a^{++}\bigr) = W^{++}.$$

It follows, in particular, that if B is an isolating block for an autonomous equation (1.1) satisfying $\partial B = B^{+} \cup B^{-}$ then for every a and b the set $W = [a,b] \times B$ is an isolating segment and

$$W^{--} = [a,b] \times B^{-}, \qquad W^{++} = [a,b] \times B^{+}.$$

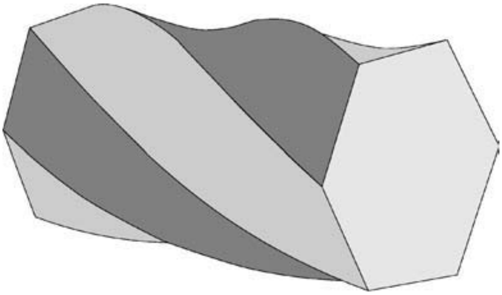


Fig. 8.

A more complicated example of an isolating segment is drawn in Figure 8. Here W is equal to the twisted prism with hexagonal base and each of the sets W^{--} and W^{++} consists of three disjoint ribbons winding around the prism.

Isolating segments are useful in proofs of results concerning the existence of solutions of two-point boundary value problems, as we show below in the case of the periodic problem. In [78], a more general notion of *isolating chain* was considered. (By definition, it is an isolating block being the union of contiguous isolating segments in the extended phase space of a nonautonomous equation.)

Let W be an isolating segment over $[a, b]$. A homeomorphism h in Definition 7.4 induces a homeomorphism

$$m\colon \bigl(W_a, W_a^{--}\bigr) \to \bigl(W_b, W_b^{--}\bigr), \qquad m(x) := \pi_2 h\bigl(b, \pi_2 h^{-1}(a, x)\bigr).$$

We call m a *monodromy homeomorphism* of the isolating segment W. It depends on h, however one can prove the monodromy homeomorphisms determined by different choices of h are homotopic each to the other. It follows that the isomorphism in homologies

$$\mu_W := H(m)\colon H\bigl(W_a, W_a^{--}\bigr) \to H\bigl(W_b, W_b^{--}\bigr)$$

is an invariant of W.

DEFINITION 7.5. An isolating segment W over $[a, b]$ is called *periodic* if $W_a = W_b$, $W_a^{--} = W_b^{--}$, and $W_a^{++} = W_b^{++}$.

The segment shown in Figure 8 is periodic. If W is periodic then μ_W is an automorphism of $H(W_a, W_a^{--})$. Let H has coefficients in $\mathbb{Q}$. If, moreover, both W_a and W_a^{--} are ANRs then the Lefschetz number $\Lambda(m)$ is defined and independent of the choice of the monodromy homeomorphism. Thus we can define

$$\Lambda_W := \Lambda(m).$$

THEOREM 7.7 (compare Theorem 7.1 in [74]). *Let W be a periodic isolating segment over $[a,b]$. Then the set*

$$K_W := \{x \in W_a\colon \Phi_{a,b}(x) = x,\ \Phi_{a,t}(x) \in W_t\ \forall t \in [a,b]\}$$

is an isolated set of fixed points of $\Phi_{a,b}$ and if W, W^{--}, and W^{++} are ENRs then

$$\operatorname{ind}(\Phi_{a,b}, K_W) = \Lambda_W.$$

PROOF. Observe first that $\partial W_a = W_a^{--} \cup W_a^{++}$. Indeed, by the open mapping theorem, $h^{-1}(\operatorname{int} W)$ is open in $\mathbb{R} \times M$. Since $\partial W = W^- \cup W^+$,

$$h^{-1}(\operatorname{int} W) = (a,b) \times \bigl(W_a \setminus \bigl(W_a^{--} \cup W_a^{++}\bigr)\bigr).$$

Thus the image of $h^{-1}(\operatorname{int} W)$ under the projection π_2 onto M is equal to $W_a \setminus (W_a^{--} \cup W_a^{++})$ and is also open, hence it is equal to the interior of W_a and the assertion follows. Since K_W does not intersect ∂W_a, it is open and compact in the set of fixed points of $\Phi_{a,b}$, hence it is isolated.

For a proof of the rest of the conclusion we define a map $m_t : W_t \to W_a$, where $a \leqslant t \leqslant b$, by

$$m_t(x) := \pi_2 h\bigl(b, \pi_2 h^{-1}(t,x)\bigr)$$

(recall that we assume $W_a = W_b$). In particular, $m_a = m$ and $m_b = \mathrm{id}$. Moreover, for $x \in W_a$ put

$$\tau(z) := a + \sigma^+(a,z) \leqslant b,$$

where σ^+ is the escape-time function for W (with respect to ϕ). We use a similar adjunction as in the proof of Theorem 6.2:

$$Z := W_a \cup_j \bigl(W_a^{--} \times S^1\bigr),$$

where $j : x \mapsto (x,1)$ for $x \in W_a^{--}$ (i.e., in the disjoint union of W_a and $W_a^{--} \times S^1$ we identify W_a^{--} with $W_a^{--} \times \{1\}$). By assumptions, Z is an ENR. Define a map $\Psi : Z \times [a,b] \to Z$ by

$$\Psi(z,t) := \begin{cases} m_t\bigl(\Phi_{a,t}(z)\bigr), & \text{if } z \in W_a,\ \tau(z) \geqslant t,\\ \bigl(m_{\tau(z)}\bigl(\Phi_{a,\tau(z)}(z)\bigr), \exp\bigl\{\pi i \frac{t-\tau(z)}{b-a}\bigr\}\bigr), & \text{if } z \in W_a,\ \tau(z) < t,\\ \bigl(m(x), \alpha \exp\bigl\{\pi i \frac{t-a}{b-a}\bigr\}\bigr), & \text{if } z = (x,\alpha) \in W_a^{--} \times S^1. \end{cases}$$

It is easy to check that Ψ is continuous. Put $\Psi_t(z) := \Psi(z,t)$. In particular,

$$\Psi_a|_{W_a^{--} \times S^1} = m|_{W_a^{--}} \times \mathrm{id}_{S^1},$$

hence $\Lambda(\Psi_a|_{W_a^{--}\times S^1}) = 0$. Since $\Psi_a \simeq \Psi_b$, it follows by the Mayer–Vietoris exact sequence

$$\begin{aligned}\Lambda(\Psi_b) = \Lambda(\Psi_a) &= \Lambda(\Psi_a|_{W_a}) + \Lambda(\Psi_a|_{W_a^{--}\times S^1}) - \Lambda(\Psi_a|_{W_a^{--}})\\ &= \Lambda(m|_{W_a}) - \Lambda(m|_{W_a^{--}}) = \Lambda_W.\end{aligned}$$

It follows by the construction of Ψ that

$$\mathrm{Fix}(\Psi_b) = K_W.$$

By the commutativity property of the fixed point index and the Lefschetz theorem (see Proposition 5.7(v) and (vi)),

$$\mathrm{ind}(\Phi_{a,b}, K_W) = \mathrm{ind}(\Psi_b, K_W) = \Lambda_W,$$

which terminates the proof of the theorem. ☐

As an immediate consequence of Proposition 5.7(i) and Theorem 7.7 we get

COROLLARY 7.3. *If W is a periodic isolating segment over $[a,b]$ and $\Lambda_W \neq 0$ then the periodic problem*

$$x(a) = x(b)$$

associated to Equation (3.8) *has a solution.*

Let $T > 0$ and $U = \mathbb{R} \times M$. An important special case of the theorem arises if w is T-periodic in t, i.e., if

$$w(t, x) = w(t + T, x)$$

for every $t \in \mathbb{R}$ and $x \in M$. In this case

$$\Phi_{s,t} = \Phi_{s+T,t+T}. \tag{7.5}$$

The map $\Phi_{0,T}$ is called the *Poincaré map* and its fixed points are initial points of T-periodic solutions of (3.8). In particular, if $\Lambda_W \neq 0$ then at least one T-periodic solution exists. If w does not depend on t, B is an isolating block for w and $W = [0, T] \times B$ then the first equation in Theorem 7.6 is a corollary of Theorem 7.7 as we have pointed out already. The following three examples illustrating the above results are taken from [74].

EXAMPLE 7.8. Consider a planar differential equation

$$\dot{z} = \mathrm{e}^{2\pi i t}\bar{z}^2 + f(t, z), \tag{7.6}$$

where $z \in \mathbb{C}$ and $f : \mathbb{R} \times \mathbb{C} \to \mathbb{C}$ is continuous, $f(t+1, z) = f(t, z)$ for all t and z, and

$$\lim_{|z|\to\infty} \frac{f(t,z)}{|z|^2} = 0 \quad \text{uniformly in } t \in \mathbb{R}.$$

One can find an isolating segment W over [0, 1] for (7.6) which is the twisted prism over a hexagonal base like on the Figure 8 (see [74]; actually it is a regular polyfacial set—compare Definition 3.2). Its monodromy homeomorphism m is equal to the rotation by the angle $\frac{2\pi}{3}$. Thus m is homotopic to the identity on the hexagon W_0 and alternates the three disjoint segments forming W_0^{--}, hence

$$\Lambda_W = \Lambda(m|_{W_0}) - \Lambda(m|_{W_0^{--}}) = 1 - 0 = 1.$$

By Corollary 7.3 we conclude that (7.6) has a 1-periodic point.

EXAMPLE 7.9. We consider a particular case of (7.6),

$$\dot{z} = \mathrm{e}^{2\pi \mathrm{i} t} \bar{z}^2 + \bar{z}. \tag{7.7}$$

Here 0 is a 1-periodic solution, hence we are looking for another one. There is another isolating segment $Z \subset W$ over [0, 1] for (7.7)

$$Z := [0,1] \times \{z \in \mathbb{C} \colon |\Re z| \leqslant \delta,\ |\Im z| \leqslant \delta\}$$

for $0 < \delta$ sufficiently small. Here Z is a prism with rectangular base and Z^{--} consists of its top and bottom sides. It is easy to observe that

$$\Lambda_Z = \chi(Z_0) - \chi(Z_0^{--}) = 1 - 2 = -1.$$

If $K_W = K_Z = \{0\}$ then

$$1 = \Lambda_W = \Lambda_Z = -1,$$

which is a contradiction, hence (7.7) has a nonzero 1-periodic solution.

EXAMPLE 7.10. We consider another equation

$$\dot{z} = \mathrm{e}^{2\pi \mathrm{i} t} \bar{z}^2 + z. \tag{7.8}$$

Here we try again to get a nonzero periodic solution. Now there is an isolating segment $U \subset W$,

$$U := [0,1] \times \{z \in \mathbb{C} \colon |z| < \varepsilon\}$$

for $0 < \varepsilon$ small. However, it is easy to observe that

$$\Lambda_U = \chi\bigl(D^2\bigr) - \chi\bigl(S^1\bigr) = 1 - 0 = 1,$$

hence we cannot predict the existence of nonzero 1-periodic solution as it was done in the previous example. In fact, we are able to prove the existence of a 3-periodic solution. Let

$$\widetilde{W} := W \cup \tau_1(W) \cup \tau_2(W),$$

where τ_T for a real number T is the translation of $\mathbb{R} \times \mathbb{C}$ given by

$$\tau_T(x, z) := (x + T, z).$$

$\widetilde{W}$ is an isolating segment over $[0, 3]$ which consists of glued three copies of W. Its monodromy homeomorphism can be obtained as the composition of three rotations by the angle $\frac{2\pi}{3}$, hence it is equal the identity and thus

$$\Lambda_{\widetilde{W}} = 1 - 3 = -2.$$

Put also

$$\widetilde{U} := [0, 3] \times \bigl\{z \in \mathbb{C}\colon\ |z| < \varepsilon\bigr\},$$

hence

$$\Lambda_{\widetilde{W}} = -2 \neq 1 = \Lambda_{\widetilde{U}}$$

and the conclusion follows.

Actually, the results given in the above examples extend to equations with Fourier–Taylor polynomials on the right-hand side, see, for example, [74]. Moreover, Theorem 7.7 can be generalized to the case of two point boundary value problems of the form

$$x(a) = g\bigl(x(b)\bigr)$$

for some function $g : M \to M$ (see [76]) and to *periodic isolating chains* (see [78]).

8. Selected applications of the Ważewski method

We provide some applications of the results stated above. We begin with an application to problems concerning asymptotic properties of solutions, then we present a proof of a classical result on the existence of solutions of two-point boundary value problems for a second-order equation, and finally we state a result on the existence of a kind of chaotic dynamics.

8.1. *Asymptotic solutions*

The most natural applications of the Ważewski method refer to the existence of solutions remaining in a given set; for instance, theorems on bounded solutions can be provided in this way. If the considered sets in the extended phase space contract as time tends to infinity, the obtained solutions represent the corresponding asymptotic properties. There are many results of this type in the literature. In particular, theorems on solutions which tend to constant mappings are contained, for example, in [48,66,84,88] (the result in [48] is reproduced also in [65]) and results on asymptotic coincidence of solutions of two systems are, among others, in [45,53,55,56,82]. Some other asymptotic results for ordinary equations are given in [10,25,39,41,46,69,83], for retarded differential equations in [27,80], for partial differential equations of the first order in [59,60], and for the second order in [61]. For an illustration of this kind of applications of the Ważewski method, here we provide a proof of an asymptotic result for a linear equation. Let $A = (a_{ij})$ and $B = (b_{ij})$ be continuous real $(n \times n)$-matrix-valued functions and we consider an equation

$$\dot{x} = A(t)x + B(t)x \tag{8.1}$$

in $\mathbb{R}^n$. We assume that $A(t)$ is diagonal for every t, i.e.,

$$a_{ij} = 0 \quad \text{if}\, i \neq j \tag{8.2}$$

and consider the question of estimating the size of B in order to get the existence of linearly independent solutions $x_1, \dots, x_n$ of (8.1), where $x_k = (x_k^1, \dots, x_k^n)$ for $k = 1, \dots, n$, such that $x_k^i(t)/x_k^k(t) \to 0$ if $t \to \infty$ and $i \neq k$. In particular, if $B(t) = 0$ for every t and $x_k(t_0)$ is equal to the kth vector of the canonical basis of $\mathbb{R}^n$ for some $t_0 \in \mathbb{R}$ then clearly x_k are linearly independent and $x_k^i = 0$ for $i \neq k$. If B is nonzero then such estimates were obtained by Perron and later they were improved by Szmydtówna in [81] and Onuchic in [55] using the method of Ważewski. Here we follow [55].

THEOREM 8.1 (compare Theorem II-1 in [55]). *Assume* (8.2). *Let* $R > 0$ *and let* $h : [R, \infty) \to (0, \infty)$ *be a continuous function such that for all* $i \neq j$,

$$\left|a_{ii}(t) - a_{jj}(t)\right| < h(t) \quad \textit{for every } t \geqslant R, \tag{8.3}$$

$$\int_R^\infty \left|b_{ij}(t)\right| \mathrm{e}^{H(t)}\, \mathrm{d}t < \infty, \tag{8.4}$$

$$\int_R^\infty \left|b_{ii}(t) - b_{jj}(t)\right| \mathrm{e}^{H(t)}\, \mathrm{d}t < \infty, \tag{8.5}$$

where $H(t) := \int_R^t h(s)\, \mathrm{d}s$. *Then there exist linearly independent solutions* $x_1, \dots, x_n$ *of* (8.1) *such that for every* $i, k = 1, \dots, n$ *and* $i \neq k$,

$$\lim_{t \to \infty} \frac{x_k^i(t)}{x_k^k(t)} = 0.$$

PROOF. Let g and ϕ be functions $\mathbb{R} \to \mathbb{R}$ given by

$$g(t) := \sum_{i \neq j} \bigl(|b_{ii}(t) - b_{jj}(t)| + |b_{ij}(t)| \bigr) + \mathrm{e}^{-H(t)-t},$$

$$\phi(t) := \mathrm{e}^{-H(t)} \int_t^\infty g(s) \mathrm{e}^{H(s)} \, \mathrm{d}s.$$

It follows that ϕ is a C^1-class function satisfying the differential equation

$$\phi'(t) + h(t)\phi(t) + g(t) = 0. \tag{8.6}$$

Moreover,

$$\phi(t) > 0 \quad \text{for every } t \in \mathbb{R} \tag{8.7}$$

since the last term in $g(t)$ is greater then 0 and

$$\lim_{t \to \infty} \phi(t) = 0 \tag{8.8}$$

since (8.4) and (8.5) hold. In particular, there exists $T > R$ such that

$$\phi(t) < 1 \quad \text{if } t > T. \tag{8.9}$$

For every $k = 1, \ldots, n$ put

$$V_k := \bigl\{ (t, v) \in \mathbb{R} \times \mathbb{R}^n \colon t > T,\ v^k \neq 0,\ |v^i| < \phi(t) |v^k|\ \forall i \neq k \bigr\}.$$

For $i \neq k$ define

$$L^i := \bigl\{ (t, v) \in \mathbb{R} \times \mathbb{R}^n \colon t > T,\ v^k \neq 0,$$
$$|v^i| = \phi(t) |v^k|,\ |v^j| \leqslant \phi(t) |v^k|\ \forall j \neq k \bigr\}.$$

Equation (8.1) generates a flow in the extended phase space $\mathbb{R} \times \mathbb{R}^n$. We show that

$$E_k := \bigcup_{i \neq k} L^i$$

is the set of egress points of V_k with respect to that flow and each egress point is a strict egress point. To this aim it suffices to prove that

$$\bigl(1, A(t)v + B(t)v\bigr) \cdot \nabla \ell^i(t, v) > 0 \quad \text{for } (t, x) \in L^i, \tag{8.10}$$

where

$$\ell^i : \mathbb{R} \times \mathbb{R}^n \ni (t, v) \to \frac{1}{2} \bigl((v^i)^2 - \phi^2(t) (v^k)^2 \bigr) \in \mathbb{R}.$$

Indeed, the only solutions passing through $\partial V_k \setminus E_k$ are the zero solution and solutions through $t = T$, hence no one of them initiates at V_k. For a proof of (8.10), let $(t, v) \in L^i$. We have

$$\begin{aligned}
&\bigl(1, A(t)v + B(t)v\bigr) \cdot \nabla \ell^i(t, v) \\
&\quad = \phi^2(t)\bigl(v^k\bigr)^2\bigl(a_{ii}(t) - a_{kk}(t) + b_{ii}(t) - b_{kk}(t)\bigr) - \phi'(t)\phi(t)\bigl(v^k\bigr)^2 \\
&\qquad - \phi^2(t)v^k \sum_{j\neq k} b_{kj}(t)v^j + v^i \sum_{j\neq i} b_{ij}(t)v^j
\end{aligned}$$

because $(v^i)^2 = \phi^2(t)(v^k)^2$. Since (8.6), (8.7), and (8.9) hold, and $|v^j| \leqslant \phi(t)|v^k|$ for every j, we get the following estimates

$$\begin{aligned}
&\bigl(1, A(t)v + B(t)v\bigr) \cdot \nabla \ell^i(t, v) \\
&\quad \geqslant \phi^2(t)\bigl(v^k\bigr)^2\bigl(a_{ii}(t) - a_{kk}(t) + b_{ii}(t) - b_{kk}(t)\bigr) - \phi'(t)\phi(t)\bigl(v^k\bigr)^2 \\
&\qquad - \phi^3(t)\bigl(v^k\bigr)^2 \sum_{j\neq k}\bigl|b_{kj}(t)\bigr| - \phi^2(t)\bigl(v^k\bigr)^2 \sum_{j\neq i}\bigl|b_{ij}(t)\bigr| \\
&\quad \geqslant \phi(t)\bigl(v^k\bigr)^2\Biggl(-\phi'(t) - \bigl|a_{kk}(t) - a_{ii}(t)\bigr|\phi(t) - \bigl|b_{kk}(t) - b_{ii}(t)\bigr| \\
&\qquad - \sum_{j\neq k}\bigl|b_{kj}(t)\bigr| - \sum_{j\neq i}\bigl|b_{ij}(t)\bigr|\Biggr) \\
&\quad > \phi^2(t)\bigl(v^k\bigr)^2\bigl(-\phi'(t) - h(t)\phi(t) - g(t)\bigr) = 0,
\end{aligned}$$

hence (8.10) is proved. Thus

$$W_k := V_k \cup E_k$$

is a Ważewski set and E_k is its exit set. Choose an $\varepsilon > 0$ such that if $M = (m_{ij})$ is an $(n \times n)$-matrix satisfying $m_{ii} = 1$ and $|m_{ij}| < \varepsilon$ for $i \neq j$ then M is nondegenerate. Let t_0 be such that $\phi(t_0) < \varepsilon$. Define

$$Z_k := \bigl\{(t, v) \in \mathbb{R} \times \mathbb{R}^n\colon\ t = t_0,\ v^k = 1,\ \bigl|v^i\bigr| \leqslant \phi(t_0)\ \forall i \neq k\bigr\}.$$

Thus $Z_k \subset W_k$, Z_k is homeomorphic to the disk D^{n-1}, and $Z_k \cap E_k$ is homeomorphic to the sphere S^{n-2}, hence it is not a retract of Z_k. Moreover, there is a retraction

$$r = (r_0, r_1, \dots, r_n)\colon E_k \to Z_k \cap E_k$$

given by

$$r_0(t, v) = t_0, \qquad r_k(t, v) = 1,$$

$$r_i(t,v) = \frac{\phi(t_0)}{\phi(t)} \frac{v^i}{|v^k|} \quad \text{for } 1 \leqslant i \neq k$$

since (8.7) holds. It follows by the original Ważewski theorem (see Corollary 2.2) that there exists a solution x_k of (8.1) such that

$$\begin{aligned} &\bigl(t_0, x_k(t_0)\bigr) \in Z_k \setminus E_k, \\ &\bigl(t, x_k(t)\bigr) \in V_k \quad \text{for } t \geqslant t_0. \end{aligned}$$

It follows by the choice of t_0 and ε that the matrix having the vectors $x_k(t_0)$ as the columns for $k = 1, \ldots, n$, is nondegenerate, hence the solutions $x_1, \ldots, x_n$ are linearly independent. Moreover, if $i \neq k$,

$$\bigl|x_k^i(t)\bigr| < \phi(t)\bigl|x_k^k(t)\bigr|$$

for every $t \geqslant t_0$, hence the theorem follows by (8.8). □

8.2. *Two-point boundary value problems*

It was already indicated in [85] how to apply the Ważewski theorem in proofs of the existence of solutions of two-point boundary value problems characterized by separate conditions at the initial and final values of time (like (1.10); compare Example 1.2). Later, there appeared several papers on those applications, for example [5,7,26,40,42]. Another approach, based on the Lefschetz fixed point theorem, was applied in [74,76] to the periodic problem and its generalizations (see Section 7.5). We illustrate those applications of the Ważewski method by providing a proof of a classical result on the existence of solutions of several boundary value problems associated to the second order scalar differential equation (3.20), i.e.,

$$x'' = g(t, x, x')$$

with continuous $g : \mathbb{R}^3 \to \mathbb{R}$, satisfying the Bernstein–Nagumo growth conditions. For $a < b$ we consider the boundary-value conditions

$$x(a) = x(b) = 0 \quad \text{(Dirichlet)}, \tag{8.11}$$

$$x'(a) = x'(b) = 0 \quad \text{(Neumann)}, \tag{8.12}$$

$$\begin{cases} -\alpha x(a) + \beta x'(a) = 0, \\ \gamma x(b) + \delta x'(b) = 0 \end{cases} \quad \text{(Sturm–Liouville; } \alpha, \beta, \gamma, \delta > 0), \tag{8.13}$$

$$\begin{cases} x(a) = x(b), \\ x'(a) = x'(b) \end{cases} \quad \text{(periodic)}. \tag{8.14}$$

THEOREM 8.2 (Bernstein–Nagumo). *If there exists an* $R > 0$ *and a function*

$$\psi : [0, \infty) \to (0, \infty)$$

such that the following conditions are satisfied:

(a) $g(t, R, 0) > 0$ *and* $g(t, -R, 0) < 0$ *for every* $t \in \mathbb{R}$,

(b) $\int_0^\infty \sigma \, \mathrm{d}\sigma / \psi(\sigma) > R$,

(c) *for every* $t \in \mathbb{R}$, $|x| \leqslant R$ *and* $v \in \mathbb{R}$,

$$\bigl|g(t, x, v)\bigr| < \psi\bigl(|v|\bigr),$$

then the boundary value problems (8.11)–(8.14) *associated to* (3.20) *have solutions whose images are contained in the interval* $[-R, R]$.

PROOF. Assume that (3.20) satisfies the uniqueness property. Let $\Psi : [0, \infty) \to [0, \infty)$ be defined as

$$\Psi(u) := \int_0^u \frac{\sigma \, \mathrm{d}\sigma}{\psi(\sigma)}.$$

Put

$$W := [a, b] \times \bigl\{(x, y) \in \mathbb{R}^2\colon\ |x| \leqslant R - \Psi\bigl(|y|\bigr)\bigr\} \subset \mathbb{R}^3.$$

We prove that W is an isolating segment for (3.21). Observe first that W is compact by (b). Define functions ℓ, m, n^+, and n^- by

$$\begin{aligned} &\ell(t, x, y) := t - b, \qquad m(t, x, y) := a - t, \\ &n^{\pm}(t, x, y) = \pm x - R + \Psi\bigl(|y|\bigr), \end{aligned}$$

where $(t, x, y) \in \mathbb{R}^3$. W is therefore a generalized polyfacial set determined by $\{\ell, m, n^+, n^-\}$. Since

$$\bigl(1, y, g(t, x, y)\bigr) \cdot \nabla n^{\pm}(t, x, y) = \frac{|y|}{\psi(|y|)}\bigl(\pm(\operatorname{sgn} y)\psi\bigl(|y|\bigr) - g(t, x, y)\bigr),$$

it follows by (a), (c), (3.22), and (3.23) that int W is a generalized polyfacial set for (3.21), and moreover, W is an isolating segment such that the proper exit and entrance sets are equal to, respectively,

$$\begin{aligned} W^{--} &= \bigl\{(t, x, y) \in W\colon\ xy \geqslant 0,\ |x| = R - \Psi\bigl(|y|\bigr)\bigr\}, \\ W^{++} &= \bigl\{(t, x, y) \in W\colon\ xy \leqslant 0,\ |x| = R - \Psi\bigl(|y|\bigr)\bigr\}. \end{aligned}$$

Both $W^{\pm\pm}$ consist of two disjoint strips, each of which is equal to an arc in $\mathbb{R}^2$ multiplied by $[a,b]$. It follows that

$$\Lambda_W = \chi(W_a) - \chi\left(W_a^{--}\right) = 1 - 2 = -1,$$

hence Corollary 7.3 implies that the periodic problem (8.14) has a solution. For the other problems let us consider the sets $Z \subset W$ and $Y \subset W^-$ such that

$$Z := \left\{(a,x,y) \in W\colon\ -\alpha x + \beta y = 0\right\},$$

$$W^- \setminus Y = \left\{(b,x,y) \in W\colon\ \gamma x + \delta y = 0\right\},$$

where $\alpha, \beta, \gamma, \delta \geqslant 0$, $\alpha + \beta > 0$, and $\gamma + \delta > 0$. It follows that Z is connected, $Z \cap W^- \subset Y$ consists of two points, and $W^- \setminus Y$ separates W^- into two disjoint parts each of which contains a point of $Z \cap Y$. Thus $Z \cap Y$ is a retract of Y and it is not a retract of Z. Since there are no saturated solutions contained in W, Corollary 2.2 implies that the problem

$$-\alpha x(a) + \beta y(a) = 0, \qquad \gamma x(b) + \delta y(b) = 0$$

has a solution. Thus all the problems (8.11), (8.12), and (8.13) have solutions. In the case (3.20) does not satisfy the uniqueness property, an approximating approach similar to the one described in the proof of Theorem 3.1 leads to the conclusion. □

8.3. *Detection of chaotic dynamics*

Isolating segments were applied in proofs of results concerning the existence of chaotic dynamics generated by a time-periodic nonautonomous equation. Like in Section 7.5, we assume $w : \mathbb{R} \times M \to TM$ is a time-dependent vector-field on the right-hand side of Equation (3.8) satisfying the uniqueness property. We assume, moreover, that $t \to w(t,x)$ is T-periodic for all $x \in M$. Recall that in this case fixed points of the Poincaré map $\Phi_{0,T}$ are the initial points of T-periodic solutions and

$$\Phi_{0,nT} = \Phi_{0,T}^n$$

by (7.5). The concept of chaotic dynamics of (3.8) refers to a chaotic behavior of the Poincaré map in the following sense. For a positive integer r put

$$\Sigma_r := \{1,\dots,r\}^{\mathbb{Z}},$$

i.e., the set of bi-infinite sequences of r symbols, and let the *shift map* be given by

$$\sigma : \Sigma_r \ni (\dots, s_{-1}.s_0, s_1, \dots) \to (\dots, s_0.s_1, s_2, \dots) \in \Sigma_r.$$

DEFINITION 8.1. Equation (3.8) is called *Σ_r-chaotic* provided there is a compact set $I \subset M$, invariant with respect to $\Phi_{0,T}$, and a map $g : I \to \Sigma_r$ such that:

(a) g is continuous and surjective,
(b) $\sigma \circ g = g \circ \Phi_{0,T}$,
(c) for every n-periodic sequence $s \in \Sigma_r$ its inverse image $g^{-1}(s)$ contains at least one n-periodic point of $\Phi_{0,T}$.

It follows by (c) that a Σ_r-chaotic equation has periodic solutions with minimal periods nT for every $n \in \mathbb{N}$.

THEOREM 8.3 (compare Theorem 2 in [79]). *Assume that compact ENRs W and Z are periodic isolating segments over* $[0, T]$ *for Equation* (3.8)*, and*
(a) $(W_0, W_0^{--}) = (Z_0, Z_0^{--})$,
(b) $Z \subset W$,
(c) $\mu_Z = \mu_W \circ \mu_W = \mathrm{id}_{H(W_0, W_0^{--})}$,
(d) $\Lambda_W \neq \chi(Z_0) - \chi(Z_0^{--}) \neq 0$.
Then (3.8) *is Σ_2-chaotic.*

PROOF. We present a sketch of the proof in [79]. Let

$$\sigma := \sigma_Z^+$$

be the escape-time function for Z. Define the set I as

$$I := \bigcap_{n=-\infty}^{\infty} \bigl\{x \in W_0 \colon \Phi_{0,nT+t}(x) \in W_t \ \forall t \in [0, T]\bigr\}.$$

It follows by (a) that $\sigma(0, x)$ is defined for every $x \in W_0$. If $x \in I$ then either $\sigma(0, x) < T$ or $\sigma(0, x) = T$ and $\Phi_{0,T} \in W_0 \setminus W_0^{--}$. Indeed, assume on the contrary that $\sigma(0, x) = T$ and $\Phi_{0,T} \in W_0^{--}$. Thus $\Phi_{0,T+\varepsilon} \notin W_\varepsilon$ for a small $\varepsilon > 0$ by (7.5), which contradicts to the definition of I. Define

$$J_Z := \bigl\{x \in I \colon \sigma(0, x) = T,\ \Phi_{0,T} \in W_0 \setminus W_0^{--}\bigr\},$$
$$J_W := \bigl\{x \in I \colon \sigma(0, x) < T\bigr\}.$$

It follows by the continuity of σ that J_Z and J_W are compact disjoint sets and

$$I = J_Z \cup J_W.$$

We use the letters Z and W also as the two symbols in the definition of Σ_2, i.e., we put

$$\Sigma_2 := \{Z, W\}^{\mathbb{Z}}.$$

To a point $x \in I$ we attach a symbol $p(x) \in \{Z, W\}$ by the rule $p(x) = Z$ if $x \in J_Z$ and $p(x) = W$ if $x \in J_W$. Define

$$g : I \ni x \to \bigl\{p\bigl(\Phi_{0,nT}(x)\bigr)\bigr\}_{n=-\infty}^{\infty} \in \Sigma_2.$$

Since J_Z and J_W are disjoint, g is continuous. Since periodic sequences are dense in Σ_2, in order to prove the theorem it suffices to show that the inverse image of each n-periodic sequence in Σ_2 contains an n-periodic point of $\Phi_{0,T}$. To a sequence

$$(V_0, \dots, V_{n-1}) \in \{Z, W\}^{\{0,\dots,n-1\}}$$

we attach a periodic isolating segment $V_0 \cdots V_{n-1}$ over $[0, nT]$ and a set $J_{V_0 \cdots V_{n-1}}$ (extending the notation J_Z and J_W) defined by

$$\begin{aligned} V_0 \cdots V_{n-1} &:= \bigl\{(kT + t, x)\colon\ k \in \{0, \dots, n-1\},\ t \in [0, T],\ x \in V_k\bigr\}, \\ J_{V_0 \cdots V_{n-1}} &:= \bigl\{x \in I\colon\ \Phi^k_{0,T}(x) \in J_{V_k}\ \forall k \in \{0, \dots, n-1\}\bigr\}. \end{aligned}$$

The sets $J_{V_0 \cdots V_{n-1}}$ over all n-element sequences from $\{Z, W\}^{\{0,\dots,n-1\}}$ form a compact and disjoint covering of I. In order to prove the theorem it suffices to prove that for every n each set $J_{V_0 \cdots V_{n-1}}$ contains a fixed point of $\Phi_{0,nT}$. Indeed, the image of that point under the map g is the n-periodic sequence

$$(\dots, V_{n-1}.V_0, \dots, V_{n-1}, V_0, \dots) \in \Sigma_2.$$

Recall, that the set K_{W^n}, where $W^n := W \cdots W$ (n times), consists of all fixed points of $\Phi_{0,nT}$ contained in I (see Theorem 7.7). Since $K_{W^n} \cap J_{V_0 \cdots V_{n-1}}$ is compact and open in the set of all fixed points of $\Phi_{0,nT}$, its fixed point index is defined, and, by (d), in order to prove the theorem it suffices to show that if W appears in the sequence $(V_0, \dots, V_{n-1}) \in \{Z, W\}^{\{0,\dots,n-1\}}$ exactly k times then

$$\operatorname{ind}(\Phi_{0,nT}, K_{W^n} \cap J_{V_0 \cdots V_{n-1}}) = \begin{cases} \chi\bigl(Z_0, Z_0^{--}\bigr), & \text{if } k = 0, \\ (-2)^{k-1}\bigl(\Lambda_W - \chi\bigl(Z_0, Z_0^{--}\bigr)\bigr), & \text{if } k \geqslant 1. \end{cases}$$

That equation is a consequence of Theorem 7.7 and some combinatorial calculations; we skip its proof here referring to the proof of Lemma 1 in [79]. □

EXAMPLE 8.1 (*compare* [79]). The planar equation

$$\dot{z} = \bigl(1 + \mathrm{e}^{\mathrm{i}\phi t}|z|^2\bigr)\bar{z} \tag{8.15}$$

is Σ_2-chaotic if $0 < \phi \leqslant 0.495$. The proof in [79] is based on the existence of two isolating segments W and Z over $[0, 2\pi/\phi]$ for (8.15) which are similar to the ones shown in Figure 9 (with W at the top and Z at the bottom). It follows by the figure that $\Lambda_W = 1$, $\chi(W_0) = 1$, $\chi(W_0^{--}) = 2$, hence Theorem 8.3 implies the result.

In [79], the parameter values $0 < \phi \leqslant 1/288$ were considered only. The extension to the other values was given in [95]. Actually, in [95] it was proved that (8.15) is Σ_3-chaotic in the considered range of ϕ. The equation was further investigated in [94,96], where the

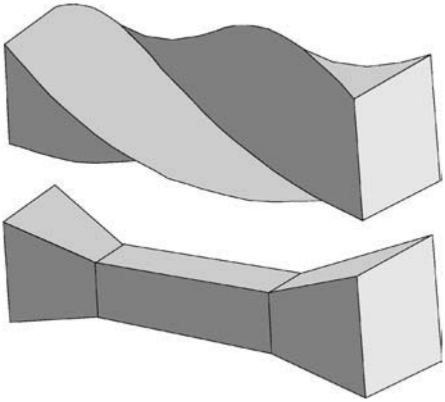


Fig. 9.

existence of infinitely many distinct homoclinic solutions to the zero one was proved. Generalizations of Theorem 8.3 and extensions of results on (8.15) to more general classes of equations were given, among others, in [91–93,97]. Another approach to chaotic dynamics based on the existence of suitably located isolating blocks or segments is presented in [29,75].

9. Conley index

Using the notion of isolating block we present the definition of the Conley index and provide its basic properties. In particular, we indicate its relation to the Ważewski method. We do not present more advanced topics of the theory of the index; for them we refer to [50].

9.1. *Isolated invariant sets and the Conley index*

Let X be a metrizable locally compact space and let ϕ be a local flow on X.

DEFINITION 9.1. A compact set $S \subset X$ is called an *isolated invariant set* provided there exists U, a neighborhood of S, such that $S = \operatorname{Inv}(U)$. Such an U is called an *isolating neighborhood for* S.

EXAMPLE 9.1. Let A be a real $n \times n$-matrix, let U be an open neighborhood of 0 in $\mathbb{R}^n$, and let $f : U \to \mathbb{R}^n$ be a C^1-class function. Denote by ϕ the local flow on U generated by the equation

$$\dot{x} = Ax + f(x).$$

Assume that

$$f(x) = \mathrm{o}(|x|) \quad \text{as } x \to 0$$

and there are no purely imaginary eigenvalues of A. Then the one-point set $\{0\}$ is an isolated invariant set for ϕ. Indeed, by the Grobman–Hartman theorem ϕ is conjugated in a neighborhood of 0 to the flow generated by the linear vector-field $x \mapsto Ax$ which does not have any bounded solutions except of the trivial one.

In an obvious way Example 9.1 generalizes to a stationary point of the local flow generated by a C^1-class vector-field $v : M \to TM$ on a Riemannian manifold M.

DEFINITION 9.2. A stationary point $x_0 \in M$ is called *hyperbolic of index k* if the differential $d_{x_0}v$ does not have eigenvalues on the imaginary axis and there are exactly k eigenvalues (counted with multiplicities) with the positive real part.

It follows that if x_0 is a hyperbolic point then $\{x_0\}$ is an isolated invariant set.

EXAMPLE 9.2. Assume that ϕ is a gradient-like local flow (see Section 6.2). It is not difficult to observe that if x_0 is a stationary point and there is a compact neighborhood U of x_0 such that there are no other stationary points in U then the one-point set $\{x_0\}$ is an isolated invariant set and U is its isolating neighborhood. It follows that isolated critical points of a smooth functional $g : M \to \mathbb{R}$ are isolated invariant sets of the local flow generated by (6.1).

If B is an isolating block then it is an isolating neighborhood for $S = \operatorname{Inv}(B)$. In this case we say that B is an *isolating block for S*. We recall in a more comprehensive version the first one of the two main theorems stated in Section 1.3:

THEOREM 9.1 (First Conley Theorem). *If U is a neighborhood of an isolated invariant set S then there exists an isolating block B for S such that $B \subset U$. Moreover, if ϕ is smooth (i.e., of C^∞-class) local flow on a smooth manifold then B can be chosen as a smooth submanifold with corners such that B^- and B^+ are smooth submanifolds with boundary.*

PROOF. We provide the idea of the proof given in the paper [90] in the smooth case. Let U be an open isolating neighborhood of S. It is proved in [90] that there exist continuous functions $\ell, m : U \to [0, \infty)$ such that

$$\operatorname{Inv}^-(U) = \ell^{-1}(0), \qquad \operatorname{Inv}^+(U) = m^{-1}(0),$$

ℓ is of C^∞-class outside of $\operatorname{Inv}^-(U)$ and m is of C^∞-class outside of $\operatorname{Inv}^+(U)$, and

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\ell\big(\phi_t(x)\big)\big) < 0, \quad \text{if } x \notin \operatorname{Inv}^-(U),$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(m\big(\phi_t(x)\big)\big) > 0, \quad \text{if } x \notin \operatorname{Inv}^+(U).$$

Let V be a compact neighborhood of S contained in U. It follows by the compactness of V,

$$S = V \cap \bigcap_{\varepsilon>0} \ell^{-1}\big([0,\varepsilon]\big) \cap \bigcap_{\delta>0} m^{-1}\big([0,\delta]\big)$$

and, moreover, there exist an $\varepsilon_0 > 0$ such that if $0 < \varepsilon \leqslant \varepsilon_0$ and $0 < \delta \leqslant \varepsilon_0$ then

$$V_{(\varepsilon,\delta)} := V \cap \ell^{-1}\big([0,\varepsilon]\big) \cap m^{-1}\big([0,\delta]\big)$$

is a compact neighborhood of S. By the Sard theorem, choose $(\varepsilon, \delta) \in (0, \varepsilon_0)^2$ which is a regular value of the map $(\ell, m) : V \to \mathbb{R}^2$. For that value, $\operatorname{int} V_{(\varepsilon,\delta)}$ is a regular polyfacial set determined by $\{\ell - \varepsilon, m - \delta\}$, hence, by Remark 3.1, the set $V_{(\varepsilon,\delta)}$ is an isolating block satisfying the required properties. For a proof in the purely topological case we refer to [18] or [70, Chapter 22]. $\square$

REMARK 9.1. In the smooth case it was additionally proved that $\operatorname{int} B$ is a regular polyfacial set, hence if v is the vector field generating ϕ and w is a vector field sufficiently close to v on B then B is also an isolating block for the local flow generated by w.

REMARK 9.2. If the phase space of ϕ is a 2-dimensional manifold then it follows by the construction provided in the proof of [72, Theorem 3.3] that there exists an isolating block B for S such that B is a topological manifold with boundary and B^- is its submanifold with boundary.

Now our aim is to prove the Second Conley Theorem which we recall below. It is preceded by a lemma which will be used in the proof. (Recall that the concept of B-invariant subset of B is given in Definition 7.2).

LEMMA 9.1. *If B is an isolating block and Z is a B-invariant compact neighborhood of* $\operatorname{Inv}(B)$ *then*

$$\big[Z/Z^-, *\big] = \big[B/B^-, *\big].$$

PROOF. By the continuity of ϕ, it is easy to prove that Z is a neighborhood of A, hence Lemma 7.1(i) implies that $B_n \subset Z$ for some $n \in \mathbb{N}$. Since the inclusion induces a homeomorphism of the compact spaces Z/Z^- and $Z \cup B^-/B^-$, it follows by (ii) in that lemma,

$$\big[Z/Z^-, *\big] = \big[Z \cup B^-/B^-, *\big] = \big[D_n/B^-, *\big] = \big[B/B^-, *\big],$$

which finishes the proof. $\square$

THEOREM 9.2 (Second Conley Theorem). *Let S be an isolated invariant set and let B and $\widetilde{B}$ be isolating blocks for S. Then*

$$\big[B/B^-, *\big] = \big[\widetilde{B}/\widetilde{B}^-, *\big].$$

PROOF. As in Section 7, we denote by A, $\sigma^{\pm}$, etc. the sets and functions corresponding to the block B. The analogous sets and functions for $\widetilde{B}$ we denote by $\widetilde{A}$, $\tilde{\sigma}^{\pm}$, etc. Without loss of generality we assume

$$B \subset \widetilde{B}.$$

(In the other case one can find an isolating block for S contained in the intersection of the blocks B and $\widetilde{B}$ by Theorem 9.1, and then compare $[B/B^-,*]$ and $[\widetilde{B}/\widetilde{B}^-,*]$ to the homotopy type corresponding to that block.) The proof will be performed in a few steps. Below we use the following notation for intervals: if $a \in \mathbb{R}$ then $[a,a]$ is the one point set $\{a\}$ and (a,a) is equal to $\emptyset$.

Step 1. For every $n \in \mathbb{N}$,

$$\bigl[B/B^-,*\bigr] = \bigl[B_n/B_n^-,*\bigr].$$

(Recall that B_n is defined in Section 7.2.) It immediately follows by Lemmas 7.1(i) and 9.1.

Step 2. There exists $n \in \mathbb{N}$, $n \geqslant 1$, such that

$$\begin{aligned} &\phi\bigl(x,\bigl(0,\tilde{\sigma}^+(x)\bigr)\bigr) \cap B_n = \emptyset, \quad \text{if } x \in B_n^-, \\ &\phi\bigl(x,\bigl(-\tilde{\sigma}^-(x),0\bigr)\bigr) \cap B_n = \emptyset, \quad \text{if } x \in B_n^+. \end{aligned}$$

By the symmetry of notation, we need to justify the first condition only. Assume on the contrary that for every n there exists $x_n \in B_n^-$ and $0 < t_n < \tilde{\sigma}^+(x_n)$ such that $\phi(x_n,t_n) \in B_n^+$. We can assume that $x_n \to x_0$ and $t_n \to t_0$ as $n \to \infty$. $A^\pm = \bigcap_n B_n^\pm$ by Lemma 7.1, hence we have

$$x_0 \in A^- \subset \operatorname{Inv}^-(B) \subset \operatorname{Inv}^-(\widetilde{B}).$$

If $t_0 = \infty$ then $x_0 \in \operatorname{Inv}^+(\widetilde{B})$ and thus $x \in S$ which is impossible. If $t_0 < \infty$ then $\phi(x_0,t_0) \in A^+$, hence again $x \in S$ and the assertion is proved.

Step 3. Since now we assume that n satisfies the conclusion of Step 2. Then $\tilde{\sigma}^\pm(x) < \infty$ for every $x \in B_n^\pm$. Indeed, in the other case the positive or negative limit set of x is contained in $\widetilde{B}$ and contains points outside of the interior of B, which is impossible.

Step 4. By Step 3, we can define compact sets

$$M := \bigcup_{x \in B_n^+} \phi\bigl(x,\bigl[-\tilde{\sigma}^-(x),0\bigr]\bigr), \qquad N := \bigcup_{x \in B_n^-} \phi\bigl(x,\bigl[0,\tilde{\sigma}^+(x)\bigr]\bigr).$$

It follows by Step 2 that $M \cap B_n = B_n^+$ and $N \cap B_n = B_n^-$. Moreover, by Lemma 9.1,

$$\bigl[\widetilde{B}/\widetilde{B}^-,*\bigr] = \bigl[B_n \cup M \cup N/N^-,*\bigr].$$

Step 5.

$$\bigl[B_n \cup M \cup N/N^-,*\bigr] = \bigl[B_n \cup N/N^-,*\bigr].$$

This is obvious since B_n^+ is a strong deformation retract of M.

Step 6. Finally,

$$\bigl[B_n \cup N/N^-,*\bigr] = \bigl[B_n/B_n^-,*\bigr].$$

In order to prove that assertion recall that $n \geqslant 1$ and observe that

$$B_n = \{x \in B_n \colon \sigma^+(x) \geqslant 1\} \cup \phi(B_n^- \times [-1, 0]),$$
$$B_n \cup N = \{x \in B_n \colon \sigma^+(x) \geqslant 1\} \cup \bigcup_{x \in B_n^-} \phi(x, [-1, \tilde{\sigma}^+(x)]).$$

Since the second compounds are homeomorphic each to the other such that the set $\phi_{-1}(B_n^-)$ remains intact and B_n^- is transformed to N^-, there exists a homeomorphism

$$(B_n, B_n^-) \to (B_n \cup N, N^-)$$

and the assertion follows. We get the conclusion of the theorem by combining the equations in Steps 1, 4, 5, and 6. ☐

As it was pointed out in Section 1.3, Theorems 9.1 and 9.2 lead to Definition 1.3 of the Conley index:

$$h(\phi, S) := [B/B^-, *]$$

for an arbitrary isolating block B for S. Actually, in order to define the Conley index, a more general notion of *index pair* (see [50, Definition 3.4]) can be used.

9.2. *Properties of the Conley index*

At first we provide the definition of the continuation (i.e., homotopy) relation between isolated invariant sets. Assume that ϕ and ψ are local flows on a metrizable locally compact space X and let S and T be isolated invariant sets for ϕ and, respectively, ψ.

DEFINITION 9.3. We say that pairs (ϕ, S) and (ψ, T) are *related by continuation* (denoted $(\phi, S) \simeq (\psi, T)$) if there exists a local flow Φ on $X \times [0, 1]$, an isolated invariant set S^* for Φ, and local flows ϕ^σ on X, $\sigma \in [0, 1]$, such that $\Phi_t(x, \sigma) = (\phi_t^\sigma(x), \sigma)$, $\phi_0 = \phi$, $\phi_1 = \psi$, $S = \{x \colon (x, 0) \in S^*\}$, and $T = \{x \colon (x, 1) \in S^*\}$.

The relation of continuity between isolated invariant sets provides no information on their topology. In particular, a nonempty set can be continued to the empty one, and a connected set can split by continuation into a nonconnected one as we see in the following examples.

EXAMPLE 9.3. Take a family of local flows ϕ^λ generated by

$$\dot{x} = x^2 + \lambda$$

on the real line. It follows that

$$(\phi^0, \{0\}) \simeq (\phi^1, \emptyset).$$

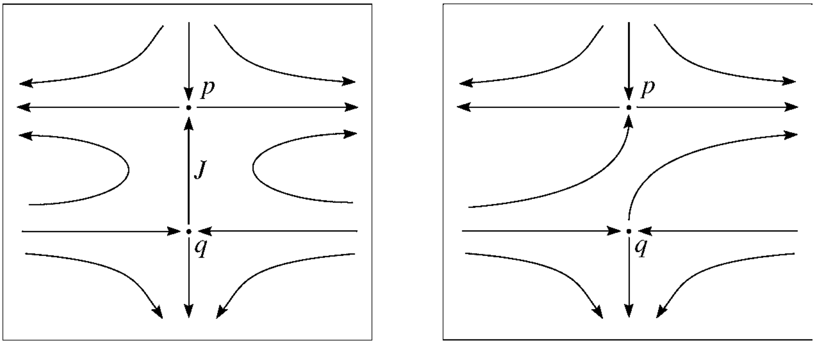


Fig. 10.

EXAMPLE 9.4. Let p and q be two stationary hyperbolic points of local flows ϕ and ψ in the plane as shown on Figure 10, where the left and right drawing represent the phase portrait of ϕ and, respectively, of ψ. Let the isolated invariant set for ϕ is equal to the set consisting of p, q, and the whole trajectory J connecting p and q (i.e., $J = \phi(y)$, $\alpha(y) = q$, and $\omega(y) = p$). The flow ψ is obtained from ϕ by a perturbation which splits J into two disjoint unbounded trajectories. Since that perturbation can by arbitrarily small,

$$\bigl(\phi, \{p, q\} \cup J\bigr) \simeq \bigl(\psi, \{p, q\}\bigr).$$

In the sequel we need also the notion of the product $\phi \times \psi$, where ψ is a local flow on X and ψ is a local flow on Y. It is given by

$$(\phi \times \psi)_t(x, y) := \bigl(\phi_t(x), \psi_t(y)\bigr).$$

In the smooth case it corresponds to the Cartesian product of the generating vector-fields.

The following theorem provides the main properties of the Conley index. They are counterparts of the properties (i)–(iv) of the fixed point index given in Proposition 5.7.

THEOREM 9.3.

(i) (Ważewski Property) *If* $h(\phi, S) \neq \overline{0}$ *then* $S \neq \emptyset$.

(ii) (Additivity) *If* S *and* T *are isolated invariant sets of* ϕ *and* $S \cap T = \emptyset$ *then*

$$h(\phi, S \cup T) = h(\phi, S) \vee h(\phi, T).$$

(iii) (Multiplicativity) *If* ϕ *is a local flow on* X *and* ψ *is a local flow on* Y, S *and* T *are isolated invariant sets of* ϕ *and* ψ, *respectively, then*

$$h(\phi \times \psi, S \times T) = h(\phi, S) \wedge (\psi, T).$$

(iv) (Continuation Property) *If* $(\phi, S) \simeq (\psi, T)$ *then* $h(\phi, S) = h(\psi, T)$.

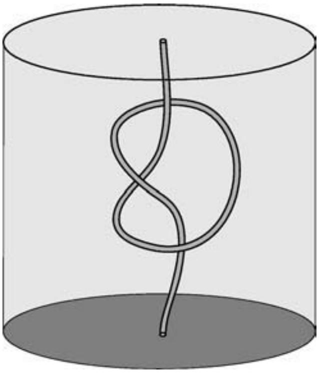


Fig. 11.

PROOF. Ad (i). This is an immediate consequence of Corollary 2.5 and Lemma 5.1.

Ad (ii). If B and $\widetilde{B}$ are isolating blocks for S and T, respectively, and $B \cap \widetilde{B} = \emptyset$ then $B \cup \widetilde{B}$ is also an isolating block for ϕ and the result follows by the definition of the index.

Ad (iii). For B and $\widetilde{B}$ as above the set $B \times \widetilde{B}$ is an isolating block for $S \times T$, hence the result follows.

Ad (iv). If the local flow ϕ is smooth, the result is a consequence of Remark 9.1, since in this case $h(\phi, S)$ is locally constant with respect to the change of the vector-field generating ϕ. For a proof in the general case we refer to [70, Chapter 23]. ☐

Theorem 9.3(i) is a version of the Ważewski theorem. It is weaker then Corollary 2.5 as it can be shown in the following example:

EXAMPLE 9.5 (*compare* II.2 *in* [21]). Let $B \subset \mathbb{R}^3$ be given by

$$B := D^2 \times [0, 1] \setminus P,$$

where D^2 denotes the disk $\{x \in R^2\colon\ |x| \leqslant 1\}$ and P is a knotted thin open pipe P as shown in Figure 11, i.e., $P \subset D^2 \times [0, 1]$ and there exists a homeomorphism

$$h : U \times [0, 1] \to P,$$

where U is equal to $\{x \in \mathbb{R}^2\colon\ |x| < r\}$ for some $0 < r < 1$, such that

$$h\big(U \times \{0\}\big) = U \times \{0\}, \qquad h\big(U \times \{1\}\big) = U \times \{1\}.$$

Assume that B is an isolating block for a local flow ϕ in $\mathbb{R}^3$ such that the bottom of B is equal to B^-. The exit set B^- is not a strong deformation retract of B since they have different fundamental groups, hence Corollary 2.5 implies that

$$S := \operatorname{Inv}(B) \neq \emptyset.$$

On the other hand, $h(\phi, S) = [B/B^-, *]$ is the trivial pointed homotopy type and therefore it does not provide any information on S. Indeed, since B is compact, the inclusion $B \hookrightarrow B \cup (D^2 \times \{0\})$ induces a homeomorphism

$$B/B^- \to B \cup \left(D^2 \times \{0\}\right)/\left(D^2 \times \{0\}\right).$$

The set $\partial U \times [0,1] \cup \overline{U} \times \{0\}$ is a strong deformation retract of $\overline{U} \times [0,1]$, hence $B \cup (D^2 \times \{0\})$ is a strong deformation retract of $D^2 \times [0,1]$. Lemma 5.1 implies

$$\begin{aligned}\left[B/B^-, *\right] &= \left[B \cup \left(D^2 \times \{0\}\right)/\left(D^2 \times \{0\}\right), *\right] \\ &= \left[D^2 \times [0,1]/\left(D^2 \times \{0\}\right), *\right] = \overline{0}.\end{aligned}$$

In applications of the Conley index it is convenient to consider its various reductions, for example $H(h(\phi, S))$ and $\chi(h(\phi, S))$ for an arbitrary homology or cohomology functor H and the corresponding Euler characteristic, the category $\operatorname{cat}(h(\phi, S))$, and the cup-length $\ell(h(\phi, S))$. In particular, by Remark 5.1,

$$\check{H}^*\left(h(\phi, S)\right) = \check{H}^*\left(B, B^-\right)$$

for arbitrary isolating block B for S. (In fact, that equation holds for every homology or cohomology functor H, compare [68, Section 1.10]; $H(h(\phi, S))$ is called the *homological* or, respectively, the *cohomological Conley index* and usually is denoted by $CH(\phi, S)$.) In the following result we gather information on the isolated invariant set provided by the above mentioned reductions of the Conley index.

THEOREM 9.4. *Let S be an isolated invariant set for a local flow ϕ. Then*

(i) $\ell(S) \geqslant \ell(h(\phi, S))$.

(ii) *If ϕ is smooth then* $\operatorname{cat}(S) \geqslant \operatorname{cat}(h(\phi, S))$.

(iii) *If the phase space of ϕ is a 2-dimensional manifold then either S has infinitely many components or $\check{\chi}(S) \geqslant \chi(h(\phi, S))$.*

(iv) *If ϕ is smooth and gradient-like then* $\#S_0 \geqslant \operatorname{cat}(h(\phi, S))$, *where S_0 denotes the set of stationary points contained in S.*

(v) *If ϕ is smooth and $T > 0$ then* $\operatorname{ind}(\phi_T, S_T) = \chi(h(\phi, S))$, *where S_T denotes the set of T-periodic points contained in S.*

(vi) *If ϕ is smooth then $i(\phi, U) = \chi(h(\phi, S))$ for every open isolating neighborhood U of S.*

PROOF. At first note that the smoothness of ϕ implies the existence of an isolating block B for S such that B and B^- are ENRs by Theorem 9.1.

Ad (i). It follows by Proposition 5.3 and Corollary 7.2.

Ad (ii). It follows by Proposition 5.2 and Theorem 7.1.

Ad (iii). The result is a consequence of Proposition 5.5, (7.2) in Example 7.2, and Remark 9.2.

Ad (iv). It follows by Theorem 7.5.

Ad (v) and (vi). These are consequences of Proposition 5.5 and Theorem 7.6. □

Obviously, the absolute homotopy type of B/B^- also does not depend on isolating block B for S, hence we can define a less accurate invariant then the Conley index h by

$$h'(\phi, S) := \left[B/B^-\right]$$

It has similar properties as h. Moreover, by Propositions 5.2, 5.3, and 5.5 we get

$$\ell\big(h(\phi, S)\big) \geqslant \ell\big(h'(\phi, S)\big) - 1$$

and, in the case of smooth ϕ,

$$\operatorname{cat}\big(h(\phi, S)\big) \geqslant \operatorname{cat}\big(h'(\phi, S)\big) - 1,$$
$$\chi\big(h(\phi, S)\big) = \chi\big(h'(\phi, S)\big) - 1,$$

hence the corresponding results to those in Theorem 9.4 hold for h'.

9.3. *Examples of Conley indices and an application*

In the following results we use the homotopy types

$$\Sigma^k := \left[S^k, s_0\right],$$
$$\Pi^k := \left[\mathbb{R}P^k, p_0\right],$$

where S^k and $\mathbb{R}P^k$ denote the k-dimensional sphere and real projective space, respectively, and s_0 and p_0 are their arbitrary points. Σ^0 is sometimes denoted by $\overline{1}$ since it is the neutral element for the smash product.

Assume that ϕ is a smooth flow on a Riemannian manifold. At first we calculate the index of a hyperbolic point (see Definition 9.2).

PROPOSITION 9.1. *Assume that x_0 is hyperbolic of index k. Then*

$$h\big(\phi, \{x_0\}\big) = \Sigma^k. \tag{9.1}$$

PROOF. By the Grobman–Hartman theorem, ϕ is conjugated in a neighborhood of x_0 to the linear flow ψ of the equation

$$\dot{x} = Ax$$

in $\mathbb{R}^n$, where $A = (a_{ij})$ is diagonal, $a_{ii} = 1$ for $i = 1, \ldots, k$ and $a_{ii} = -1$ for $i = k + 1, \ldots, n$. Thus isolating blocks for ϕ in a neighborhood of x_0 and for ψ in a neighborhood of 0 are homeomorphic, hence it suffices to calculate $h(\psi, \{0\})$.

$$B = D^k \times D^{n-k}$$

is an isolating block for $\{0\}$ with respect to ψ and

$$B^- = S^{k-1} \times D^{n-k},$$

hence B/B^- has the same homotopy type as S^k and the result follows. $\square$

EXAMPLE 9.6. Let $g : M \to \mathbb{R}$ be a Morse function on an n-dimensional Riemannian manifold M. By definition, it means that each its critical point x_0 is nondegenerate, hence, by the Morse lemma, there exists a chart $f : U \to \mathbb{R}^n$ in a neighborhood of x_0 such that $f(x_0) = 0$ and

$$g\bigl(f^{-1}(v)\bigr) = g(x_0) - \sum_{i=1}^{k} v_i^2 + \sum_{i=k+1}^{n} v_i^2 \tag{9.2}$$

for $v = (v_1, \dots, v_n) \in f(U)$. The number k is independent of the choice of the chart and is called the *Morse index of* x_0. It follows by (9.2) that x_0 is a hyperbolic point of index k of the local flow ϕ generated by (6.1), hence (9.1) holds.

By the example, the Conley index is a direct generalization of the Morse index (up to the correspondence k vs. Σ^k): it also applies to degenerate isolated critical points of a functional, since they are isolated invariant sets as we have mentioned in Example 9.2. In spite of the fact that functionals used in variational methods operate in infinite-dimensional spaces, hence not locally compact ones, the Conley index finds its application there thanks to the saddle-point reduction method from [2] (compare also Remark 7.2).

Let Γ be a nontrivial periodic orbit of ϕ, i.e., the trajectory $\phi(x)$ of a periodic, nonstationary point x. It is called *hyperbolic of index* k if $d_x P$, the differential at x of the Poincaré map P associated to a section of Γ in a neighborhood of x does not have eigenvalues on the unit circle and there are exactly k eigenvalues (counted with multiplicities) outside of the unit disc. Then the unstable manifold $W^u(\Gamma)$ of Γ is of dimension $k + 1$ and we call Γ *untwisted* (respectively, *twisted*) provided $W^u(\Gamma)$ is orientable (respectively, nonorientable).

PROPOSITION 9.2. *Assume that Γ is hyperbolic of index k. Then Γ is an isolated invariant set and*

(i) *if $k = 0$ then $h(\phi, \Gamma) = [S^1/\emptyset, *]$,*
(ii) *if $k \geqslant 1$ and Γ is untwisted then $h(\phi, \Gamma) = \Sigma^k \vee \Sigma^{k+1}$,*
(iii) *if $k \geqslant 1$ and Γ is twisted then $h(\phi, \Gamma) = \Pi^2 \wedge \Sigma^{k-1}$.*

PROOF. We present a sketch of a proof. In the untwisted case one can find an isolating block B for Γ which is homeomorphic to $S^1 \times D^k \times D^{n-k-1}$ such that its exit set B^- corresponds to $S^1 \times S^{k-1} \times D^{n-k-1}$. In the case $k = 0$ the result follows directly. Assume that $k \geqslant 1$. It follows that $(B/B^-, *)$ has the homotopy type of $(S^1 \times S^k/(S^1 \times \{s_0\}, *)$, i.e., $[S^1/\emptyset, *] \wedge \Sigma^k$. By the geometric considerations which are shown on Figure 12 in the case $k = 1$, that homotopy type appears to be equal to $\Sigma^k \vee \Sigma^{k+1}$. In drawing 1 on

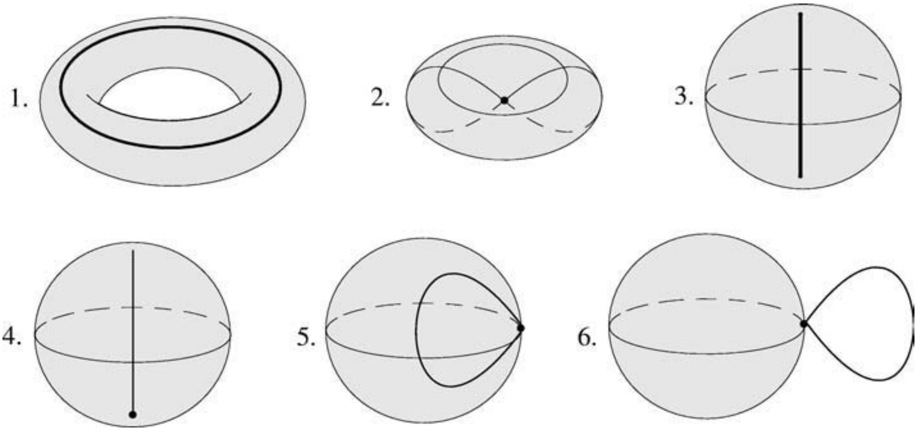


Fig. 12.

the figure, there is the Cartesian product $S^1 \times S^1$ having marked the circle $S^1 \times \{s_0\}$. By collapsing the circle to a point one gets what is seen in drawing 2. The resulting quotient space is homeomorphic to the union of the two-dimensional sphere S^2 with the bar joining its poles divided by the bar. This is marked in drawing 3. In drawing 4, there is a pointed space with the same homotopy type, but here the bar is not collapsed and the south pole is the base point. Without changing the pointed homotopy type, the bar is deformed to the circle inside the sphere in drawing 5. In order to better see that this is in fact the pointed wedge sum $S^1 \vee S^2$, the homeomorphic space with the circle attached outside of the sphere is shown in drawing 6. That argument generalizes to the higher values of k.

If Γ is twisted then B is homeomorphic to $M \times D^{k-1} \times D^{n-k-1}$ such that B^- corresponds to $\partial M \times S^{k-2} \times D^{n-k-1}$, where M is the compact Möbius band and ∂M denotes its geometric boundary. Thus $[B/B^-, *]$ is equal to $\Pi^2 \wedge \Sigma^{k-1}$. □

Recall that in Example 5.1 we have shown that $\Sigma^0 \vee \Sigma^1$ is not the same as $[S^1/\emptyset, *]$, hence (i) in Proposition 9.2 is not of the form (ii) for $k = 0$. Since the Conley index of an attracting hyperbolic orbit Γ is equal to $[S^1/\emptyset, *]$ by (i), and the Conley index of an isolated invariant set S consisting of two hyperbolic points of index 0 and index 1 is equal to $\Sigma^0 \vee \Sigma^1$ by Theorem 9.3(ii)), it follows by (iv) of that theorem that there is no continuation between (ϕ, Γ) and (ψ, S) for any two local flows ϕ and ψ. Notice that we essentially used the Conley index h in order to prove that assertion. Its reduction to h' is useless since both the absolute homotopy types $h'(\phi, \Gamma)$ and $h'(\psi, S)$ are equal to $[S^1/\emptyset]$. Moreover, for every homology or cohomology functor H with coefficients in R,

$$H_q\big(h(\phi, \Gamma)\big) \cong H_q\big(h(\psi, S)\big) \cong \begin{cases} R, & \text{if } q = 0 \text{ or } q = 1, \\ 0, & \text{if } q \neq 0, 1. \end{cases}$$

Surprisingly, the corresponding result in the case of nonattracting hyperbolic periodic orbits is not true. We provide an example of a hyperbolic periodic orbit of index 1 which continues to the set consisting of two hyperbolic stationary points of index 1 and 2.

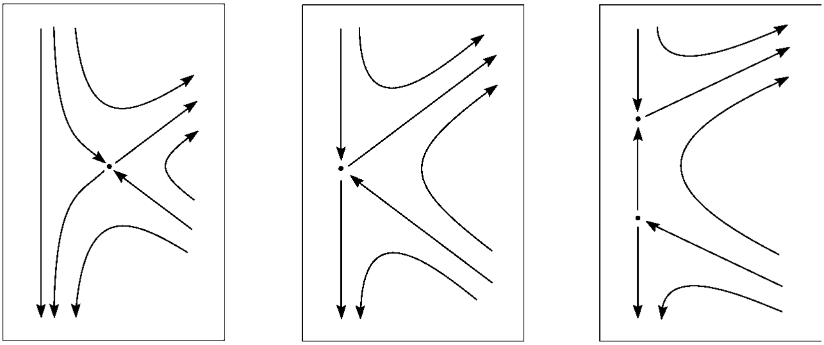


Fig. 13.

EXAMPLE 9.7. We consider a family of local flows ϕ^λ in $\mathbb{R}^3$, where $\lambda \in [0, 1]$. For $\lambda \in [0, 2/3]$ let ϕ^λ be generated by the equation

$$\dot{\theta} = 1, \qquad \dot{r} = f(r, z, \lambda), \qquad \dot{z} = g(r, z, \lambda)$$

in the cylindrical coordinates θ, r, z, where $\theta \in [0, 2\pi)$, $r \geqslant 0$, and $z \in \mathbb{R}$. Assume that the vector-field (f, g) in $[0, \infty) \times \mathbb{R}$ has the phase portrait given in Figure 13 (where r is the horizontal coordinate and z is the vertical one) for the parameter values $\lambda = 0$ (left), $\lambda = 1/3$ (middle), and $\lambda = 2/3$ (right). Assume that $1/3$ is the only parameter value of $\lambda \in [0, 2/3]$ in which the phase portrait changes, which implies that the right drawing represents also phase portraits at $\lambda \in [0, 1/3)$ and the right one at $\lambda \in (1/3, 2/3]$. It follows that for $\lambda \in [0, 1/3)$ there is a periodic hyperbolic orbit S_λ of index 1 represented by the dot in the left drawing. It collapses at $\lambda = 1/3$ to the one-point set $S_{1/3}$ consisting of the stationary point on the line $r = 0$ which is seen in the middle drawing. For $\lambda \in (1/3, 2/3]$, that point bifurcates into the isolated invariant set S_λ consisting of two hyperbolic points of index 1 (the lower point) and of index 2 (the upper point) and the trajectory connecting them. For the parameter values $\lambda \in (2/3, 1]$ we no longer assume any rotating and we split the connecting trajectory into two unbounded ones in a similar way as it was done for the planar system in Example 9.4 and shown on Figure 10. Thus S_λ for $\lambda \in (2/3, 1]$ consists of two hyperbolic points of index 1 and 2 and

$$\left(\phi^0, S_0\right) \simeq \left(\phi^1, S_1\right).$$

As an example of application of various properties of the Conley index we present a proof of a result on existence of a bounded solution.

PROPOSITION 9.3 (see I.5 in [21]). *The scalar equation*

$$x^{(n)} = x^2 - 1 \tag{9.3}$$

has a nonconstant solution x such that $x, x', \dots, x^{(n-1)}$ are bounded.

PROOF. We rewrite the proof in [21]—it is also reproduced in [70], where we refer for more detailed calculations. By the change of variables $\tau = t/\varepsilon$ and $z = \varepsilon^n x$ Equation (9.3) is equivalent to

$$\frac{\mathrm{d}^n z}{\mathrm{d}\tau^n} = z^2 - \varepsilon^{2n},$$

hence to the system

$$z_i' = z_{i+1}, \quad \text{if } i = 1, \dots, n-1, \qquad z_n' = z_1^2 - \varepsilon^{2n}. \tag{9.4}$$

In particular, for $\varepsilon = 0$, (9.4) becomes

$$z_i' = z_{i+1}, \quad \text{if } i = 1, \dots, n-1, \qquad z_n' = z_1^2 \tag{9.5}$$

and we assert that the constant solution 0 is the only bounded solution of (9.5). In order to prove that assertion assume that $z = (z_1, \dots, z_n)$ is a bounded solution. Since $z_n'(t) \geqslant 0$, the map z_n is constant or increasing. In both cases there exist $\alpha^-, \alpha^+ \in \mathbb{R}$ such that

$$\lim_{t \to \pm\infty} z_n(t) = \alpha^\pm.$$

Let $\alpha^+ \neq 0$. Then $z_{n-1}'(t)$ is close to α^+ for t near ∞, hence z_{n-1} is unbounded which contradicts to the assumption. The same holds if $\alpha^- \neq 0$, hence $\alpha^- = \alpha^+ = 0$ and z_n is the constant zero function. It follows that z_{n-1} is a constant function and by a similar argument one concludes that the constant is equal to 0. By repeating that argument we conclude $z_n(t) = \cdots = z_1(t) = 0$ for every t. Let ϕ^ε be the local flow generated by (9.4) if $\varepsilon \geqslant 0$ and by

$$z_i' = z_{i+1}, \quad \text{if } i = 1, \dots, n-1, \qquad z_n' = z_1^2 + \varepsilon^{2n}$$

if $\varepsilon \leqslant 0$. We have just proved that $\{0\}$ is an isolated invariant set of ϕ^0. Take an arbitrary isolating block B for $\{0\}$ such that conclusion of Remark 9.1 is satisfied, hence B is also an isolating block for local flows ϕ^ε, where $|\varepsilon| \leqslant \varepsilon_0$ for some $\varepsilon_0 > 0$. For such an ε define the isolated invariant set S_ε as the invariant part of B with respect to the local flow ϕ^ε. It is easy to see that if $\varepsilon < 0$ then there are no bounded trajectories for ϕ^ε, hence $S_\varepsilon = \emptyset$. By Theorem 9.3(i) and (iv),

$$h(\phi^\varepsilon, S_\varepsilon) = \overline{0} \tag{9.6}$$

if $|\varepsilon| < \varepsilon_0$. Assume now that $\varepsilon > 0$. There are two stationary points p^+ and p^-,

$$p^\pm = (\pm\varepsilon^n, 0, \dots, 0),$$

and the eigenvalues of the derivative of the vector-field generating ϕ^ε at $p^\pm$ satisfy the characteristic equation

$$\lambda^n = \pm 2\varepsilon.$$

Assume on the contrary that $p^{\pm}$ are the only bounded trajectories, hence S_ε is equal to the union $\{p^-, p^+\}$ and Theorem 9.3(ii) implies

$$h(\phi^\varepsilon, S_\varepsilon) = h(\phi^\varepsilon, \{p^-\}) \vee h(\phi^\varepsilon, \{p^+\}). \tag{9.7}$$

If n is odd then there are no eigenvalues on imaginary axis for both p^+ and p^-, and if n is even then the same holds for one of $p^{\pm}$. Thus, by Proposition 9.1, at least one summand on the right-hand side of (9.7) is not equal to $\overline{0}$ which contradicts to (9.6) by Remark 5.2. □

9.4. *Concluding remarks*

We provided only basic facts concerning the Conley index; a more comprehensive introductions to that notion are given in [21] and [70, Chapters 22 and 23]. For some more advanced topics we refer to [50], where the index or its generalizations are defined both for continuous and discrete-time flows, and results related to attractor–repealer pairs, Morse decompositions, connection and transition matrices, singular isolating neighborhoods, and computer assisted proofs are considered. Finally, we would like to mention some extensions and modifications of the Conley index for the continuous-time local flows:

- An improvement due to Conley in which the index is a connected simple system generated by all blocks (and more generally, by all index pairs) for an isolated invariant set S and connecting maps are naturally given by pushing along trajectories (compare [21]).
- A structure of the module over $\check{H}^*(X)$ for the cohomological Conley index, where X is the phase space, and its generalization to the equivariant case due to Floer in [30].
- An infinite-dimensional version given in [68], which applies to semilinear partial differential equations with the linear parts generated by sectorial operators having compact resolvents.
- Another infinite-dimensional version defined in [36] for flows generated by vectorfields of the form $L + K$ in a Hilbert space, where L is a bounded linear Fredholm operator and K is a completely continuous nonlinear operator. It particularly well fits to results on the existence of critical points in variational problems.
- Multivalued versions given in [51] and, in a simpler case of equations without uniqueness, in [44].
- The index over a base space defined in [52] which provides a better recognition of isolated invariant sets then the index h in the case of noncontractible phase space. In particular, it is nontrivial for the invariant part of the block considered in Example 9.5 if the knotted orbit is removed from the phase space.

References

[1] F. Albrecht, *Remarques sur un theoreme de T. Ważewski relatif a l'allure asymptotique des intégrales des équations différentielles*, Bull. Acad. Polon. Sci. Cl. III **2** (1954), 315–318.

[2] H. Amann, E. Zehnder, *Nontrivial solutions for a class of nonresonance problems and applications to nonlinear differential equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **7** (4) (1980), 539–603.

[3] J.P. Aubin, *Viability Theory*, Birkhäuser, Boston (1991).

[4] J.P. Aubin, A. Cellina, *Differential Inclusions*, Springer-Verlag, Berlin (1984).

[5] I. Barbălat, *Applications du principe topologique de T. Ważewski aux équations différentielles du second ordre*, Ann. Polon. Math. **5** (1958/1959), 303–317.

[6] J. Bebernes, *Positive invariance and a Ważewski theorem*, Ordinary and Partial Differential Equations (Proc. Conf., Univ. Dundee, Dundee, 1974), Lecture Notes in Math., Vol. 415, Springer, Berlin (1974), 35–39.

[7] J.W. Bebernes, W. Kelley, *Some boundary value problems for generalized differential equations*, SIAM J. Appl. Math. **25** (1973), 16–23.

[8] J.W. Bebernes, J.D. Schuur, *The Ważewski topological method for contingent equations*, Ann. Mat. Pura Appl. (4) **87** (1970), 271–279.

[9] J.W. Bebernes, J.D. Schuur, *Investigations in the topological method of Ważewski*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8) **49** (1970), 39–42.

[10] M. Bertolino, *Framing and the asymptotic solutions of a first order differential equation*, Mat. Vesnik **2** (17) (1965), 223–231.

[11] A. Bielecki, *Remarques sur la méthode de T. Ważewski dans l'étude qualitative des équations différentielles ordinaires*, Bull. Acad. Polon. Sci. Cl. III **4** (1956), 493–495.

[12] A. Bielecki, *Extension de la methode du rétracte de T. Ważewski aux équations au paratingent*, Ann. Univ. Mariae Curie-Sklodowska Sect. A. **9** (1955), 37–61.

[13] K. Borsuk, *Sur les rétractes*, Fund. Math. **17** (1931), 171–209.

[14] K. Borsuk, *Theory of Retracts*, Monografie Matematyczne, Vol. 44, PWN—Polish Scientific Publishers, Warszawa (1967).

[15] K. Borsuk, *Theory of Shape*, Monografie Matematyczne, Vol. 59, PWN—Polish Scientific Publishers, Warszawa (1975).

[16] K. Borsuk, *On the Lusternik–Schnirelmann category in the theory of shape*, Fund. Math. **99** (1) (1978), 35–42.

[17] G. Bredon, *Topology and Geometry*, Springer-Verlag, New York (1993).

[18] R.C. Churchill, *Isolated invariant sets in compact metric spaces*, J. Differential Equations **12** (1972), 330–352.

[19] C. Conley, *On a generalization of the Morse index*, Ordinary Differential Equations (Proc. Conf., Math. Res. Center, Naval Res. Lab., Washington, DC, 1971), L. Weiss, ed., Academic Press, New York (1972), 27–33.

[20] C. Conley, *A new statement of Ważewski theorem and an example*, Ordinary and Partial Differential Equations (Proc. Fourth Conf., Univ. Dundee, Dundee, 1976), W. Everitt, ed., Lecture Notes in Math., Vol. 564, Springer-Verlag, Berlin (1976), 61–71.

[21] C. Conley, *Isolated Invariant Sets and the Morse Index*, CBMS Regional Conference Series in Mathematics, Vol. 38, Amer. Math. Soc., Providence, RI (1978).

[22] C. Conley, R. Easton, *Isolated invariant sets and isolating blocks*, Trans. Amer. Math. Soc. **158** (1971), 35–61.

[23] C.C. Conley, E. Zehnder, *The Birkhoff–Lewis fixed point theorem and a conjecture of V.I. Arnold*, Invent. Math. **73** (1983), 33–49.

[24] J. Diblík, *On the question of the applicability of Ważewski's topological method*, Soviet Math. Dokl. **31** (2) (1985), 314–317.

[25] J. Diblík, *On existence of δ-bounded solutions of a nonhomogeneous linear system of differential equations*, Funkcjalaj Ekvacioj **34** (1991), 1–18.

[26] J. Diblík, *The singular Cauchy–Nicoletti problem for the system of two ordinary differential equations*, Math. Bohem. **117** (1) (1992), 55–67.

[27] J. Diblík, *Asymptotic behaviour of solutions of linear differential equations with delay*, Ann. Polon. Math. **58** (2) (1993), 131–137.

[28] A. Dold, *Lectures on Algebraic Topology*, Springer-Verlag, Berlin (1980).

[29] R. Easton, *Isolating blocks and symbolic dynamics*, J. Differential Equations **17** (1975), 96–118.

[30] A. Floer, *A refinement of the Conley index and an application to the stability of hyperbolic invariant set*, Ergodic Theory Dynamical Systems **7** (1987), 93–103.

[31] G. Fournier, D. Lupo, M. Ramos, M. Willem, *Limit relative category and critical point theory*, Dynamics Reported. Expositions in Dynamical Systems, New Series, Vol. 3, C.K.R.T. Jones, U. Kirchgraber, H.-O. Walther, eds., Springer-Verlag, Berlin (1994), 1–24.

[32] A. Granas, J. Dugundji, *Fixed Point Theory, Vol. 1*, Monografie Matematyczne, Vol. 61, PWN—Polish Scientific Publishers, Warszawa (1982).

[33] L. Górniewicz, *Homological methods in fixed-point theory of multi-valued maps*, Dissertationes Math. (Rozprawy Mat.) **129** (1976), 1–71.

[34] D.H. Gottlieb, *A De Moivre like formula for fixed point theory*, Contemp. Math. **72** (1988), 99–105.

[35] G. Gabor, M. Quincampoix, *On existence of solutions to differential equations or inclusions remaining in a prescribed closed subset of a finite-dimensional space*, J. Differential Equations **185** (2002), 483–512.

[36] K. Gęba, M. Izydorek, A. Pruszko, *The Conley index in Hilbert spaces and its applications*, Stud. Math. **134** (3) (1999), 217–233.

[37] P. Hartman, *Invariant sets and on a theorem of Ważewski*, Proc. Amer. Math. Soc. **32** (1972), 511–520.

[38] H.M. Hastings, *A higher-dimensional Poincaré–Bendixson theorem*, Glas. Mat. Ser. III **14** (34) (2) (1979), 263–268.

[39] A.F. Izé, *Asymptotic behavior and existence of periodic solutions of a class of non linear ordinary differential equations of the second order*, An. Acad. Brasil. Cienc. **41** (1969), 297–302.

[40] L.K. Jackson, G. Klaasen, *A variation of the topological method of Ważewski*, SIAM J. Appl. Math. **20** (1971), 124–130.

[41] J. Kalas, J. Osicka, *Bounded solutions of dynamical systems in the plane under the conditions of instability*, Math. Nachr. **170** (1994), 133–147.

[42] J.L. Kaplan, A. Lasota, J.A. Yorke, *An application of the Ważewski retract method to boundary value problems*, Zeszyty Nauk. Uniw. Jagiello. Prace Mat. **16** (1974), 7–14.

[43] C. Kluczny, *Sur certaines familles de courbes en relation avec la théorie des équations différentielles ordinaires, I, II*, Ann. Univ. Mariae Curie-Sklodowska Sect. A **15** (1961) 13–40, **16** (1962) 5–18.

[44] M. Kunze, T. Küpper, Y. Li, *On Conley index theory for non-smooth dynamical systems*, Differential Integral Equations **13** (4–6) (2000), 479–502.

[45] V. Lakshmikantham, N. Onuchic, *On the comparison between the solutions of ordinary differential systems*, Bol. Soc. Mat. Sao Paulo **15** (1960) 25–34.

[46] L. Malaguti, V. Taddei, *Asymptotic properties of an ordinary differential equation via topological methods*, Atti Sem. Mat. Fis. Univ. Modena **46** (2) (1998), 479–497.

[47] R. McGehee, *Charles C. Conley, 1933–1984*, Ergodic Theory Dynamical Systems **8*** (Charles Conley Memorial Issue) (1988), 1–7.

[48] Z. Mikołajska, *Sur une propriete asymptotique des intégrales d'une équation différentielle du second ordre*, Bull. Acad. Polon. Sci. Cl. III **2** (1954), 113–116.

[49] Z. Mikołajska, *Une remarque sur les solutions bornées d'une équation différodifférentielle non linéaire*, Ann. Polon. Math. **15** (1964), 23–32.

[50] K. Mischaikow, M. Mrozek, *Conley index*, Handbook of Dynamical Systems, Vol. 2, B. Fiedler, ed., Elsevier, Amsterdam (2002), 393–460.

[51] M. Mrozek, *A cohomological index of Conley type for multi-valued admissible flows*, J. Differential Equations **84** (1) (1990), 15–51.

[52] M. Mrozek, J.F. Reineck, R. Srzednicki, *The Conley index over a base*, Trans. Amer. Math. Soc. **352** (9) (2000), 4171–4194.

[53] C. Olech, *On the asymptotic behaviour of the solutions of a system of ordinary non-linear differential equations*, Bull. Acad. Polon. Sci. Cl. III **4** (1956), 555–561.

[54] C. Olech, J. Szarski, Z. Szmydt, *Tadeusz Ważewski (1896–1972)*, Ann. Polon. Math. **29** (1974), 1–13.

[55] N. Onuchic, *Applications of the topological method of Wazewski to certain problems of asymptotic behavior in ordinary differential equations*, Pacific J. Math. **11** (1961), 1511–1527.

[56] N. Onuchic, *Relationships among the solutions of two systems of ordinary differential equations*, Michigan Math. J. **10** (1963), 129–139.

[57] N. Onuchic, *On the asymptotic behavior of the solutions of functional-differential equations*, Differential Equations and Dynamical Systems (Proc. Internat. Sympos., Mayaguez, PR, 1965), J.K. Hale, J.P. LaSalle, eds., Academic Press, New York (1967), 223–233.

[58] A. Pelczar, *On some extensions of the retract theorem of T. Ważewski*, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. **17** (1969), 693–698.

[59] A. Pelczar, *Some generalizations of the retract theorem of T. Ważewski with applications to ordinary and partial differential equations of the first order*, Ann. Polon. Math. **29** (1974), 15–59.

[60] A. Pelczar, *A local version of the generalized retract theorem of Ważewski with applications in the theory of partial differential equations of first order*, Ann. Polon. Math. **36** (1) (1979), 11–28.

[61] A. Pelczar, *Some qualitative problems in the theory of second order partial differential equations*, Ann. Polon. Math. **52** (1) (1990), 1–17.

[62] A. Pliś, *On a topological method for studying the behaviour of the integrals of ordinary differential equations*, Bull. Acad. Polon. Sci. Cl. III **2** (1954), 415–418.

[63] M. Poźniak, *Lusternik–Schnirelmann category of an isolated invariant set*, Univ. Iagel. Acta Math. **31** (1994), 129–139.

[64] C.C. Pugh, *A generalized Poincaré index formula*, Topology **7** (1968), 217–226.

[65] R. Reissig, G. Sansone, R. Conti, *Qualitative Theorie nichtlinearer Differentialgleichungen*, Edizioni Cremonese, Rome (1963).

[66] J. Radzikowski, *Proprietes asymptotiques des solutions d'un systeme de deux équations différentielles lineaires homogenes du premier ordre*, Demonstratio Math. **11** (1) (1978), 83–103.

[67] K.P. Rybakowski, *Ważewski principle for retarded functional differential equations*, J. Differential Equations **36** (1980), 117–138.

[68] K.P. Rybakowski, *The Homotopy Index and Partial Differential Equations*, Springer-Verlag, Berlin–Heidelberg (1987).

[69] J.D. Schuur, *The asymptotic behavior of a solution of the third order linear differential equation*, Proc. Amer. Math. Soc. **18** (1967), 391–393.

[70] J. Smoller, *Shock waves and reaction-diffusion equations*, Springer-Verlag, New York–Berlin (1983).

[71] E.H. Spanier, *Algebraic Topology*, McGraw–Hill, New York–Toronto–London (1966).

[72] R. Srzednicki, *On rest points of dynamical systems*, Fund. Math. **126** (1) (1985), 69–81.

[73] R. Srzednicki, *Periodic and constant solutions via topological principle of Ważewski*, Univ. Iagel. Acta Math. **26** (1987), 183–190.

[74] R. Srzednicki, *Periodic and bounded solutions in blocks for time-periodic nonautonomous ordinary differential equations*, Nonlinear Anal. **22** (6) (1994), 707–737.

[75] R. Srzednicki, *On geometric detection of periodic solutions and chaos*, Non-linear Analysis and Boundary Value Problems for Ordinary Differential Equations (Udine), F. Zanolin, ed., CISM Courses and Lectures, Vol. 371, Springer, Vienna (1996), 197–209.

[76] R. Srzednicki, *On solutions of two-point boundary value problems inside isolating segments*, Topol. Methods Nonlinear Anal. **13** (1) (1999), 73–89.

[77] R. Srzednicki, *On cohomology of the invariant part of an isolating block*, Topol. Methods Nonlinear Anal. **14** (2) (1999), 257–260.

[78] R. Srzednicki, *On periodic solutions inside isolating chains*, J. Differential Equations **165** (1) (2000), 42–60.

[79] R. Srzednicki, K. Wójcik, *A geometric method for detecting chaotic dynamics*, J. Differential Equations **135** (1) (1997), 66–82.

[80] Z. Svoboda, *Asymptotic behaviour of solutions of a delayed differential equation*, Demonstratio Math. **28** (1) (1995), 9–18.

[81] Z. Szmydtówna, *Sur l'allure asymptotique des intégrales des équations différentielles ordinaires*, Ann. Soc. Polon. Math. **24** (2) (1951), 17–34.

[82] Z. Szmydt, *Sur l'allure asymptotique des intégrales de certains systemes d'équations différentielles non lineaires*, Ann. Polon. Math. **1** (1955), 253–276.

[83] K. Tatarkiewicz, *Quelques exemples de l'allure asymptotique des solutions d'équations différentielles*, Ann. Univ. Mariae Curie-Sklodowska Sect. A **8** (1954), 105–133.

[84] B. Vrdoljak, *New possibilities for the application of the retraction method in the qualitative analysis of differential equations*, Mat. Vesnik **10** (25) (1973), 45–58.

[85] T. Ważewski, *Sur un principe topologique de l'examen de l'allure asymptotique des intégrales des équations différentielles ordinaires*, Ann. Soc. Polon. Math. **20** (1947), 279–313.

[86] T. Ważewski, *Une méthode topologique de l'examen du phénomène asymptotique relativement aux équations différentielles ordinaires*, Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat. (8) **3** (1947), 210–215.

[87] T. Ważewski, *Sur les intégrales asymptotiques des équations différentielles ordinaires*, Soc. Sci. Lett. Varsovie. C. R. Cl. III Sci. Math. Phys. **40** (1947), 38–42.

[88] T. Ważewski, *Sur certaines conditions de coincidence asymptotique des intégrales des deux systèmes d'équations différentielles*, Soc. Sci. Lett. Varsovie. C. R. Cl. III Sci. Math. Phys. **42** (1949), 198–203.

[89] T. Ważewski, *Sur une méthode topologique de l'examen de l'allure asymptotique des intégrales des équations différentielles*, Proceedings of the International Congress of Mathematicians 1954, Amsterdam, Vol. III, North-Holland, Amsterdam (1956), 132–139.

[90] F.W. Wilson, J.A. Yorke, *Lyapunov functions and isolating blocks*, J. Differential Equations **13** (1973), 106–123.

[91] K. Wójcik, *Isolating segments and symbolic dynamics*, Nonlinear Anal. **33** (6) (1998), 575–591.

[92] K. Wójcik, *On some nonautonomous chaotic system on the plane*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **9** (9) (1999), 1853–1858.

[93] K. Wójcik, *On detecting periodic solutions and chaos in the time-periodically forced ODEs*, Nonlinear Anal. **45** (1) (2001), 19–27.

[94] K. Wójcik, P. Zgliczyński, *On existence of infinitely many homoclinic solutions*, Monatsh. Math. **130** (2) (2000), 155–160.

[95] K. Wójcik, P. Zgliczyński, *Isolating segments, fixed point index, and symbolic dynamics*, J. Differential Equations **161** (2) (2000), 245–288.

[96] K. Wójcik, P. Zgliczyński, *Isolating segments, fixed point index, and symbolic dynamics, II: Homoclinic solutions*, J. Differential Equations **172** (1) (2001), 189–211.

[97] K. Wójcik, P. Zgliczyński, *Isolating segments, fixed point index, and symbolic dynamics, III: Applications*, J. Differential Equations **183** (1) (2002), 262–278.

Author Index

Roman numbers refer to pages on which the author (or his/her work) is mentioned. Italic numbers refer to reference pages. Numbers between brackets are the reference numbers. No distinction is made between first and co-author(s).

Adje, A. 93, *157* [1]
Agarwal, R.P. 3, *67* [1]; *67* [2]; *67* [3]; *67* [4]; *67* [5]; *67* [6]; 327, 338, *349* [1]; *349* [2]; *349* [3]; *349* [4]; *357* [218]; *357* [219]
Ahlbrandt, C.D. 338, *349* [5]; *349* [6]
Akô, K. 93, *157* [2]; *157* [3]
Albrecht, F. 599, *680* [1]
Alessio, F. *585* [1]
Alexander, J.C. 404, *431* [1]
Allegretto, W. 333, *349* [7]; *349* [8]; *349* [9]; *349* [10]; 364, *431* [2]
Al'mukhamedov, M.I. 481, *528* [1]
Alonso, I.P. 409, *431* [6]
Alonso, J.M. 551, *585* [2]
Amann, H. 71, 102, 114, 115, 133, 139, 155, *157* [4]; *157* [5]; *157* [6]; *157* [7]; *157* [8]; 364, *431* [3]; 651, 676, *680* [2]
Ambrosetti, A. 115, *157* [8]; 547, 576, *585* [3]; *585* [4]
Amine, Z. *585* [5]
Anane, A. 331, 332, 334–336, *349* [11]; *349* [12]; *350* [13]; 365, 370, *431* [4]
Andrade, R.F.S. 439, *528* [2]
Andres, J. *350* [14]
Andronov, A.A. 540, 541, *585* [6]
Ángel Cid, J. 93, *157* [22]
Aris, R. 3, *67* [7]
Arnold, V. 515, *528* [3]
Atkinson, F.V. *350* [15]
Aubin, J.P. 615, *681* [3]; *681* [4]
Aubin, T. 405, *431* [5]
Aubry, S. 574, *585* [7]
Aynlar, B. 322, *350* [16]
Azorero, J.G. 409, *431* [6]

Babkin, B.N. 155, *157* [9]
Badiale, M. 363, *431* [7]; 576, *585* [3]
Balibrea, F. 481, *528* [4]
Bamon, R. 480, *528* [5]
Bandle, C. 363, 405, 412, *431* [8]; *431* [9]; *431* [10]
Bangert, V. *585* [8]
Barbălat, I. 662, *681* [5]
Bartsch, T. 134, *157* [10]
Bates, P.W. 552, *585* [9]
Bautin, N.N. 452, *528* [6]
Baxley, J.V. 3, *67* [8]
Bebernes, J.W. 405, 406, 409, *431* [11]; 615, 662, *681* [6]; *681* [7]; *681* [8]; *681* [9]
Belley, J.M. 583, *585* [10]; *585* [11]; *585* [12]; *585* [13]
Bellman, R. *350* [17]
ben Naoum, A. 401, *431* [12]
Bennewitz, Ch. 334, *350* [18]
Berger, M.S. 364, *431* [13]
Bernstein, S. 93, *157* [11]
Bertolino, M. 659, *681* [10]
Besicovitch, A.S. *350* [19]; 583, *585* [14]
Bhattacharya, T. 370, *431* [14]
Bielecki, A. 599, 604, 614, *681* [11]; *681* [12]
Bihari, I. *350* [20]; *350* [21]; *350* [22]; *350* [23]
Binding, P.A. *350* [24]
Birkhoff, G.D. 562, 573, *585* [15]; *585* [16]
Blot, J. 580, 584, *585* [17]; *585* [18]; *585* [19]; *585* [20]
Bobisud, L.E. 3, *67* [9]; *67* [10]
Bognár, G. 337, *350* [25]; *350* [26]; *350* [27]; *350* [28]
Bohner, M. 349, *350* [29]; *350* [30]
Bohr, H. *350* [31]
Bongsoo Ko 115, *159* [65]
Borsuk, K. 594, 629, 630, 642, 648, 649, *681* [13]; *681* [14]; *681* [15]; *681* [16]
Borůvka, O. *350* [32]

Bosetto, E. 575, 576, *585* [21]; *585* [22]; *585* [23]
Bott, R. 564, *585* [24]
Bountis, T.C. 439, *528* [7]
Bratu, G. 362, *431* [15]
Bredon, G. 610, 624, *681* [17]
Brenig, L. *528* [8]
Brezis, H. 110, 134, *157* [12]; *157* [13]; 418, 425, *431* [16]

Cabada, A. 156, *157* [14]; *157* [15]
Cairó, L. 439, 448, 452, 486, 491, 492, 503–505, 509–511, *528* [9]; *528* [10]; *528* [11]
Calanchi, M. 560, *585* [1]; *585* [25]; *585* [26]
Callegari, A. 3, *67* [11]; *67* [12]
Calvert, J.E. 3, *67* [10]
Campillo, A. 454, *528* [12]
Campos, J. 557, *585* [27]
Cañada, A. 537, *586* [28]; *586* [29]; *586* [30]; *586* [31]
Cantrijn, F. 439, *528* [13]
Carleman, T. 439, *528* [14]
Carnicer, M.M. 454, *528* [12]; *528* [15]
Castro, A. 364, 381, 388, *431* [17]; *431* [18]; *431* [19]; 536, 551, *586* [32]
Cecchi, M. 322, 326, 327, *350* [33]; *350* [34]; *350* [35]; *350* [36]; *350* [37]; *350* [38]; *350* [39]; *350* [40]; *350* [41]
Cellina, A. 615, *681* [4]
Čepička, J. 558, *586* [33]
Cerveau, D. 454, *528* [16]
Chandrasekhar, S. 362, 405, 407, *431* [20]
Chang, K.C. 71, 134, *157* [17]; *157* [18]; *157* [19]; 549, *586* [34]
Chantladze, T. *351* [42]
Chanturia, T.A. 326, *351* [43]; *351* [44]; *354* [127]
Chaplygin, S.A. 155, *157* [16]
Charkone, O. 334, *349* [12]
Chavarriga, J. 439, 446, 448, 461, 470, 475, 487, 501, 515, 517, *528* [17]; *529* [18]; *529* [19]; *529* [20]; *529* [21]; *529* [22]; *529* [23]; *529* [24]; *529* [25]; *529* [26]
Chen, Sh. *351* [45]
Chenciner, A. *586* [35]
Cheng, S.S. *355* [157]; *355* [158]; *357* [227]
Cherpion, M. 155, 156, *157* [20]; *157* [21]
Chow, S.N. 369, *431* [21]
Christopher, C.J. 439, 446, 448, 455, 457, 478, 485, 487, 491, 492, 500, 502, 503, 510, *529* [27]; *529* [28]; *529* [29]; *529* [30]; *529* [31]; *529* [32]; *529* [33]; *529* [34]; *529* [35]; *530* [61]; *531* [84]
Churchill, R.C. 600, 643, 644, 669, *681* [18]
Cieutat, P. 580, *585* [19]
Clément, P. 362, 380, 409, 410, 418, 419, *431* [22]; *431* [23]
Coffman, C.V. 363, 425, *431* [9]; *431* [24]; *431* [25]
Coles, W.J. *351* [46]
Collatz, L. 156, *157* [23]
Conley, C.C. 599, 600, 606, 608, 651, 673, 678–680, *681* [19]; *681* [20]; *681* [21]; *681* [22]; *681* [23]
Conti, M. 134, *157* [24]
Conti, R. 659, *683* [65]
Coppel, W.A. 489, *529* [36]
Courant, R. 155, *157* [25]
Cozma, D. 487, *529* [37]
Crandall, M.G. 409, *432* [26]
Cuesta, M. 334, *351* [47]; *351* [48]

Damascelli, L. 428, *432* [27]; *432* [28]
Dancer, E.N. 134, *157* [26]; 536, 544, 546, 557, 559, *585* [27]; *586* [36]; *586* [37]
Dang, H. 363, 364, 371, 379, 391, 393, 400, 401, *432* [29]; *432* [30]; *432* [31]
Darboux, G. 439, 448, 460, 466, 515, *529* [38]
De Coster, C. 72, 80, 90, 91, 93, 94, 101, 115, 155, 156, *157* [20]; *157* [21]; *157* [27]; *157* [28]; *157* [29]; *158* [30]; *158* [31]; *158* [32]; *158* [33]; *158* [34]; *158* [35]; *158* [36]; 401, *431* [12]
de Figueiredo, D.G. 71, 133, 134, *158* [37]; *158* [38]; 334, *351* [47]; *351* [48]; 362, 409, 410, 418, 419, *431* [22]
de Thélin, F. 316, *353* [101]; 381, *432* [37]
de Valleé Poussin, Ch. *357* [214]
Debnath, L. 322, *354* [144]
Deimling, K. 13, *67* [13]; 364, 369, *432* [32]
Del Pino, M.A. 313, 316–320, *351* [49]; *351* [50]; *351* [51]; *351* [52]; *351* [53]; 365, 369–371, 400, *432* [33]; *432* [34]
Delanoë, P. 363, *432* [35]
Denzler, J. 566, 567, *586* [38]
Díaz, J.I. *351* [54]
Diblík, J. 618, 659, 662, *681* [24]; *681* [25]; *681* [26]; *681* [27]
Ding, T. 119, *158* [39]; 478, *532* [105]
Dold, A. 610, 630, 632, 634, 635, 647, *681* [28]
Donati, F. 554, 558, 560, *586* [40]; *586* [41]
Dong Zhenxi 478, *532* [105]
Dorizzi, B. 439, *528* [7]
Došlá, Z. 322, 326, 327, *350* [33]; *350* [34]; *350* [35]; *350* [36]; *350* [37]
Došlý, O. 334, 337, 342, 345, *350* [28]; *351* [55]; *351* [56]; *351* [57]; *351* [58]; *351* [59]; *351* [60]; *351* [61]; *351* [62]; *351* [63];

351 [64]; *351* [65]; *351* [66]; *351* [67]; *351* [68]; *352* [69]; *352* [70]; *352* [71]; *352* [72]
Drábek, P. 313–318, 321, 323, 334, *350* [24]; *351* [49]; *352* [73]; *352* [74]; *352* [75]; *352* [76]; *352* [77]; *352* [78]; *352* [79]; *352* [80]; *352* [81]; *352* [82]; 365, *432* [36]; 558, *586* [33]
Du, Y. 134, *157* [26]
Dugundji, J. 638, *682* [32]
Dunninger, D.R. 333, *352* [83]

Easton, R. 600, 667, *681* [22]; *681* [29]
Eberhard, W. *352* [84]
Eberly, D. 405, 406, 409, *431* [11]
Écalle, J. 480, 517, *529* [39]
el Hachimi, A. 381, *432* [37]
Elbert, Á. *350* [15]; *351* [65]; *351* [66]; *352* [84]; *352* [85]; *352* [86]; *352* [87]; *352* [88]; *352* [89]; *352* [90]; *352* [91]; *352* [92]; *352* [93]; *353* [94]
Elgueta, M. 313, 316, 319, 320, *351* [51]; *351* [52]; 400, *432* [33]
Eloe, P.W. 3, *67* [14]
Emden, R. *353* [95]
Emden, V. 405, 407, *432* [38]
Epheser, H. 93, *158* [40]
Erbe, L.H. 349, *351* [45]; *353* [96]; *353* [97]; 401, *432* [39]
Evdokimenco, R.M. 475, *529* [40]; *529* [41]; *529* [42]
Evtuchov, V.M. *353* [98]

Fabry, C. 93, *158* [41]
Fan, X. *353* [99]
Feix, M.R. 439, 448, 452, 486, 491, 509, *528* [9]
Filiptsov, V.F. 475, *529* [43]
Fink, A.M. 582, 583, *586* [43]; *586* [44]
Fleckinger, J. 316, *353* [100]; *353* [101]
Floer, A. 645, 647, 680, *681* [30]
Fonda, A. 562, 563, 565, *586* [45]; *586* [46]
Forni, G. *587* [73]
Fournier, G. 544, 579, 583, 584, *585* [10]; *585* [11]; *585* [12]; *586* [47]; *586* [48]; 630–633, *682* [31]
Fowler, R.H. *353* [102]
Frank-Kamenetskii, D.A. 362, 405, 407, *432* [40]
Friedlander, L. 370, *432* [41]
Frommer, M. 453, *529* [44]
Fučík, S. 314, 332, *353* [103]; 536, *586* [49]
Fulton, W. 462, *530* [45]
Furta, S.D. 456, 474, 492, *530* [46]

Gabor, G. 615, 616, *682* [35]
Gaines, R.E. 543, 544, *586* [50]
Gallouët, T. 409, *432* [42]
García Azorero, J. 362, 370, 409, 425, *432* [43]; *432* [44]
García-Huidobro, M. 363–365, 371, 379–381, 383, 385, 387, *431* [23]; *432* [45]; *432* [46]; *432* [47]; *432* [48]
Gatica, J.A. 3, *68* [15]
Gaudenzi, M. 115, *158* [42]
Gęba, K. 680, *682* [36]
Gelfand, I.M. 362, 405, 407, *432* [49]
Gendzhoyan, G.V. 156, *158* [43]
Ghoussoub, N. 547, *586* [51]
Giacomini, H.J. 439, 446, 475, 482, 487, 492, 501, 503–505, 510, 511, 515, 517, *528* [10]; *529* [18]; *529* [19]; *529* [20]; *529* [21]; *530* [47]; *530* [48]; *530* [49]; *530* [50]; *530* [51]
Gidas, B. 361, 362, 425, 428, *432* [50]
Giné, J. 439, 487, 501, 515, 517, *529* [18]; *529* [19]; *529* [20]
Girg, P. 313, *352* [76]; *352* [77]; 537, *586* [52]
Goriely, A. 505, *528* [8]; *530* [52]
Górniewicz, L. 615, 629, *682* [33]
Gossez, J.-P. 101, 115, *158* [44]; *158* [45]; 334, *349* [12]; *351* [47]; *351* [48]; *353* [100]
Gottlieb, D.H. 652, *682* [34]
Grace, S.R. *349* [2]; *349* [3]
Grammaticos, B. 439, 505, *528* [7]; *530* [53]
Granas, A. 638, *682* [32]
Grossinho, M.R. 93, 115, *157* [29]; *353* [104]
Gudkov, V.V. 93, *158* [46]
Guillou, L. 554, *586* [53]
Gustafson, G.B. 393, 396, *432* [51]
Gustafson, K. 405, *432* [52]
Gutierrez, C. *530* [54]

Habets, P. 3, 4, *68* [16]; 72, 80, 90, 91, 93, 94, 101, 115, 134, 155, 156, *157* [14]; *157* [20]; *157* [21]; *157* [29]; *158* [30]; *158* [31]; *158* [32]; *158* [33]; *158* [41]; *158* [42]; *158* [47]; *158* [48]; *158* [49]; *158* [50]; *158* [51]; *158* [52]; *158* [53]; *159* [54]; 537, *586* [54]
Hai, D. 363, 388, *433* [53]; *433* [54]
Hai Huang 570, 573, 584, *587* [56]; *587* [57]; *587* [58]; *587* [59]
Hale, J.K. 369, *431* [21]
Hamel, G. 535, 546, *586* [55]
Hardy, G.H. *353* [105]
Hartman, P. *353* [106]; 368, *433* [55]; 616, 617, *682* [37]
Hasselblatt, B. *587* [62]
Hastings, H.M. 642, *682* [38]
Hatvani, L. *353* [107]; *353* [108]
Hayes, J. *585* [10]
Heinonen, J. 362, *433* [56]
Henderson, J. 3, *67* [14]

Henrard, M. 115, *158* [34]
Hernández, J. 316, *353* [101]
Hess, P. 93, *159* [55]
Hietarinta, J. 439, 480, *530* [55]
Hilbert, D. 155, *157* [25]; 439, *530* [56]
Hilger, S. 349, *353* [97]; *353* [109]
Hille, E. *353* [110]
Hinton, D.B. *349* [5]
Hofer, H. 133, 134, *159* [56]; *159* [57]
Hong, H.L. *353* [111]
Hoshino, H. *353* [112]
Huang, Ch. *353* [113]
Huang, Q. *351* [45]
Huang, Y.X. 333, *349* [9]; *349* [10]; *350* [24]; *353* [114]
Huang Wenzao 478, *532* [105]

Iannacci, R. 119, *158* [39]
Ilyashenko, Yu. 480, 517, *530* [57]
Imabayashi, R. *353* [112]
Izé, A.F. 659, *682* [39]
Izydorek, M. 680, *682* [36]

Jackson, L.K. 93, *159* [58]; 615, 662, *682* [40]
Jacobsen, J. 362, 409–411, 417, *433* [57]; *433* [58]; *433* [59]
Jäger, W. 425, *433* [60]
Jaroš, J. 333, *351* [67]; *353* [115]; *353* [116]; *353* [117]; *353* [118]
Jenšiková, J. 558, *586* [33]
Jianchu, J. *353* [119]
Jiangong You 572, *589* [130]; *589* [131]; *589* [132]; *589* [133]; *589* [134]
Jimenez, V. 481, *528* [4]
Jingfa, W. 327, *353* [120]; *353* [121]
Jingxian Sun 134, *159* [68]
Joseph, D.D. 362, *433* [61]
Jouanolou, J.P. 439, 446, 461, 492, *530* [58]

Kalas, J. 659, *682* [41]
Kandelaki, N. *351* [42]; *354* [122]
Kannan, R. 544, 557, *587* [60]; *587* [61]
Kantorovich, L. 156, *159* [59]
Kaper, H. 393, 400, *433* [62]
Kaplan, J.L. 662, *682* [42]
Kapteyn, W. 452, *530* [59]; *530* [60]
Karsai, J. *354* [123]
Katok, A. *587* [62]
Katriel, G. 536, *587* [63]; *587* [64]
Kazdan, J.L. 405, *433* [63]
Kelley, J.L. 78, *159* [60]
Kelley, W.G. 338, *354* [124]; 615, 662, *681* [7]
Khaikin, S.E. 540, 541, *585* [6]
Kiguradze, I.T. 93, *159* [61]; *159* [62]; *159* [63]; *354* [125]; *354* [126]; *354* [127]; *354* [128]
Kilpeläinen, T. 362, *433* [56]
Kiomura, J. *354* [129]
Kitano, M. 322, *354* [132]
Klaasen, G. 615, 662, *682* [40]
Kluczny, C. 615, *682* [43]
Knaap, M. 393, 400, *433* [62]
Knobloch, H.W. 93, *159* [64]
Kolesov, Y.S. 115, *159* [66]
Kong, Q. *354* [130]
Kooij, R. 491, 492, *530* [61]
Krasnosel'skii, M. 369, 396, *433* [64]; *433* [65]
Kufner, A. *354* [131]
Kunze, M. 680, *682* [44]
Küpper, T. 680, *682* [44]
Kurepa, A. 381, *431* [17]; *431* [18]
Kusano, T. 322, 327, 333, 334, *352* [90]; *352* [91]; *352* [92]; *352* [93]; *353* [112]; *353* [115]; *353* [116]; *353* [117]; *353* [118]; *354* [129]; *354* [132]; *354* [133]; *354* [134]; *354* [135]; *354* [136]; *354* [137]; *354* [138]; *354* [139]; *354* [140]; *354* [141]; *354* [142]
Kvinikadze, G. 326, *354* [128]; *354* [143]
Kwong, M. 393, 400, *433* [62]

Lakshmikantham, V. 3, *67* [4]; *67* [5]; 659, *682* [45]
Laloy, M. 93, *158* [47]
Landesman, E.M. 321, *354* [145]
Lasota, A. 662, *682* [42]
Lax, P.D. 439, *530* [62]
Lazer, A.C. 321, *354* [145]
Le, V. 391, 401, *433* [66]
Le Daeron, P.Y. 574, *585* [7]
Leela, S. 3, *67* [4]
Leighton, W. *354* [146]; *354* [147]; *355* [148]
Lepin, A.J. 93, *158* [46]
Levi, M. 571, 572, *587* [65]; *587* [66]; *587* [67]
Lewis, R.T. *349* [5]
Li, H.J. *355* [149]; *355* [150]; *355* [151]; *355* [152]; *355* [153]; *355* [154]; *355* [155]
Li, S.J. 134, *159* [67]
Li, W.-T. *353* [99]; *355* [156]; *355* [157]; *355* [158]
Li, Y. 680, *682* [44]
Lian, W.Ch. *353* [111]
Lin, S.S. 363, 425, *433* [67]; *433* [68]
Lindqvist, P. 330, 331, *355* [159]; *355* [160]; *355* [161]; 365, *433* [69]
Lins Neto, A. 454, 455, *528* [16]; *530* [65]
Lions, P.L. 417, *433* [70]
Liouville, J. 362, *433* [71]
Littlewood, J.E. *353* [105]

Llibre, J. 439, 446, 448, 452, 455–457, 461, 470, 475, 478, 480–482, 485–487, 491, 492, 500, 501, 503–505, 509–511, 515, 517, 518, 528, *528* [9]; *528* [10]; *528* [11]; *529* [19]; *529* [20]; *529* [21]; *529* [22]; *529* [23]; *529* [24]; *529* [25]; *529* [26]; *529* [30]; *529* [31]; *529* [32]; *529* [33]; *529* [34]; *529* [35]; *530* [47]; *530* [48]; *530* [49]; *530* [50]; *530* [54]; *530* [63]; *530* [64]; *530* [66]; *530* [67]; *530* [68]; *530* [69]; *530* [70]; *531* [71]
Lloyd, N.G. 86, *159* [69]; 364, *433* [72]; 455, *531* [84]
Lomtatidze, A. *351* [42]; *351* [68]; *352* [69]; *354* [122]; *355* [162]
Lotka, A.J. 442, 504, *531* [72]
Lundgren, T.S. 362, *433* [61]
Luning, C.D. 3, *68* [17]
Luo, J. 322, *354* [144]
Lupo, D. 630–633, *682* [31]

Maccallum, M.A.H. 455, *531* [73]
Malaguti, L. 659, *682* [46]
Manásevich, R.F. 313, 316–320, 323, *351* [49]; *351* [50]; *351* [51]; *351* [52]; *351* [53]; *352* [76]; *352* [78]; *355* [164]; *355* [165]; *355* [166]; *355* [167]; *355* [168]; 362–365, 369–371, 379–381, 383, 385, 387, 393, 395, 398–400, *431* [23]; *432* [29]; *432* [33]; *432* [34]; *432* [36]; *432* [45]; *432* [46]; *432* [47]; *432* [48]; *433* [73]; *433* [74]
Mancini, G. 115, *157* [8]
Manojlović, J.V. *355* [163]
Marcus, M. 363, *431* [9]; *431* [25]
Mařík, R. 334, *352* [70]; *355* [169]; *355* [170]; *355* [171]
Marini, M. 322, 326, 327, *350* [33]; *350* [34]; *350* [35]; *350* [36]; *350* [37]; *350* [38]; *350* [39]; *350* [40]; *350* [41]; *355* [172]; *355* [173]; *355* [174]
Markus, L. 517, *531* [74]
Martinez-Amores, P. 554, 556, 565, *587* [68]
Martio, O. 362, *433* [56]
Mather, J.N. 566, 574, *587* [69]; *587* [70]; *587* [71]; *587* [72]; *587* [73]
Matucci, S. *356* [179]
Mawhin, J. 94, *159* [70]; 313, *355* [164]; 422, *433* [75]; *433* [76]; 536, 537, 543, 544, 547, 549, 551, 552, 554, 556, 563–566, 578–580, 584, *585* [19]; *586* [47]; *586* [50]; *587* [68]; *587* [74]; *587* [75]; *587* [76]; *587* [77]; *587* [78]; *587* [79]; *587* [80]; *587* [81]; *587* [82]; *588* [83]; *588* [84]; *588* [85]; *588* [86]
McGehee, R. 602, *682* [47]
McGough, J.S. 412, 413, *433* [77]
Merizzi, L. 134, *157* [24]
Mignot, F. 409, *432* [42]; *433* [78]; *433* [79]
Mikołajska, Z. 621, 659, *682* [48]; *682* [49]
Milloux, H. *355* [175]
Mirzov, J.D. 322, 326, *355* [176]; *356* [177]; *356* [178]
Mischaikow, K. 593, 667, 671, 680, *682* [50]
Mitidieri, E. 362, 409, 410, 418, 419, *431* [22]
Mizukami, M. 327, *356* [180]
Mlak, W. 93, *159* [71]
Morse, M. *355* [148]
Moser, J. 566, 571, 572, *588* [87]; *588* [88]; *588* [89]; *588* [90]; *588* [91]; *588* [92]; *588* [93]; *588* [94]; *588* [95]
Moulin Ollagnier, J. 446, 457, 461, 492, 503–505, *529* [24]; *530* [53]; *531* [75]; *531* [76]; *531* [77]; *531* [78]; *531* [79]
Mrozek, M. 593, 667, 671, 680, *682* [50]; *682* [51]; *682* [52]
Müller, M. 71, *159* [72]

Nabana, E. 363, *431* [7]
Nachman, A. 3, *67* [11]; *67* [12]
Nagabuchi, Y. *356* [181]
Nagasaki, K. 363, 409, 425, 427, *434* [80]; *434* [81]; *434* [82]
Nagumo, M. 93, *159* [73]; *159* [74]
Naito, M. 327, *352* [91]; *352* [92]; *354* [129]; *354* [133]; *354* [134]; *354* [137]; *356* [180]; *356* [182]
Naito, Y. 334, *354* [135]; *354* [136]; *356* [182]
Nečas, J. 314, 332, *353* [103]
Nehari, Z. 428, 429, *434* [83]
Nemytskii, V.V. *531* [80]
Neuman, F. *356* [183]
Neumann, D.A. 517, *531* [81]
Nhashama, M.N. 3, *68* [18]
Ni, W.M. 361, 362, 425, 428, 430, *432* [50]; *434* [84]
Nicolau, M. 515, 517, *530* [63]
Nirenberg, L. 134, *157* [13]; 361, 362, 418, 425, 428, *431* [16]; *432* [50]
Nistri, P. 364, *431* [2]
Njoku, F.I. 313, *355* [165]; 393, *433* [73]
Nowicki, A. 446, *531* [79]

Odani, K. 446, *531* [82]
Offin, D. 563, 576, *588* [96]; *588* [97]
Ogata, A. 334, *354* [136]
Ohriska, J. *356* [184]
Olech, C. *356* [185]; 602, 659, *682* [53]; *682* [54]
Oliker, V. 3, *68* [15]
Olver, P.J. 439, *531* [83]

Omari, P. 93, 115, 134, 156, *158* [35]; *158* [45]; *158* [48]; *158* [49]; *159* [75]; *159* [76]; *353* [104]
Onuchic, N. 621, 659, *682* [45]; *682* [55]; *682* [56]; *682* [57]
Opial, Z. *356* [185]; 578, *588* [98]
O'Regan, D. 3, 43, 48, 50, *67* [1]; *67* [2]; *67* [3]; *67* [4]; *67* [5]; *67* [6]; *68* [19]; *349* [3]
Ortega, R. 544, 551, 552, 554, 556–561, 565, 566, 579, 584, *585* [5]; *585* [27]; *586* [37]; *587* [60]; *587* [61]; *587* [68]; *588* [99]; *588* [100]; *588* [101]; *588* [102]; *588* [103]; *588* [104]; *588* [105]; *588* [106]; *588* [107]; *588* [108]; *588* [109]
Osicka, J. 659, *682* [41]
Otaga, A. 327, *354* [138]; *354* [139]; *354* [140]
Otani, M. *352* [79]; *356* [186]; *356* [187]

Pacella, F. 428, *432* [27]; *432* [28]
Pantazi, C. 491, 492, *529* [33]
Parasyuk, I.O. *589* [135]
Peano, G. 93, *159* [77]
Pearson, J.M. 455, *531* [84]
Peitgen, H.-O. 420, *434* [85]; *434* [86]; *434* [87]
Pelczar, A. 599, 606, 659, *683* [58]; *683* [59]; *683* [60]; *683* [61]
Peletier, L.A. 363, 430, *431* [10]; *434* [88]
Peña, S. *356* [188]
Pennequin, D. 584, *585* [20]
Peral Alonso, I. 362, 370, 409, 425, *432* [43]; *432* [44]
Percival, I.C. *589* [110]; *589* [111]
Pereira, J.V. 439, 448, 502, 503, 510, *529* [34]; *531* [85]
Pérez del Río, J.S. 491, *530* [66]
Perron, O. 71, 93, *159* [78]
Perry, W.L. 3, *68* [17]
Persson, L.-E. *354* [131]
Peterson, A.C. 338, 349, *349* [6]; *350* [29]; *350* [30]; *354* [124]
Philos, Ch. *356* [189]
Picard, E. 71, 93, 154, 155, *159* [79]; *159* [80]; *160* [81]; *160* [82]; *160* [83]; *160* [84]
Picone, M. *356* [190]
Piros, M. *356* [191]
Pliś, A. 599, 606, 608, *683* [62]
Pohozaev, S. 412, 429, *434* [89]
Poincaré, H. 439, 454, 456, *531* [86]; 562, *589* [112]
Pólya, G. *353* [105]
Poppenberg, M. 428, 430, *434* [90]
Pouso, R. 80, 94, 156, *157* [14]; *158* [50]
Poźniak, M. 636, 642, *683* [63]
Preiss, D. 547, *586* [51]
Prelle, M.J. 439, 501, *531* [87]
Prodi, G. 94, *160* [86]
Protter, M.H. 139, 155, *160* [85]
Pruszko, A. 680, *682* [36]
Pucci, P. 414–416, 419, *434* [91]; *434* [92]; 547, *589* [113]
Puel, J.-P. 409, *431* [6]; *432* [42]; *433* [78]; *433* [79]
Pugh, C.C. 652, *683* [64]

Qin Yuan-Xun 475, 478, *531* [88]
Quincampoix, M. 615, 616, *682* [35]

Ráb, M. *356* [192]
Rabinowitz, P.H. 369, 409, 420, *432* [26]; *434* [93]; *434* [94]; 547, 549, 574, *585* [4]; *589* [114]; *589* [115]
Rabtsevich, V.A. *356* [193]
Rachùnková, I. 115, *160* [87]
Radzikowski, J. 659, *683* [66]
Ramani, A. 439, 505, *528* [7]; *530* [53]
Ramos, M. 630–633, *682* [31]
Rao, M. 365, *434* [95]
Rauh, A. 439, *528* [2]
Řehák, P. 338, 340, 345, 349, *352* [71]; *356* [194]; *356* [195]; *356* [196]; *356* [197]; *356* [198]; *356* [199]; *356* [200]; *356* [201]
Reichel, W. *356* [202]
Reineck, J.F. 680, *682* [52]
Reissig, R. 659, *683* [65]
Ren, Z. 365, *434* [95]
Repetto, C.E. 439, *530* [51]
Řezníčková, J. *352* [72]
Robinson, B. 313–315, 321, 334, *352* [80]; *352* [81]
Roca, F. 537, *586* [52]
Rockafeller, R.T. *356* [203]
Rodríguez, G. 480, 481, 518, *530* [67]; *530* [68]; *530* [69]
Rodríguez, J.A. 491, *530* [66]
Rogovchenko, S.P. 322, 327, *356* [205]
Rogovchenko, Y.V. 322, 327, *356* [204]; *356* [205]
Rong Yuan *587* [59]
Roselli, P. 537, *589* [116]
Rosenblatt, A. 94, *160* [88]
Rousseau, C. 486, *531* [89]
Royalty, W.D. 3, *67* [10]
Ruiz, D. 537, *586* [28]; *586* [29]
Rybakowski, K.P. 621–625, 674, 680, *683* [67]; *683* [68]

Saadi Drissi, K. 583, *585* [11]; *585* [12]
Sagdeev, R.Z. 536, *589* [117]

Saitö, Y. 334, *350* [18]
Sanchez, L. 93, 156, *157* [15]; *158* [51]
Sansone, G. *357* [206]; 659, *683* [65]
Sarlet, W. 439, *528* [13]
Satō, T. 93, *160* [89]; *160* [90]
Sattinger, D.H. 115, *160* [91]; 405, *434* [96]
Schaaf, R. 420, 422, 424, *434* [97]; *434* [98]; 537, *589* [118]
Schecter, S. 481, *531* [90]
Schlomiuk, D. 439, 453, 454, 486, 503, *530* [70]; *531* [89]; *531* [91]; *531* [92]; *531* [93]
Schmitt, K. 93, *160* [92]; 362–365, 371, 379–381, 383, 385, 387, 388, 391, 393, 395, 396, 398–401, 410–413, 420, 422, 424, 425, 428, 430, *431* [23]; *432* [29]; *432* [30]; *432* [31]; *432* [39]; *432* [45]; *432* [46]; *432* [47]; *432* [51]; *433* [53]; *433* [54]; *433* [59]; *433* [60]; *433* [66]; *433* [74]; *433* [75]; *433* [76]; *434* [85]; *434* [86]; *434* [87]; *434* [90]; *434* [97]; *434* [98]; *434* [99]; *434* [100]; *434* [101]; 537, *589* [118]
Schneider, A. *353* [94]
Schröder, J. 156, *157* [23]; *160* [93]
Schuur, J.D. 615, 659, *681* [8]; *681* [9]; *683* [69]
Scorza Dragoni, G. 71, 92, 93, *160* [94]; *160* [95]; *160* [96]; *160* [97]
Sedziwy, S. 313, *355* [166]
Serra, E. 550, 552, 554, 556, 557, 565, 574–576, *585* [1]; *585* [21]; *585* [22]; *585* [23]; *585* [25]; *588* [107]; *589* [119]; *589* [120]; *589* [121]; *589* [122]
Serrin, J. 414–416, 419, 428, 430, *434* [84]; *434* [88]; *434* [91]; *434* [92]; *434* [102]; 547, *589* [113]
Shapiro, V.L. 370, *434* [103]
Shekhter, B.L. 93, *159* [63]
Shieh, S.L. *349* [4]
Shivaji, R. 115, *160* [98]; 363, 364, 388, 391, 401, *431* [19]; *432* [31]; *433* [54]
Shubé, A.S. 487, *531* [94]
Silva, E.A.D.B. 409, *434* [104]
Singer, M.F. 439, 481, 501–504, *531* [87]; *531* [90]; *531* [95]
Smoller, J. 425, *434* [105]; 669, 673, 679, 680, *683* [70]
Soares, S.H.M. 409, *434* [104]
Solimini, S. 71, 133, 134, *158* [38]
Sorolla, J. 475, *529* [25]
Sotomayor, J. 439, 448, 515, 516, *529* [26]; *531* [96]
Souček, J. 314, 332, *353* [103]
Souček, V. 314, 332, *353* [103]
Spanier, E.H. 627, 642, *683* [71]
Srikanth, P.N. 427, *434* [106]
Srzednicki, R. 637, 645, 652, 654–658, 662, 665–667, 669, 680, *682* [52]; *683* [72]; *683* [73]; *683* [74]; *683* [75]; *683* [76]; *683* [77]; *683* [78]; *683* [79]
Stampacchia, G. 93, *160* [99]
Stepanov, V.V. *531* [80]
Strelcyn, J.M. 439, 446, 505, *530* [53]; *531* [79]; *531* [97]
Struwe, M. *357* [207]
Şubă, A. 487, *529* [37]
Suzuki, T. 363, 409, 425, 427, *434* [80]; *434* [81]; *434* [82]; *435* [107]
Sverdlove, R. 481, *531* [98]
Svoboda, Z. 659, *683* [80]
Swanson, C.A. *357* [208]
Świrszcz, G. 478, *529* [35]
Szarski, J. 71, *160* [100]; 602, *682* [54]
Szmydt, Z. 602, 659, *682* [54]; *683* [82]
Szmydtówna, Z. 659, *683* [81]
Szulkin, A. 579, 584, *586* [48]

Taddei, V. 659, *682* [46]
Takáč, P. 313, 316, *352* [77]; *352* [82]; *353* [100]; *353* [101]; *355* [167]
Talenti, G. *357* [209]
Taliaferro, S. 3, *68* [20]
Tanigawa, T. 327, *352* [93]; *353* [112]; *353* [117]; *354* [137]; *357* [210]
Tarallo, M. 115, *158* [36]; 550, 552, 554, 556, 557, 560, 565, 574, 584, *585* [26]; *588* [107]; *588* [108]; *588* [109]; *589* [119]; *589* [120]; *589* [121]; *589* [122]
Tarantello, G. 558, *589* [123]
Tatarkiewicz, K. 659, *683* [83]
Terracini, S. 134, *157* [24]; 554, 556, 565, 574, 576, *585* [22]; *585* [23]; *589* [121]; *589* [122]
Thomas, L.H. *357* [211]
Tineo, A. 3, *68* [21]; 577, *589* [124]
Tipler, F.J. *357* [212]
Tiryaki, A. 322, *350* [16]
Torres, P.J. 94, 115, *158* [52]; 537, *586* [54]
Tricomi, F. 536, 541, *589* [125]; *589* [126]
Trombetta, M. 156, *159* [75]
Trudinger, N.S. 330, *357* [213]; 362, *435* [108]
Tso, K. 416, 417, *435* [109]
Tsouli, N. 335, 336, *350* [13]
Tvrdý, M. 115, *160* [87]

Ugulava, D. *354* [122]
Ulm, M. 313, *352* [77]
Ureña, A.J. 537, 561, 562, *586* [30]; *586* [31]; *589* [127]

Usami, H. 327, *354* [140]; *356* [180]; *356* [182]
Usikov, P.A. 536, *589* [117]

Vaillancourt, F. *585* [13]
Valeeva, R.T. 481, *531* [99]
Viano, M. 482, *530* [47]; *530* [48]; *530* [49]; *530* [50]
Villari, G. *350* [38]; *350* [39]; *350* [40]; *350* [41]
Vitt, A.A. 540, 541, *585* [6]
Volterra, V. 442, 504, *532* [100]
Vrdoljak, B. 659, *683* [84]

Walcher, S. 454, *532* [101]
Walter, W. 71, *160* [101]; *356* [202]
Waltman, P. 3, *68* [15]
Wang, J. *354* [142]
Wang, Q.-R. *357* [216]
Wang, T. 134, *159* [67]
Wang, X.-J. 362, 417, *435* [108]; *435* [110]
Wang, Z.-Q. 134, *157* [10]; 420, 428, 430, *434* [90]; *434* [101]
Ward, J.R., Jr. 537, *589* [128]
Warner, F.W. 405, *433* [63]
Wasserman, A. 425, *434* [105]
Ważewski, T. *356* [185]; 595, 596, 599, 606, 607, 611, 612, 614, 659, 662, *683* [85]; *683* [86]; *684* [87]; *684* [88]; *684* [89]
Weigu Li 456, 515, 517, *530* [63]; *530* [64]
Weil, J.A. *532* [102]
Weinberger, H.F. 139, 155, *160* [85]
Weiyue Ding 562, *586* [39]
Willem, M. 428, 429, *435* [111]; 536, 544, 546, 547, 549, 551, 552, 554, 556, 563–565, 579, 584, *586* [45]; *586* [48]; *587* [68]; *588* [85]; *588* [86]; *589* [129]; 630–633, *682* [31]
Willet, D. *357* [217]
Wilson, F.W. 668, *684* [90]
Wojciechowski, S. 439, 505, *530* [53]; *531* [97]
Wójcik, K. 665–667, *683* [79]; *684* [91]; *684* [92]; *684* [93]; *684* [94]; *684* [95]; *684* [96]; *684* [97]
Wong, J.S. *357* [215]
Wong, P.J.Y. 3, *67* [6]; 327, *357* [218]; *357* [219]

Xiang Zhang 456, *530* [64]

Yablonskii, A.I. 475, *532* [103]
Yamamoto, M. *356* [181]
Yan, J. *357* [220]; *357* [221]
Yang, X. *357* [222]; *357* [223]; *357* [224]; *357* [225]
Ye Yanqian 475, 478, 488, *532* [104]
Yeh, C.C. *349* [4]; *353* [111]; *355* [150]; *355* [151]; *355* [152]; *355* [153]; *355* [154]; *355* [155]; *357* [226]
Yihong Du 547, *586* [42]
Yiu-Kwong Man 455, *531* [73]
Yorke, J.A. 662, 668, *682* [42]; *684* [90]
Yoshida, N. 327, 333, *353* [118]; *354* [141]
Yu Hongan 576, *588* [97]

Zakharin, S.F. *589* [135]
Zandron, O.P. 439, *530* [51]
Zanolin, F. 3, 4, *68* [16]; 90, 91, 94, 115, 119, 134, *158* [39]; *158* [42]; *158* [53]; *159* [54]; *159* [76]; 313, *355* [165]; *355* [168]; 371, 379, 380, 393, *432* [48]; *433* [73]; 562, 565, *586* [46]
Zaslavsky, G.M. 536, *589* [117]
Zecca, P. 364, *431* [2]
Zehnder, E. 572, *587* [67]; 651, 676, *680* [2]; *681* [23]
Zezza, P. *355* [174]
Zgliczyński, P. 666, 667, *684* [94]; *684* [95]; *684* [96]; *684* [97]
Zhang, G. *357* [227]
Zhang, X. 491, 492, 515, 517, 528, *529* [33]; *530* [63]; *531* [71]
Zhang Zhifen 478, *532* [105]
Zhaoli Liu 134, *159* [68]
Zhong, C. *353* [99]
Żołądek, H. 491, 502, *532* [106]; *532* [107]
Zou, H. 428, 430, *434* [102]

Subject Index

2-torus 523
α-limit 516
ω-limit 516

$\mathcal{A}$-Carathéodory function 90
absolute 631
action functional 546
action integral 535
almost periodic coefficient 270
almost periodic function 582, 583
annular domain 361, 364, 392
annular system 516
ANR (absolute neighborhood retract) 629, 630, 632, 634–638, 642, 648–650, 654
Armellini–Tonelli–Sansone theorem 279, 281
Arrhenius equation 405
asymptotic part 604, 636, 638, 639
asymptotic properties 215
Aubry–Mather's twist theorem 574
average 577
averaging technique 256

Banach space 372, 420
base point 627, 629, 677
Besicovitch almost periodic function 271
bifurcation 364
bifurcation equation 543
bifurcation from infinity 420
bifurcation of solutions 369
bifurcation point 369
bifurcation problems 369
bifurcation results 419
Birkhoff point 573
Birkhoff's twist theorem 573
blow-up 405
Bott's iteration formula 564
boundary value problem 1, 3, 8, 21, 41, 42, 47, 361, 363, 396, 399
bounded function 577
bounded Palais–Smale condition (BPS) 547
bounded solution 577, 578

C^k-equivalent systems 516
C^k-parallel systems 516
Calabi invariant 572
canonical region 517
Carathéodory conditions 72
Carathéodory functions 72
category 594, 601, 602, 630, 631, 633, 636, 642, 648, 674
category, relative 630, 650
center 485
chaotic dynamics 574
chemical kinetics 405
circular cylinder 522
circular paraboloid 521
class $\mathbb{M}^+$ 215, 216, 222, 223
class $\mathbb{M}^-$ 215, 216, 222–224, 226
class $\mathbb{M}_0^-$ 217, 220, 221
class $\mathbb{M}_B^+$ 217, 218, 227
class $\mathbb{M}_B^-$ 217, 218, 221
class $\mathbb{M}_\infty^+$ 217, 220, 227
closed systems 442
cofactor 520
cofactor of an exponential factor 447
cofactor of an invariant algebraic curve 444
coincidence degree 544
compact embedding 428
complete continuity 374, 377
completely continuous 365, 420
completely continuous mapping 404
completely continuous operator 369, 372, 376, 388, 393
complex polynomial system 440
complex solution 440
conditionally oscillatory equation 283
cone 522
configuration of limit cycles 480
conjugate equation 176
conjugate number of p 167
conjugate points 176, 241
Conley index 593, 600–602, 667, 671, 672, 674–678, 680
Conley index, cohomological 674

Conley index, homological 674
Conley theorem, first 600, 668
Conley theorem, second 600, 669
consecutive minimizers 574
constant torque 536, 540, 541
constraint 428
continuation methods 401, 427
continuation theorem 390, 394, 404
continuous embedding 415
contraction mapping principle 367
coupled lower and upper quasi-solutions 152
critical dimension 418, 419
critical exponent 416–419
critical group 565
critical orbit 549
critical point 314, 408, 416, 424, 428, 546
critical point theorems 370
critical point theory 430
critical value 314, 546
cup-length 601, 602, 632, 633, 636, 647, 674
curvature problems 405

Darboux lemma 461, 466
Darboux proposition 468
Darbouxian function 452
Darbouxian theory of integrability 449
degeneracy 556
degeneracy problem 554
degree of a homogeneous ideal 462
degree of a polynomial system 440
degree of a polynomial vector field 518
degree of an algebraic limit cycle 475
degree theory 364, 381, 427
density of the sequence of intervals 279
derivation 500
differential field 500
diffusion coefficients 405
diffusion problems 363
direct method of the calculus of variations 535, 546
Dirichlet boundary conditions 406
Dirichlet boundary value problem 330
disconjugate equation 175
discrete Picone identity 338
discrete quadratic functional 338
distance of consecutive zeros 275
divergence 442
divergence theorem 414
dominant solution 300
Dulac function 488

egress 604
egress point 593, 595–599, 604, 660
egress point, strict 595–597, 599, 600, 604, 611, 612, 619, 660
egress point, strong 599
eigenfunction 293, 330, 424
eigenvalue 293, 330, 365
eigenvalue problem 365, 417
elementary first integral 501
ellipsoid 521
elliptic cylinder 522
elliptic paraboloid 521
Emden–Fowler differential equation 168
Emden–Fowler equation 326
energy functional 174, 242, 316, 336, 414, 425
ENR (Euclidean neighborhood retract) 630, 634, 635, 644, 651, 652, 655, 665, 674
entrance set 639, 640, 653
entrance set, proper 653, 663
equilibrium points 408
escape-time function 599, 605, 607, 616, 623, 625, 638, 655, 665
escape-time function, extended 606
Euler beta function 324
Euler–Lagrange equation 416
Euler-type differential equation 188, 200
Euler-type equation 234
eventually minimal solution 189
evolutionary operator 653
exact system 442
excision property 382, 385
existence principles 4, 6
exit set 599, 600, 604, 609, 611–613, 616, 619, 622, 638, 646–648, 653, 661, 673, 676
exit set, proper 653, 663
exponential factor 447
exponential factor on a surface 520
exterior domain 361

fine focus 485
fine focus of order *k* 485
first integral 441, 519
first integral associated to an integrating factor 442
fixed point 372, 376, 378, 381, 393, 395, 396
fixed point index 594, 600–602, 634, 635, 651, 652, 656, 666, 672
forced half-linear differential equation 284
Frank-Kamenetskii approximation 405
Fredholm alternative 182, 317
free conservative pendulum equation 537
free damped pendulum 539

G-invariant 549
gap of an orthogonal system 463
Gelfand problem 427

Gelfand type problems 404
generalized polyfacial set 644, 663
generalized Prüfer transformation 168, 169, 275
generalized Pythagorian identity 168
generalized sine function 169
generalized zero 339
geometrically distinct 543
global bifurcation 378
gradient of a function 518
gradient-like 602, 637, 648, 649, 651, 668, 674
graininess 348
ground states 363, 428, 430

H-function averaging technique 272
half-linear cosine function 183, 232, 275, 296
half-linear cotangent 169
half-linear difference equation 338
half-linear differential equation 167
half-linear dynamic equations 348
half-linear Prüfer transformation 232, 294, 296
half-linear sine function 183, 192, 232, 275, 296
half-linear tangent 169
Hamiltonian 481
Hamiltonian vector field 481
Harnack inequality 330
Hartman–Wintner theorem 203, 257, 264
Hessian matrix 362
heteroclinic orbit 408, 539
heteroclinic solutions 575
Hilbert's Nullstellensatz 492
Hille–Wintner comparison theorem 207, 212
homoclinic orbits 539
homogeneity conditions 365
homogeneous polynomial vector field 468
homotopic deformation along p 320
homotopy type 600, 601, 628, 631, 633, 670, 675–677
homotopy type, absolute 628, 629, 631, 675, 677
homotopy type, pointed 628, 629, 631, 633, 677
homotopy types 601
Hopf bifurcation 485
hyperbolic cylinder 522
hyperbolic paraboloid 521
hyperboloid of one sheet 522
hyperboloid of two sheets 521

indefinite weight 295
independent points 448, 525
index calculations 365
index of stationary points 651, 652
ingress point 595, 596
ingress point, strict 595, 600, 611
initial value problem 1, 3, 51, 65, 66, 367
integrable system 441
integral characterization 236, 240
integral curve 440
integral equation 376
integral J_1 217
integral J_2 217
integral J_c 216
integral J_r 216
integral R_1 223
integral R_2 223
integrating factor 442
intermittently tending function 279
intersection index 462
intersection index at a point 462
invariant 441
invariant algebraic curve 444
invariant algebraic curve on a surface 519
invariant part 593, 600, 638–640, 642, 643, 647, 648, 652, 679, 680
invariant part, negative 639
invariant part, positive 639
irreducible invariant algebraic curve 446
irreducible orthogonal system 463
isolated invariant set 600–602, 667–669, 671, 672, 674, 676–680
isolated set of fixed points 635, 651, 655
isolating block 593, 600, 602, 638–640, 642–649, 651–654, 656, 667–671, 673–676, 679
isolating neighborhood 600, 667, 668, 674, 680
isolating segment 602, 652–654, 657, 658, 663, 664, 666, 667
isolating segment, periodic 654–656, 665, 666
isothermal gas 362, 408

k-Hessian 419
k-Hessian equation 416
k-Hessian operator 362, 410
Kamenev criterion 263
Kapteyn–Bautin theorem 452
Krasnoselskii genus 315, 335

L^p-Carathéodory condition 72
L^p-Carathéodory function 72
Lagrange multiplier 428
Lagrange stability 572
Lagrangian 414–417
Landesman–Lazer conditions 321
Landesman–Lazer results 421
Lefschetz fixed point theorem 601, 635, 638, 662
Lefschetz number 634, 654
left disfocal equation 250
left jump operator 348

Leighton–Wintner criterion 303
Leighton–Wintner oscillation criterion 178
Leray–Schauder alternative 4, 6, 7
Leray–Schauder degree 319, 366, 382, 384, 385, 389
Liapunov function 486, 487, 610, 637, 648–650
Liapunov quantity 485
limit characterization 222, 241
limit cycle 475
linear differential equations 443
Liouville theorem 501
Liouville–Gelfand equation 405, 406
Liouville–Gelfand problem 361, 405, 409
Liouvillian first integral 452, 501, 503
local flow 595, 598, 600–603, 608–611, 617, 619, 621, 624, 636–638, 642–644, 646, 648, 651–653, 667–669, 671–674, 676–680
local ring 462
local semiflow 601–604, 610, 620–622, 625, 636–638, 652
Lotka–Volterra system 442, 503
lower semicontinuous function 430
lower solution 73, 83, 89, 391, 404
Lusternik–Schnirelman 549
Lusternik–Schnirelman procedure 314
Lyapunov inequality 246
Lyapunov stability 558
Lyapunov–Schmidt's decomposition 543

Mather set 570, 574
maximum principle 361, 391
measure chain 348
Milloux theorem 279
minimal and maximal solution 77
minimal function 566
minimal solution 231, 256
minus gradient flow 122
mode 426
Monge–Ampére operator 362, 417
Monge–Ampére version of the Liouville–Gelfand equation 409
monodromy homeomorphism 654, 657, 658
monotone increasing map 136
monotone operator 377
monotone twist homeomorphism 573
Morse index 564, 601, 676
Morse index, generalized 601
Morse theory 427
Moser's twist theorem 571
mountain pass lemma 547
mountain pass theorem 428, 429
multi-bump solutions 575
multiple point 469
multiplicity 13
multiplicity of a curve at a point 469
multiplicity of a point 462
multiplicity results 362

Nagumo condition 81, 83
negative half-orbit 516
negative semitrajectory 596, 603
Neumann lemma 517
nodal contour 335
nodal domain 335
nonoscillation theorems 363
nonoscillatory equation 176, 197, 341
nonradial solution 425–427
nonresonant case 182
nonuniform nonresonance 321
normal cone 135

one-sided Nagumo condition 82, 83
open question 460, 480
orbit 440
order 135
order cone 135
Orlicz–Sobolev conjugate 385
orthogonal complement 422
orthogonal system of polynomials 462
orthogonal system without projective zero 463
oscillation 367
oscillation constant 283
oscillation theorems 363
oscillatory equation 176, 197, 341
oscillatory solutions 367, 376
oscillatory term 422
outward tangency point 595, 600, 611, 612, 619

p-Laplace operator 379
p-Laplacian 318, 329, 334, 381
p-Laplacian operator 410
Palais–Smale condition (PS) 126, 314, 318, 335, 428, 547
Palais–Smale sequence 547
parabolic cylinder 521
parallel neighborhood 516
partial order 377
partially ordered Banach space 376
periodic coefficient 268
periodic solution of the second kind 539, 565
periodically forced pendulum 541
perturbation of two-term half-linear equation 302
phase plane 407
phase portrait 441
Picone's identity 172, 243, 333
planar polynomial differential system 440

Pohozaev identity 415
Poincaré map 656, 664, 676
Poincaré mapping 553
Poincaré–Birkhoff fixed point theorem 562
pointed space 600, 627, 628, 677
points 611
polyfacial set 596, 601, 610–614, 616, 618, 621–624, 626
polyfacial set, generalized 601, 617–620
polyfacial set, regular 612, 613, 616
polynomial 1-form 440
polynomial system 440
polynomial vector field 440, 518
polynomial vector field on a surface 519
positive half-orbit 516
positive semitrajectory 596, 599, 600, 603, 604, 609, 610, 622
positively invariant set 124
positone 3, 8
positone problems 400
primary limit cycle 480
principal eigenfunction 422
principal eigenvalue 370, 371, 378, 418, 420
principal solution 197, 222, 229, 231, 233, 236, 242, 305
projective homogeneous 1-form 468
proper solution 326
Prüfer transformation 192, 280
$(\mathrm{PS})_c$-condition 547
$(\mathrm{PS})_G$-condition 549
pseudolaplacian 336
Pythagorean identity 297

quasi-isotopic deformation retract 601, 608–610, 614
quasi-jumping function 279

radial eigenvalues 361
radial solutions 361
radial symmetry 361
Rayleigh quotient 330
Rayleigh–Ritz inequality 43
real polynomial system 440
real solution 440
realization of a configuration of limit cycles 481
reciprocal equation 178, 240
reciprocity principle 178, 215
reduction method of Liapunov–Schmidt's type 550
regular growth 279
regular half-linear equation 240
regular polyfacial set 657, 669
regular Sturm–Liouville problem 295
regular surface 519
regular value 556
regularly varying function 229
Rellich–Pohozaev identity 412
repelling node 408
resonant case 321
retarded half-linear equation 286
retract 593, 594, 596–598, 606, 607, 621, 623, 624, 626, 629, 630, 650, 661, 664
retract method 593, 594, 596
Riccati difference equation 338
Riccati inequality 200, 201, 343, 345
Riccati integral equation 207, 265
Riccati integral inequality 210
Riccati substitution 172, 249
Riccati technique 180, 197, 343
Riccati-type differential equation 171
right disfocal equation 250
right focal point 249
right jump operator 348
rotating solution 566
rotation number 566, 568, 573
roundabout theorem 174, 197, 338

Schauder 4, 6, 7, 25, 28, 30, 46, 47, 54, 57, 59
Schauder–Tychonov fixed point theorem 208, 212, 326
Schauder's fixed point theorem 376
second eigenvalue of p-Laplacian 335
secondary bifurcations 427
semitrajectory, negative 595
semitrajectory, positive 595
separatrix 517
sign changing nonlinearities 22
simple point 469
singular 1, 3, 8, 13, 21, 22, 51
singular point 242, 449
singular solution 168, 407, 408
singular solution of the first kind 326
singular solution of the second kind 326
singular Sturm–Liouville problem 300
slowly varying function 229
smash product 627, 628, 675
Sobolev exponent 415
Sobolev space 370, 428
solution continuum 364, 373, 404
sphere 521
spiral node 408
spiral system 516
star-shaped domain 409, 412, 416
starlike domain 412, 413
state-space 407
strict lower solution 94, 102
strict upper solution 94, 102

strictly convex domain 409
strip system 516
strong deformation retract 594, 599, 606–608, 610, 622, 624, 629, 630, 632, 636, 637, 639, 641, 645, 649, 670, 673, 674
strong first integral 515
strongly increasing solution 327
strongly nonoscillatory equation 283
strongly oscillatory equation 283
Sturm type separation properties 368
Sturm–Liouville difference equation 337
Sturm–Liouville differential equation 167, 201
Sturm–Liouville problem 293
Sturmian majorant 234, 255, 295
Sturmian oscillation theory 167
subcritical exponents 417
subcritical growth 380
subdominant solution 300
subharmonic solution 543, 563
symmetry breaking 363, 424, 427
symmetry breaking bifurcations 361

T-periodic solution 542
thermal ignition problems 362
time scale 348
topological degree 427
topologically equivalent configurations of limit 481
toral system 516
total intersection index 462
total multiplicity 462
trajectory 440, 595, 603, 617, 637, 672, 676, 678
Tricomi's equation 536
trigonometric transformation 181, 252
trivial pointed homotopy type 628, 674
true minimizer 576
Tychonov fixed point theorem 220, 222

unbounded continuum 378
upper and lower solutions 364, 401, 543, 583
upper solution 73, 84, 89, 391, 404

Vallée Poussin-type inequality 248
variational approach 429
variational characterization of eigenvalues 314, 322
variational methods 370, 409, 425, 427
variational principle 197

Ważewski lemma 599, 601, 605, 606
Ważewski method 593, 594, 596, 599–602, 604, 610, 613, 614, 620, 621, 626, 629, 630, 636, 658, 659, 662, 667
Ważewski set 593, 599–602, 604–607, 610, 612, 617, 619, 622, 625, 636–639, 661
Ważewski theorem 596–601, 604, 606–608, 614, 616, 626, 662, 673
weak first integral 515
weak singular point 449
weak solutions 428
weakly closed sets 428
weakly increasing solution 327
wedge sum 627, 628, 677
weight function 256
Willet's criteria 266
Wirtinger inequality 198
Wronskian identity 180, 230

Young inequality 173, 274, 286

HANDBOOK OF DIFFERENTIAL EQUATIONS

ORDINARY DIFFERENTIAL EQUATIONS

VOLUME II

Contents

Preface v
List of Contributors vii
Contents of Volume 1 xi

1. Optimal control of ordinary differential equations 1
V. Barbu and C. Lefter
2. Hamiltonian systems: periodic and homoclinic solutions by variational methods 77
T. Bartsch and A. Szulkin
3. Differential equations on closed sets 147
O. Cârjă and I.I. Vrabie
4. Monotone dynamical systems 239
M.W. Hirsch and H. Smith
5. Planar periodic systems of population dynamics 359
J. López-Gómez
6. Nonlocal initial and boundary value problems: a survey 461
S.K. Ntouyas

Author index 559
Subject index 565

Contents of Volume 1

Preface v
List of Contributors vii

1. A survey of recent results for initial and boundary value problems singular in the dependent variable 1
   R.P. Agarwal and D. O'Regan
2. The lower and upper solutions method for boundary value problems 69
   C. De Coster and P. Habets
3. Half-linear differential equations 161
   O. Došlý
4. Radial solutions of quasilinear elliptic differential equations 359
   J. Jacobsen and K. Schmitt
5. Integrability of polynomial differential systems 437
   J. Llibre
6. Global results for the forced pendulum equation 533
   J. Mawhin
7. Ważewski method and Conley index 591
   R. Srzednicki

Author Index 685
Subject Index 693

CHAPTER 1

Optimal Control of Ordinary Differential Equations

Viorel Barbu and Cătălin Lefter

University "Al.I. Cuza", Iaşi, Romania, and
Institute of Mathematics "Octav Mayer", Romanian Academy, Romania

Contents

1. Introduction . . . 3
    1.1. The calculus of variations . . . 4
    1.2. General form of optimal control problems . . . 9
2. Preliminaries . . . 10
    2.1. Elements of convex analysis . . . 10
    2.2. Ekeland's variational principle . . . 15
    2.3. Elements of differential geometry and exponential representation of flows . . . 16
3. The Pontriaghin maximum principle . . . 29
    3.1. The main theorem . . . 29
    3.2. Proof of the maximum principle . . . 32
    3.3. Convex optimal control problems . . . 38
    3.4. Examples . . . 47
    3.5. Reachable sets and optimal control problems . . . 55
    3.6. Geometric form of Pontriaghin maximum principle . . . 57
    3.7. Free time optimal control problems . . . 59
4. The dynamic programming equation . . . 62
    4.1. Optimal feedback controllers and smooth solutions to Hamilton–Jacobi equation . . . 62
    4.2. Linear quadratic control problems . . . 65
    4.3. Viscosity solutions . . . 68
    4.4. On the relation between the two approaches in optimal control theory . . . 72
References . . . 74

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 2
Edited by A. Cañada, P. Drábek and A. Fonda
© 2005 Elsevier B.V. All rights reserved

1. Introduction

The theory of control of differential equations has developed in several directions in close relation with the practical applications of the theory. Its evolution has shown that its methods and tools are drawn from a large spectrum of mathematical branches such as ordinary differential equations, real analysis, calculus of variations, mechanics, geometry. Without being exhaustive we just mention, as subbranches of the control theory, the controllability, the stabilizability, the observability, the optimization of differential systems and of stochastic equations or optimal control. For an introduction to these fields, and not only, see [22,33,38], as well as [2,25] and [26] for a geometric point of view.

The purpose of this work is to discuss the first order necessary and sufficient conditions of optimality in control problems governed by ordinary differential systems. We do not treat the optimal control of partial differential equations although all basic questions of the finite dimensional theory (existence of optimal control, maximum principle, dynamic programming) remain valid but the treatment requires more sophisticated methods because of the infinite dimensional nature of the problems (see [4,27,38]).

In Section 1 we present some aspects and ideas in the classical Calculus of variations that lead later, in the fifties, to the modern theory of optimal control for differential equations.

Section 2 presents some preliminary material. It contains elements of convex analysis and the generalized differential calculus for locally Lipschitz functionals, introduced by F.H. Clarke [10]. This will be needed for the proof of the maximum principle of Pontriaghin, under general hypotheses, in Section 3.1. We then discuss the exponential representation of flows, introduced by A. Agrachev and R. Gamkrelidze in order to give a geometric formulation to the maximum principle that we will describe in Sections 3.5, 3.6.

Section 3 is concerned with the Pontriaghin maximum principle for general Bolza problems. There are several proofs of this famous classical result and here, following F.H. Clarke's ideas (see [11]), we have adapted the simplest one relying on Ekeland's variational principle. Though the maximum principle given here is not in its most general form, it is however sufficiently general to cover most of significant applications. Some examples are treated in detail in Section 3.4. Since geometric control theory became in last years an important branch of mathematics (for an introduction to the theory see [2,26]), it is useful and interesting to give a geometric formulation of optimal control problems and, consequently, a geometric form of the maximum principle. Free time optimal problems are also considered as a special case.

In the last section we present the dynamic programming method in optimal control problems based on the partial differential equation of dynamic programming, or Bellman equation (see [7]). The central result of this chapter says that the value function is a viscosity solution to Bellman equation and that, if a classical solution exists, then an optimal control, in feedback form, is obtained. Applications to linear quadratic problems are given. We discuss also the relationship between the maximum principle and the Bellman equation and we will see in fact that the dynamic programming equation is the Hamilton–Jacobi equation for the Hamiltonian system given by the maximum principle.

1.1. *The calculus of variations*

In this section we point out the fundamental lines of development in the Calculus of variations. We will not impose rigorous assumptions on the functions entering the described problems, they will be as regular as needed. The main purpose is just to emphasize some fundamental ideas that will be reencountered, in a metamorphosed form, in the theory of optimal control for differential equations. For a rigorous presentation of the theory a large literature may be cited, however we restrict for instance to [8,24] and to a very nice survey of extremal problems in mathematics, including the problems of Calculus of variations, in [34].

Let M be an n-dimensional manifold, and M_0, M_1 be subsets (usually submanifolds) of M. $L : \mathbb{R} \times TM \to \mathbb{R}$ is the Lagrangean function, TM being the tangent bundle of M. The generic problem of the classical Calculus of variations consists in finding a curve, y^*, which minimizes a certain integral

$$J(y) = \int_{t_0}^{t_1} L\big(t, y(t), y'(t)\big)\,\mathrm{d}t \tag{1.1}$$

in the space of curves

$$Y = \big\{y : [t_0, t_1] \to M;\ y(t_j) \in M_j,\ j = 1, 2,\ y \text{ continuous and piecewise } C^1\big\}.$$

The motivation for studying such problems comes from both geometry and classical mechanics.

EXAMPLES. *1. The brachistocrone.* The classical brachistocrone problem proposed by Johann Bernoulli in 1682, asks to find the curve, in a vertical plane, on which a material point, moving without friction under the action of its weight, is reaching the lower end of the curve in minimum time. More precisely, if the curve is joining two points $y(t_0) = y_0, y(t_1) = y_1$, then the time necessary for the material point to reach y_1 from y_0 is

$$T = \int_{t_0}^{t_1} \big(2g\big|y(t) - y_0\big|\big)^{-1/2} \sqrt{1 + \big(y'(x)\big)^2}\,\mathrm{d}t.$$

The curve with this property is a cycloid.

2. The minimal surface of revolution. One is searching for the curve $y : [t_0, t_1] \to \mathbb{R}$, $y(t_0) = y_0$, $y(t_1) = y_1$, which generates the surface of revolution of least area. The functional to be minimized is

$$J(y) = 2\pi \int_{t_0}^{t_1} y(x)\sqrt{1 + \big(y'(t)\big)^2}\,\mathrm{d}t.$$

The solution is the catenary.

3. Lagrangean mechanics. A mechanical system with a finite number of degrees of freedom is mathematically modelled by a manifold M and a Lagrangean function

$L : \mathbb{R} \times TM \to \mathbb{R}$ (see [3]). The manifold M is the configuration space of the mechanical system. The points $y \in M$ are generalized coordinates and the $y' \in TM$ are generalized speeds. The principle of least action of Maupertuis–d'Alembert–Lagrange states that the trajectories of the mechanical system are *extremal* for the functional J defined in (1.1). Consider the case of a system of N material points in the 3 dimensional space, moving under the action of mutual attraction forces. In this case the configuration space is $(\mathbb{R}^3)^N$, while the Lagrangean is

$$L = T - U \tag{1.2}$$

where T is the kinetic energy

$$T = \sum_{i=1}^{N} \frac{1}{2} m_i |x_i'|^2$$

and U is the potential energy

$$U(x_1, \dots, x_N) = \sum_{i=1}^{N} \frac{k m_i m_j}{|x_i - x_j|},$$

k is an universal constant.

To make things clear we consider the simplest problem in the Calculus of variations when $M = \mathbb{R}^n$ and y_0, y_1 are fixed.

We consider the space of variations $\mathcal{Y} = \{h : [t_0, t_1] \to \mathbb{R}^n;\ h(t_0) = h(t_1) = 0, h \in C^1\}$; if y^* is a minimum of J in Y, then the first variation

$$\delta J(y^*)h := \frac{\mathrm{d}}{\mathrm{d}s} J(y^* + sh)\bigg|_{s=0} = 0. \tag{1.3}$$

A curve that satisfies (1.3) is called *extremal* and this is only a *necessary condition* for a curve to realize the infimum of J. One easily computes

$$\delta J(y)h = \int_{t_0}^{t_1} L_y\big(t, y^*(t), (y^*)'(t)\big) \cdot h + L_{y'}\big(t, y^*(t), (y^*)'(t)\big) \cdot h' \,\mathrm{d}t$$

where L_y, $L_{y'}$ are the gradients of L with respect to y and y', respectively. If y^* is C^2, an integration by parts in the previous formula gives

$$L_y\big(t, y^*(t), (y^*)'(t)\big) - \frac{\mathrm{d}}{\mathrm{d}t} L_{y'}\big(t, y^*(t), (y^*)'(t)\big) = 0 \tag{1.4}$$

which are the *Euler–Lagrange* equations. It is a system of n differential equations of second order.

It may be proved that if the Hessian matrix $(L_{y'y'}) > 0$, then the regularity of L is inherited by the extremals, for instance if $L \in C^2$ then the extremals are C^2 and thus satisfy the Euler–Lagrange system. The proof of this fact is based on the first of the *Weierstrass–Erdmann necessary conditions* which state that, along each extremal, $L_{y'}$ and the Hamiltonian defined below in (1.6) are continuous.

Another necessary condition for the extremal y^* to realize the infimum of J is that $(L_{y'y'}) \geqslant 0$ along y^*. This is the *Legendre necessary condition.*

Suppose from now on that $(L_{y'y'})$ is a nondegenerate matrix at any point (t, y, y'). We set

$$p = L_{y'}(t, y, y'). \tag{1.5}$$

Since $(L_{y'y'})$ is nondegenerate, formula (1.5) defines a change of coordinates $(t, y, y') \to (t, y, p)$. From the geometric point of view it maps TM locally onto T^*M, the cotangent bundle. In mechanics p is called the generalized momentum of the system and in most applications its significance is of adjoint (or dual) variable. We consider the *Hamiltonian*

$$H(t, y, p) = (p, y') - L(t, y, y'). \tag{1.6}$$

If, moreover, L is convex in y' then $H = L^*$, the *Legendre transform* of L:

$$H(t, y, p) = \sup_{y'}\bigl\{(p, y') - L(t, y, y')\bigr\}.$$

For example, if L is given by (1.2) then $H = T + U$ and it is just the total energy of the system. If we compute the differential of H along an extremal, taking into account the Euler–Lagrange equations, we obtain

$$\mathrm{d}H = -L_t\,\mathrm{d}t - \frac{\mathrm{d}}{\mathrm{d}t}L_{y'}\,\mathrm{d}y + y'\,\mathrm{d}p.$$

Thus, through these transformations we obtain the *Hamiltonian equations*

$$\begin{cases} y' = \dfrac{\partial H}{\partial p}(t, y, p), \\ p' = -\dfrac{\partial H}{\partial y}(t, y, p). \end{cases} \tag{1.7}$$

Solutions of the Hamiltonian system are in fact extremals corresponding to the Lagrangean $\widetilde{L}(t, (y, p), (y', p')) = p \cdot y' - H(t, y, p)$ in T^*M. The projections on M are extremals for J. Roughly speaking, solving the Euler–Lagrange system is equivalent to solving the Hamiltonian system of $2n$ differential equations of first order. From the mechanics point of view these transforms give rise to the Hamiltonian mechanics which study the mechanical phenomena in the phase space T^*M while in mathematics this is the start point for the symplectic geometry (see for example [3,28]).

Consider now the more general case of end points lying on two submanifolds M_0, M_1. It may be shown that the first variation of J in y computed in an admissible variation h (assume that also t_0, t_1 are free) is

$$\delta J(y)h = \int_{t_0}^{t_1} L_y(t, y, y') \cdot h + L_{y'}(t, y, y') \cdot h' \, \mathrm{d}t + \{p\delta y - H\delta t\}\big|_{t_0}^{t_1} \tag{1.8}$$

where

$$H\delta t\big|_{t_0}^{t_1} = H\big(t_1, y(t_1), p(t_1)\big)\delta t_1 - H\big(t_0, y(t_0), p(t_0)\big)\delta t_0,$$

$$p\delta y\big|_{t_0}^{t_1} = p(t_1)\delta y_1 - p(t_2)\delta y_0.$$

It turns out that $y^* \in C^2$ is extremal for J if y^* satisfies the Euler–Lagrange equations (1.4) and in addition

$$\{p\delta y - H\delta t\}\big|_{t_0}^{t_1} = 0. \tag{1.9}$$

These are *transversality conditions*. In case t_0, t_1 are fixed, these become

$$p(t_0) \perp M_0, \qquad p(t_1) \perp M_1.$$

Since $(L_{y'y'})$ is supposed to be nondegenerate, the Euler–Lagrange equations form a second order nondegenerate system of equations and this implies that the family of extremals starting at moment t_0 from a given point of $y_0 \in M$ cover a whole neighborhood $\mathcal{V}$ of (t_0, y_0) (we just vary the value of $y'(t_0)$ in the associated Cauchy problem and use some result on the differentiability of the solution with respect to the initial data, coupled with the inverse function theorem). We consider now the function $S : \mathcal{V} \to \mathbb{R}$ defined by

$$S(t, y) = \int_{t_0}^{t} L\big(s, x(s), x'(s)\big) \, \mathrm{d}s$$

where the integral is computed along the extremal $x(s)$ joining the points (t_0, y_0) and (t, y). It may be proved that S satisfies the first order nonlinear partial differential equation

$$S_t + H(t, y, S_y) = 0. \tag{1.10}$$

This is the *Hamilton–Jacobi* equation. This is strongly related to the Hamiltonian system (1.7) which is the system of characteristics associated to the partial differential equation (1.10) (see [16]).

A partial differential equation is usually a more complicated mathematical object than an ordinary differential system. Solving a first order partial differential system reduces to solving the corresponding characteristic system. This is the method of characteristics (see [16]).

However, this duality may be successfully used in a series of concrete situations to integrate the Hamiltonian systems appearing in mechanics or in the calculus of variations. This result, belonging to Hamilton and Jacobi, states that if a general solution for the Hamilton–Jacobi equation (1.10) is known, then the Hamiltonian system may be integrated (see [3,16,24]). More precisely, we assume that a general solution of (1.10) is $S = S(t, y_1, \ldots, y_n, \alpha_1, \ldots, \alpha_n)$ such that the matrix $(\frac{\partial^2 S}{\partial y_i \partial \alpha_j})$ is nondegenerated. Then $\frac{\partial S}{\partial \alpha_j}$ are prime integrals and a general solution of the Hamiltonian system (1.7) is given by the $2n$ system of implicit equations:

$$\beta_i = \frac{\partial S}{\partial \alpha_i}, \qquad p_i = \frac{\partial S}{\partial y_i}.$$

In fact S is a generating function for the symplectic transform $(y_i, p_i) \to (\beta_i, \alpha_i)$ and in the new coordinates the system (1.7) has a simple form for which the Hamiltonian function is $\equiv 0$. A last remark is that a general solution to equation (1.10) may be found if variables of H are separated (see [3,24]).

We considered previously first order necessary conditions. Suppose that $(L_{y'y'}) > 0$. Let us take now the second variation

$$\delta^2 J(y)h := \frac{\mathrm{d}^2}{\mathrm{d}s^2} J(y + sh)\bigg|_{s=0}.$$

This is a quadratic form denoted by

$$Q^y(h) = \int_{t_0}^{t_1} \Omega^y(t, h, h')\,\mathrm{d}t$$

where the new Lagrangean

$$\Omega^y(t, h, h') = \big(L_{yy}(t, y, y')h, h\big) + 2\big(L_{yy'}(t, y, y')h, h'\big) \\ + \big(L_{y'y'}(t, y, y')h', h'\big).$$

Here $(\cdot, \cdot)$ denotes the scalar product in $\mathbb{R}^n$ and we assumed that the matrix $(L_{yy'})$ is symmetric (for $n = 1$ this is trivial, in higher dimensions the hypothesis simplifies computations but may be omitted). Clearly, if y^* realizes a global minimum of J, then the quadratic form $Q(y^*) \geqslant 0$. The positivity of Q is related to the notion of *conjugate point*. A point $\bar{t}$ is conjugate to t_0 along the extremal y^* if there exists a non trivial solution $h : [t_0, \bar{t}] \to \mathbb{R}^n$, $h(t_0) = h(\bar{t}) = 0$ of the second Euler equation:

$$\Omega_h^{y^*} - \frac{\mathrm{d}}{\mathrm{d}t}\Omega_{h'}^{y^*} = 0.$$

The *Jacobi necessary condition* states that if y^* realizes the infimum of J then the open interval (t_0, t_1) does not contain conjugate points to t_0. If y^* is just an extremal and the closed interval $[t_0, t_1]$ does not contain conjugate points to t_0, then y^* is a local weak minimum of J (in C^1 topology).

1.2. *General form of optimal control problems*

We consider the controlled differential equation

$$y'(t) = f\big(t, y(t), u(t)\big), \quad t \in [0, T]. \tag{1.11}$$

The input function $u : [0, T] \to \mathbb{R}^m$ is called *controller* or *control* and $y : [0, T] \to \mathbb{R}^n$ is the *state of the system*. We will assume that $u \in \mathcal{U}$ where $\mathcal{U}$ is the set of measurable, locally integrable functions which satisfy the *control constraints*:

$$u(t) \in U(t) \quad \text{a.e. } t \in [0, T] \tag{1.12}$$

where $U(t) \subset \mathbb{R}^n$ are given closed subsets. The differential system (1.11) is called the *state system*. We also consider a Lagrangean L and the *cost functional*

$$J(y, u) = \int_0^T L\big(t, y(t), u(t)\big)\, \mathrm{d}t + g\big(y(0), y(T)\big). \tag{1.13}$$

A pair (y, u) is said to be *admissible pair* if it satisfies (1.11), (1.12) and $J(y, u) < +\infty$.

The *optimal control problem* we consider is

$$\min\big\{J(y, u);\ \big(y(0), y(T)\big) \in C,\ (y, u) \text{ verifies (1.11)}\big\} \tag{1.14}$$

Here $C \subset \mathbb{R}^n \times \mathbb{R}^n$ is a given closed set.

A controller u^* for which the minimum in (1.14) is attained is called *optimal controller*. The corresponding states y^* are called *optimal states* while (y^*, u^*) will be referred as *optimal pairs*. By solution to (1.11) we mean an absolutely continuous function $y : [0, T] \to \mathbb{R}$ (i.e., $y \in AC([0, T]; \mathbb{R}^n)$) which satisfies almost everywhere the system (1.11). In the special case $f(t, y, u) \equiv u$, problem (1.14) reduces to the classical problem of calculus of variations that was discussed in Section 1.1. For different sets C we obtain different types of control problems. For example, if C contains one element, that is the initial and final states are given, we obtain a *Lagrange problem*. If the initial state of the system is given and the final one is free, $C = \{y_0\} \times \mathbb{R}$, one obtains a *Bolza problem*. A Bolza problem with the Lagrangean $L \equiv 0$ becomes a *Mayer problem*.

An optimal controller u^* is said to be a *bang-bang controller* if $u^* \in \partial U(t)$ a.e. $t \in (0, T)$ where ∂U stands for the topological boundary of U.

It should be said that the control constraints (1.12) as well as end point constraints $(y(0), y(T)) \in C$ can be implicitely incorporated into the cost functional J by redefining L and g as

$$\widetilde{L}(t, y, u) = \begin{cases} L(t, u) & \text{if } u \in U(t), \\ +\infty & \text{otherwise,} \end{cases}$$

$$\tilde{g}(y_1, y_2) = \begin{cases} g(y_1, y_2) & \text{if } (y_1, y_2) \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

Moreover, integral (isoperimetric) constraints of the form

$$\int_0^T h_i\big(t, y(t), u(t)\big)\,\mathrm{d}t \leqslant \alpha_i, \quad i = 1, \ldots, l,$$

$$\int_0^T h_i\big(t, y(t), u(t)\big)\,\mathrm{d}t = \alpha_i, \quad i = l+1, \ldots, m$$

can be implicitly inserted into problem (1.14) by redefining new state variables $\{z_1 \ldots, z_m\}$ and extending the state system (1.11) to

$$\begin{cases} y'(t) = f\big(t, y(t), u(t)\big), \quad t \in (0, T), \\ z'(t) = h\big(t, y(t), u(t)\big), \\ z(0) = 0, \quad z_i(T) \leqslant \alpha_i \text{ for } i = 1, \ldots, l, \quad z_i(T) = \alpha_i \text{ for } i = l+1, \ldots, m \end{cases}$$

where $h = \{h_i\}_{i=1}^m$. For the new state variable $X = (y, z)$ we have the end point constraints

$$\big(X(0), X(T)\big) \in K,$$

where

$$K = \Big\{\big((y_0, 0, \ldots, 0), (y_1, z)\big) \in \mathbb{R}^{n+m} \times \mathbb{R}^{n+m}, (y_0, y_1) \in C,$$
$$z_i \leqslant \alpha_i, i = 1, \ldots, l,\ z_i = \alpha_i, i = l+1, \ldots, m\Big\}.$$

2. Preliminaries

2.1. *Elements of convex analysis*

Here we shall briefly recall some basic results pertaining convex analysis and generalized gradients we are going to use in the formulation and in proof of the maximum principle.

Let X be a real Banach space with the norm $\|\cdot\|$ and dual X^*. Denote by $(\cdot, \cdot)$ the pairing between X and X^*.

The function $f : X \to \overline{\mathbb{R}} =]-\infty, +\infty]$ is said to be *convex* if

$$f\big(\lambda x + (1-\lambda)y\big) \leqslant \lambda f(x) + (1-\lambda) f(y), \quad 0 \leqslant \lambda \leqslant 1,\ x, y \in X. \tag{2.1}$$

The set $D(f) = \{x \in X;\ f(x) < \infty\}$ is called the *effective domain* of f and

$$E(f) = \big\{(x, \lambda) \in X \times \mathbb{R};\ f(x) \leqslant \lambda\big\} \tag{2.2}$$

is called the *epigraph* of f. The function f is said to be *lower semicontinuous* (l.s.c.) if

$$\liminf_{x \to x_0} f(x) \geqslant f(x_0).$$

The function f is said to be *proper* if $f \not\equiv +\infty$.

It is easily seen that a convex function is l.s.c. if and only if it is weakly lower semicontinuous. Indeed, f is l.s.c. if and only if every level set $\{x \in X;\ f(x) \leqslant \lambda\}$ is closed. Moreover, the level sets are also convex, by the convexity of f; the conclusion follows by the coincidence of convex closed sets and weakly closed sets.

Note also that, by Weierstrass theorem, if X is a reflexive Banach space and if f is convex, l.s.c. and $\lim_{\|x\|\to\infty} f(x) = +\infty$, then f attains its infimum on X.

We note without proof (see, e.g., [9,6]) the following result:

PROPOSITION 2.1. *Let $f : X \to \overline{\mathbb{R}}$ be a l.s.c. convex function. Then f is bounded from below by an affine function and f is continuous on* $\operatorname{int} D(f)$.

Given a l.s.c. convex function $f : X \to \overline{\mathbb{R}}$, the mapping $\partial f : X \to X^*$ defined by

$$\partial f(x) = \{w \in X^*;\ f(x) \leqslant f(u) + (w, x - u),\ \forall u \in X\} \tag{2.3}$$

is called the *subdifferential* of f. An element of $\partial f(x)$ is called *subgradient* of f at x.

The mapping ∂f is generally multivalued. The set

$$D(\partial f) = \{x;\ \partial f(x) \neq \phi\}$$

is the domain of ∂f. It is easily seen that x_0 is a minimum point for f on X if and only if $0 \in \partial f(x_0)$.

We note also, without proof, some fundamental properties of ∂f (see, e.g., [6,9,31]).

PROPOSITION 2.2. *Let $f : X \to \overline{\mathbb{R}}$ be convex and l.s.c. Then* $\operatorname{int} D(f) \subset D(\partial f)$.

Let C be a closed convex set and let $\mathbb{I}_C(x)$ be the indicator function of C, i.e.,

$$\mathbb{I}_C(x) = \begin{cases} 0, & x \in C, \\ +\infty, & x \notin C. \end{cases}$$

Clearly, $\mathbb{I}_C(x)$ is convex and l.s.c. Moreover, we have $D(\partial \mathbb{I}_C(x)) = C$ and

$$\partial \mathbb{I}_C(x) = \{w \in X^*;\ (w, x - u) \geqslant 0,\ \forall u \in C\}. \tag{2.4}$$

$\partial \mathbb{I}_C(x)$ is precisely the *normal cone* to C at x, denoted $N_C(x)$.

If $F : X \to Y$ is a given function, X, Y Banach spaces, we set

$$F'(x, y) = \lim_{\lambda \to 0} \frac{F(x + \lambda y) - F(x)}{\lambda}$$

called the *directional derivative* of F in direction y.

By definition F is Gâteaux differentiable in x if $\exists DF(x) \in L(X, Y)$ such that

$$F'(x, v) = DF(x)v, \quad \forall v \in X.$$

In this case, DF is the Gâteaux derivative (differential) at x.

If $f : X \to \mathbb{R}$ is convex and Gâteaux differentiable in x, then it is subdifferentiable at x and $\partial f(x) = \nabla f(x)$.

In general, we have

PROPOSITION 2.3. *Let $f : X \to \overline{\mathbb{R}}$ be convex, l.s.c. and proper. Then, for each $x_0 \in D(\partial f)$*

$$\partial f(x_0) = \bigl\{w \in X^*;\ f'(x_0, u) \geqslant (w, u),\ \forall w \in X\bigr\}. \tag{2.5}$$

If f is continuous at x_0, then

$$f'(x_0, u) = \sup\bigl\{(w, u);\ w \in \partial f(x_0)\bigr\}, \quad \forall u \in X. \tag{2.6}$$

Given $f : X \to \overline{\mathbb{R}}$, the function $f^* : X^* \to \overline{\mathbb{R}}$

$$f^*(p) = \sup\bigl\{(p, x) - f(x);\ x \in X\bigr\}$$

is called the *conjugate* of f, or the *Legendre transform* of f.

PROPOSITION 2.4. *Let $f : X \to \overline{\mathbb{R}}$ be convex, proper, l.s.c. Then the following conditions are equivalent*:

1. $x^* \in \partial f(x)$,
2. $f(x) + f^*(x^*) = (x^*, x)$,
3. $x \in \partial f^*(x^*)$.

In particular, $\partial f^* = (\partial f)^{-1}$ and $f = f^{**}$. In general, $\partial(f + g) \supset \partial f + \partial g$ and the inclusion is strict. We have, however,

PROPOSITION 2.5 (Rockafellar). *Let f and g be l.s.c. and convex on D. Assume that $D(f) \cap$ int $D(g) \neq \phi$. Then*

$$\partial(f + g) = \partial f + \partial g. \tag{2.7}$$

We shall assume now that $X = H$ is a Hilbert space. Let $f : H \to \overline{\mathbb{R}}$ be convex, proper and l.s.c. Then ∂f is maximal monotone. In other words,

$$(y_1 - y_2, x_1 - x_2) \geqslant 0, \quad \forall (x_i, y_i) \in \partial f,\ i = 1, 2 \tag{2.8}$$

and

$$R(I + \lambda \partial f) = H, \quad \forall \lambda > 0. \tag{2.9}$$

$R(I + \lambda \partial f)$ is the range of $I + \lambda \partial f$.

The mapping

$$(\partial f)_\lambda = \lambda^{-1}\bigl(I - (I + \lambda \partial f)^{-1}\bigr), \quad \lambda > 0 \tag{2.10}$$

is called the *Yosida approximation* of f.

Denote by $f_\lambda : H \to \mathbb{R}$ the function

$$f_\lambda(x) = \inf\left\{ \frac{|x-y|^2}{2\lambda} + f(y);\ y \in H \right\}, \quad \lambda > 0$$

which is called the *regularization* of f (see [29]).

PROPOSITION 2.6 (Brezis [9]). *Let* $f : H \to \overline{\mathbb{R}}$ *be convex and l.s.c. Then* f_λ *is Fréchet differentiable on* H, $\partial f_\lambda = \{\nabla f_\lambda\}$ *and*

$$f_\lambda(x) = \frac{\lambda}{2}\left|\partial f_\lambda(x)\right|^2 + f\left(\left(I + \lambda \partial f(x)\right)^{-1}\right), \tag{2.11}$$

$$\lim_{\lambda \to 0} f_\lambda(x) = f(x), \quad \forall x \in H. \tag{2.12}$$

Consider the function $\mathbb{I}_g : L^p(\Omega) \to \overline{\mathbb{R}}$ defined by

$$\mathbb{I}_g(y) = \begin{cases} \displaystyle\int_\Omega g\big(x, y(x)\big)\mathrm{d}x & \text{if } g\big(\cdot, y(\cdot)\big) \in L^1(\Omega), \\ +\infty & \text{otherwise} \end{cases} \tag{2.13}$$

where $g : \Omega \times \mathbb{R}^m \to \mathbb{R}$ is a function satisfying (Ω is a measurable subset of $\mathbb{R}^n$)

1. $g(x, \cdot) : \mathbb{R}^m \to \overline{\mathbb{R}}$ is convex and l.s.c. for a.e. $x \in \omega$.
2. g is $\mathcal{L} \times \mathcal{B}$ measurable, i.e. g is measurable with respect to the σ-algebra of subsets of $\Omega \times \mathbb{R}^m$ generated by products of Lebesgue sets in Ω and Borelian sets in $\mathbb{R}^m$.
3. $g(x, y) \geqslant (\alpha(x), y) + \beta(x)$, a.e. $x \in \Omega$, $y \in \mathbb{R}^m$, where

$$\alpha \in L^q(\Omega), \quad \beta \in L^1(\Omega), \quad \frac{1}{p} + \frac{1}{q} = 1.$$

4. $\exists y_0 \in L^p(\Omega)$ such that $\mathbb{I}_g(y_0) < +\infty$.

PROPOSITION 2.7. *Let* $1 \leqslant p < \infty$. *Then* $\mathbb{I}_g$ *is convex, l.s.c. and* $\not\equiv +\infty$. *Moreover,*

$$\partial \mathbb{I}_g(y) = \left\{ w \in L^q(\Omega);\ w(x) \in \partial g\big(x, y(x)\big) \text{ a.e. } x \in \Omega \right\}. \tag{2.14}$$

EXAMPLE 2.1. Let

$$C = \left\{ y \in L^p(\Omega);\ a \leqslant y(x) \leqslant b,\ \text{a.e. } x \in \Omega \right\}.$$

Then

$$g(y) = \begin{cases} 0 & \text{if } a \leqslant y \leqslant b, \\ +\infty & \text{otherwise} \end{cases}$$

and so by (2.14)

$$\begin{aligned} N_C(y) &= \big\{w \in L^q(\Omega);\ w(x) \in N_{[a,b]}\big(y(x)\big) \text{ a.e. } x \in \Omega\big\} \\ &= \big\{w \in L^q(\Omega);\ w(x) = 0 \text{ if } a < y(x) < b, \\ &\qquad w(x) \leqslant 0 \text{ if } y(x) = a, \\ &\qquad w(x) \geqslant 0 \text{ if } y(x) = b\big\}. \end{aligned} \tag{2.15}$$

The case $\mathbb{I}_g : L^\infty(\Omega) \to \overline{\mathbb{R}}$ is more delicate since in this case $\partial \mathbb{I}_g(y)$ takes values in a measure space on Ω (see [32]).

Generalized gradients Let X be a Banach space of norm $\|\cdot\|$ and dual X^*. The function $f : X \to \mathbb{R}$ is said to be *locally Lipschitz continuous* if for any bounded subset M of X there exists a constant L_M such that

$$\big\|f(x) - f(y)\big\| \leqslant L_M \|x - y\|, \quad \forall x, y \in M.$$

The *directional derivative* of f in x is defined by

$$f^0(x, v) = \limsup_{\substack{y \to x \\ \lambda \to 0}} \lambda^{-1}\big(f(y + \lambda v) - f(y)\big). \tag{2.16}$$

The function f^0 is finite, positively homogeneous in v and subadditive. Then, by the Hahn–Banach theorem, $\exists \eta \in X^*$ such that

$$(\eta, v) \leqslant f^0(x, v), \quad \forall v \in X.$$

By definition, the *generalized gradient* of f in x, denoted $\partial f(x)$, is the set

$$\partial f(x) = \big\{\eta \in X^*;\ (\eta, v) \leqslant f^0(x, v),\ \forall v \in X\big\}. \tag{2.17}$$

PROPOSITION 2.8 (See [10,12]). *For each $x \in X$, $\partial f(x)$ is a convex and w^*-compact subset of X^*. Moreover,*

$$f^0(x, v) = \sup\big\{(\eta, v);\ \eta \in \partial f(x)\big\}, \quad \forall v \in X$$

and the map $\partial f : X \to 2^{X^*}$ is weakly star upper semicontinuous, i.e. if $x_n \to x$ and $\eta_n \to \eta$ weakly star in X^*, then $\eta \in \partial f(x)$.

If f is locally Lipschitz and Gâteaux differentiable, then $\partial f = Df$. Moreover, if f is convex and locally Lipschitz, then ∂f is precisely the subdifferential of f.

Given a closed subset C of X, denote by d_C the distance function

$$d_C(x) = \inf\big\{\|x - y\|;\ y \in C\big\}, \quad \forall x \in X.$$

We can see that d_C is Lipschitzian

$$\left|d_C(x) - d_C(y)\right| \leqslant \|x - y\|, \quad \forall x, y \in X.$$

Let $x \in C$. The element $v \in X$ is said to be tangent to C in x if

$$d_C^0(x, v) = 0.$$

The set of all tangent elements v is denoted $T_C(x)$ (the *tangent cone* to C at x). The *normal cone* $N_C(x)$ to C at x is by definition

$$N_C(x) = \left\{\eta \in X^*;\ (\eta, v) \leqslant 0,\ \forall v \in T_C(x)\right\}.$$

PROPOSITION 2.9. *The vector $h \in X$ is tangent to C in x if and only if $\forall\{x_n\} \subset C$ convergent to x and each $\{\lambda_n\} \to 0$, there is $\{h_n\} \to h$ such that*

$$x_n + \lambda_n h_n \in C, \quad \forall n.$$

PROPOSITION 2.10. *If f, g are locally Lipschitzian, then*

$$\partial(f + g)(x) \subset \partial f(x) + \partial g(x), \quad \forall x \in X.$$

If C is a closed subset of X and if f attains its minimum on C in x, then

$$0 \in \partial f(x) + N_C(x).$$

We refer to the book [12] for further properties of generalized gradients.

2.2. *Ekeland's variational principle*

Here we shall briefly recall, without proof, an important result known in literature as *Ekeland variational principle* [21].

THEOREM 2.1. *Let X be a complete metric space and $F : X \to \mathbb{R}$ be a l.s.c. function, $\not\equiv +\infty$ and bounded from below. Let $\varepsilon > 0$ and $x \in X$ be such that*

$$F(x) \leqslant \inf\left\{F(y);\ y \in X\right\} + \varepsilon. \tag{2.18}$$

Then there exists $x_\varepsilon \in X$ such that

$$F(x_\varepsilon) \leqslant F(x), \tag{2.19}$$

$$d(x_\varepsilon, x) \leqslant \sqrt{\varepsilon}, \tag{2.20}$$

$$F(x_\varepsilon) < F(y) + \sqrt{\varepsilon}\, d(x_\varepsilon, y), \quad \forall y \neq x_\varepsilon. \tag{2.21}$$

Roughly speaking, Theorem 2.1 *says that x_ε is a minimum point of the function*

$$y \to F(y) + \sqrt{\varepsilon} d(x_\varepsilon, y).$$

COROLLARY 2.1. *let X be a Banach space and $F : X \to \mathbb{R}$ be Gâteaux differentiable and bounded from below. Then $\forall \varepsilon > 0$, $\exists x_\varepsilon \in X$ such that*

$$F(x) \leqslant \inf\{F(y);\ y \in X\} + \varepsilon, \tag{2.22}$$

$$|\nabla F(x_\varepsilon)| \leqslant \sqrt{\varepsilon}. \tag{2.23}$$

One may thus construct a minimizing sequence of almost critical points.

2.3. *Elements of differential geometry and exponential representation of flows*

In what follows we present some basic facts concerning the operator calculus introduced by A. Agrachev and R. Gamkrelidze (see [1,2,23]) called exponential representation of flows or chronological calculus. This is a very elegant tool that allows to replace nonlinear objects such as manifolds, tangent vector fields, flows, diffeomorphisms with linear ones which will be functionals and operators on the algebra $C^\infty(M)$ of real infinitely differentiable functions on M. At the end of the section a variation of parameters formula will be given; this formula will show to be very useful in proving the geometric form of Pontriaghin maximum principle. We follow essentially the description in [2].

Differential equations on manifolds In what follows M is a smooth n-dimensional manifold, $TM = \bigcup_{y \in M} T_y M$ is the tangent bundle.

We consider the Cauchy problem for the nonautonomous ordinary differential equation:

$$\begin{cases} y' = f^t(y) := f(t, y), \\ y(0) = y_0 \end{cases} \tag{2.24}$$

where f^t is a *nonautonomous vector field* on M, that is $f^t(y) \in T_y M$ for any $y \in M$, $t \in \mathbb{R}$. In the case $M = \mathbb{R}^n$ or a subdomain of $\mathbb{R}^n$ we have the following classical theorem of Carathéodory (see [15, Chapter 2, Theorem 1.1]):

THEOREM 2.2. *If f is measurable in t for each fixed y and continuous in y for every fixed t and there exists a L^1 function m_0 such that in a neighborhood of* $(0, y_0)$

$$|f(t, y)| \leqslant m_0(t),$$

then problem (2.24) *has a local solution in the extended sense* (*see Section* 1.2).

If for any fixed t, $f_i(t, \cdot)$ is C^1 and for any $(\bar{t}, \bar{y})$ there exists an L^1 function m_1 and neighborhood of $(\bar{t}, \bar{y})$ such that for any (t, y) in this neighborhood

$$\left| \frac{\partial f_i}{\partial y_j}(t, y) \right| \leqslant m_1(t),$$

then the solution is unique. Moreover, under this assumption the solution is C^1 *with respect to the initial data.*

In order to solve equation (2.24) in the case of a general manifold M, we represent it in local coordinates. Let $\varphi : \mathcal{N}(y_0) \subset M \to \mathcal{N}(x_0) \subset \mathbb{R}^n$, a local chart. In these coordinates the vector field f^t is represented as:

$$\left(\varphi_* f^t\right)(x) = \sum_i^n \tilde{f}_i(t,x) \frac{\partial}{\partial x_i} = \tilde{f}^t(x).$$

Here φ_* is the tangent map, or differential of φ. Solving problem (2.24) is equivalent to solving the following Cauchy problem in $\mathbb{R}^n$:

$$\begin{cases} x' = \tilde{f}(t,x), \\ x(0) = x_0. \end{cases} \tag{2.25}$$

In order to insure existence and uniqueness of a local solution, we will assume that $\tilde{f}$ satisfies the hypotheses of Theorem 2.2 which are in fact hypothesis on f since they do not depend on the choice of the local chart. Under these hypothesis, by the theorem of Carathéodory, problem (2.25) has a unique local solution $x(t, x_0)$ which is absolutely continuous with respect to t and C^1 with respect to the initial data x_0 and satisfies the equation almost everywhere. The solution of (2.24) is $y(t, y_0) = \varphi^{-1}(x(t, x_0))$ and one may prove that this is independent of the local chart. The solution of the Cauchy problem (2.24) is defined on a maximal interval that we will suppose to be $\mathbb{R}$ for all initial data. Such vector fields that determine global flows are called *complete*. This always happens if the manifold M is compact.

If we denote by F^t the *flow* defined by the equation (2.24): $F^t(y_0) = y(t, y_0)$, then $F^t \in \mathrm{Diff}(M)$ the set of diffeomorphisms of the manifold M and equation (2.24) may be written

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} F^t(y) = f^t \circ F^t(y), \quad y \in M, \\ F^0 = \mathrm{Id}. \end{cases} \tag{2.26}$$

HYPOTHESES. We will suppose from now on that M is a C^∞ manifold and $\mathrm{Diff}(M)$ denotes the set of C^∞ diffeomorphisms of M. Moreover, we will suppose that the nonautonomous vector field f^t is complete and in any local chart $\tilde{f}(t,x)$ is measurable with respect to t for any fixed x and C^∞ with respect to x for every fixed t and there exist locally integrable functions $m_k(t)$ such that locally

$$\left|D_x^k \tilde{f}(t,x)\right| \leqslant m_k(t).$$

These hypotheses insure that the Cauchy problem (2.24) has unique solution depending C^∞ on the initial data.

Exponential representation of flows We describe in the sequel how the chronological exponential is defined and we will see that topological and differential structures are translated in the new language into the weak convergence of functionals and operators.

Points are represented as algebra homomorphisms from $C^\infty(M)$ to $\mathbb{R}$. If $y \in M$ then it defines an algebra homomorphism $\hat{y} : C^\infty(M) \to \mathbb{R}$, $\hat{y}(\alpha) = \alpha(y)$. One may prove that for any algebra homomorphism $\psi : C^\infty(M) \to \mathbb{R}$, there exists an unique $y \in M$ such that $\psi = \hat{y}$ (see [2]).

Diffeomorphisms of the manifold M are represented as automorphisms of the algebra $C^\infty(M)$. More precisely, if $F \in \mathrm{Diff}(M)$ we define $\widehat{F} : C^\infty(M) \to C^\infty(M)$ as $\widehat{F}(\alpha) = \alpha \circ F$. More generally, if $F : M \to N$ is a smooth map between two manifolds, then it defines an algebra homomorphism $\widehat{F} : C^\infty(N) \to C^\infty(M)$ as $\widehat{F}(\beta) = \beta \circ F$ with $\beta \in C^\infty(N)$. Observe that if $F, G \in \mathrm{Diff}(M)$ then $\widehat{F \circ G} = \widehat{G} \circ \widehat{F}$.

Tangent vectors. Let $f \in T_y M$. Then, as is well known f may be seen either as tangent vector in y to a curve passing through y or as directional derivative, or Lie derivative, of functions in the point y in the direction f. For the first point of view one considers a smooth curve $y(t)$, $y(0) = y$, $y'(0) = v$. The second point of view is to consider the Lie derivative $L_f \alpha = \frac{\mathrm{d}}{\mathrm{d}t} \alpha(y(t))|_{t=0}$. Through the representation described above, we may construct $\hat{f} : C^\infty(M) \to \mathbb{R}$, $\hat{f}(\alpha) := \frac{\mathrm{d}}{\mathrm{d}t}[\hat{y}(t)(\alpha)]|_{t=0} = L_f \alpha$. Obviously, $\hat{f}$ is a linear functional on $C^\infty(M)$ and satisfies the Leibnitz rule

$$\hat{f}(\alpha\beta) = \alpha(y)\hat{f}(\beta) + \hat{f}(\alpha)\beta(y). \tag{2.27}$$

Any linear functional on $C^\infty(M)$ satisfying (2.27) corresponds in this way to a tangent vector.

Vector fields. Let $\mathrm{Vec}(M)$ be the set of smooth vector fields on M and let $f \in \mathrm{Vec}(M)$. Then f defines a linear operator $\hat{f} : C^\infty(M) \to C^\infty(M)$, $\hat{f}(\alpha)(y) = \widehat{f(y)}(\alpha)$. This operator satisfies the Leibnitz rule

$$\hat{f}(\alpha\beta) = \alpha\hat{f}(\beta) + \hat{f}(\alpha)\beta. \tag{2.28}$$

Any linear functional of $C^\infty(M)$ satisfying (2.28) is called *derivation* and corresponds to a unique vector field.

We study now the behaviour of tangent vectors and vector fields under the action of diffeomorphisms.

Let $F \in \mathrm{Diff}(M)$ and $g \in T_y M$ such that $g = \frac{\mathrm{d}}{\mathrm{d}t} y(t)|_{t=0}$. Then $F_* g \in T_{F(y)} M$ and is defined as $F_* g = \frac{\mathrm{d}}{\mathrm{d}t} F(y(t))|_{t=0}$. So, if $\alpha \in C^\infty(M)$, then

$$\widehat{F_* g}(\alpha) = \frac{\mathrm{d}}{\mathrm{d}t} \widehat{F\big(y(t)\big)}(\alpha)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} \alpha\big(F\big(y(t)\big)\big)\bigg|_{t=0} = \hat{g}(\alpha \circ F) = \hat{g} \circ \widehat{F}(\alpha).$$

So,

$$\widehat{F_* g} = \hat{g} \circ \widehat{F}. \tag{2.29}$$

In the same way, if $g \in \mathrm{Vec}(M)$, since $\widehat{g(y)} = \hat{y} \circ \hat{g}$, $F_*\widehat{g(F(y))} = \widehat{F(y)} \circ \widehat{F_*g} = \hat{y} \circ \widehat{F} \circ \widehat{F_*g}$. On the other hand $F_*\widehat{g(F(y))} = \widehat{F_*(g(y))} = \hat{y} \circ \hat{g} \circ \widehat{F}$. As y is arbitrary, $\widehat{F} \circ \widehat{F_*g} = \hat{g} \circ \widehat{F}$ so

$$\widehat{F_*g} = \widehat{F}^{-1} \circ \hat{g} \circ \widehat{F} = \mathrm{Ad}\, \widehat{F}^{-1} \hat{g}. \tag{2.30}$$

REMARK 2.1. The notation Ad comes from the theory of Lie groups where it stands for the adjoint representation of the group in the space of linear operators of the associated Lie algebra. In our situation the group of diffeomorphisms of the manifold stands for the Lie group and the associated Lie algebra is the algebra of vector fields. Through the described representation one obtains the group of automorphisms of $C^\infty(M)$ and the associated Lie algebra is the algebra of derivations of $C^\infty(M)$ (see, e.g., [3,28]).

The equation (2.24) becomes, through the described representation:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} \widehat{y(t)} = \widehat{y(t)} \circ \hat{f}^t, \\ \hat{y}(0) = \hat{y}_0 \end{cases} \tag{2.31}$$

so the flow defined by the equation satisfies

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} \widehat{F^t} = \widehat{F^t} \circ \hat{f}^t, \\ F^0 = \mathrm{Id}. \end{cases} \tag{2.32}$$

The flow $\widehat{F^t}$ is called *the right chronological exponential* and, in accordance with the linear case, is denoted by

$$\widehat{F^t} = \overrightarrow{\exp} \int_0^t \hat{f}^s \, \mathrm{d}s. \tag{2.33}$$

In order to simplify notations, we will omit from now on the hat $\widehat{\ }$ unless confusion is possible and, usually, when we refer to diffeomorphisms and vector fields we mean their representations.

We observe however that at this point equations (2.31), (2.32) are not completely rigorous since we have not yet defined a topology in the corresponding spaces of functionals or operators on $C^\infty(M)$.

Topology We consider on $C^\infty(M)$ the topology of uniform convergence on compacta of all derivatives. More precisely, if $M = \Omega \subset \mathbb{R}^n$, for $\alpha \in C^\infty(M)$, $K \Subset M$ and $k = (k_1, \ldots, k_n)$, $k_i \geqslant 0$, we define the seminorms:

$$\|\alpha\|_{s,K} = \sup\bigl\{\bigl|D^k \alpha(y)\bigr|;\ |k| = k_1 + \cdots + k_n \leqslant s, y \in K\bigr\}.$$

This family of seminorms determines a topology on $C^\infty(M)$ which becomes a Fréchet space (locally convex topological linear space with a complete metric topology given by a translation invariant metric). In this topology $\alpha_m \to \alpha$ iff $\|\alpha_m - \alpha\|_{s,K} \to 0$ for all $s \geqslant 0$ and $K \Subset M$.

In the case of a general manifold, we choose a locally finite covering of M with charts $(V_i, \varphi_i)_{i\in I}$, $\varphi : Vi \to \mathcal{O}_i \subset \mathbb{R}^n$ diffeomorphisms and let $\{\alpha_i\}_{i\in I}$ be a partition of unity subordinated to this covering. We define the family of seminorms

$$\|\alpha\|_{s,K} = \sup\bigl\{D^k\bigl[(\alpha_i\alpha)\circ\varphi^{-1}\bigr](y) \bigm| |k| \leqslant s, \varphi^{-1}(y) \in K, i \in I\bigr\}.$$

This family of seminorms depends on the choice of the atlas but the topology defined on $C^\infty(M)$ is independent of this choice. One could also proceed by using the Whitney theorem and considering M as a submanifold of some Euclidean space.

Once we have defined the topology on $C^\infty(M)$ we consider the space of linear continuous operators $L(C^\infty(M))$. The spaces $\mathrm{Diff}(M)$ and $\mathrm{Vec}(M)$, through the representation are linear subspaces. Indeed, one may easily verify that for $f \in \mathrm{Vec}(M)$ and $F \in \mathrm{Diff}(M)$

$$\|\hat{f}\alpha\|_{s,K} \leqslant C_1\|\alpha\|_{s+1,K}, \qquad \|\widehat{F}\alpha\|_{s,K} \leqslant C_2\|\alpha\|_{s,K}$$

where the constants $C_1 = C_1(s, K, f)$, $C_2 = C_2(s, K, F)$. We thus define a family of seminorms on $\mathrm{Vec}(M)$, respectively $\mathrm{Diff}(M)$:

$$\|f\|_{s,K} = \sup\bigl\{\|\hat{f}\alpha\|_{s+1,K} \bigm| \|\alpha\|_{s,K} = 1\bigr\},$$
$$\|F\|_{s,K} = \sup\bigl\{\|\widehat{F}\alpha\|_{s,K} \bigm| \|\alpha\|_{s,K} = 1\bigr\},$$

which define locally convex topologies. On these spaces we also may consider the weak topology induced from $C^\infty(M)$: $F_n \to F$ iff $F_n\alpha \to F\alpha$ for all $\alpha \in C^\infty(M)$ (the same for a sequence of vector fields).

Differentiability and integrability of families of functions or operators First of all we define these properties on $C^\infty(M)$ which is a Fréchet space. In general, let X be a Fréchet space whose topology is defined by the family $\{p_k\}_{k\in\mathbf{N}}$ of seminorms. The metric on X is defined by

$$d(x, y) = \sum_{k\in\mathbf{N}} \frac{1}{2^k}\frac{p_k(x-y)}{1 + p_k(x-y)}.$$

Let $h : J \subset \mathbb{R} \to X$. The function h is differentiable in t_0 if there exists in X the limit

$$\lim_{t\to t_0} \frac{h(t) - h(t_0)}{t - t_0}.$$

The function h is Lipschitz continuous if $p_k \circ h$ is Lipschitz for all p_k. Differentiability and Lipschitz continuity may also be defined using the metric structure of X.

The function h is bounded if $p_k \circ h$ is bounded for all p_k.

For measurability and integrability we adapt the plan of development for the *Bochner integral* (see, e.g., [35]). A function h is called a step function if it may be represented as

$$h = \sum_{n \in \mathbf{N}} x_n \chi_{J_n}$$

where χ_{J_n} is the characteristic function of a measurable subset $J_n \subset J$. We call such a representation of h a σ-representation and it is obvious that this is not unique. We say that the function h is *strongly measurable* if h is the limit a.e. of a sequence of step functions. The function h is *weakly measurable* if $x^* \circ h$ is measurable for all $x^* \in X^*$. One may prove that if X is separable the two notions of measurability coincide (see Pettis theorem in [35] in the case X is a Banach space). If h is a step function then h is *integrable* if

$$\sum_n \mu(J_n) p_k(x_n) \leqslant \infty$$

for all the seminorms p_k. The integral of h is then defined as

$$\int_J h(t)\, \mathrm{d}t = \sum_n \mu(J_n) x_n$$

and it may be shown that it is independent of the order of summation and of the σ-representation of h.

If h is a measurable function we say that it is *integrable* if there exists a sequence of integrable step functions $\{h_n\}_{n \in \mathbf{N}}$ such that for all k

$$\lim_{n \to +\infty} \int_J p_k\bigl(h(t) - h_n(t)\bigr)\, \mathrm{d}t = 0.$$

In this case one may show that there exists

$$\lim_{n \to +\infty} \int_J h_n(t)\, \mathrm{d}t$$

and this is independent of the sequence $\{h_n\}_{n \in \mathbf{N}}$ with the given properties. The limit is denoted by

$$\int_J h(t)\, \mathrm{d}t$$

and is the integral of h on J.

For a family P^t, $t \in J \subset \mathbb{R}$ of linear continuous operators or linear continuous functionals on $C^\infty(M)$ the above notions (continuity, differentiability, boundedness, measurability, integrability) will be considered in the weak sense, that is the function $t \to P^t$ has one of these properties if $P^t \circ \alpha$ has the corresponding property for all $\alpha \in C^\infty(M)$. We will not discuss here the relation between the strong and weak properties.

At this point we see that the operator equation (2.32) makes sense and it can be easily proved that it has a unique solution. We point out the Leibnitz rule:

$$\frac{\mathrm{d}}{\mathrm{d}t} P^t \circ Q^t \bigg|_{t=t_0} = \frac{\mathrm{d}}{\mathrm{d}t} P^t \bigg|_{t=t_0} \circ Q^{t_0} + P^{t_0} \circ \frac{\mathrm{d}}{\mathrm{d}t} Q^t \bigg|_{t=t_0}$$

for two functions $t \to P^t$, $t \to Q^t$ differentiable at t_0.

Consider now the flow F^t defined by (2.24) and $G^t = (F^t)^{-1}$. If we differentiate the identity $F^t \circ G^t = I$ we obtain

$$F^t \circ f^t \circ G^t + F^t \circ \frac{\mathrm{d}}{\mathrm{d}t} G^t = 0$$

and thus

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} G^t = -f^t \circ G^t, \\ G^0 = \mathrm{Id}. \end{cases} \tag{2.34}$$

We define thus the *left chronological exponential*:

$$G^t = \overleftarrow{\exp} \int_0^t -f^t \, \mathrm{d}t.$$

Further properties and extensions We have seen that $\widehat{F_* g} = \mathrm{Ad}\, \widehat{F}^{-1} \hat{g}$ for $F \in \mathrm{Diff}(M)$, $g \in \mathrm{Vec}(M)$. We compute now the differential $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \mathrm{Ad}(\widehat{F^t})$ for a flow F^t on M such that

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} F^t \bigg|_{t=0} = f \in \mathrm{Vec}\, M, \\ F^0 = \mathrm{Id}. \end{cases}$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \big(\mathrm{Ad}\, F^t\big) g = f \circ g - g \circ f = [f, g] =: (\mathrm{ad}\, f) g.$$

In the particular case

$$F^t = \overrightarrow{\exp} \int_0^t f^s \, \mathrm{d}s$$

we obtain:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} (\mathrm{Ad}\, F^t) g = \big(\mathrm{Ad}\, F^t\big)\, \mathrm{ad}\, f^t g, \\ \mathrm{Ad}\, F^0 = \mathrm{Id} \end{cases}$$

so we may write, formally:

$$\mathrm{Ad}\left(\overrightarrow{\exp}\int_0^t f^s\,\mathrm{d}s\right)=\overrightarrow{\exp}\left(\int_0^t \mathrm{ad}\,f^s\,\mathrm{d}s\right).$$

Let now $F\in\mathrm{Diff}(M)$ and g^t a nonautonomous vector field. Then

$$F\circ\overrightarrow{\exp}\int_0^t g^s\,\mathrm{d}s\circ F^{-1}=\overrightarrow{\exp}\int_0^t (\mathrm{Ad}\,Fg^s)\,\mathrm{d}s. \tag{2.35}$$

Indeed, the both sides of the equality verify the same Cauchy problem for the operator equation

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}t}q^t=q^t\circ\left(\mathrm{Ad}\,F\ g^t\right),\\ q^0=\mathrm{Id}\end{cases}$$

and thus, by uniqueness of the solution, they coincide.

Now if we take again $G^t=(F^t)^{-1}$ and if we differentiate the identity $G^t\circ F^t=\mathrm{Id}$ we obtain that $\frac{\mathrm{d}}{\mathrm{d}t}G^t\circ F^t=-G^t\circ F^t\circ f^t$ and thus

$$\frac{\mathrm{d}}{\mathrm{d}t}G^t=-G^t\circ\left(\mathrm{Ad}\,F^t\right)f^t.$$

This gives the relationship between left and right chronological exponentials:

$$\overleftarrow{\exp}\int_0^t f^s\,\mathrm{d}s=\overrightarrow{\exp}\int_0^t\left(\mathrm{Ad}\,F^s\right)f^s\,\mathrm{d}s. \tag{2.36}$$

If $F\in\mathrm{Diff}(M)$, as we have seen, it defines an algebra automorphism of $C^\infty(M):\widehat{F}\alpha=\alpha\circ F=F^*\alpha$, where F^* is the pull back of C^∞ differential forms defined by F. This suggests the fact that $\widehat{F}$ may be extended, as algebra automorphism to the graded algebra $\Lambda(M)=\bigoplus\Lambda^k(M)$ of differential forms. If $\omega\in\Lambda^k(M)$ then we define

$$\widehat{F}\omega:=F^*\omega.$$

It is well known that F^* commutes with exterior differential:

$$F^*\circ d=d\circ F^*$$

and for $\omega_i\in\Lambda^{k_i}(M)$,

$$F^*(\omega_1\wedge\omega_2)=F^*(\omega_1)\wedge F^*(\omega_2).$$

So $\widehat{F}$ is an algebra automorphism for $\Lambda(M)$. We now consider a vector field:

$$\frac{\mathrm{d}}{\mathrm{d}t}F^t|_{t=0}=f,\qquad F^0=\mathrm{Id}\,.$$

The action on $\Lambda^k(M)$ is the Lie derivative of differential forms:

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{F^t}\omega\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\big(F^t\big)^*\omega\bigg|_{t=0} = L_f\omega.$$

We may thus understand the action of a vector field $f \in \mathrm{Vec}(M)$ as the Lie derivative of differential forms:

$$\hat{f} = L_f.$$

The chronological exponential may be thus written

$$F^t = \overrightarrow{\exp}\int_0^t L_{f^s}\,\mathrm{d}s.$$

We point out two fundamental properties of the Lie derivative:

Since $\widehat{F^t} \circ d = d \circ \widehat{F^t}$ one obtains that

$$\hat{f} \circ d = d \circ \hat{f} \quad \text{(equivalently } L_f \circ d = d \circ L_f\text{)}.$$

Denote by i_f the interior product of a differential form ω with a vector field f: $i_f\omega(f_1, \ldots, f_k) = \omega(f, f_1, \ldots, f_k)$, for $\omega \in \Lambda^k(M)$, $f_i \in \mathrm{Vec}\,M$. Then the classical *Cartan's formula* reads:

$$\hat{f} = d \circ i_f + i_f \circ d. \tag{2.37}$$

Variation of parameters formula Consider the Cauchy problem for the linear differential equation in $\mathbb{R}^n$:

$$\begin{cases} y' = Ay + b(t), \\ y(0) = y_0. \end{cases}$$

The solution of the homogeneous equation ($b \equiv 0$) is $y(t) = \mathrm{e}^{At}y_0$. For the nonhomogeneous equation a solution may be found by the *variation of constants* or *variation of parameters method*. This consists in searching a solution of the form $y(t) = \mathrm{e}^{At}c(t)$ and an equation for $c(t)$ is obtained: $c'(t) = A(t)b(t)$. The solution is given by the *variation of constants formula*

$$y(t) = \mathrm{e}^{At}y_0 + \int_0^t \mathrm{e}^{A(t-s)}b(s)\,\mathrm{d}s. \tag{2.38}$$

We consider now the nonlinear differential equation

$$y' = f^t(y)$$

which generates the flow $F^t = \overrightarrow{\exp}\int_0^t f^s\,\mathrm{d}s$. We consider also the perturbed equation

$$y' = f^t(y) + g^t(y)$$

which generates the flow $H^t = \overrightarrow{\exp}\int_0^t f^s + g^s\,\mathrm{d}s$, depending on the perturbation g^t. We want to find an expression for this dependence. For this purpose one proceeds as in the linear case and search H^t in the form

$$H^t = G^t \circ F^t$$

where the flow G^t has to be deduced. Differentiating this equality we find

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}H^t &= G^t \circ F^t \circ \left(f^t + g^t\right)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}G^t \circ F^t + G^t \circ \frac{\mathrm{d}}{\mathrm{d}t}F^t = \frac{\mathrm{d}}{\mathrm{d}t}G^t \circ F^t + G^t \circ F^t \circ f^t .\end{aligned}$$

So,

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}t}G^t = G^t \circ F^t \circ f^t \circ \left(F^t\right)^{-1} = G^t \circ \mathrm{Ad}\,F^t g^t ,\\ G^0 = \mathrm{Id}\,.\end{cases}$$

We thus obtained

$$G^t = \overrightarrow{\exp}\int_0^t \mathrm{Ad}\,F^s g^s\,\mathrm{d}s$$

and the first form of variations formula

$$H^t = \overrightarrow{\exp}\int_0^t \mathrm{Ad}\,F^s g^s\,\mathrm{d}s \circ F^t . \tag{2.39}$$

We also obtain

$$H^t = F^t \circ \mathrm{Ad}\big(F^t\big)^{-1}\left(\overrightarrow{\exp}\int_0^t \mathrm{Ad}\,F^s g^s\,\mathrm{d}s\right)$$

and by (2.35)

$$H^t = F^t \circ \overrightarrow{\exp}\int_0^t \mathrm{Ad}\big[\left(F^t\right)^{-1} \circ F^s\big]g^s\,\mathrm{d}s = F^t \circ \overrightarrow{\exp}\int_0^t \left(F^t_s\right)_* g^s\,\mathrm{d}s \tag{2.40}$$

where $F^t_s = \overrightarrow{\exp}\int_s^t f^\tau\,\mathrm{d}\tau$.

The second form of variations formula may be thus written

$$\overrightarrow{\exp}\int_0^t f^s + g^s\,\mathrm{d}s = \overrightarrow{\exp}\int_0^t f^s\,\mathrm{d}s \circ \overrightarrow{\exp}\int_0^t \overrightarrow{\exp}\int_t^s \mathrm{ad}\, f^\tau\,\mathrm{d}\tau g^s\,\mathrm{d}s.$$

In the case when f, g are autonomous vector fields the formula becomes (compare with (2.38)):

$$\mathrm{e}^{t(f+g)} = \overrightarrow{\exp}\int_0^t \mathrm{Ad}\,\mathrm{e}^{sf} g\,\mathrm{d}s \circ \mathrm{e}^{tf} = \mathrm{e}^{tf} \circ \overrightarrow{\exp}\int_0^t \mathrm{Ad}\,\mathrm{e}^{(s-t)f} g\,\mathrm{d}s.$$

Elements of symplectic geometry. Hamiltonian formalism

DEFINITION 2.1. A symplectic structure on a (necessarily odd dimensional) manifold N is a nondegenerate closed differential 2-form. A manifold with a symplectic structure ω is called a symplectic manifold

Let M be a manifold and $T^*M = \bigcup_{y\in M} T_q^*M$ be the cotangent bundle. If $(x_1, \dots, x_n)$ are local coordinates on M then if $p \in T_y^*M$, $p = \sum_{i=1}^n p_i\,\mathrm{d}x_i$, $(p_1, \dots, p_n, x_1, \dots, x_n)$ define the canonical local coordinates on T^*M. Define

$$\omega = \sum_{i=1}^n \mathrm{d}p_i \wedge \mathrm{d}x_i. \tag{2.41}$$

To see that the definition is independent of the local coordinates let $\pi : T^*M \to M$ be the canonical projection and the canonical 1-form on T^*M:

$$\omega_{1\xi}(w) = \xi \circ \pi_*(w), \quad \text{for } w \in T_\xi\big(T^*M\big).$$

If $w \in T(T^*M)$ then

$$w = \sum_{i=1}^n \xi_i \frac{\partial}{\partial p_i} + v_i \frac{\partial}{\partial x_i}.$$

Since

$$\pi_*\left(\frac{\partial}{\partial p_i}\right) = 0 \quad \text{and} \quad \pi_*\left(\frac{\partial}{\partial x_i}\right) = \frac{\partial}{\partial x_i}$$

one finds that

$$\omega_{1\xi}(w) = \sum_{i=1}^n p_i v_i$$

so

$$\omega_1 = \sum_{i=1}^{n} p_i \, \mathrm{d}x_i$$

and

$$\omega = \mathrm{d}\omega_1 .$$

It is easy to see now that ω is a symplectic structure on T^*M which becomes a symplectic manifold.

Let now (N, ω) be a general symplectic manifold. Functions in $C^\infty(N)$ are called *Hamiltonians*. Let H be such a Hamiltonian. Then there exists a unique vector field on N denoted $\overrightarrow{H}$ such that

$$-i_{\overrightarrow{H}}\,\omega = \omega\big(\cdot, \overrightarrow{H}\big) = \mathrm{d}H .$$

$\overrightarrow{H}$ is called the *Hamiltonian vector field* of H and the corresponding flow is the *Hamiltonian flow*. The *Hamiltonian equation* is

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi(t) = \overrightarrow{H}\big(\xi(t)\big) \tag{2.42}$$

and the Hamiltonian flow is

$$\Xi^t = \overrightarrow{\exp} \int_0^t \overrightarrow{H} \, \mathrm{d}s .$$

The *Poisson bracket* of the Hamiltonians α, β is defined as

$$\{\alpha, \beta\} = L_{\overrightarrow{\alpha}}\,\beta = \mathrm{d}\beta(\overrightarrow{\alpha}) = \omega\big(\overrightarrow{\alpha}, \overrightarrow{\beta}\big) = -\{\beta, \alpha\}.$$

One may prove that $(C^\infty(N), \{\cdot, \cdot\})$ is a Lie algebra and the map $H \to \overrightarrow{H}$ is a Lie algebra homomorphism from $C^\infty(N)$ to $\mathrm{Vec}(N)$. Bilinearity and antisymmetry are immediate. Jacobi identity as well as the fact that $\overrightarrow{\{\alpha, \beta\}} = [\overrightarrow{\alpha}, \overrightarrow{\beta}\,]$ are easy to prove if in local coordinates ω has the canonic form (2.41). We conclude since, by *Darboux theorem* (see [3]), there exists indeed a symplectic atlas on N such that ω in local coordinates is in canonical form. In these coordinates

$$\overrightarrow{H} = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\frac{\partial}{\partial x_i} - \frac{\partial H}{\partial x_i}\frac{\partial}{\partial p_i}$$

and

$$\{\alpha, \beta\} = \sum_{i=1}^{n} \frac{\partial \alpha}{\partial p_i}\frac{\partial \beta}{\partial x_i} - \frac{\partial \alpha}{\partial x_i}\frac{\partial \beta}{\partial p_i}.$$

Moreover, the Hamiltonian system (2.42) is written in the form (1.7).

Now, if $F \in \operatorname{Diff}(N)$ preserves the symplectic structure, that is $F^*\omega = \omega$, then $\operatorname{Ad} F\overrightarrow{H} = \overrightarrow{FH}$. Indeed,

$$\begin{aligned}\omega\big(\cdot, \overrightarrow{FH}\big) &= \mathrm{d}(FH) = \mathrm{d}(H \circ F) = H_* \circ F_* = \mathrm{d}H \circ F_* \\ &= \omega\big(F_* \cdot, \overrightarrow{H}\big) = F^*\omega\big(\cdot, (F_*)^{-1}\overrightarrow{H}\big) = \omega\big(\cdot, \operatorname{Ad} F\overrightarrow{H}\big).\end{aligned}$$

Prime integrals of Hamiltonian systems are those which commute with the Hamiltonian. Indeed, consider the equation (2.42). Then $\alpha \in C^\infty(M)$ is a prime integral iff $\mathrm{e}^{t\overrightarrow{H}}\alpha = \text{const}$ which is equivalent, by differentiation, to $\overrightarrow{H}\alpha = 0$ or $\{H, \alpha\} = 0$.

We consider again the case of the symplectic manifold T^*M. Given a nonautonomous vector field $f^t \in \operatorname{Vec}(M)$ we consider the Hamiltonian

$$\big(f^t\big)^{\#}(\xi) = \big(\xi, f^t\big).$$

In canonical symplectic coordinates (p_i, x_i), if $f^t = \sum_{i=1}^n f_i(t,x)\frac{\partial}{\partial x_i}$ then $(f^t)^{\#}(\xi) = \sum_{i=1}^n p_i f_i(t,x)$. Using the canonical coordinates it is easy to see that for $f, g \in \operatorname{Vec}(M)$

$$\big\{f^{\#}, g^{\#}\big\} = [f,g]^{\#}.$$

For a nonautonomous vector field f^t the Hamiltonian vector field on T^*M defined by $\overrightarrow{(f^t)^{\#}}$ is named the *Hamiltonian lift*. It satisfies

$$\pi_* \overrightarrow{\big(f^t\big)^{\#}} = f^t.$$

We want to establish now the relation between the flows determined by f^t, respectively by $\overrightarrow{(f^t)^{\#}}$. Let $F_\tau^t = \overrightarrow{\exp} \int_\tau^t f^s \, \mathrm{d}s$. Then $(F_\tau^t)^* \in \operatorname{Diff}(T^*M)$. Let

$$g^\tau = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=\tau} \big(F_\tau^t\big)^*$$

PROPOSITION 2.11. *$g^t = -\overrightarrow{(f^t)^{\#}}$ and*

$$\big(F_\tau^t\big)^* = \overrightarrow{\exp} \int_t^\tau \overrightarrow{\big(f^s\big)^{\#}} \, \mathrm{d}s \tag{2.43}$$

PROOF. First of all, since, in the exponential representation on T^*M, $(F_\tau^{t+\varepsilon})^* = (F_t^{t+\varepsilon})^* \circ (F_\tau^t)^*$, it follows by differentiating with respect to ε in 0 that

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(F_\tau^t\big)^* = g^t \circ \big(F_\tau^t\big)^*$$

so

$$\big(F_\tau^t\big)^* = \overleftarrow{\exp} \int_\tau^t g^s \, \mathrm{d}s.$$

Since

$$\pi \circ \big(F_\tau^t\big)^* = \big(F_\tau^t\big)^{-1} \circ \pi$$

it follows by differentiation that

$$\pi_* g^t = -f^t. \tag{2.44}$$

On the other hand, the flow $(F_\tau^t)^*$ preserves the 1-form ω_1 and thus the symplectic form ω. By Cartan's formula (2.3)

$$0 = L_{g^t}\omega_1 = i_{g^t}\omega + \mathrm{d}\omega_1\big(g^t\big)$$

which implies that

$$g^t = \overrightarrow{\omega_1\big(g^t\big)}$$

which means that the field g^t is Hamiltonian. By (2.44) and observing that g^t is linear homogeneous on the tangent space to the fibers, the first conclusion follows. Equality (2.43) is now immediate taking into account the relationship between the left and right chronological exponential. □

3. The Pontriaghin maximum principle

The Pontriaghin maximum principle, developed by L. Pontriaghin and his collaborators (see [30]), is a set of first order *necessary conditions* of optimality in optimal control problems. This is expressed in terms of the dual linearized state system and reduces the problem to solving a two point boundary problem for a differential system, which is in fact a Hamiltonian system, composed of the state equation and the dual equation.

3.1. *The main theorem*

Throughout this section the following conditions will be assumed on the optimal control problem (1.14):

(i) The functions

$$L : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}; \quad f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$$

and L_y, f_y are measurable in t, continuous in (y,u) and

$$\|f_y(t,y,u)\| + \|L_y(t,y,u)\| \leqslant \alpha(t,u), \quad \forall t \in [0,T],\ \forall u \in U(t).$$

(ii) For each $t \in [0,T]$, $U(t) \subset \mathbb{R}^m$ is closed and for each closed set $D \subset \mathbb{R}^m$ the set $\{t;\ D \cap U(t) \neq \emptyset\}$ is measurable.

(iii) $g \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$ and C is a closed convex subset of $\mathbb{R}^n \times \mathbb{R}^n$.

Here we have denoted by the same symbol $\|\cdot\|$ the norm in $L(\mathbb{R}^n, \mathbb{R}^n)$ and in $\mathbb{R}^n$.

Theorem 3.1 below is the celebrated Pontryagin's maximum principle for problem (1.14).

THEOREM 3.1. *Let* (y^*, u^*) *be an optimal pair in problem* (1.14) *such that* $\alpha(t,u^*) \in L^1(0,T)$. *Then there exists* $p \in AC([0,T];\mathbb{R}^n)$ *and a constant* λ *which is equal to* 0 *or* 1 *such that*

$$\|p(t)\| + |\lambda| \neq 0, \quad \forall t \in [0,T]$$

and

$$p' = -f_y^*(t,y^*,u^*)p + \lambda L_y(t,y^*,u^*), \quad a.e.\ t \in [0,T], \tag{3.1}$$

$$\{p(0), -p(T)\} \in \lambda \nabla g\big(y^*(0), y^*(T)\big) + N_C\big(y^*(0), y^*(T)\big), \tag{3.2}$$

$$\begin{aligned}&\big(p(t), f\big(t, y^*(t), u^*(t)\big)\big) - \lambda L\big(t, y^*(t), u^*(t)\big)\\ &\quad = \max_{u \in U(t)} \big\{\big(p(t), f\big(t, y^*(t), u\big)\big) - \lambda L\big(t, y^*(t), u\big)\big\} \quad a.e.\ t \in [0,T].\end{aligned} \tag{3.3}$$

If $\lambda = 1$, the problem is called *normal* otherwise it is *abnormal*. If $C = \{y_0\} \times \mathbb{R}^n$ and $g(y_1, y_2) \equiv g_1(y_2)$ this is the *Bolza problem* with initial condition $y(0) = y_0$ and (3.2) reduces to

$$p(T) = -\lambda \nabla g_1\big(y(T)\big). \tag{3.4}$$

If $C = \{(y_1, y_2);\ y_1 = y_2\}$ and $g \equiv 0$, this is the periodic optimal control problem and (3.2) has the form

$$p(0) = p(T). \tag{3.5}$$

The end point boundary conditions (3.2) are also called *transversality conditions*.

An elementary approach to the maximum principle In order to understand better how the optimality system (3.1)–(3.3) appears, as well as the proof we give in the next section, we discuss first the special case where $L(t,\cdot,\cdot)$, $f(t,\cdot,\cdot)$ are smooth (of class C^1 for instance), $g(y_1,y_2) = g_2(y_2) \in C^1$, $C = \{y_0\} \times \mathbb{R}^n$ and $U(t) \equiv U$ is a closed convex subset of $\mathbb{R}^m$. We shall assume also $f_y(t,y^*,u^*) \in L^\infty(0,T,\mathbb{R}^{2n})$, $f_u(t,y^*,u^*) \in L^\infty(0,T,\mathbb{R}^{2n})$, where

(y^*, u^*) is an optimal pair. By optimality we have for $\lambda \in (0,1)$ and $\forall v \in L^2(0,T;\mathbb{R}^m)$, $u^*(t) + \lambda v(t) \in U$ *a.e.* $t \in (0,T)$

$$\begin{aligned}&\int_0^T L(t, y^*, u^*)\,\mathrm{d}t + g_2\big(y^*(T)\big)\\ &\quad\leqslant \int_0^T L\big(t, y^{u^*+\lambda v}, u^* + \lambda v\big)\,\mathrm{d}t + g_2\big(y^{u^*+\lambda v}(T)\big).\end{aligned}$$

Here y^v is the solution of the Cauchy problem $y' = f(t, y, v)$, $y(0) = y_0$. Noticing that

$$z(t) = \lim_{\lambda \to 0} \frac{1}{\lambda}\big(y^{u^*+\lambda v}(t) - y^*(t)\big)$$

is the solution to the system in variations

$$\begin{cases} z' = f_y(t, y^*, u^*)z + f_u(t, y^*, u^*)v, & t \in (0,T),\\ z(0) = 0, \end{cases} \tag{3.6}$$

we find that

$$\int_0^T \big(L_y(t, y^*, u^*), z(t)\big) + \big(L_u(t, y^*, u^*), v(t)\big)\,\mathrm{d}t + \big(\nabla g_2\big(y^*(T)\big), z(T)\big) \geqslant 0. \tag{3.7}$$

Next we define p the solution to the Cauchy problem

$$\begin{cases} p' = -f_y^*(t, y^*, u^*)p + L_y(t, y^*, u^*), & \text{a.e. } t \in (0,T),\\ p(T) = -\nabla g_2\big(y^*(T)\big). \end{cases}$$

Multiplying the latter by z and integrating on $(0,T)$ we find by (3.6), (3.7) that

$$\int_0^T \big(L_u(t, y^*, u^*) - f_u(t, y^*, u^*), v(t)\big)\,\mathrm{d}t \geqslant 0.$$

Equivalently,

$$\int_0^T \big(L_u(t, y^*, u^*) - f_u(t, y^*, u^*), w(t) - u^*(t)\big)\,\mathrm{d}t \geqslant 0$$

for all $w \in L^2(0,T;\mathbb{R}^m)$, $w(t) \in U$ a.e. $t \in (0,T)$. This means (see (2.4)) that

$$f_u\big(t, y^*(t), u^*(t)\big) - L_u\big(t, y^*(t), u^*(t)\big) \in N_U\big(u^*(t)\big) \quad \text{a.e. } t \in (0,T).$$

If f is linear in u and $u \to L(t, y, u)$ is convex the latter is equivalent with (3.3).

3.2. *Proof of the maximum principle*

The main idea of proof given here is essentially due to F. Clarke [11]. We may assume that

$$\left\| f(t, y^*, u)\right\| + \left\| L_y(t, y, u)\right\| + \left\| f_y(t, y, u)\right\| \leqslant \beta(t) \tag{3.8}$$

for a.e. $t \in [0, T]$ and all $(y, u) \in \mathbb{R}^n \times U(t)$ where $\beta \in L^1[0, T]$. This clearly implies that

$$\begin{aligned} &\left\| f(t, x, u) - f(t, y, u)\right\| \leqslant \beta(t)\|x - y\|, \\ &\left\| f(t, y, u)\right\| \leqslant \beta(t)\big(\|y - y^*\| + 1\big) \end{aligned} \tag{3.9}$$

and so (standard existence theory) the state system (1.11) has a unique absolutely continuous solution $y \in \mathrm{AC}([0, T]; \mathbb{R}^n)$ for each function $u(t) \in U(t)$, $\forall t \in (0, T)$. This can be achieved by replacing $U(t)$ by

$$\begin{aligned} U_n(t) = \big\{u \in U(t);\ &\left\| f(t, y^*, u) - f(t, y^*, u^*)\right\| \leqslant n, \\ &\left\| f_y(t, y, u) - f_y(t, y, u^*)\right\| \leqslant n\ \forall y \in \mathbb{R}^n\big\}. \end{aligned}$$

Indeed, if (y^*, u^*) is optimal then clearly it is optimal in problem (1.14) with control constraints $u(t) \in U_n(t)$ a.e. $t \in (0, T)$ and so if the maximum principle (3.1)–(3.3) is true in this case, there are $\{p_n, \lambda_n\}$ satisfying the conditions of Theorem 3.1 where U is replaced by $U_n(t)$.

By (3.1) and by assumption (ii) we see that

$$\left\| p_n'(t)\right\| \leqslant \alpha_0\big(t, u^*(t)\big), \quad \text{a.e. } t \in (0, T).$$

If $\{|p_n(0)|\}$ is bounded this implies that $\{p_n\}$ is compact in $C([0, T]; \mathbb{R}^n)$ and so on a subsequence

$$p_n \to p \quad \text{in } C\big([0, T]; \mathbb{R}^n\big)$$

where p satisfies Eqs. (3.1)–(3.3). Otherwise, we set $q_n = p_n(t)/\|p_n(0)\|$ and conclude as above that

$$p_n \to p \quad \text{in } C\big([0, T]; \mathbb{R}^n\big)$$

where p satisfies Eqs. (3.1)–(3.3) with $\lambda = 0$.

Now we come back to the proof of the maximum principle. We set

$$X = \big\{u : [0, T] \to \mathbb{R}^m \text{ measurable};\ u(t) \in U(t),\ \text{a.e. } t \in [0, T]\big\}. \tag{3.10}$$

It is easily seen that this is a complete metric space with the Ekeland distance

$$d(u, v) = m\big\{t \in [0, T];\ u(t) \neq v(t)\big\}. \tag{3.11}$$

Let $Y = X \times \mathbb{R}^n$ endowed with the standard product metric and consider the function

$$\varphi : Y \to \mathbb{R}, \quad \varphi(u, y_0) = \int_0^T L\big(t, y^u(t), u(t)\big)\,\mathrm{d}t + g\big(y_0, y^u(T)\big) \tag{3.12}$$

where

$$\begin{cases} (y^u)' = f(t, y^u, u) & \text{a.e. in } (0, T), \\ y^u(0) = y_0. \end{cases}$$

We define the function

$$\varphi_\varepsilon(u, y_0) = \Big[\Big(\big(\varphi(u, y_0) - \varphi\big(u^*, y^*(0)\big) + \varepsilon\big)^+\Big)^2 + d_C^2\big(y_0, y^u(T)\big)\Big]^{1/2} \tag{3.13}$$

for all $(u, y_0) \in Y$. Here d_C is the distance to C.

By Ekeland's principle, Theorem 2.1, for each $\varepsilon > 0$ there exists $(u_\varepsilon, y_0^\varepsilon) \in Y$ such that $\varphi_\varepsilon(u_\varepsilon, y_0^\varepsilon) \leqslant \varepsilon$,

$$d(u_\varepsilon, u^*) + \big\| y_0^\varepsilon - y^*(0) \big\| \leqslant \sqrt{\varepsilon} \tag{3.14}$$

and

$$\varphi_\varepsilon(u_\varepsilon, y_0^\varepsilon) \leqslant \varphi_\varepsilon(u, y_0) + \sqrt{\varepsilon}\big(d(u_\varepsilon, u) + \| y_0^\varepsilon - y_0 \|\big), \quad \forall (u, y_0) \in Y. \tag{3.15}$$

We take $y_0^\rho = y_0^\varepsilon + \rho z_0$ and set $y_\varepsilon = y^{u_\varepsilon}$, $y_\varepsilon(0) = y_0^\varepsilon$ with $\|z_0\| = 1$. In (3.15) we take the "spike" admissible control

$$u_\rho(t) = \begin{cases} u(t), & t \in [t_0 - \rho, t_0], \\ u_\varepsilon(t), & t \in [0, T] \setminus (t_0 - \rho, t_0) \end{cases}$$

where $t_0 \in [0, T]$ is arbitrary but fixed and $u \in X$. Then we have

$$y_\rho(t) = y^{u_\rho}(t) = y_\varepsilon(t) + \rho z_\varepsilon(t) + \mathrm{o}(\rho)$$

where $z_\varepsilon \in BV([0, T]; \mathbb{R}^n)$ is the solution to the equation in variations

$$\begin{cases} z_\varepsilon' = f_y(t, y_\varepsilon, u_\varepsilon) z_\varepsilon & \text{a.e. } t \in (0, t_0) \cup (t_0, T), \\ z_\varepsilon(0) = z_0, \\ z_\varepsilon^+(t_0) - z_\varepsilon^-(t_0) = f\big(t_0, y_\varepsilon(t_0), u(t_0)\big) - f\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big). \end{cases} \tag{3.16}$$

(Here $BV([0, T]; \mathbb{R}^n)$ is the space of functions of bounded variation on $[0, T]$.)

By (3.15) we have

$$-2\sqrt{\varepsilon} \leqslant \frac{\varphi_\varepsilon(u_\rho, y_0^\rho) - \varphi_\varepsilon(u_\varepsilon, y_0^\varepsilon)}{\rho}$$

$$= \frac{1}{\varphi_\varepsilon(u_\rho, y_0^\rho) + \varphi_\varepsilon(u_\varepsilon, y_0^\varepsilon)}\Bigg[\frac{((\varphi(u_\rho, y_0^\rho) - \varphi(u^*, y^*(0)) + \varepsilon)^+)^2}{\rho}$$
$$- \frac{((\varphi(u_\varepsilon, y_0^\varepsilon) - \varphi(u^*, y^*(0)) + \varepsilon)^+)^2}{\rho}$$
$$+ \frac{d_C^2(y_0^\rho, y_\rho(T)) - d_C^2(y_0^\varepsilon, y_\varepsilon(T))}{\rho}\Bigg].$$

Letting ρ tend to zero, we obtain after some calculation that

$$-2\sqrt{\varepsilon} \leqslant \frac{(\varphi(u_\varepsilon, y_0^\rho) - \varphi_\varepsilon(u^*, y^*(0)) + \varepsilon)^+}{\varphi_\varepsilon(u_\varepsilon, y_0^\varepsilon)}\Bigg[\int_0^T L_y(t, y_\varepsilon, u_\varepsilon) z_\varepsilon \,\mathrm{d}t$$
$$+ \big(\nabla g\big(y_0^\varepsilon, y_\varepsilon(T)\big), \big(z_0, z_\varepsilon(T)\big)\big) + L\big(t_0, y_\varepsilon(t_0), u(t_0)\big)$$
$$- L\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big)\Bigg]$$
$$+ \frac{(\nabla d_C(y_0^\varepsilon, y_\varepsilon(T)), (z_0, z_\varepsilon(T)))}{\varphi_\varepsilon(u_\varepsilon, y_0^\varepsilon)} d_C\big(y_0^\varepsilon, y_\varepsilon(T)\big).$$

Here t_0 is a Lebesgue point for $L(t, y_\varepsilon(t), u_\varepsilon(t))$ and $L(t, y_\varepsilon(t), u(t))$.

We remind that s is a *Lebesgue point* for a function $v : I \subset \mathbb{R} \to \mathbb{R}$ if

$$\lim_{t\to s} \frac{1}{t-s} \int_s^t \big|v(\tau) - v(s)\big| \,\mathrm{d}\tau = 0 \tag{3.17}$$

and if $v \in L^1_{\mathrm{loc}}$ then almost all points $s \in I$ are Lebesgue points of v. We have

$$\begin{cases} \big|\nabla d_C\big(y_0^\varepsilon, y_\varepsilon(T)\big)\big| = 1 & \text{if } (y_0^\varepsilon, y_\varepsilon(T)) \notin C, \\ d_C\big(y_0^\varepsilon, y_\varepsilon(T)\big)\big|\nabla d_C\big(y_0^\varepsilon, y_\varepsilon(T)\big)\big| = 0 & \text{if } (y_0^\varepsilon, y_\varepsilon(T)) \in C. \end{cases}$$

We set

$$\begin{cases} \lambda_\varepsilon = \dfrac{(\varphi(u_\varepsilon, y_0^\varepsilon) - \varphi(u^*, y^*(0)) + \varepsilon)^+}{\varphi_\varepsilon(u_\varepsilon, y_0^\varepsilon)}, \\ \mu_\varepsilon = \dfrac{d_C(y_0^\varepsilon, y_\varepsilon(T))}{\varphi_\varepsilon(u_\varepsilon, y_0^\varepsilon)} \nabla d_C\big(y_0^\varepsilon, y_\varepsilon(T)\big). \end{cases}$$

This yields

$$\lambda_\varepsilon\Bigg[\int_0^T L_y(t, y_\varepsilon, u_\varepsilon) z_\varepsilon \,\mathrm{d}t + \big(\nabla g\big(y_0^\varepsilon, y_\varepsilon(T)\big), \big(z_0, z_\varepsilon(T)\big)\big) + L\big(t_0, y_\varepsilon(t_0), u(t_0)\big)$$
$$- L\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big)\Bigg] + \big(\mu_\varepsilon, \big(z_0, z_\varepsilon(T)\big)\big) \geqslant -2\sqrt{\varepsilon}, \tag{3.18}$$

where $\lambda_\varepsilon > 0$ and

$$\lambda_\varepsilon^2 + \|\mu_\varepsilon\|^2 = 1, \quad \forall \varepsilon > 0. \tag{3.19}$$

Let $p_\varepsilon \in AC([0, T]; \mathbb{R}^n)$ be the solution to the backward differential system

$$\begin{cases} p_\varepsilon'(t) = -f_y^*(t, y_\varepsilon, u_\varepsilon)p_\varepsilon + \lambda_\varepsilon L_y(t, y_\varepsilon, u_\varepsilon), & t \in [0, T], \\ p_\varepsilon(T) = -P\big(\lambda_\varepsilon \nabla g\big(y_0^\varepsilon, y_\varepsilon(T)\big) + \mu_\varepsilon\big) \end{cases} \tag{3.20}$$

where P is the projection $P(y_1, y_2) = y_2$.

Substituting into (3.18), we get (via (3.16))

$$\begin{aligned} &\lambda_\varepsilon \int_0^T L_y(t, y_\varepsilon, u_\varepsilon)z_\varepsilon \,\mathrm{d}t \\ &\quad = -\int_0^T \big(p_\varepsilon(t), z_\varepsilon'(t) - f_y(t, y_\varepsilon, u_\varepsilon)z_\varepsilon\big)\,\mathrm{d}t \\ &\qquad - \big(p_\varepsilon(t_0), f\big(t_0, y_\varepsilon(t_0), u(t_0)\big) - f\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big)\big) \\ &\qquad - \big(p_\varepsilon(0), z_0\big) - \big(P\big(\lambda_\varepsilon \nabla g\big(y_0^\varepsilon, y_\varepsilon(T)\big) + \mu_\varepsilon\big), z_\varepsilon(T)\big). \end{aligned}$$

Substituting into (3.18) we obtain

$$\begin{aligned} -2\sqrt{\varepsilon} \leqslant{} & \lambda_\varepsilon\big(L\big(t_0, y_\varepsilon(t_0), u(t_0)\big) - L\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big)\big) \\ & - \big(p_\varepsilon(t_0), f\big(t_0, y_\varepsilon(t_0), u(t_0)\big) - f\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big)\big) \\ & + \lambda_\varepsilon\big(\nabla g\big(y_0^\varepsilon, y_\varepsilon(T)\big), \big(z_0, z_\varepsilon(T)\big)\big) + \big(\mu_\varepsilon, \big(z_0, z_\varepsilon(T)\big)\big) \\ & - \big(p_\varepsilon(0), z_0\big) - \big(P\big(\lambda_\varepsilon \nabla g\big(y_0^\varepsilon, y_\varepsilon(T)\big) + \mu_\varepsilon\big), z_\varepsilon(T)\big) \end{aligned} \tag{3.21}$$

for all $z_0 \in \mathbb{R}^n$ and $u(t_0) \in U(T_0)$. This yields

$$\begin{aligned} -2\sqrt{\varepsilon} \leqslant{} & \lambda_\varepsilon\big(L\big(t_0, y_\varepsilon(t_0), u(t_0)\big) - L\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big)\big) \\ & - \big(p_\varepsilon(t_0), f\big(t_0, y_\varepsilon(t_0), u(t_0)\big) - f\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big)\big) \\ & - \big(\big(p_\varepsilon(0), -(I - P)\big(\lambda_\varepsilon \nabla g\big(y_0^\varepsilon, y_\varepsilon(T)\big) + \mu_\varepsilon\big)\big), \big(z_0, z_\varepsilon(T)\big)\big). \end{aligned}$$

For $z_0 = 0$ we get

$$\begin{aligned} & 2\sqrt{\varepsilon} - \lambda_\varepsilon L\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big) + \big(f\big(t_0, y_\varepsilon(t_0), u_\varepsilon(t_0)\big), p_\varepsilon(t_0)\big) \\ & \quad \geqslant \max_{u \in U(t_0)} \big\{-\lambda_\varepsilon L\big(t_0, y_\varepsilon(t_0), u\big) + \big(f\big(t_0, y_\varepsilon(t_0), u\big), p_\varepsilon(t_0)\big)\big\} \end{aligned} \tag{3.22}$$

and for $u = u_\varepsilon$ we see by (3.21) that

$$\big(p_\varepsilon(0), -p_\varepsilon(T)\big) = \lambda_\varepsilon \nabla g\big(y_\varepsilon(0), y_\varepsilon(T)\big) + \mu_\varepsilon + \eta_\varepsilon \tag{3.23}$$

where $\|\eta_\varepsilon\| \leqslant 2\sqrt{\varepsilon}$.

Then, on a subsequence, again denoted ε, we have

$$\begin{aligned}
&\lambda_\varepsilon \to \lambda,\\
&\mu_\varepsilon \to \mu,\\
&p_\varepsilon \to p \qquad \text{in } C\big([0,T];\mathbb{R}^n\big) \quad \text{(by Arzela theorem)},\\
&y_0^\varepsilon \to y^*(0) \quad \text{in } C\big([0,T];\mathbb{R}^n\big),\\
&m\big\{t \in [0,T];\ u_\varepsilon(t) \neq u^*(t)\big\} \to 0.
\end{aligned} \tag{3.24}$$

Recalling that by (3.9)

$$\big\| f(t, y_\varepsilon, u_\varepsilon)\big\| \leqslant \beta(t)\big(\|y_\varepsilon - y^*\| + 1\big)$$

we infer that $\{y_\varepsilon\}$ is compact in $C([0,T];\mathbb{R}^n)$ and $\{y'_\varepsilon\}$ is weakly compact in $L^1(0,T;\mathbb{R}^n)$. We have

$$y_\varepsilon(t) = y_\varepsilon(0) + \int_0^t f\big(s, y_\varepsilon(s), u_\varepsilon(s)\big)\,\mathrm{d}s \tag{3.25}$$

and

$$\begin{aligned}
&\big\| f\big(t, y_\varepsilon(t), u_\varepsilon(t)\big) - f\big(t, y^*(t), u^*(t)\big)\big\|\\
&\quad \leqslant \big\| f\big(t, y_\varepsilon(t), u_\varepsilon(t)\big) - f\big(t, y^*(t), u_\varepsilon(t)\big)\big\|\\
&\qquad + \big\| f\big(t, y^*(t), u_\varepsilon(t)\big) - f\big(t, y^*(t), u^*(t)\big)\big\|\\
&\quad \leqslant \beta(t)\big\|y_\varepsilon(t) - y^*(t)\big\| + \big\| f\big(t, y^*(t), u_\varepsilon(t)\big) - f\big(t, y^*(t), u^*(t)\big)\big\|.
\end{aligned} \tag{3.26}$$

On the other hand, since f is continuous in (y,u) we have

$$\big\| f\big(t, y^*(t), u_\varepsilon(t)\big) - f\big(t, y^*(t), u^*(t)\big)\big\| \to 0 \quad \text{a.e. in } (0,T).$$

Since

$$\big\| f\big(t, y^*(t), u_\varepsilon(t)\big)\big\| \leqslant \beta(t) \qquad \text{a.e. } t \in [0,T]$$

we infer by (3.26) and the Lebesgue dominated convergence theorem that

$$f(t, y_\varepsilon, u_\varepsilon) \to f(t, y^*, u^*) \quad \text{in } L^1(0,T).$$

Thus, by (3.25) we conclude that

$$y_\varepsilon \to y^* \quad \text{in } C\big([0,T];\mathbb{R}^n\big).$$

By (3.23) we see that

$$\big(p(0), -p(T)\big) = \lambda \nabla g\big(y^*(0), y^*(T)\big) + \mu \tag{3.27}$$

where $\mu \in N_C(y^*(0), y^*(T))$.

Indeed, since C is convex, d_C^2 is convex and we have that

$$\big(\mu_\varepsilon, \big(y_\varepsilon(0), y_\varepsilon(T)\big) - w\big) \geqslant 2\big(d_C^2\big(y_\varepsilon(0), y_\varepsilon(T)\big) - d_C^2(w)\big), \quad \forall w \in C$$

and passing to the limit we obtain

$$\big(\mu, \big(y^*(0), y^*(T)\big) - w\big) \geqslant 0, \quad \forall w \in C.$$

Since

$$\begin{aligned} &f_y^*\big(t, y_\varepsilon(t), u_\varepsilon(t)\big) \to f_y^*\big(t, y^*(t), u^*(t)\big) \quad \text{a.e. } t \in (0,T),\\ &L_y\big(t, y_\varepsilon(t), u_\varepsilon(t)\big) \to L_y\big(t, y^*(t), u^*(t)\big) \quad \text{a.e. } t \in (0,T) \end{aligned}$$

by (3.8) and (3.20) it follows that

$$p' = -f_y^*(t, y^*, u^*)p + \lambda L_y(t, y^*, u^*) \quad \text{a.e. } t \in (0,T) \tag{3.28}$$

and by (3.22)

$$\begin{aligned} &-\lambda L\big(t_0, y^*(t_0), u^*(t_0)\big) + \big(f\big(t_0, y^*(t_0), u^*(t_0)\big), p(t_0)\big)\\ &\quad = \max_{u \in U(t_0)} \big\{-\lambda L\big(t_0, y^*(t_0), u\big) + \big(f\big(t_0, y^*(t_0), u\big), p(t_0)\big)\big\} \quad \text{a.e. } t_0 \in [0,T]. \end{aligned} \tag{3.29}$$

Recall also that $\lambda \geqslant 0$ and

$$|\lambda| + \|\mu\| = 1. \tag{3.30}$$

If $\lambda > 0$, then replacing p by p/λ we get (3.1)–(3.3) with $\lambda = 1$.

If $\lambda = 0$, then by (3.30) we see that $\mu \neq 0$ and so

$$\big\|p(0)\big\| + \big\|p(T)\big\| \neq 0.$$

Clearly, this implies that

$$\big\|p(t)\big\| \neq 0 \quad \forall t \in [0,T].$$

The proof of the theorem is complete.

REMARK 3.1. The problem (1.14) with state constraints

$$y(t) \in K \subset \mathbb{R}^m \quad \forall t \in [0,T] \tag{3.31}$$

where K is a closed convex subset of $\mathbb{R}^n$ can be treated similarly. In this case we take

$$\varphi_\varepsilon(u, y_0) = \biggl[\Bigl(\bigl(\varphi(u, y_0) - \varphi\bigl(u^*, y^*(0)\bigr) + \varepsilon\bigr)^+\Bigr)^2 + d_C^2\bigl(y_0, y^u(T)\bigr) + \int_0^T d_K^2\bigl(y(t)\bigr)\,\mathrm{d}t\biggr]^{1/2}.$$

We propose to the reader to obtain a optimality theorem of the type of Theorem 3.1 in the present situation.

REMARK 3.2. In Theorem 3.1 the condition that C is convex can be removed. (See [4] for a proof in this general case.)

3.3. *Convex optimal control problems*

We shall study here the problem (P):

$$\min\biggl\{\int_0^T L\bigl(t, y(t), u(t)\bigr)\,\mathrm{d}t + g\bigl(y(0), y(T)\bigr);\ \bigl(y(0), y(T)\bigr) \in C\biggr\} \tag{3.32}$$

subject to

$$\begin{aligned} &y'(t) = A(t)y(t) + B(t)u(t) + f_0(t), \quad u \in L^1\bigl(0, T; \mathbb{R}^m\bigr) \\ &u(t) \in U(t) \quad \text{a.e. } t \in [0, T] \end{aligned}$$

where L is convex in (y, u), measurable in t, $g \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$ is convex, C is closed and convex, $A(t) \in L(\mathbb{R}^n \times \mathbb{R}^n)$, $B(t) \in L(\mathbb{R}^m, \mathbb{R}^n)$, $A(\cdot)$, $B(\cdot)$, $f_0(\cdot)$ are integrable and $U(t) \subset \mathbb{R}^m$ is closed and convex.

We have by Theorem 3.1 the following sharpening of the maximum principle in this case:

COROLLARY 3.1. *If (y^*, u^*) is optimal and $u^* \in L^1(0, T)$, then there exists $p \in AC([0, T]; \mathbb{R}^n)$, $\lambda \in \{0, 1\}$ such that $\lambda + |p(t)| \neq 0$ and*

$$\begin{cases} p' = -A^*(t)p + \lambda L_y\bigl(t, y^*(t), u^*(t)\bigr) \quad \text{a.e. } t \in (0, T), \\ \bigl(p(0), -p(T)\bigr) - \lambda\nabla g\bigl(y^*(0), y^*(T)\bigr) \in N_C\bigl(y^*(0), y^*(T)\bigr), \end{cases} \tag{3.33}$$

$$B^* p(t) \in \lambda L_u\bigl(t, y^*(t), u^*(t)\bigr) + N_{U(t)}\bigl(u^*(t)\bigr) \quad \text{a.e. } t \in (0, T) \tag{3.34}$$

where $\partial L = (\partial_y L, \partial_u L)$ is the subdifferential of L.

In fact, (3.34) comes from the maximum principle

$$\begin{aligned}&\lambda L\big(t, y^*(t), u^*(t)\big) - \big(f\big(t, y^*(t), u^*(t)\big), p(t)\big)\\ &\quad = \max_{u\in U(t)} \big\{\lambda L\big(t, y^*(t), u\big) - \big(f\big(t, y^*(t), u\big), p(t)\big)\big\}\end{aligned} \tag{3.35}$$

where $f(t, y, u) \equiv A(t)y + B(t)u + f_0(t)$ and, because $\lambda L(t, y^*, \cdot)$ is convex, the maximum in (3.35) is attained for u^* satisfying (3.34).

COROLLARY 3.2. *Assume that* $C = C_1 \times C_2$ *and there exists* (y, u) *admissible such that* $y(T) \in \operatorname{int} C_2$ *or* $y(0) \in \operatorname{int} C_1$. *Then the problem is normal.*

PROOF. Assume that $\lambda = 0$ and $y(T) \in \operatorname{int} C_2$. Then $|p(t)| \neq 0$ and

$$\begin{cases} p' = -A^*(t)p \quad \text{a.e. } t \in [0, T],\\ p(0) \in N_{C_1}\big(y^*(0)\big), \quad -p(T) \in N_{C_2}\big(y^*(T)\big), \end{cases} \tag{3.36}$$

$$B^*(t)p(t) \in N_{U(t)}\big(u^*(t)\big), \quad \text{a.e. } t \in [0, T]. \tag{3.37}$$

We have

$$\begin{cases} y' = A(t)y + B(t) + f_0(t)u \quad \text{a.e. } t \in [0, T],\\ y(T) \in \operatorname{int} C_2. \end{cases} \tag{3.38}$$

This yields

$$(y^* - y)' = A(t)(y - y^*) + B(t)(u - u^*) \quad \text{a.e. } t \in [0, T]$$

and

$$-\big(p(T), y^*(T) - \rho w - y(T)\big) \geqslant 0, \quad \forall |w| = 1.$$

Then using (3.37) we get

$$\begin{aligned}\rho\big|p(T)\big| &= -\big(p(T), y^*(T) - y(T)\big)\\ &= -\big(p(0), y^*(0) - y_0\big) + \int_0^T \big(u - u^*, B^*(t)p(t)\big)\,\mathrm{d}t \leqslant 0\end{aligned}$$

(because $p(0) \in N_{C_1}(y^*(0))$). Hence $p(T) = 0$, contradiction. The corollary is proved. The case $y(0) \in \operatorname{int} C_1$ can be treated similarly. □

COROLLARY 3.3. *Under assumption of Corollary* 3.2, *the system* (3.33)–(3.34) *is also sufficient for optimality.*

PROOF. As seen earlier, $\lambda = 1$. Let (y, u) be an arbitrary pair such that

$$y' = A(t)y + B(t)u + f_0(t) \quad \text{a.e. } t \in [0, T],\ u(t) \in U(t),$$
$$\big(y(0), y(t)\big) \in C_1 \times C_2.$$

We have (by convexity of C)

$$L(t, y^*, u^*) \leqslant L(t, y, u) + \big(L_y(t, y^*, u^*), y^* - y\big) + \big(L_u(t, y^*, u^*), u^* - u\big),$$
$$g\big(y^*(0), y^*(T)\big) \leqslant g\big(y(0), y(T)\big) + \big(y^*(0) - y(0), y^*(T) - y(T)\big)$$
$$\times \nabla g\big(y^*(0), y^*(T)\big).$$

Integrating on $[0, T]$ and using (3.33) and (3.34), we see that

$$\int_0^T L(t, y^*, u^*)\,\mathrm{d}t + g\big(y^*(0), y^*(T)\big) \leqslant \int_0^T L(t, y, u)\,\mathrm{d}t + g\big(y(0), y(T)\big)$$

as claimed. □

We shall prove now a maximum principle for the *convex control problem of Bolza* under more general conditions on L. We shall assume that

1. $L(t, \cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is convex, continuous and the Hamiltonian function
$$H(t, x, p) = \sup\big\{(p, u) - L(t, x, u):\ u \in U(t)\big\}$$
is finite and summable in t for each $(x, p) \in \mathbb{R}^n \times \mathbb{R}^m$. $U(t) \subset \mathbb{R}^m$ is closed and convex for each t. Moreover, assumption (ii) in Section 3.1 holds.
2. $g \in C(\mathbb{R}^n \times \mathbb{R}^n)$ is convex and $C = C_1 \times C_2$ where C_1, C_2 are convex and closed.
3. $f(t, y, u) = A(t)y + B(t)u + f_0(t)$ where $A(t) \in L^1(0, T; \mathbb{R}^n \times \mathbb{R}^n)$, $B(t) \in L^1(0, T; \mathbb{R}^m \times \mathbb{R}^n)$, $f_0 \in L^1(0, T)$.
4. There is (y, u) admissible such that either $y(0) \in \operatorname{int} C_1$ or $y(T) \in \operatorname{int} C_2$.

THEOREM 3.2. *Assume that* $(y^*, u^*) \in C([0, T]; \mathbb{R}^n) \times L^2([0, T]; \mathbb{R}^n)$ *is optimal. Then there exists* $p \in AC([0, T]; \mathbb{R}^n)$ *such that*

$$p' + A^*(t)p \in \partial_u L\big(t, y^*(t), u^*(t)\big) \quad \textit{a.e. } t \in (0, T), \tag{3.39}$$
$$\big(p(0), -p(T)\big) \in \partial g\big(y^*(0), y^*(T)\big) + N_C\big(y^*(0), y^*(T)\big), \tag{3.40}$$
$$B^* p \in \partial_u L\big(t, y^*(t), u^*(t)\big) + N_{U(t)}\big(u^*(t)\big) \quad \textit{a.e. } t \in (0, T). \tag{3.41}$$

Moreover, conditions (3.39)–(3.41) *are also sufficient for optimality.*

PROOF. For $\lambda > 0$ consider the problem

$$\min\bigg\{\int_0^T L_\lambda(t, y, u)\,\mathrm{d}t + g_\lambda\big(y(0), y_\lambda(T)\big) + \frac{1}{2}\int_0^T \big|u(t) - u^*(t)\big|^2\,\mathrm{d}t$$

$$+\frac{1}{2}\big|y(0)-y^*(0)\big|^2+\frac{1}{2\lambda}\Big(d_{C_1}^2\big(y_\lambda(0)\big)+d_{C_2}^2\big(y_\lambda(T)\big)\Big),\ u(t)\in U(t)\Bigg\} \quad (\mathrm{P}_\lambda)$$

subject to

$$y'=A(t)y+B(t)u+f_0(t)\quad \text{a.e. } t\in(0,T).$$

Problem (P_λ) has a unique solution (y_λ,u_λ).

LEMMA 3.1. *For* $\lambda\to 0$ *we have*

$$y_\lambda\to y^*\quad \textit{in } C\big([0,T];\mathbb{R}^n\big), \tag{3.42}$$

$$u_\lambda\to u^*\quad \textit{in } L^2\big([0,T];\mathbb{R}^n\big). \tag{3.43}$$

PROOF. Recall that (see Section 2.1)

$$L_\lambda(t,y,u)=\inf_{(z,v)}\left\{\frac{|y-z|^2}{2\lambda}+\frac{|u-v|^2}{2\lambda}+L(t,z,v)\right\},$$

$$g_\lambda(y_1,y_2)=\inf_{(z_1,z_2)}\left\{\frac{|y_1-z_1|^2}{2\lambda}+\frac{|y_2-z_2|^2}{2\lambda}+g(z_1,z_2)\right\}$$

and d_{C_1} (d_{C_2}) is the distance to C_1 (and C_2 respectively). We have

$$\begin{aligned}&\int_0^T L_\lambda\big(t,y_\lambda(t),u_\lambda(t)\big)\,\mathrm{d}t+g_\lambda\big(y_\lambda(0),y_\lambda(T)\big)+\frac{1}{2}\int_0^T\big|u_\lambda(t)-u^*(t)\big|^2\,\mathrm{d}t\\&\qquad+\frac{1}{2}\big|y_\lambda(0)-y^*(0)\big|^2+\frac{1}{2\lambda}\Big(d_{C_1}^2\big(y_\lambda(0)\big)+d_{C_2}^2\big(y_\lambda(T)\big)\Big)\\&\quad\leqslant\int_0^T L\big(t,y^*(t),u^*(t)\big)\,\mathrm{d}t+g\big(y^*(0),y^*(T)\big)\end{aligned} \tag{3.44}$$

because $L_\lambda\leqslant L$, $g_\lambda\leqslant g$.

Let for $\lambda\to 0$,

$$\begin{aligned}&u_\lambda\to\bar u\quad \text{weakly in } L^2\big(0,T;\mathbb{R}^m\big),\\&y_\lambda\to\bar y\quad \text{strongly in } C\big(0,T;\mathbb{R}^n\big).\end{aligned}$$

Then by the Fatou lemma

$$\liminf_{\lambda\to 0}\int_0^T L_\lambda\big(t,y_\lambda(t),u_\lambda(t)\big)\,\mathrm{d}t\geqslant\int_0^T L\big(t,\bar y(t),\bar u(t)\big)\,\mathrm{d}t$$

and by the lower semicontinuity of g

$$\liminf_{\lambda\to 0}g_\lambda\big(_\lambda(0),y_\lambda(T)\big)\geqslant g\big(\bar y(0),\bar u(T)\big).$$

Then by (3.44) we deduce that

$$y_\lambda(0) \to y^*(0), \quad u_\lambda \to u^* \quad \text{in } L^2\big(0,T;\mathbb{R}^m\big). \qquad \square$$

Now the maximum principle in (P_λ) yields ((P_λ) is smooth and so Corollary 3.2 applies)

$$\begin{cases} p_\lambda' + A^*(t)p_\lambda = \nabla_y L_\lambda(t, y_\lambda, u_\lambda) \quad \text{a.e. } t \in (0,T), \\ \big(p_\lambda(0), -p_\lambda(T)\big) = \nabla g_\lambda\big(y_\lambda(0), y_\lambda(T)\big) + \big(y_\lambda(0) - y^*(0), 0\big) \\ \qquad\qquad + \dfrac{1}{\lambda}\big(\nabla d_{C_1}^2\big(y_\lambda(0)\big), \nabla d_{C_2}^2\big(y_\lambda(T)\big)\big), \\ B^* p_\lambda = \partial_u L_\lambda(y_\lambda, u_\lambda) + u_\lambda - u^* + N_{U(t)}\big(u_\lambda(t)\big) \quad \text{a.e. } t \in (0,T). \end{cases} \tag{3.45}$$

We shall prove now that $\{p_\lambda(t)\}$ is bounded in $\mathbb{R}^n$. We shall use the same argument as in the proof of Corollary 3.2. Indeed, we have

$$\big(p_\lambda(T), y_\lambda(T) - y(0) - \rho w\big) \geqslant \big(\nabla_2 g_\lambda\big(y_\lambda(0), y_\lambda(T)\big), y_\lambda(T) - y(0)\big) \geqslant M$$

because

$$\frac{1}{2}\nabla d_{C_2}^2(y) = \frac{1}{\lambda}\big(I - (I + \lambda\partial I_{C_2})^{-1}\big)(y).$$

Hence

$$\rho\big|p_\lambda(T)\big| \leqslant \big(p_\lambda(T), y_\lambda(T) - y(0)\big).$$

On the other hand, by (3.45), we have

$$\begin{aligned} &-\big(p_\lambda(0), y_\lambda(0) - y(0)\big) + \big(p_\lambda(T), y_\lambda(T) - y(T)\big) \\ &\quad = \int_0^T \big(\nabla_y L_\lambda(t, y_\lambda, u_\lambda), y_\lambda - y\big)\,\mathrm{d}t + \int_0^T \big(B(u_\lambda - u), p_\lambda\big)\,\mathrm{d}t \end{aligned}$$

where

$$y' = Ay + Bu + f_0.$$

Then again by (3.45) we see that

$$\begin{aligned} &\big(p_\lambda(T), y_\lambda(T) - y(T)\big) \\ &\quad \geqslant \big(p_\lambda(0), y_\lambda(0) - y(0)\big) + \frac{1}{2\lambda}\big(d_{C_1}^2\big(y_\lambda(0)\big) - d_{C_2}^2\big(y(0)\big)\big). \end{aligned}$$

Substituting the latter in the previous inequalities we see that $\{p_\lambda(T)\}$ is bounded in $\mathbb{R}^n$.

On the other hand, the Hamiltonian function $H(t, y, p)$ is concave in y and convex in p. By assumption 1, it follows that for each $y_0 \in \mathbb{R}^n$ there is a neighborhood $\mathcal{V}(y_0)$ of y_0 and $\alpha \in L^1(0, T)$ such that

$$-H(t, y, 0) \leqslant \alpha(t), \quad \forall y \in \mathcal{V}(y_0),\ t \in (0, T).$$

Indeed, we may choose $\mathcal{V}(y_0)$ a simplex generated by $\{y_1, \dots, y_{N+1}\}$. Then

$$-H(t, y, 0) \leqslant -\sum_i \lambda_i H(t, y_i, 0) =: \alpha(t).$$

By the inequality

$$H(t, y, 0) - H(t, y, \rho w) \leqslant \rho(v, w), \quad \forall v \in \partial_p H(t, y, 0)$$

it follows that

$$\sup\bigl\{\|v\|;\ v \in \partial_p H(t, y, 0)\bigr\} \leqslant \beta(t), \quad \text{a.e. } t \in (0, T) \tag{3.46}$$

where $\beta \in L^1(0, T)$. We have

$$L_\lambda\bigl(t, y^*(t) + \rho w, v_0(t)\bigr) \leqslant L\bigl(t, y^*(t) + \rho w, v_0(t)\bigr) \leqslant \alpha(t)$$

for all $|w| = 1$ and $v_0(t) \in \partial_p H(t, y^*(t) + \rho w, 0)$. Because

$$L(t, y, v_0) = \sup\bigl\{(v_0, p) - H(t, y, p)\bigr\}$$

and

$$L(t, y, v_0) + H(t, y, 0) = 0.$$

We have, by (3.45) and the convexity of $L_\lambda(t, y, \cdot)$, that

$$\begin{aligned}&\bigl(p_\lambda' + A^* p_\lambda, y_\lambda - y^* - \rho w\bigr) + \bigl(B^* p_\lambda + u_\lambda - u^*, u_\lambda - v_0\bigr)\\&\quad \geqslant L_\lambda(t, y_\lambda, u_\lambda) - L_\lambda(t, y^* + \rho w, v_0) \quad \text{a.e. } t \in (0, T).\end{aligned}$$

Hence

$$\begin{aligned}\rho\bigl|p_\lambda' + A^* p_\lambda\bigr| &\leqslant \alpha(t) - L_\lambda(t, y_\lambda, u_\lambda) + \bigl(p_\lambda' + A^* p_\lambda, y_\lambda - y\bigr)\\&\quad + \bigl(B^* p_\lambda, u_\lambda - v_0\bigr) + (u_\lambda - u^*, u_\lambda - v_0).\end{aligned} \tag{3.47}$$

This yields (L_λ is bounded from below by an affine function)

$$\begin{aligned}\rho\int_0^T \big|p_\lambda' + A^* p_\lambda\big|\,\mathrm{d}t \leqslant C + \big(p_\lambda(T), y_\lambda(T) - y^*(T)\big) \\ - \big(p_\lambda(0), y_\lambda(0) - y^*(0)\big) + C\big|p_\lambda(t)\big|\big|u_\lambda(t) - v_0(t)\big| \\ + |u_\lambda - u^*|^2 + |u_\lambda - u^*||u_\lambda - v_0| \quad \text{a.e. } t \in (0, T).\end{aligned}$$

Since $\int_0^T |u_\lambda|^2\,\mathrm{d}t$ and $|p_\lambda(0)|$, $|p_\lambda(T)|$ are bounded it follows by Gronwall's lemma that

$$\begin{aligned}&|p_\lambda(t)| \leqslant C, \quad \forall t \in [0, T],\\ &\int_0^T \big|p_\lambda'(t) + A^* p_\lambda(t)\big|\,\mathrm{d}t \leqslant C.\end{aligned}$$

As a matter of fact, by (3.47) we see that $\{p_\lambda'\}$ is weakly compact in $L^1(0, T; \mathbb{R}^n)$ (the Dunford–Pettis theorem). Then on a subsequence, again denoted λ, we have

$$\begin{aligned}&p_\lambda(t) \to p(t) && \text{uniformly on } [0, T],\\ &p_\lambda' \to p' && \text{weakly in } L^1\big(0, T; \mathbb{R}^n\big).\end{aligned} \tag{3.48}$$

Moreover, letting $\lambda \to 0$ into (see the second equation in (3.45))

$$\begin{aligned}&\big(p_\lambda(T), y_\lambda(0) - \xi\big) - \big(p_\lambda(T), y_\lambda(T) - \eta\big)\\ &\quad \geqslant g_\lambda\big(y_\lambda(0), y_\lambda(T)\big) - g_\lambda(\xi, \eta) + \frac{1}{\lambda}\big(d_{C_1}^2\big(y_\lambda(0)\big) + d_{C_2}^2\big(y_\lambda(T)\big)\big)\\ &\qquad + \big(y_\lambda(0) - y^*(0), y_\lambda(0) - \xi\big), \quad \forall(\xi, \eta) \in C\end{aligned}$$

we see that

$$\begin{aligned}&\big(p(0), y_\lambda(0) - \xi\big) - \big(p(T), y^*(T) - \eta\big) \geqslant g\big(y^*(0), y^*(T)\big) - g(\xi, \eta),\\ &\quad \forall(\xi, \eta) \in C.\end{aligned}$$

Then letting $\lambda \to 0$ we see that the transversality condition (3.40) holds.

Next, we let $\lambda \to 0$ into the inequality (see (3.45))

$$\begin{aligned}&\int_0^T \big(p_\lambda' + A^* p_\lambda, y_\lambda - y\big)\,\mathrm{d}t + \int_0^T \big(B^* p_\lambda, u_\lambda - u\big)\,\mathrm{d}t\\ &\quad \geqslant \int_0^T L_\lambda\big(t, y_\lambda, u_\lambda\big)\,\mathrm{d}t + \frac{1}{2}\int_0^T |u_\lambda - u^*|^2\,\mathrm{d}t - \int_0^T L_\lambda(t, y, u)\,\mathrm{d}t\\ &\qquad - \frac{1}{2}\int_0^T |u - u^*|^2\,\mathrm{d}t.\end{aligned}$$

By weak lower semicontinuity of the convex integrand we know that (see Lemma 3.1)

$$\liminf_{\lambda\to 0}\int_0^T L_\lambda(t,y_\lambda,u_\lambda)\,\mathrm{d}t \geqslant \int_0^T L(t,y^*,u^*)\,\mathrm{d}t.$$

Hence

$$\begin{aligned}
&\int_0^T \big(p'+A^*p, y^*-y\big)\,\mathrm{d}t + \int_0^T \big(B^*p, u^*-u\big)\,\mathrm{d}t\\
&\quad\geqslant \int_0^T L(t,y^*,u^*)\,\mathrm{d}t - \int_0^T L(t,y,u)\,\mathrm{d}t,\\
&\quad\forall (y,u)\in L^\infty\big(0,T;\mathbb{R}^n\big)\times L^2\big(0,T;\mathbb{R}^m\big).
\end{aligned}$$

The latter implies that

$$\begin{aligned}
&p'(t)+A^*(t)p(t)\in\partial_p L\big(t,y^*(t),u^*(t)\big) &&\text{a.e. } t\in(0,T),\\
&B^*(t)p(t)\in\partial_u L\big(t,y^*(t),u^*(t)\big)+N_{U(t)}\big(u^*(t)\big) &&\text{a.e. } t\in(0,T)
\end{aligned}$$

as claimed.

The sufficiency of conditions (3.39)–(3.41) for optimality is immediate. It relies on the obvious inequalities

$$\begin{aligned}
L(t,y^*,u^*)\leqslant{}& L(t,y,u)+\big(L_y(t,y^*,u^*),y^*-y\big)\\
&+\big(L_u(t,y^*,u^*)+\eta,u^*-u\big)\quad\text{a.e. } t\in(0,T),\ \eta\in N_{U(t)}\big(u^*(t)\big)
\end{aligned}$$

where (y,u) is any pair of functions such that $u(t)\in U(t)$. If we take (y,u) such that $y(0)\in C_1$ and $y(T)\in C_2$

$$y' = A(t)y+B(t)u+f_0(t)$$

and we integrate on $(0,T)$ we get by (3.40)–(3.41)

$$\begin{aligned}
\int_0^T L(t,y^*,u^*)\,\mathrm{d}t \leqslant{}& \int_0^T L(t,y,u)\,\mathrm{d}t + \big(p(T),y^*(T)-y(T)\big)\\
&-\big(p(0),y^*(0)-y(0)\big)\\
\leqslant{}& \int_0^T L(t,y,u)\,\mathrm{d}t - g\big(y^*(0),y^*(T)\big)+g\big(y(0),y(T)\big).
\end{aligned}$$

Hence (y^*,u^*) is optimal in problem (P_λ). □

The dual problem Define the functions

$$\begin{aligned}&M(t,q,w)=\sup\bigl\{(q,v)+(w,y)-L(t,y,v);\ v\in U(t)\bigr\},\\&m(p_1,p_2)=\sup\bigl\{(p_1,x_1)-(p_2,x_2)-g(x_1,x_2);\ x_1\in C_1,\ x_2\in C_2\bigr\}.\end{aligned}\tag{3.49}$$

The problem

$$\begin{aligned}\min\biggl\{&\int_0^T M\bigl(t,B^*(t)p,w(t)\bigr)\,\mathrm{d}t+\int_0^T\bigl(f_0(t),p(t)\bigr)\,\mathrm{d}t\\&+m\bigl(p(0),p(T)\bigr);\ p'+A^*(t)p=w\text{ a.e. in }(0,T)\biggr\}\end{aligned}\tag{P*}$$

is called the dual of (P). We have

THEOREM 3.3. *Under assumptions of Theorem* 3.2 *the pair* (y^*,u^*) *is optimal in problem* (P) *if and only if* (P*) *has a solution* (p^*,w^*). *We have*

$$\inf \mathrm{P}+\inf \mathrm{P}^*=0.\tag{3.50}$$

The proof is immediate. It relies on the conjugacy relationship between L and M.

It turns out that in certain situations the dual problem (P*) is simpler than the primal problem (P). In particular, control constraints $u(t)\in U(t)$ disappear in the dual problem. Let us illustrate this on the following simple example:

$$\begin{aligned}\min\biggl\{&\frac{1}{2}\int_0^T\|y(t)\|^2\,\mathrm{d}t;\ y'(t)=A(t)y(t)+u(t),\ y(0)=y_0,\\&\bigl\|y(T)\bigr\|\leqslant 1,\ \bigl\|u(t)\bigr\|\leqslant\rho\text{ a.e. }t\in(0,T)\biggr\}\end{aligned}$$

where $\|\cdot\|$ is the Euclidean norm.

In this case

$$\begin{aligned}&M(t,q,w)=\frac{1}{2}\|w\|^2+\rho\|q\|\quad\forall(q,w)\in\mathbb{R}^n\times\mathbb{R}^n,\\&m(p_1,p_2)=(p_1,y_0)-\|p_2\|\quad\forall(p_1,p_2)\in\mathbb{R}^n\times\mathbb{R}^n\end{aligned}$$

and so the dual control problem (P*) is

$$\begin{aligned}\min\biggl\{&\int_0^T\Bigl(\rho\bigl\|p(t)\bigr\|+\frac{1}{2}\bigl\|w(t)\bigr\|^2\Bigr)\,\mathrm{d}t+\bigl(p(0),y_0\bigr)+\bigl\|p(T)\bigr\|;\\&p'+A^*(t)p=w\text{ a.e. in }(0,T)\biggr\}.\end{aligned}$$

Comments The results presented above are essentially due to R.T. Rockafellar (see [32]). In infinite dimensional spaces such results were established in [6].

3.4. *Examples*

3.4.1. *The optimal control of the prey–predator system* We shall treat here some specific problems from different areas of interest. Consider the Volterra prey–predator system

$$\begin{cases} x'(t) = x(t)\big(\lambda_1 - \mu_1 u(t) y(t)\big), & t \in [0,T], \\ y'(t) = y(t)\big(-\lambda_2 + \mu_2 u(t) x(t)\big), & t \in [0,T], \\ x(0) = x_0, \quad y(0) = y_0, & x_0, y_0 > 0 \end{cases} \tag{3.51}$$

where x is the prey, y the predator and $0 \leqslant u(t) \leqslant 1$ is the segregation rate; λ_i, μ_i are positive constants.

Consider the optimal control problem (see [4,37]):

$$\min\big\{-\big(x(T) + y(T)\big);\ 0 \leqslant u \leqslant 1\big\}. \tag{3.52}$$

This is a Bolza optimal control problem where

$$L \equiv 0, \quad g(x,y) = -(x+y),$$
$$U = \{u \in \mathbb{R};\ 0 \leqslant u \leqslant 1\},$$
$$f(x,y,u) = \begin{pmatrix} x(\lambda_1 - \mu_1 u y) \\ y(-\lambda_2 + \mu_2 u x) \end{pmatrix}.$$

The maximum principle (see Theorem 3.1) yields

$$\begin{cases} p_1' = -(\lambda_1 - \mu_1 u^* y^*) p_1 - \mu_2 u^* y^* p_2, & t \in [0,T], \\ p_2' = \mu_1 u^* x^* p_1 - (-\lambda_2 + \mu_2 u^* x^*) p_2, & t \in [0,T], \\ p_1(T) = 1, \quad p_2(T) = 1, \end{cases} \tag{3.53}$$

$$u^*(t) = \operatorname*{argmax}_{u \in U}\big\{x^*(t)\big(\lambda_1 - \mu_1 u y^*(t)\big) p_1(t) + y^*(t)\big(-\lambda_2 + \mu_2 u x^*(t)\big) p_2(t)\big\}$$
$$\text{a.e. } t \in [0,T]. \tag{3.54}$$

Equivalently

$$u^*(t) = \begin{cases} 0 & \text{if } \mu_2 p_2(t) - \mu_1 p_1(t) < 0, \\ 1 & \text{if } \mu_2 p_2(t) - \mu_1 p_1(t) > 0. \end{cases} \tag{3.55}$$

(Since x_0, $y_0 > 0$, we have $x^*(t) > 0$, $y^*(t) > 0$.)

We shall discuss the form of the optimal control according to the sign of $\mu_2 - \mu_1$. We note first that always u^* is a *bang–bang controller* because the set of zeros of the function $\mu_2 p_2 - \mu_1 p_1$ consists of a finite number of points.

1. $\mu_2 - \mu_1 < 0$.

   In this case

$$(\mu_2 p_2 - \mu_1 p_1)(T) = \mu_2 - \mu_1 < 0. \tag{3.56}$$

Hence

$$\mu_2 p_2 - \mu_1 p_1 < 0$$

in a maximal interval $[T - \delta, T]$. On this interval we have by (3.56) that $u^*(t) = 0$ and so

$$\begin{cases} p_1' = -\lambda_1 p, & t \in (T - \delta, T), \\ p_2' = \lambda_2 p_2, & t \in (T - \delta, T), \\ p_1(T) = p_2(T) = 1. \end{cases}$$

Hence

$$p_1(t) = \mathrm{e}^{\lambda_1(T-t)}, \quad p_2(t) = \mathrm{e}^{\lambda_2(t-T)} \quad \text{on } [T - \delta, T].$$

Since the function $\mu_2 \mathrm{e}^{\lambda_2(t-T)} - \mu_1 \mathrm{e}^{-\lambda_1(T-t)}$ is increasing, it follows that $(T - \delta, T) = (0, T)$. In other words, in this case $u^*(t) = 0$, $\forall t \in (0, T)$.

2. $\mu_2 - \mu_1 > 0$.

   Then

$$\mu_2 p_2(t) - \mu_1 p_1(t) > 0 \quad \text{for } t \in [T - \varepsilon, T]$$

and

$$u^*(t) = 0 \quad \text{for } t \in [T - \varepsilon, T].$$

(We may assume that $t \in [T - \varepsilon, T]$ is maximal with this property.) Let us prove that $t_1 = T - \varepsilon$ is a switching point for u^*, i.e.,

$$\mu_2 p_2(t) - \mu_1 p_1(t) < 0 \quad \text{for } 0 \leqslant t \leqslant T - \varepsilon.$$

We have on $(T - \varepsilon, T)$

$$\begin{cases} p_1' = -p_1(\lambda_1 - \mu_1 y^*) - \mu_2 y^* p_2, & t \in (T - \varepsilon, T), \\ p_2' = -p_2(\mu_2 x^* - \lambda_2) + \mu_1 x^* p_1 \end{cases} \tag{3.57}$$

i.e.,

$$\begin{cases} p_1' = -\lambda_1 p_1 + y^*(\mu_1 p_1 - \mu_2 p_2), & t \in (t_1, T), \\ p_2' = \lambda_2 p_2 + x^*(\mu_1 p_1 - \mu_2 p_2), & t \in (t_1, T). \end{cases}$$

Hence

$$p_1(t) \geqslant e^{\lambda_1(T-t)} \geqslant 1, \quad p_1(t_1) = \frac{\mu_2}{\mu_1} p_2(t_1) \geqslant \frac{\mu_2}{\mu_1} > 0.$$

Note that $\mu_1 p_1(t) - \mu_2 p_2(t) = \varphi(t)$ satisfies the equation

$$\varphi'(t) = u^* \varphi(t)(\mu_2 x^* - \mu_1 y^*) - \lambda_2 \mu_2 p_2 - \lambda_1 \mu_1 p_1, \quad t \in (0, t_1).$$

Hence

$$\varphi(t) = C e^{\int_t^{t_1} u^*(\mu_2 x^* - \mu_1 y^*)\,\mathrm{d}s} + \int_t^{t_1} e^{\int_t^s u^*(\mu_2 x^* - \mu_1 y^*)\,\mathrm{d}r} (\lambda_2 \mu_2 p_2 + \lambda_1 \mu_1 p_1)\,\mathrm{d}s \quad \text{for } t \in (0, t_1).$$

Since $\varphi(t_1) = 0$ we see that $C = 0$ and

$$\lambda_2 \mu_2 p_2(t_1) + \lambda_1 \mu_1 p_1(t_1) \geqslant \lambda_2 \mu_2 + \lambda_1 \mu_1 > 0$$

we conclude that $\varphi(t) > 0$ in a left neighborhood of t_1. Hence

$$u^*(t) = 0 \quad \text{for } t \in (t_1 - \varepsilon, t_1) = (t_2, t_1).$$

But as seen above, $\mu_2 p_2 - \mu_1 p_1$ is increasing on (t_2, t_1) and so

$$\mu_2 p_2(t) - \mu_1 p_1(t) < 0$$

for all $t < t_1$. Hence $t_1 = 0$ and so

$$u^*(t) = \begin{cases} 1 & \text{for } t_1 < t \leqslant T, \\ 0 & \text{for } 0 \leqslant t < t_1 \end{cases} \tag{3.58}$$

where the switching point t_1 can be computed from the equation

$$\mu_2 p_2(t) - \mu_1 p_1(t) = 0.$$

3. $\mu_2 - \mu_1 = 0$.
   In this case it follows that $u^*(t) = 0$ for all $t \in (0, T)$.

3.4.2. *Periodic solutions to Hamiltonian systems* Consider the Hamiltonian system

$$\begin{cases} y'(t) = \partial_p H\big(y(t), p(t)\big) + f_1(t), & t \in (0, T), \\ p'(t) = -\partial_y H\big(y(t), p(t)\big) + f_2(t), & t \in (0, T) \end{cases} \tag{3.59}$$

with periodic conditions

$$y(0) = y(T), \quad p(0) = p(T). \tag{3.60}$$

Assumptions:

1. $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is convex, of class C^1 and

$$\gamma_1 |y|^2 + \gamma_2 |p|^2 + C_1 \leqslant H(y, p) < \frac{\pi}{T}\big(|y|^2 + |p|^2\big) + C_2 \quad \forall y, p \in \mathbb{R}$$

   where $\gamma_1, \gamma_2 > 0$.

2. $f_1, f_2 \in L^2(0, T; \mathbb{R}^n)$.

A special case of (3.59) is $f_1 = 0$, $H(y, p) = g(y) + \frac{1}{2}|p|^2$. In this case, (3.59) reduces to

$$y' = p, \quad p' = -g'(y) + f_2(t)$$

i.e.,

$$y'' + g'(y) = f_2(t), \quad y(0) = y(t), \quad y'(0) = y'(T).$$

Following Clarke and Ekeland [13], we may reduce problem (3.59), (3.60) to the optimal control problem

$$\min\left\{\int_0^T \Big(G\big(v(t) - f_2(t), u(t) - f_1(t)\big) - \big(y(t), v(t)\big)\Big)\,\mathrm{d}t\right\} \tag{3.61}$$

subject to

$$\begin{aligned} &y'(t) = u(t), \quad z'(t) = -v(t) \quad \text{a.e. } t \in (0, T), \\ &y(0) = y(T), \quad z(0) = z(T) \end{aligned} \tag{3.62}$$

where $G = H^*$ is the conjugate of H, i.e.,

$$G(q_1, q_2) = \sup_{(y,p)} \big\{(y, p_1) + (p, q_2) - H(y, p)\big\}.$$

If (y^*, z^*, u^*, v^*) is optimal the maximum principle yields (see Theorem 3.1)

$$\begin{cases} q_1'(t) = -\lambda v^*(t) & \text{a.e. } t \in (0, T), \\ q_2'(t) = 0 & \text{a.e. } t \in (0, T), \\ q_1(0) = q_1(T), \end{cases} \tag{3.63}$$

$$\begin{aligned} \big\{u^*(t), v^*(t)\big\} = \operatorname*{argmax}_{\{u,v\}} \big\{&\big(u, q_1(t)\big) - \big(v, q_2(t)\big) \\ &- \lambda G\big(v - f_2(t), u - f_1(t)\big) + \lambda\big(y^*(t), v\big)\big\}. \end{aligned} \tag{3.64}$$

Clearly, $\lambda = 1$ because otherwise (i.e. if $\lambda = 0$), $q_1 \equiv q_2 \equiv 0$ which contradicts (3.64).

By (3.62)–(3.64) we have $q_2 \equiv C_2$, $q_1 - z^* \equiv C_1$ and by (3.64) we see that

$$\big\{y^*(t) - q_2(t), q_1(t)\big\} \in \partial G\big(v^*(t) - f_2(t), u^*(t) - f_1(t)\big) \quad \text{a.e. } t \in [0, T].$$

Since $(\partial G)^{-1} = \partial H$ we get (see Proposition 2.4)

$$\big\{v^*(t) - f_2(t), u^*(t) - f_1(t)\big\} \in \partial H\big(y^*(t) - C_2, z^*(t) + C_1\big).$$

We set

$$\bar{y} = y^*(t) - C_2, \quad p = z^*(t) + C_1. \tag{3.65}$$

We have (see (3.63))

$$\begin{cases} \bar{y}' \in \partial_p H(\bar{y}, p) + f_1(t) & \text{a.e. } t \in (0, T), \\ p' \in -\partial_y H(\bar{y}, p) + f_2(t) & \text{a.e. } t \in (0, T), \\ \bar{y}(0) = \bar{y}(T), \quad p(0) = p(T) \end{cases}$$

i.e., $(\bar{y}, p)$ is a solution of (3.59). We note also that, by assumption 1, we have

$$\begin{aligned} &G(q_1, q_2) \leqslant \sup_{\{y,p\}} \big\{(y, p_1) + (p, q_2) - \gamma_1|y|^2 - \gamma_2|p|^2\big\} - C_1 < \infty, \\ &\quad \forall (q_1, q_2) \in \mathbb{R}^n \times \mathbb{R}^n \end{aligned}$$

i.e., G is continuous on $\mathbb{R}^n \times \mathbb{R}^n$.

To conclude the proof, it remains to show that problem (3.61) has at least one solution.

Let $\{y_n, z_n, u_n, v_n\}$ be such that

$$\begin{aligned} &y_n' = u_n, \quad z_n' = -v_n, \quad y_n(0) = y_n(T), \quad z(0) = z_n(T), \\ &d \leqslant \int_0^T G(v_n - f_2, u_n - f_1)\,\mathrm{d}t - \int_0^T (y_n, v_n)\,\mathrm{d}t \leqslant d + \frac{1}{n}. \end{aligned} \tag{3.66}$$

where d is the infimum in (3.61).

By definition of G and by assumption 1, we have

$$\begin{aligned} &G(v_n - f_2, u_n - f_1) - (y_n, u_n) \\ &\quad \geqslant (y, v_n - f_2) + (p, u_n - f_1) - (y_n, v_n) - \omega\big(|y|^2 + |p|^2\big) \\ &\qquad \forall (y, p) \in \mathbb{R}^n \times \mathbb{R}^n \end{aligned}$$

where $\omega < \frac{\pi}{T}$. This yields

$$G(v_n - f_2, u_n - f_1) - (y_n, u_n) \geqslant \frac{1}{4\omega}\big((v_n - f_2)^2 + (u_n - f_1)^2\big) - (y_n, v_n). \tag{3.67}$$

On the other hand, the periodic solution to $y' = u$ is expressed as

$$y(t) = \sum_{m \neq 0} \frac{u^m \mathrm{e}^{\mathrm{i}\mu_m t}}{\mathrm{i}\mu_m}, \quad \mu_m = \frac{2m\pi}{T}.$$

Hence

$$y_n(t) = \sum_{m \neq 0} \frac{u_n^m \mathrm{e}^{\mathrm{i}\mu_m t}}{\mathrm{i}\mu_m}, \quad u_n^m = \frac{1}{T} \int_0^T u_n(t) \mathrm{e}^{-\mathrm{i}\mu_m t}\, \mathrm{d}t.$$

This yields

$$\int_0^T (y_n, v_n)\, \mathrm{d}t = \sum_{m \neq 0} \frac{u_n^m v_n^m}{\mathrm{i}\mu_m}, \quad v_n^m = \frac{1}{T} \int_0^T v_n \mathrm{e}^{-\mathrm{i}\mu_m t}\, \mathrm{d}t.$$

Hence by the Parseval formula

$$\int_0^T (y_n, v_n)\, \mathrm{d}t \leqslant \frac{T}{4\pi} \int_0^T \left(u_n^2 + v_n^2\right) \mathrm{d}t.$$

Then, by (3.67) and by assumption 1, we have

$$\int_0^T \big(G(v_n - f_2, u_n - f_1) - (y_n, u_n)\big)\, \mathrm{d}t \geqslant \alpha \int_0^T \left(|u_n|^2 + |v_n|^2\right) \mathrm{d}t + C.$$

Hence the sequences $\{u_n\}$, $\{v_n\}$ are bounded in $L^2(0, T; \mathbb{R}^n)$.

On a subsequence we have

$$\begin{aligned} u_n &\to u^* \quad \text{weakly in } L^2(0, T; \mathbb{R}^n), \\ v_n &\to v^* \quad \text{weakly in } L^2(0, T; \mathbb{R}^n). \end{aligned} \tag{3.68}$$

We set

$$\begin{aligned} \bar{y}_n(t) &= y_n(t) - \int_0^T y_n(t)\, \mathrm{d}t, \\ \bar{z}_n(t) &= z_n(t) - \int_0^T z_n(t)\, \mathrm{d}t. \end{aligned}$$

Then

$$\int_0^T \bar{y}_n\, \mathrm{d}t = \int_0^T \bar{z}_n\, \mathrm{d}t = 0, \quad \bar{y}_n' = u_n, \quad \bar{y}_n' = -v_n$$

and $\{\bar{y}_n\}$, $\{\bar{z}_n\}$ are bounded in $H^1(0,T;\mathbb{R}^n)$. Hence

$$\begin{aligned}\bar{y}_n \to y^* \quad & \text{strongly in } C([0,T];\mathbb{R}^n)\\ & \text{weakly in } H^1(0,T;\mathbb{R}^n),\\ \bar{z}_n \to z^* \quad & \text{strongly in } C([0,T];\mathbb{R}^n)\\ & \text{weakly in } H^1(0,T;\mathbb{R}^n)\end{aligned} \tag{3.69}$$

where $(y^*)' = u^*$, $(z^*)' = -v^*$, a.e. $t \in [0,T]$.

Clearly, $\{y^*, z^*, u^*, v^*\}$ is optimal in (3.61). This follows by (3.66) using (3.68), (3.69) and the fact that the convex integrand

$$(u,v) \to \int_0^T G(v - f_2, u - f_1)\,\mathrm{d}t$$

is weakly lower semicontinuous in $L^2(0,T;\mathbb{R}^n) \times L^2(0,T;\mathbb{R}^n)$. We have proved therefore

THEOREM 3.4. *Under assumptions* 1 *and* 2 *there exists a solution* $(y,p) \in H^1(0,T;\mathbb{R}^n) \times H^1(0,T;\mathbb{R}^n)$ *to the periodic problem* (3.59), (3.60).

3.4.3. *An application to resonant systems* Consider the problem

$$\min\left\{\int_0^T u(t)\,\mathrm{d}t;\ u \in U\right\}$$

where

$$\begin{aligned}U = \{u : (0,1) \to \mathbb{R},\ \text{measurable},\ 0 \leqslant u \leqslant B,\ \text{meas}[t\colon\ u(t) > 0] > 0,\\ y'' + uy = 0,\ y'(0) = y'(1) = 0 \text{ has at least one nontrivial solution}\}.\end{aligned}$$

THEOREM 3.5. *Assume that* $B > \pi^2$. *Then every optimal pair* (y^*, u^*) *satisfies*

$$u^*(t) = \begin{cases} B, & 0 \leqslant t \leqslant t_1,\\ 0, & t_1 < t < t_2,\\ B, & t_2 \leqslant t < 1.\end{cases} \tag{3.70}$$

PROOF. The state system is

$$y' = z, \quad z' = -uy, \quad z(0) = z(1) = 0.$$

By the maximum principle, Theorem 3.1, there exists $\lambda = 0, 1$ and (p_1, p_2) such that

$$\begin{aligned} & p_1' = up_2, \quad p_2' = -p_1 \quad \text{a.e. } t \in [0,1],\\ & p_2'(0) = 0, \quad p_2'(1) = 0,\end{aligned} \tag{3.71}$$

$$u^*(t) = \underset{u \in U}{\operatorname{argmax}}\left\{-uy^*(t)p_2(t) - \lambda u\right\} \quad \text{a.e. } t \in [0,T]. \tag{3.72}$$

We note that $p_2 = p$ is the solution to

$$\begin{cases} p'' + u^* p = 0 \quad \text{a.e. } t \in [0, T], \\ p(0) = p(1) = 0. \end{cases}$$

Hence p and y^* are both solutions to $y'' + u^* y = 0$ and therefore $p = Cy^*$ on $(0, 1)$. If $\lambda = 0$, by (3.72) we see that

$$u^*(t) = \operatorname*{argmax}_{u \in U} \bigl\{-C\bigl(y^*(t)\bigr)^2 u\bigr\}.$$

Hence, $u^*(t) = 0$ if $C > 0$ and $u^*(t) = B$ if $C < 0$. However $u^* \neq 0$ and $u^*(t) \neq B$ because we know that $u = \pi^2$ is admissible in our problem and $B > \pi^2$. Hence $\lambda \neq 0$ and so $\lambda = 1$. Then by (3.72) it follows that

$$-\bigl(y^*(t)p(t) + 1\bigr) \in N_U\bigl(u^*(t)\bigr) \subset N_{[0,B]}\bigl(u^*(t)\bigr) \quad \text{a.e. } t \in [0, 1].$$

Hence (see (2.15))

$$u^*(t) = \begin{cases} 0 & \text{if } y^*(t)p(t) + 1 > 0, \\ B & \text{if } y^*(t)p(t) + 1 < 0. \end{cases}$$

Equivalently

$$u^*(t) = \begin{cases} 0 & \text{if } C(y^*(t))^2 + 1 > 0, \\ B & \text{if } C(y^*(t))^2 + 1 < 0. \end{cases} \tag{3.73}$$

(This follows by an elementary argument as in [36].)

On the other hand, if u^* is optimal, then $|(y^*)'(t)| \neq 0$, $\forall t \in (0, 1)$. One can compute exactly (t_1, t_2) and find

$$v = \int_0^T u^*(t)\, \mathrm{d}t.$$

From here, we may conclude that for any $u \in L^\infty(0, T)$ with $0 \leqslant u \leqslant B$, if

$$\int_0^T u(t)\, \mathrm{d}t < v,$$

then the equation

$$y'' + uy = 0, \quad y'(0) = y'(1) = 0$$

has only the trivial solution $y = 0$.

This result is relevant in the theory of nonlinear resonant problems of the form

$$y'' + f(y) = 0, \quad y'(0) = y'(1) = 0. \qquad \square$$

3.5. *Reachable sets and optimal control problems*

In what follows we consider the controlled equation (1.11) supposed for simplicity to be autonomous:

$$y' = f^u(y) := f(y, u)$$

with the initial data

$$y(0) = y_0.$$

The hypotheses on f and the controller set $\mathcal{U}$ are such that for any $u \in \mathcal{U}$ the nonautonomous vector field $f^{u(t)}$ satisfies the hypotheses in Section 2.3. For example we will suppose that the controller set $\mathcal{U} = \{u : [0, T] \to U\colon\ u \in L^\infty(0, T)\}$ and $U \subset \mathbb{R}^m$ is closed.The vector field f is assumed to be C^∞ in y for any fixed $u \in U$ and f, $\frac{\partial f}{\partial y}$ are continuous in (y, u).

The *reachable set* (or *attainable set*) at moment t is defined as follows:

$$\mathcal{R}_{y_0}(t) = \left\{ y_0 \circ \overrightarrow{\exp} \int_0^t f^{u(\tau)}\,\mathrm{d}\tau \;\middle|\; u \in \mathcal{U} \right\}.$$

We denote by $y^u(t) = \overrightarrow{\exp} \int_0^t f^{u(\tau)}\,\mathrm{d}\tau$ and the flow corresponding to the equation $F_s^t = \overrightarrow{\exp} \int_s^t f^{u(\tau)}\,\mathrm{d}\tau$.

We prove in this section that optimal control problems reduce to the study of *reachable sets*.

THEOREM 3.6. *If $y^{u^*(T)} \in \partial \mathcal{R}_{y_0}(T)$, the boundary of $\mathcal{R}_{y_0}(T)$, then, for all $0 < s < T$, $y^{u^*(s)} \in \partial \mathcal{R}_{y_0}(s)$.*

PROOF. It is clear that $F_s^T(\mathcal{R}_{y_0}(s)) \subset \mathcal{R}_{y_0}(T)$ and, since F_s^T is a diffeomorphism, $F_s^T(\operatorname{int} \mathcal{R}_{y_0}(s)) \subset \operatorname{int} \mathcal{R}_{y_0}(T)$ and the conclusion is immediate. □

We will still call such trajectories *optimal* and the corresponding controller *optimal control.*

Optimal control problems We consider the following controlled equation

$$\begin{cases} y' = f^u(y), \\ y(0) = y_0. \end{cases} \tag{3.74}$$

The optimal control problem is to find the admissible strategy $u : [0, T] \to U$ such that the pair (u, y^u) minimizes a cost functional J. Depending on the form of the cost functional J we distinguished in Section 1.2 three types of optimal control problems: Lagrange, Mayer and Bolza.

We consider first of all the case of the *Lagrange problem* with fixed end points y_0, y_1 and cost functional $J(y,u) = \int_0^T L(y(t), u(t))\,\mathrm{d}t$, with the Lagrangean L satisfying similar hypotheses with f. We introduce the new variable j and consider the new system:

$$\begin{cases} j' = L(t,y,u), \\ y' = f^u(y) \end{cases} \tag{3.75}$$

with initial data

$$y(0) = y_0, \quad j(0) = 0.$$

We consider the set of controllers $u \in \mathcal{U}$ such that $y^u(T) = y_1$. Then it is clear that, if (u^*, y^{u^*}) is an optimal pair then

$$\big(j(T), y^{u^*}(T)\big) \in \partial\mathcal{R}_{(0,y_0)}(T)$$

where $R_{(0,y_0)}$ refers to the reachable set for system (3.75). This reduction has some inconvenient in the sense that it does not distinguish between trajectories realizing the minimum of J and those realizing the maximum. That is why one considers the extended set of controllers $\widetilde{\mathcal{U}} = \{(u,v) \mid u \in \mathcal{U}, v \in [0,+\infty)\}$ and the problem

$$\begin{cases} j' = L(t,y,u) + v, \\ y' = f^u(y), \\ y(0) = q_0, \quad j(0) = 0. \end{cases} \tag{3.76}$$

It is now clear that if u^* is optimal for the Lagrange problem then $(\tilde{u}, 0)$ is optimal in (3.76).

Since *Mayer problem* is just a particular case of a *Bolza problem*, we describe the reduction for the latter. In this case the cost functional is $J(y,u) = g(y(T)) + \int_0^T L(t, y(t), u(t))\,\mathrm{d}t$ and the functions g, L satisfy similar hypotheses to f. We introduce the new variable j, the extended set of controllers $\widetilde{U}$ and consider the new system:

$$\begin{cases} j' = L(t,y,u) + \displaystyle\sum_{i=1}^{n} \frac{\partial l}{\partial y_i}(y) f_i(y,u) + v, \\ y' = f^u(y), \\ q(0) = q_0, \quad j(0) = 0. \end{cases} \tag{3.77}$$

As in the case of the Lagrange problem, if u^* is optimal for the Bolza problem, then $(u^*, 0)$ is optimal in (3.77).

We have thus seen that the study of optimal control problems reduces to the study of reachable sets, more precisely to the study of those trajectories whose final points belong to the boundary of the reachable set at that moment.

3.6. *Geometric form of Pontriaghin maximum principle*

In this section we present a coordinate free version of the maximum principle (see also Section 3.1). The proof we give here follows the ideas in [2].

For $u \in \mathcal{U}$ we consider the Hamiltonian $H^u := (f^u)^\#$.

THEOREM 3.7. *Suppose that for the admissible control* $u^*(t), t \in [0,T]$, $y^{u^*}(T) \in \partial \mathcal{R}_{y_0}(T)$. *Then there exists a non-zero Lipschitz curve in the cotangent bundle* $\xi(t) \in T^*_{y^{u^*}(t)}M$ *solution of the Hamiltonian equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi(t) = \overrightarrow{H}^{u^*(t)}\bigl(\xi(t)\bigr)$$

and the following maximality condition is satisfied

$$\overrightarrow{H}^{u^*(t)}\bigl(\xi(t)\bigr) = \max_{u \in U} \overrightarrow{H}^{u(t)}\bigl(\xi(t)\bigr).$$

PROOF. The idea of the proof is to consider as curve $\xi(t)$ a curve of covectors "orthogonal" at each point to the boundary of the reachable set: $\partial\mathcal{R}_{y_0}(t)$.

Step 1. Let $\mathbf{T}: \mathbb{R}^N \to \mathbb{R}^n$ be Lipschitz continuous, $\mathbf{T}(0) = 0$ and differentiable at 0. Denote by $\mathbf{T}_0 = D\mathbf{T}(0)$ and by

$$\mathbb{R}^N_+ = \bigl\{(x_1, \dots, x_N) \in \mathbb{R}^N \mid x_i \geqslant 0\bigr\}.$$

We prove that if $\mathbf{T}_0(\mathbb{R}^N_+) = \mathbb{R}^n$, then for any neighborhood $\mathcal{V}$ of 0 in $\mathbb{R}^N$, $0 \in \operatorname{int} \mathbf{T}(\mathcal{V} \cap \mathbb{R}^N_+)$.

By the convexity of $\mathbb{R}^N_+$ it turns out that $\mathbf{T}_0|_{\operatorname{int}\mathbb{R}^N_+}$ is surjective and let $y \in \operatorname{int}\mathbb{R}^N_+$ with $\mathbf{T}_0(y) = 0$ and $\delta > 0$ such that $y + B_\delta \subset \operatorname{int}\mathbb{R}^N_+$.

It is clear that one may find an n-dimensional linear subspace $X \subset \mathbb{R}^N$ such that $\mathbf{T}_0(X) = \mathbb{R}^n$ so $\mathbf{T}_0|_X$ is invertible. Denote by $D = B_\delta \cap X$ and for $\varepsilon > 0$ the continuous functions $\mathbf{T}_\varepsilon : D \to \mathbb{R}^n$:

$$\mathbf{T}_\varepsilon(v) = \frac{1}{\varepsilon}\mathbf{T}\bigl(\varepsilon(y + v)\bigr).$$

One may easily check that $\mathbf{T}_\varepsilon \to \mathbf{T}_0$ uniformly on $\overline{D}$ and since $0 \in \operatorname{int}\mathbf{T}_0(D)$ it turns out, using an usual argument based on degree theory, that $0 \in \operatorname{int}\mathbf{T}_\varepsilon(D)$ for $\varepsilon > 0$ small enough or, equivalently, $0 \in \operatorname{int}\mathbf{T}(\varepsilon(y + B_\delta))$ which concludes the Step 1.

Step 2. We consider an admissible control $u(t)$ and compute the end point of the trajectory using the second form of variations of parameters formula (2.40):

$$\begin{aligned} y^u(T) &= y_0 \circ \overrightarrow{\exp} \int_0^T f^{u^*(t)} + \bigl(f^{u(t)} - f^{u^*(t)}\bigr)\,\mathrm{d}t \\ &= y^{u^*}(T) \circ \overrightarrow{\exp} \int_0^T \bigl(F_t^T\bigr)_* \bigl(f^{u(t)} - f^{u^*(t)}\bigr)\,\mathrm{d}t. \end{aligned} \tag{3.78}$$

Denote by

$$g^{t,u} = \left(F_t^T\right)_*\left(f^u - f^{u^*(t)}\right)$$

and let $\mathcal{L}$ be the set of Lebesgue points of u^* (see (3.17)). We suppose that the convex cone $\mathcal{W}$ generated by $\{g^{t,u}(y^{u^*}(T)) \mid t \in \mathcal{L}, u \in U\}$ is the whole tangent space $T_{y^{u^*}(T)}M$. We prove that, in this case, necessarily $y_1 := y^{u^*}(T) \in \operatorname{int}\mathcal{R}_{y_0}(T)$.

Indeed, one may find a finite set of points $0 < t_1 < \cdots < t_N < T$ and $u_1, \ldots, u_N \in U$ such that $\mathcal{W}$ is generated by the finite set $\{g^{t_i,u_i} \mid i = 1, \ldots, N\}$. For $x = (x_1, \ldots, x_N) \in \mathbb{R}_+^N$ we consider the strategy:

$$u^x(t) = \begin{cases} u_i, & t \in [t_i, t_i + x_i], \\ u^*(t), & t \in [0, T] \setminus \bigcup\limits_{i=1}^{N} [t_i, t_i + x_i]. \end{cases}$$

Variation of parameters formula (3.78) gives:

$$\begin{aligned} y^{u^x}(T) &= y_0 \circ \overrightarrow{\exp} \int_0^T f^{u^x(t)}\,\mathrm{d}t \\ &= y^{u^*}(T) \circ \overrightarrow{\exp} \int_{t_1}^{t_1+x_1} g^{t,u_1}\,\mathrm{d}t \circ \cdots \circ \overrightarrow{\exp} \int_{t_N}^{t_N+x_N} g^{t,u_n}\,\mathrm{d}t. \end{aligned} \tag{3.79}$$

If one considers the map:

$$\mathbf{T}(x_1, \ldots, x_N) = y^{u^x}(T), \quad x_1, \ldots, x_N \in \mathbb{R}$$

then it is easy to verify that $\mathbf{T}$ is Lipschitz continuous, $\mathbf{T}(0) = y_1$ and

$$\left.\frac{\partial \mathbf{T}}{\partial x_i}\right|_{(0,\ldots,0)} = g^{t_i,u_i}(y_1).$$

The hypotheses of Step 1 are verified and, consequently,

$$y_1 \in \operatorname{int}\mathbf{T}\left(\mathcal{V} \cap \mathbb{R}_+^N\right)$$

for any neighborhood $\mathcal{V}$ of 0 in $\mathbb{R}^N$ and, since u^x is an admissible strategy we find that

$$y_1 \in \operatorname{int}\mathcal{R}_{y_0}(T).$$

Step 3. Suppose now that $y^{u^*}(T) \in \partial\mathcal{R}_{y_0}(T)$. Since $0 \in \partial\mathcal{W}$, there exists a support hyperplane defined by $\xi(T) \in T^*_{y^{u^*}(T)}M$, $\xi(T) \neq 0$ such that

$$\xi(T)\left(g^{T,u}\left(y^{u^*}(T)\right)\right) \leqslant 0 \quad \text{a.e. } t \in [0, T], u \in U.$$

This means exactly that

$$\big[\big(F_t^T\big)^*\xi(T)\big]\big(f^{u^*}\big(y^{u^*}(t)\big)\big) \geqslant \big[\big(F_t^T\big)^*\xi(T)\big]\big(f^{u}\big(y^{u}(t)\big)\big).$$

Denote by

$$\xi(t) = \big(F_t^T\big)^*\xi(T).$$

The maximality condition is satisfied and, by (2.43)

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi(t) = \overrightarrow{H}^{u^*(t)}\big(\xi(t)\big)$$

which concludes the proof. □

It is now easy to recover the maximum principle for optimal control problems, as is expressed in Section 3.1 by using the transformation described in Section 3.5 and the geometric form of maximum principle, Theorem 3.7.

3.7. *Free time optimal control problems*

The free time optimal control problem

$$\min_{(u,T)}\left\{\int_0^T L\big(t, y(t), u(t)\big)\,\mathrm{d}t + g\big(y(0), y(T)\big);\ \big(y(0), y(T)\big) \in C\right\} \tag{3.80}$$

subject to

$$y' = f(y, u) \quad \text{a.e. } t \in (0, T)$$

can be reduced to a fixed time optimal control problem of the form (1.14) by substitution

$$t = \int_0^s w^2(\tau)\,\mathrm{d}\tau, \quad 0 \leqslant s \leqslant 1,$$
$$z(s) = y\big(t(s)\big), \qquad v(s) = u\big(t(s)\big).$$

This yields

$$\min_{(v,w)}\left\{\int_0^1 L\big(t(s), z(s), v(s)\big)w^2(s)\,\mathrm{d}s + g\big(z(0), z(1)\big);\ \big(z(0), z(1)\big) \in C\right\}$$

subject to

$$z'(s) = f\big(z(s), v(s)\big)w^2(s) \quad \text{a.e. } s \in (0, 1).$$

In this case, by Theorem 3.1 we find that the maximum principle has the form

$$\begin{aligned}
&p'(t) = -f_y(y^*, u^*)p + \lambda L_y(y^*, u^*),\\
&\big(p(0), -p(T)\big) \in \lambda \nabla g\big(y^*(0), y^*(T)\big) + N_C\big(y^*(0), y^*(T)\big),\\
&\big(p(t), f\big(t, y^*(t), u^*(t)\big)\big) - \lambda L\big(t, y^*(t), u^*(t)\big)\\
&\quad = \max_{(u\in U(t)}\big\{\big(p(t), f\big(y^*(t), u^*\big)\big) - \lambda L\big(y^*(t), u\big)\big\},\\
&\big(p(t), f\big(t, y^*(t), u^*(t)\big)\big) - \lambda L\big(t, y^*(t), u^*(t)\big) = 0 \quad \text{a.e. } t \in [0, T]
\end{aligned} \tag{3.81}$$

where $\lambda = 0, 1$ and $\|p(t)\| + \lambda \neq 0$, $\forall t \in [0, T]$.

In the special case, $L \equiv 1$, $g \equiv 0$, (3.80) reduces to the optimal time problem

$$\min_{(T,u)}\big\{T;\ \big(y(0), y(T)\big) \in C,\ u \in U(t),\ \text{a.e. } t \in (0, T)\big\}. \tag{3.82}$$

We leave to the reader to deduce the maximum principle (from (3.81)) in this case.

We note, however, that if $g = 0$, $C = \{y_0\} \times \{y_1\}$, then it reads as

$$\begin{aligned}
&p'(t) = -f_y(t, y^*, u^*)p && \text{a.e. in } (0, T),\\
&\big(p(t), f\big(t, y^*(t), u^*(t)\big)\big) - \lambda = 0 && \text{a.e. } t \in (0, T),\\
&\big(p(t), f\big(t, y^*(t), u^*(t)\big)\big) = \max_{u\in U(t)}\big(p(t), f\big(t, y^*(t), u\big)\big)
\end{aligned}$$

where $\lambda \in \{0, 1\}$ and $|p(t)| + \lambda \neq 0$, $\forall t \in [0, T]$. Hence $|p(t)| \neq 0$, $\forall t \in [0, T]$.

From the geometric point of view, in the case of Lagrange or Bolza problems considered in Section 3.5, if the final time T is free the conclusion is that for the equivalent problems (3.75), respectively (3.77), the optimal pair $((u^*, 0), (j^*, y^{u^*})$ satisfy, for $\varepsilon > 0$ small enough

$$\big(j^*(T), y^{u^*}(T)\big) \in \partial\Bigg(\bigcup_{|T-t|<\varepsilon} \mathcal{R}_{(0,y_0)}(t)\Bigg). \tag{3.83}$$

So, in general we obtain the following maximum principle for free time problems:

THEOREM 3.8. *Suppose that for the admissible control* $u^*(t)$, $t \in [0, T]$,

$$y^{u^*}(T) \in \partial\Bigg(\bigcup_{|T-t|<\varepsilon} \mathcal{R}_{y_0}(t)\Bigg)$$

Then there exists a non-zero Lipschitz curve in the cotangent bundle $\xi(t) \in T^*_{y^{u^*}(t)}$ *solution of the Hamiltonian equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi(t) = \overrightarrow{H}^{u^*(t)}\big(\xi(t)\big)$$

and the following maximality condition is satisfied

$$H^{u^*(t)}\big(\xi(t)\big) = \max_{u\in U} H^{u(t)}\big(\xi(t)\big). \tag{3.84}$$

Moreover, the following condition holds:

$$H^{u^*(t)}\big(\xi(t)\big) = 0 \quad \text{a.e. } t \in [0,T]. \tag{3.85}$$

PROOF. As before, the case of free time may be reduced to the fixed time situation by considering reparameterizations of the trajectories of the initial system and introduce a supplementary control related to the time rescaling in the way described below. First of all, one considers a reparameterization (a new time scale):

$$t = \gamma(s), \quad \gamma' > 0.$$

Then the solution $y^u(s) = y^u(\gamma(s))$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}s} y^u(s) = \gamma'(s) f^{u(\gamma(s))}(y).$$

The modified system is

$$\begin{cases} y' = \varphi f^u(y), & u \in U, |\varphi - 1| < \dfrac{\varepsilon}{T} < 1, \\ y(0) = y_0. \end{cases}$$

Admissible controls are measurable, bounded functions of the form $v(t) = (\varphi(t), u(t))$, $|\varphi - 1| \leqslant \frac{\varepsilon}{T}$ and the corresponding solution is denoted by $y^v(t)$. If we denote by $v^*(t) = (1, u^*(t))$ then it is clear that, since $y^{u^*}(T) \in \partial(\bigcup_{|T-t|<\varepsilon} \mathcal{R}_{y_0}(t))$ it follows that $y^{v^*}(T) \in \partial\mathcal{R}_{y_0}(T)$ and, at this point, the maximum principle may be applied. The Hamiltonian is

$$H^v(\xi) = \varphi H^u(\xi), \quad v = (\varphi, u).$$

The Hamiltonian system for H^{v^*} is the same as for H^{u^*}. The maximality condition becomes

$$H_{v^*}\big(\xi(t)\big) = H^{u^*(t)}\big(\xi(t)\big) = \max_{|\varphi-1|<\varepsilon T, u\in U} \varphi H^u\big(\xi(t)\big)$$

from which conditions (3.84), (3.85) follow. □

Transversality conditions If in the optimal control problem of Lagrange we consider the initial and final states belonging to some submanifolds $y_0 \in M_0$, $y_1 \in M_1$, then supplementary transversality conditions appear, that is, the adjoint flow should satisfy (compare with (3.2))

$\xi(0)$ orthogonal to M_0,

$\xi(T)$ orthogonal to M_1.

4. The dynamic programming equation

The optimal controller u^* determined on $[0,T]$ via the optimality system (maximum principle) is also referred as *open loop optimal controller*. An alternative way is to look for an optimal controller expressed in feedback form. Such a controller is called *closed loop optimal controller* and it is determined from a Hamilton–Jacobi equation.

4.1. *Optimal feedback controllers and smooth solutions to Hamilton–Jacobi equation*

Consider the optimal control problem

$$\min\left\{\int_0^T L\big(t,y(t),u(t)\big)\,\mathrm{d}t+\ell\big(y(T)\big);\ y'(t)=f\big(t,y(t),u(t)\big),\right.$$
$$\left. y(0)=y_0,\ u(t)\in U(t),\ \text{a.e. } t\in[0,T]\right\}. \tag{4.1}$$

This is a special case of problem (1.14) where $g(y_1,y_2)=l(y_2)$ and $C=\{y_0\}\times\mathbb{R}^n$. The control $u:[0,T]\to\mathbb{R}^m$ is of course measurable. We shall assume as usually that

$$L:[0,T]\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R},\quad f:[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R},\quad \ell:\mathbb{R}^n\to\mathbb{R}$$

satisfy the assumption

$$\big\|L_y(t,y,u)\big\|+\big\|f_y(t,y,u)\big\|\leqslant\alpha(t,u),\quad \forall u\in U(t),\ t\in[0,T]$$

and ℓ, L, f are measurable in t, continuous in u while $U(t)$ is closed and $t\to U(t)$ measurable.

A function

$$V:[0,T]\times\mathbb{R}^n\to\mathbb{R}^m$$

Borel measurable is said to be a *feedback controller* for the system $y'=f(t,y,u)$ if for each $(t_0,y_0)\in[0,T]\times\mathbb{R}^n$ the Cauchy problem

$$\begin{cases} y'(t)=f\big(t,y(t),V\big(t,y(t)\big)\big), & \text{a.e. } t\in(t_0,T),\\ y(t_0)=y_0 \end{cases} \tag{4.2}$$

has at least one absolutely continuous solution $y\in C([0,T];\mathbb{R}^n)$. The system (4.2) is called a *closed loop system*.

A control $u(t)=V(t,y(t))$ where y is the solution to (4.2) is called *feedback controller*. In other words, a *feedback controller* (or *control*) is an input function (control) which is expressed as a function of time and of the state of the system in the present time.

If $u(t)=V(t,y(t))$ is optimal for problem (4.1), then this feedback control is called *optimal feedback control*.

The representation of an optimal control in the feedback form (i.e. as a feedback controller) is called the *synthesis problem* of optimal control. The function V is also called *synthesis function* for the corresponding optimal control problem.

The synthesis problem is closely related to a first order partial differential equation. This is the *Hamilton–Jacobi–Bellman* or the *dynamic programming equation* associated to problem (4.1) and has the following form:

$$\begin{cases} \varphi_t(t,x) - \sup\limits_{u\in U(t)} \big\{-\big(\varphi_x(t,x), f(t,x,u)\big) - L(t,x,u)\big\} = 0 \\ \qquad\qquad\qquad\qquad \forall x \in \mathbb{R}^n,\ t \in [0,T], \\ \varphi(T,x) = \ell(x) \quad \forall x \in \mathbb{R}^n \end{cases} \tag{4.3}$$

where

$$\varphi_t = \frac{\partial\varphi}{\partial t}, \quad \varphi_x = \frac{\partial\varphi}{\partial x}.$$

If we denote by $H : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ the Hamiltonian function

$$H(t,x,p) = \sup\big\{-\big(p, f(t,x,u)\big) - L(t,x,u);\ u \in U(t)\big\} \tag{4.4}$$

then we may rewrite (4.3) as (compare with Eq. (1.10)):

$$\begin{cases} \varphi_t(t,x) - H\big(t,x,\varphi_x(t,x)\big) = 0, & t \in (0,T), x \in \mathbb{R}^n, \\ \varphi(T,x) = \ell(x), & x \in \mathbb{R}^n. \end{cases} \tag{4.5}$$

In order to make clear the relationship between this equation and problem (4.1) we introduce the function

$$\varPhi(t,x) = \operatorname*{argsup}_{u\in U(t)}\big\{-\big(\varphi_x(t,x), f(t,x,u)\big) - L(t,x,u)\big\}. \tag{4.6}$$

Let $\psi : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ be the optimal value function associated with the optimal control problem, i.e.,

$$\psi(t,x) = \inf_u \bigg\{ \int_t^T L\big(x, y(s), u(s)\big)\,\mathrm{d}s + \ell\big(y(T)\big); \\ y'(s) = f\big(s, y(s), u(s)\big),\ s \in (t,T),\ y(t) = x \bigg\}. \tag{4.7}$$

THEOREM 4.1. *Let $\varphi \in C^1([0,T] \times \mathbb{R}^n)$ be a solution to Hamilton–Jacobi equation, with the Cauchy condition $\varphi(T) = \ell(x)$. Assume that $\varPhi$ is a feedback controller. Then*

$$\varphi(t,x) = \psi(t,x), \quad \forall (t,x) \in (0,T) \times \mathbb{R}^n$$

and $\varPhi$ is an optimal feedback controller.

PROOF. Let (t, x) be fixed and let y^t be the solution to

$$\big(y^t\big)'(s) = f\big(s, y^t, \Phi\big(s, y^t(s)\big)\big), \quad s \in (t, T), \qquad y^t(t) = x. \tag{4.8}$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi\big(s, y^t(s)\big) = \varphi_s\big(s, y^t(s)\big) + \Big(\varphi_y\big(s, y^t(s)\big), f\big(s, y^t(s), \Phi\big(s, y^t(s)\big)\big)\Big)$$
$$\text{a.e. } s \in (t, T), \tag{4.9}$$

Then by (4.6) we see that

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi\big(s, y^t(s)\big) = \varphi_s\big(s, y^t(s)\big) - H\big(s, y^t(s), \varphi_y\big(s, y^t(s)\big)\big)$$
$$- L\big(s, y^t(s), \Phi\big(s, y^t(s)\big)\big) \quad \text{a.e. } s \in (t, T).$$

By (4.5) we get

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi\big(s, y^t(s)\big) = -L\big(s, y^t(s), \Phi\big(s, y^t(s)\big)\big) \quad \text{a.e. } s \in (t, T).$$

Hence

$$\varphi(t, x) = \int_t^T L\big(s, y^t(s), \Phi\big(s, y^t(s)\big)\big)\,\mathrm{d}s + \ell\big(y^t(T)\big) \geqslant \psi(t, x). \tag{4.10}$$

On the other hand, if (y, v) is any admissible pair into problem (4.1), i.e.,

$$v(s) \in U(s) \quad \text{a.e. } s \in (t, T), \quad L\big(s, y(s), v(s)\big) \in L^1(t, T)$$

and

$$\begin{cases} y'(s) = f\big(s, y(s), v(s)\big) \quad \text{a.e. } s \in (t, T), \\ y(t) = x \end{cases}$$

we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi\big(s, y(s)\big) = \varphi_s\big(s, y(s)\big) + \big(\varphi_x\big(s, y(s)\big), f\big(s, y(s), v(s)\big)\big)$$
$$\geqslant \varphi_s\big(s, y(s)\big) - H\big(s, y(s), \varphi_x\big(s, y(s)\big)\big) - L\big(s, y(s), v(s)\big)$$
$$\text{a.e. } s \in (t, T).$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}s}\varphi\big(s, y(s)\big) \geqslant -L\big(s, y(s), v(s)\big) \quad \text{a.e. } s \in (t, T)$$

and so

$$\varphi(t,x) \leqslant \int_t^T L\big(s, y(s), v(s)\big)\,\mathrm{d}s + \ell\big(y(T)\big). \tag{4.11}$$

Since (y,v) was arbitrary we conclude by (4.10) and (4.11) that $\varphi = \psi$ as claimed. Moreover, by (4.10) it follows that (for $t=0$) $u = \Phi(t, y(t))$ is an optimal feedback controller. This completes the proof. □

In this way, if the Hamilton–Jacobi equation has a C^1 solution then the synthesis problem reduces to this equation. Let us consider, in the next section, a simple special case.

4.2. *Linear quadratic control problems*

Let $U(t) \equiv \mathbb{R}^m$ and let

$$\begin{cases} f(t,y,u) \equiv b(t) + A(t)x + B(t)u \quad \forall x \in \mathbb{R}^n,\ u \in \mathbb{R}^m,\ t \in (0,T), \\ L(t,y,u) \equiv \dfrac{1}{2}\big(Q(t)x, x\big) + \dfrac{1}{2}|u|^2, \\ \ell(x) = \dfrac{1}{2}(P_0 x, x) \end{cases} \tag{4.12}$$

where

$$\begin{aligned} &A(t) \in L^\infty\big(0,T; L\big(\mathbb{R}^n, \mathbb{R}^n\big)\big), \quad B(t) \in L^\infty\big(0,T; L\big(\mathbb{R}^m, \mathbb{R}^n\big)\big), \\ &P_0 \in L\big(\mathbb{R}^n, \mathbb{R}^n\big), \quad b \in L^\infty\big(0,T;\mathbb{R}^n\big), \quad Q \in L^\infty\big(0,T; L\big(\mathbb{R}^n, \mathbb{R}^n\big)\big). \end{aligned}$$

Here we have denoted by $(\cdot,\cdot)$ the scalar product and by $|\cdot|$ the Euclidean norm in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. We shall assume further that P_0, $Q(t)$ are symmetric positive matrices, i.e.,

$$(P_0 x, x) \geqslant 0 \qquad \forall x \in \mathbb{R}^n, \tag{4.13}$$

$$\big(Q(t)x, x\big) \geqslant 0 \quad \forall x \in \mathbb{R}^n,\ t \in [0,T]. \tag{4.14}$$

In this case the Hamiltonian function is (see (4.4))

$$\begin{aligned} H(t,x,p) &= \sup_{u \in \mathbb{R}^m} \left\{ -\big(p, A(t)x + B(t)u + b\big) - \frac{1}{2}|u|^2 - \frac{1}{2}\big(Q(t)x, x\big) \right\} \\ &= -\frac{1}{2}\big(Q(t)x, x\big) - \big(p, b(t)\big) - \big(p, A(t)x\big) \\ &\quad + \sup_{u \in \mathbb{R}^m} \left\{ -\big(B^*(t)p, u\big) - \frac{1}{2}|u|^2 \right\} \\ &= -\frac{1}{2}\big(Q(t)x, x\big) + \frac{1}{2}\big|B^*(t)p\big|^2 - \big(p, b(t)\big) - \big(A^*(t)p, x\big) \end{aligned}$$

and so the Hamilton–Jacobi equation (4.5) becomes

$$\begin{cases} \varphi_t(t,x) - \frac{1}{2}\big|B^*(t)\varphi_x(t,x)\big|^2 + \big(\varphi_x(t,x), b(t)\big) \\ \quad + \big(A(t)x, \varphi_x(t,x)\big) + \frac{1}{2}\big(Q(t)x, x\big) = 0, \quad t \in (0,T),\ x \in \mathbb{R}^n, \\ \varphi(T,x) = \frac{1}{2}(P_0 x, x). \end{cases} \tag{4.15}$$

Moreover, by (4.6) we see that

$$\Phi(t,x) = -B^*(t)\varphi_x(t,x) \quad \forall t \in (0,T),\ x \in \mathbb{R}^n. \tag{4.16}$$

The form of Eq. (4.15) suggests that φ should be a quadratic function in x. Indeed, if we look for φ under the form

$$\varphi(t,x) = \frac{1}{2}\big(P(t)x, x\big) + \big(r(t), x\big) \quad \forall t \in (0,T),\ x \in \mathbb{R}^n$$

we obtain by (4.15) (note that $\varphi_x = P(t)x + r(t)$)

$$\begin{aligned} &\frac{1}{2}\big(P'(t)x, x\big) + \big(r(t), x\big) - \frac{1}{2}\big|B^*(t)P(t)x + B^*(t)r(t)\big|^2 \\ &\quad + \big(A(t)x, P(t)x + r(t)\big) + \big(P(t)x + r(t), b(t)\big) + \frac{1}{2}\big(Q(t)x, x\big) = 0, \\ &P(T) = P_0, \quad r(T) = 0. \end{aligned}$$

Differentiating with respect to x, we get

$$\begin{aligned} &P'(t)x + r'(t) - P(t)B(t)B^*(t)P(t)x + A^*(t)P(t)x + P(t)A(t)x \\ &\quad + A^*(t)r - P(t)B(t)B^*(t)r(t) + P(t)b(t) + Q(t)x = 0. \end{aligned}$$

Hence $P(t)$ must be a solution to the Riccati equation

$$\begin{cases} P'(t) + A^*(t)P(t) + P(t)A(t) - P(t)B(t)B^*(t)P(t) + Q(t) = 0, \\ \quad t \in (0,T), \\ P(T) = P_0 \end{cases} \tag{4.17}$$

while r satisfies the linear equation

$$\begin{cases} r'(t) + A^*(t)r(t) - P(t)B(t)B^*(t)r(t) + P(t)b(t) = 0 \\ \quad \text{a.e. } t \in (0,T), \\ r(T) = 0. \end{cases} \tag{4.18}$$

The optimal feedback control is expressed as (see (4.16))

$$u(t) = \Phi(t,y) = -B^*(t)P(t)x - B^*(t)r(t), \quad t \in (0,T). \tag{4.19}$$

In this way the synthesis problem for the linear quadratic optimal control problem (4.1), (4.12) reduces to the Riccati equation (4.17).

It turns out that under our assumptions the Cauchy problem (4.17) has a unique solution, globally defined (see [4,38])

$$P \in W^{1,\infty}\big([0,T];\mathbb{R}^n\big), \quad P(t) = P^*(t), \quad P(t) \geqslant 0 \quad \forall t \in [0,T]$$

and this fact proves that a feedback operator for this problem exists.

REMARK 4.1. The hypothesis $Q \geqslant 0$ is essential for existence of optimal pairs. If Q is not positive, then it is possible that an optimal pair does not exist and the dynamic programming equation (4.3) does not have a solution defined for $t \in [0,T]$. More precisely, if we still search for a solution as a quadratic form, then the Riccati equation (4.17) may not have a globally defined solution. This is deeply related to the existence of *conjugate points*, almost as in the classical calculus of variations. For more details on conjugate points see [26,2].

EXAMPLE. Let us consider the problem

$$\min\left\{\int_0^T \big(y^2+u^2\big)\,\mathrm{d}t;\; y' = y+u,\; y(0) = y_0\right\}. \tag{4.20}$$

In this case the optimal feedback control is given by

$$u(t) = -P(t)y$$

where P is the solution of the equation

$$\begin{cases} P'(t) + 2P(t) - P^2(t) + 2 = 0, \quad t \in (0,T), \\ P(T) = 0. \end{cases}$$

We get

$$P(t) = 2\sqrt{3}C\frac{\mathrm{e}^{-2t/\sqrt{3}}}{1 - C\mathrm{e}^{-2t/\sqrt{3}}} + 1 + \sqrt{3}$$

where C is to be determined from the condition $P(T) = 0$.

The above discussion remains true for infinite dimensional linear quadratic problems of the form (4.1), (4.2) where $\mathbb{R}^n$ and $\mathbb{R}^m$ are replaced by the Hilbert spaces H and U and

$$B(t) \in L^\infty\big(0,T;L(U,H)\big), \quad P_0 \in L(H,H)$$

and

$$A(t) : H \to H$$

is a family of linear closed operators which generate an evolution operator in H. In this case, however the Riccati equation (4.17) should be understood in a weak sense.

4.3. *Viscosity solutions*

Coming back to Eq. (4.5) it must be said, however, that in general it has not a classical solution. Thus a new concept of solution has been introduced by M.G. Crandall and P.L. Lions in [20] (see also [17]).

Consider the equation

$$\varphi_t(t,x) - H(t,x,\varphi_x) = 0, \quad (t,x) \in Q = [0,T] \times \Omega. \tag{4.21}$$

The function $\varphi \in C(Q)$ is said to be a viscosity solution to Eq. (4.1) if for each $\chi \in C^1(\overline{Q})$ the following conditions hold:

1. If $\varphi - \chi$ has local maximum at $(t_0, x_0) \in Q$ then
$$\chi_t(t_0,x_0) - H\big(t_0,x_0,\chi_x(t_0,x_0)\big) \geqslant 0.$$
2. If $\varphi - \chi$ has a local minimum at (t_0, x_0) then
$$\chi_t(t_0,x_0) - H\big(t_0,x_0,\chi_x(t_0,x_0)\big) \leqslant 0.$$

We have

THEOREM 4.2. *The optimal value function ψ defined by* (4.7) *is a viscosity solution to Hamilton–Jacobi equation* (4.21).

PROOF. Let $(t_0, x_0) \in Q = (0,T) \times \mathbb{R}^n$ and let $\chi \in C^1([0,T] \times \mathbb{R}^n)$ such that $\psi - \chi$ has a local maximum at (t_0, x_0), i.e.,

$$\psi(t_0,x_0) - \chi(t_0,x_0) \geqslant \psi(t,x) - \chi(t,x), \quad \forall (t,x) \in V \tag{4.22}$$

where V is a neighborhood of (t_0, x_0).

Let (y,u) be an admissible pair on $[t_0, T]$, such that $y(t_0) = x_0$ and u is continuous on $[t_0, T]$. We have

$$\begin{cases} y'(s) = f\big(s, y(s), u(s)\big) & \text{a.e. } s \in (t_0,T),\ u(s) \in U(s), \\ y(t_0) = x_0. \end{cases} \tag{4.23}$$

In (4.22) we take $t \geqslant t_0$ and $x = y(t)$ and we have

$$\begin{aligned} \chi(t_0,x_0) - \chi\big(t,y(t)\big) &\leqslant \psi(t_0,x_0) - \psi\big(t,y(t)\big) \\ &\leqslant \int_{t_0}^t L\big(\tau, y(\tau), u(\tau)\big)\,\mathrm{d}\tau, \quad t \geqslant t_0. \end{aligned} \tag{4.24}$$

Indeed for each $0 \leqslant t \leqslant s \leqslant T$ we have

$$\psi(t,x) = \inf\left\{\int_t^s L\big(\tau, y(\tau,t,x), u(\tau)\big)\,\mathrm{d}\tau + \psi\big(s, y(s,t,x,u)\big);\ u(\tau) \in U\right\}. \tag{4.25}$$

Here (s,t,x,u) is the solution to system (4.23) with the initial point (t,x). This equality is the so called dynamic programming principle and follows by a direct calculation. Now if we divide (4.24) by $t - t_0$ and let $t \to t_0$ we get

$$-\chi_t(t_0,x_0) - \big(\chi_x(t_0,x_0), f\big(t_0,x_0,u(t_0)\big)\big) \leqslant L\big(t_0,x_0,u(t_0)\big).$$

Since $u(t_0)$ is arbitrary in $U(t_0)$ we obtain

$$-\chi_t(t_0,x_0) + H\big(t_0,x_0,\chi_x(t_0,x_0)\big) \leqslant 0$$

as claimed.

Assume now that $\psi - \chi$ has a local minimum at (t_0,x_0). We have

$$\chi(t_0,x_0) - \chi(t,x) \geqslant \psi(t_0,x_0) - \psi(t,x) \quad \forall (t,x) \in V$$

where V is a neighborhood of (t_0,x_0). Let (y,u) be an admissible pair such that $y(t_0) = x_0$. For $\varepsilon > 0$ sufficiently small we have

$$\begin{aligned}
&\psi\big(t_0+\varepsilon, y(t_0+\varepsilon)\big) - \psi(t_0,x_0) \geqslant \chi\big(t_0+\varepsilon, y(t_0+\varepsilon)\big) - \chi(t_0,x_0)\\
&\quad = \int_{t_0}^{t_0+\varepsilon} \frac{\mathrm{d}}{\mathrm{d}t}\chi\big(t,y(t)\big)\,\mathrm{d}t\\
&\quad = \int_{t_0}^{t_0+\varepsilon} \chi_t\big(t,y(t)\big) + \big(\chi_y\big(t,y(t)\big), f\big(t,y(t),u(t)\big)\big)\,\mathrm{d}t.
\end{aligned} \tag{4.26}$$

If

$$\chi_t(t_0,x_0) - H\big(t_0,x_0,\chi_x(t_0,x_0)\big) \geqslant \lambda > 0$$

then we have

$$\chi_t(t_0,x_0) + \big(\chi_x(t_0,x_0), f(t_0,x_0,u)\big) + L\big(t,y(t),u(t)\big) \geqslant \lambda, \quad \forall u \in U(t_0).$$

Hence

$$\chi_t\big(t,y(t)\big) + \big(\chi_y\big(t,y(t)\big), f\big(t,y(t),u(t)\big)\big) + L\big(t,y(t),u(t)\big) \geqslant \frac{\lambda}{2}$$

for $t_0 \leqslant t \leqslant t_0 + \varepsilon$, ε sufficiently small.

Then by (4.26) we see that

$$\psi\big(t_0+\varepsilon, y(t_0+\varepsilon)\big)-\psi(t_0,x_0)\geqslant\frac{\varepsilon\lambda}{2}-\int_{t_0}^{t_0+\varepsilon}L\big(t,y(t),u(t)\big)\,\mathrm{d}t.$$

The latter contradicts the principle of dynamic programming (4.25). This completes the proof.

In particular, Theorem 4.2 implies the existence of a viscous solution for the Hamilton–Jacobi equation (4.5) with final Cauchy condition.

The uniqueness of such a viscosity solution is more delicate and follows by some sharp arguments developed by Crandall, Lions and Ishii (see [17–20]). □

Let us conclude this section with a representation result for the viscosity solution to Hamilton–Jacobi equation in the case of a convex Hamiltonian, i.e. (see [4,5])

$$\begin{aligned} &f(t,x,u)\equiv Ax+Bu, && U \text{ convex},\\ &L(t,x,u)\equiv L(x,u), && L \text{ convex}. \end{aligned}$$

Denote by M the conjugate function

$$M(q,w)=\sup\big\{(q,v)+(w,y)-L(y,v);\ v\in U,\ y\in\mathbb{R}^n\big\}.$$

Then, as seen earlier, we have

$$\begin{aligned}\psi(t,x)&=-\inf\bigg\{\int_t^T M\big(B^*p,w\big)\,\mathrm{d}s+\ell^*\big(-p(T)\big)+\big(p(t),x\big);\\ &\qquad\qquad p'=-A^*p+w \text{ a.e. in } (t,T)\bigg\}\\ &=-\inf_q\bigg\{\ell^*(-q)+\inf_w\bigg\{\int_t^T M\big(B^*p,w\big)\,\mathrm{d}s+\big(p(t),x\big);\\ &\qquad\qquad p'=-A^*p+w,\ p(T)=q\bigg\}\bigg\}\\ &=-\inf_q\big\{\ell^*(-q)+\theta(t,q)\big\}\end{aligned}$$

where

$$\begin{aligned}\theta(t,q)&=\inf\bigg\{\int_t^T M\big(B^*p,w\big)\,\mathrm{d}s+\big(p(t),x\big);\ p'=-A^*p+w,\ p(T)=q\bigg\}\\ &=\inf\bigg\{\int_t^{T-t} M\big(B^*z,v\big)\,\mathrm{d}s+\big(z(T-t),x\big);\ z'=A^*z+v,\ z(0)=q\bigg\}\\ &=\inf_w\bigg\{\int_t^T M\big(B^*z,w\big)\,\mathrm{d}s+\big(z(T),x\big);\ z'=A^*z+w,\ z(t)=q\bigg\}.\end{aligned}$$

In other words, θ is the solution to Hamilton–Jacobi equation

$$\begin{cases} \theta_t(t,q) + \big(A^*q, \theta_q(t,q)\big) - \sup\limits_{w}\big\{-\big(\theta_q(t,q), Bw\big) - M\big(B^*q, w\big);\, w \in \mathbb{R}^n\big\} = 0, \\ \theta(T,q) = (x,q). \end{cases} \tag{4.27}$$

We have proved therefore

THEOREM 4.3. *The viscosity solution to Hamilton–Jacobi equation*

$$\begin{cases} \varphi_t(t,x) + \big(Ax, \varphi_x(t,x)\big) - \sup\limits_{u\in U}\big\{(-\varphi_x, Bu) - L(x,u)\big\} = 0, \\ \varphi(T,x) = \ell(x) \end{cases} \tag{4.28}$$

is given by

$$\varphi(t,x) = -\inf_q\big\{\ell^*(-q) + \theta(t,q)\big\} \tag{4.29}$$

where θ is solution to (4.27).

Let us consider the special case

$$L(x,u) \equiv h(u) + g(x), \quad A \equiv 0, \quad B \equiv I.$$

Then Eq. (4.28) reduces to (h^* is the conjugate of h)

$$\begin{cases} \varphi_t(t,x) - h^*\big(-\varphi_x(t,x)\big) + g(x) = 0, \\ \varphi(T,x) = \ell(x) \end{cases} \tag{4.30}$$

and so Theorem 4.3 yields

$$\varphi(t,x) = -\inf_q\big\{\ell^*(-q) + \theta(t,q)\big\} \tag{4.31}$$

where

$$\begin{cases} \theta_t(t,q) - g^*\big(-\theta_q(t,q)\big) + h^*\big(-B^*q\big) = 0, \\ \theta(T,q) = (x,q) \quad \forall t \in [0,T],\ q \in \mathbb{R}^n. \end{cases} \tag{4.32}$$

For $g = 0$ the solution φ to the problem

$$\begin{cases} \varphi_t(t,x) - h^*\big(-\varphi(t,x)\big) = 0 \\ \varphi(T,x) = \ell(x) \end{cases} \tag{4.33}$$

is therefore given by (Lax–Hopf formula)

$$\varphi(t,x) = -\inf_q\bigl\{\ell^*(-q) + (x,q) + (T-t)h^*(-q)\bigr\}$$
$$= \inf_y\left\{\ell(y) + (T-t)h\left(\frac{y-x}{T-t}\right)\right\}, \quad t \in (0,T). \tag{4.34}$$

As is easily seen, this formula remains true for a non convex function ℓ.

EXAMPLE. The equation

$$\begin{cases} \varphi_t - c\|\varphi_x\| = 0, & x \in \mathbb{R}^n, t \in (0,T), \\ \varphi(0,x) = \varphi_0(x) \end{cases} \tag{4.35}$$

is called the *eikonal equation* and models the flame propagation ($\varphi(t,x)$ is the characteristic equation of the burnt region Ω_t at moment t).

For $\psi = -\varphi$ we get

$$\begin{cases} \psi_t + c\|\psi_x\| = 0, \\ \psi(0,x) = \ell(x), & \ell = -\varphi_0 \end{cases}$$

and by (4.34) we get

$$\psi(t,x) = \inf_y\left\{\ell(y) + th\left(\frac{y-x}{t}\right)\right\}$$
$$= \inf_y\bigl\{\ell(y);\ \|y-x\| \leqslant ct\bigr\}$$
$$= \inf_y\bigl\{-\varphi_0(y);\ \|y-x\| \leqslant ct\bigr\}$$

where

$$h(u) = \begin{cases} 0, & |u| \leqslant c, \\ +\infty, & |u| > c. \end{cases}$$

Finally

$$\varphi(t,x) = \sup\bigl\{\varphi_0(y);\ \|x-y\| \leqslant ct\bigr\}.$$

4.4. *On the relation between the two approaches in optimal control theory*

We make now more precise the relationship between the maximum principle and the dynamic programming equation.

We consider the optimal problem (4.1). As we have seen in Section 3.1, if u^* is an optimal control, then there exist $\lambda \in \{0, 1\}$ and $p(t)$ (not both 0), solution to the following system:

$$\begin{cases} y' = f(t, y^*, u^*) = \dfrac{\partial}{\partial p} H^{u^*}(t, y^*, p), \\ p' = -\big(p, f_y(t, y^*, u^*)\big) + \lambda L_y(t, y^*, u^*) = -\dfrac{\partial}{\partial y} H^{u^*}(t, y^*, p), \\ p(T) = -\nabla l\big(y^*(T)\big) \end{cases} \tag{4.36}$$

where $H^u(t, y, p) = (p, f(t, y, u)) - \lambda L$ and

$$H^{u^*(t)}\big(t, y^*(t), p(t)\big) = \max_{u \in U} H^u\big(t, y * (t), p(t)\big).$$

In Sections 4.1, 4.3, we derived a partial differential equation satisfied in the viscosity sense by the value function ψ:

$$\begin{cases} \psi_t - H(t, x, \psi_x) = 0, \\ \psi(T, x) = l(x) \end{cases} \tag{4.37}$$

and $H(t, x, p) = \max_u H^u(t, x, -p)$. In fact, if we denote by $S = -\psi$, then

$$S_t + H^{u^*}(t, x, S_x) = 0$$

and this is just the Hamilton–Jacobi equation (see Section 1.1) corresponding to the Hamiltonian system (4.36) that appears in the maximum principle of Pontriaghin.

The next theorem states for this control problem the maximum principle by making use of the Bellman equation. The hypotheses are more restrictive than in Theorem 3.1. Anyhow, this makes more precise the relationship between the two branches of control theory and it is in a certain sense the analogous of the Jacobi theorem in calculus of variations. For a general setting of this problem we refer to [14] (see also [22]).

THEOREM 4.4. *Suppose that u^* is an optimal control and the value function ψ is C^2. Then the problem is normal and if we denote by*

$$p = -\psi_x\big(t, y^*(t)\big)$$

then p satisfies (4.36) *with* $\lambda = 1$.

PROOF. As seen in Theorem 4.2, ψ is a viscosity solution of (4.37) and since it is C^1, it is a classical solution. Thus

$$\frac{\mathrm{d}}{\mathrm{d}t} p(t) = -\psi_{tx}\big(t, y^*(t)\big) - \psi_{xx}\big(t, y^*(t)\big) f(t, y^*, u^*). \tag{4.38}$$

By the other hand, by the dynamic programming principle

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\psi\big(t, y^*(t)\big) &= -L\big(t, y^*(t), u^*(t)\big)\\ &= \psi_t\big(t, y^*(t)\big) + \big(\psi_x\big(t, y^*(t)\big), f\big(t, y^*(t), u^*(t)\big)\big).\end{aligned}$$

This, combined with Bellman equation tells us that $(y^*(t), u^*(t))$ realizes the infimum of $(x, u) \to \psi_t(t, x) + (\psi_x(t, x), f(t, x, u)) + L(t, x, u)$. This implies that the derivative with respect to x computed in $(t, y^*(t), u^*(t))$ is 0 that is

$$\begin{aligned}&\psi_{xt}\big(t, y^*(t)\big) + \psi_{xx}\big(t, y^*(t)\big) f\big(t, y^*(t), u^*(t)\big)\\ &\quad + \big(\psi_x\big(t, y^*(t)\big), f_y\big(t, y^*(t), u^*(t)\big)\big) + L_y\big(t, y^*(t), u^*(t)\big) = 0\end{aligned}$$

which combined with (4.38) concludes the proof. □

References

[1] A.A. Agrachev and R.V. Gamkrelidze, *Exponential representation of flows and a chronological enumeration*, Mat. Sb. (N.S.) **107(149)** (1978), 467–532, 639.

[2] A.A. Agrachev and Y.L. Sachkov, *Control Theory from the Geometric Viewpoint*, Encyclopaedia Math. Sci., Vol. 87, Springer-Verlag, Berlin (2004).

[3] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, Grad. Texts in Math., Vol. 60, second edition, Springer-Verlag, New York (1989). Translated from Russian by K. Vogtmann and A. Weinstein.

[4] V. Barbu, *Mathematical Methods in Optimization of Differential Systems*, Math. Appl., Vol. 310, Kluwer Academic, Dordrecht (1994). Translated and revised from the 1989 Romanian original.

[5] V. Barbu and G. Da Prato, *Hamilton–Jacobi Equations in Hilbert Spaces*, Pitman Res. Notes Math. Ser., Vol. 86, Pitman, Boston, MA (1983).

[6] V. Barbu and Th. Precupanu, *Convexity and Optimization in Banach Spaces*, Math. Appl. (East European Series), Vol. 10, Romanian edition, D. Reidel, Dordrecht (1986).

[7] R. Bellman, *Dynamic Programming*, Princeton Univ. Press, Princeton, NJ (1957).

[8] G.A. Bliss, *Calculus of Variations*, Carus Math. Monogr., Vol. 1, sixth edition, Math. Assoc. America, Washington; The Open Court Publishing Company, La Salle, IL (1971).

[9] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Math. Stud., Vol. 5, North-Holland, Amsterdam (1973).

[10] F.H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.

[11] F.H. Clarke, *The maximum principle under minimal hypotheses*, SIAM J. Control Optim. **14** (1976), 1078–1091.

[12] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Math. Soc. Ser. Monogr. Adv. Texts, Wiley-Interscience, New York (1983).

[13] F.H. Clarke and I. Ekeland, *Nonlinear oscillations and boundary value problems for Hamiltonian systems*. Arch. Rational Mech. Anal. **78** (1982), 315–333.

[14] F.H. Clarke and R.B. Vinter, *The relationship between the maximum principle and dynamic programming*, SIAM J. Control Optim. **25** (1987), 1291–1311.

[15] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York (1955).

[16] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. II: *Partial Differential Equations* (Vol. II by R. Courant), Wiley-Interscience, New York (1962).

[17] M.G. Crandall, L.C. Evans and P.-L. Lions, *Some properties of viscosity solutions of Hamilton–Jacobi equations*, Trans. Amer. Math. Soc. **282** (1984), 487–502.

[18] M.G. Crandall, H. Ishii and P.-L. Lions, *Uniqueness of viscosity solutions of Hamilton–Jacobi equations revisited*, J. Math. Soc. Japan **39** (1987), 581–596.

[19] M.G. Crandall, H. Ishii and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), 1–67.

[20] M.G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton–Jacobi equations*, Trans. Amer. Math. Soc. **277** (1983), 1–42.

[21] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. **47** (1974), 324–353.

[22] W.H. Fleming and R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Appl. Math. (N.Y.), Vol. 1, Springer-Verlag, New York (1975).

[23] R. Gamkrelidze, *Exponential representation of solutions of ordinary differential equations*, Proceedings of Equadiff IV, Prague 1977, Lecture Notes in Math., Vol. 703, Springer-Verlag (1977), 118–129.

[24] I.M. Gelfand and S.V. Fomin, *Calculus of Variations*, translated from Russian and edited by R.A. Silverman, reprint of the 1963 original, Dover, Mineola, NY (2000).

[25] A. Isidori, *Nonlinear Control Systems*, Comm. Control Engrg. Ser., third edition, Springer-Verlag, Berlin (1995).

[26] V. Jurdjevic, *Geometric Control Theory*, Cambridge Stud. Adv. Math., Vol. 52, Cambridge Univ. Press, Cambridge (1997).

[27] J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, translated from French by S.K. Mitter, Grundlehren Math. Wiss., Band 170, Springer-Verlag, New York (1971).

[28] J.E. Marsden and T.S. Ratiu, *Introduction to Mechanics and Symmetry*, Texts Appl. Math., Vol. 17, second edition, Springer-Verlag, New York (1999). A basic exposition of classical mechanical systems.

[29] J. Moreau, *Fonctionnelles convexes, Séminaire sur les Équations aux Dérivées Partielles (1966–1967). II*, Collège de France, Paris (1967).

[30] L. Pontriaghin, V. Boltianski, R. Gamkrelidze and E. Mishchenko, *Théorie mathématique des processus optimaux*, Mir, Moscow (1974). Traduit du russe par D. Embarek.

[31] R.T. Rockafellar, *Convex Analysis*, Princeton Math. Ser., Vol. 28, Princeton Univ. Press, Princeton, NJ (1970).

[32] R.T. Rockafellar, *Integral functionals, normal integrands and measurable selections*, Nonlinear Operators and the Calculus of Variations (Summer School, Univ. Libre Bruxelles, Brussels, 1975), Lecture Notes in Math., Vol. 543, Springer-Verlag, Berlin (1976), 157–207.

[33] E.D. Sontag, *Mathematical Control Theory*, Texts Appl. Math., Vol. 6, second edition, Springer-Verlag, New York (1998). Deterministic finite-dimensional systems.

[34] V.M. Tikhomirov, *Stories about Maxima and Minima*, translated from Russian by A. Shenitzer, Math. World, Vol. 1, Amer. Math. Soc., Providence, RI; Math. Assoc. America, Washington, DC (1990).

[35] I.I. Vrabie, *C_0-Semigroups and Applications*, North-Holland Math. Stud., Vol. 191, North-Holland, Amsterdam (2003).

[36] H.Z. Wang and Y. Li, *Neumann boundary value problems for second-order ordinary differential equations across resonance*, SIAM J. Control Optim. **33** (1995), 1312–1325.

[37] S. Yosida, *An optimal control problem of the prey-predator system*, Funkcial. Ekvac. **25** (1982), 283–293.

[38] J. Zabczyk, *Mathematical Control Theory: An Introduction*, Systems Control Found. Appl., Birkhäuser, Boston, MA (1992).

This page intentionally left blank

CHAPTER 2

Hamiltonian Systems: Periodic and Homoclinic Solutions by Variational Methods

Thomas Bartsch

Mathematisches Institut, Universität Giessen, Arndtstr. 2, 35392 Giessen, Germany

Andrzej Szulkin

Department of Mathematics, Stockholm University, 106 91 Stockholm, Sweden

Contents

0. Introduction . . . 79
1. Critical point theory . . . 81
    - 1.1. Basic critical point theory . . . 81
    - 1.2. Critical point theory for strongly indefinite functionals . . . 90
2. Periodic solutions . . . 98
    - 2.1. Variational setting for periodic solutions . . . 98
    - 2.2. Periodic solutions near equilibria . . . 106
    - 2.3. Fixed energy problem . . . 110
    - 2.4. Superlinear systems . . . 115
    - 2.5. Asymptotically linear systems . . . 119
    - 2.6. Spatially symmetric Hamiltonian systems . . . 121
3. Homoclinic solutions . . . 125
    - 3.1. Variational setting for homoclinic solutions . . . 125
    - 3.2. Existence of homoclinics . . . 130
    - 3.3. Multiple homoclinic solutions . . . 135
    - 3.4. Multibump solutions and relation to the Bernoulli shift . . . 138

References . . . 142

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 2
Edited by A. Cañada, P. Drábek and A. Fonda

© 2005 Elsevier B.V. All rights reserved

0. Introduction

The complex dynamical behavior of Hamiltonian systems has attracted mathematicians and physicists ever since Newton wrote down the differential equations describing planetary motions and derived Kepler's ellipses as solutions. Hamiltonian systems can be investigated from different points of view and using a large variety of analytical and geometric tools. The variational treatment of Hamiltonian systems goes back to Poincaré who investigated periodic solutions of conservative systems with two degrees of freedom using a version of the least action principle. It took however a long time to turn this principle into a useful tool for finding periodic solutions of a general Hamiltonian system

$$\begin{cases} \dot{p} = -H_q(p,q,t), \\ \dot{q} = H_p(p,q,t) \end{cases} \tag{HS}$$

as critical points of the Hamiltonian action functional

$$\Phi(p,q) = \int_0^{2\pi} p \cdot \dot{q}\,\mathrm{d}t - \int_0^{2\pi} H(p,q,t)\,\mathrm{d}t$$

defined on a suitable space of 2π-periodic functions $(p,q) : \mathbb{R} \to \mathbb{R}^N \times \mathbb{R}^N$. The reason is that this functional is unbounded from below and from above so that the classical methods from the calculus of variations do not apply. Even worse, the quadratic form

$$(p,q) \mapsto \int_0^{2\pi} p \cdot \dot{q}\,\mathrm{d}t$$

has infinite-dimensional positive and negative eigenspaces. Therefore Φ is said to be strongly indefinite. For strongly indefinite functionals refined variational methods like Morse theory or Lusternik–Schnirelmann theory still do not apply. These were originally developed for the closely related problem of finding geodesics and extended to many other ordinary and partial differential equations, in particular to the second order Hamiltonian system

$$\ddot{q} = -V_q(q,t)$$

where the associated Lagrangian functional

$$J(q) = \frac{1}{2}\int_0^{2\pi} \dot{q}^2\,\mathrm{d}t - \int_0^{2\pi} V(q,t)\,\mathrm{d}t$$

is not strongly indefinite.

A major breakthrough was the pioneering paper [78] of Rabinowitz from 1978 who obtained for the first time periodic solutions of the first order system (HS) by the above mentioned variational principle. Some general critical point theory for indefinite functionals was subsequently developed in the 1979 paper [19] by Benci and Rabinowitz. Since then the number of papers on variational methods for strongly indefinite functionals and on applications to Hamiltonian systems has been growing enormously. These methods are not restricted to periodic solutions but can also be used to find heteroclinic or homoclinic orbits and to prove complex dynamics. In fact, they can even be applied to infinite-dimensional Hamiltonian systems and strongly indefinite partial differential equations having a variational structure.

The goal of this chapter is to present an introduction to variational methods for strongly indefinite functionals like Φ and its applications to the Hamiltonian system (HS). The chapter is divided into three sections. Section 1 is concerned with critical point theory, Section 2 with periodic solutions, and Section 3 with homoclinic solutions of (HS). We give proofs or sketches of proofs for selected basic theorems and refer to the literature for more advanced results. No effort is being made to be as general as possible. Neither did we try to write a comprehensive survey on (HS). The recent survey [81] of Rabinowitz in Volume 1A of the Handbook of Dynamical Systems facilitated our task considerably. We chose our topics somewhat complementary to those treated in [81] and concentrated on the first order system (HS), though a certain overlap cannot and should not be avoided. As a consequence we do not discuss second order systems nor do we discuss convex Hamiltonian systems where one can work with the dual action functional which is not strongly indefinite. One more topic which we have not included—though it has recently attracted attention of many researchers—is the problem of finding heteroclinic solutions by variational methods. These and many more topics are being treated in a number of well written monographs dealing with variational methods for Hamiltonian systems, in particular [1,5,33,52,69,73,80]. Further references can be found in these books and in Rabinowitz' survey [81]. Naturally, the choice of the topics is also influenced by our own research experience.

Restricting ourselves to variational methods we do not touch upon the dynamical systems approach to Hamiltonian systems which includes perturbation theory, normal forms, stability, KAM theory, etc. An introduction to these topics can be found for instance in the textbooks [47,75]. Also we do not enter the realm of symplectic topology and Floer homology dealing with Hamiltonian systems on symplectic manifolds. Here we refer the reader to the monograph [74] and the references therein.

We conclude this introduction with a more detailed description of the contents. In Section 1 we consider pertinent results in critical point theory. Particular emphasis is put on a rather simple and direct approach to strongly indefinite functionals.

Section 2 is concerned with periodic solutions of (HS). We present a unified approach, via a finite-dimensional reduction in order to show the existence of one solution, and via a Galerkin-type method in order to find more solutions. Subsection 2.2 concerns the existence of periodic solutions near equilibria (Lyapunov-type results) and in Subsection 2.3 the fixed energy problem is considered (finding solutions of a priori unknown period which lie on a prescribed energy surface). The remaining subsections consider the existence and the number of periodic solutions under different growth conditions on the Hamiltonian and for spatially symmetric Hamiltonians.

Section 3 deals with homoclinic solutions for (HS) with time-periodic Hamiltonian. Here we present a few basic existence and multiplicity results and discuss a relation to the

Bernoulli shift and complicated dynamics. The proofs are more sketchy than in Section 2 because we did not want to enter too much into technicalities which are more complex than in the periodic case. Moreover, the subject of this section is still rapidly developing and has not been systematized in the same way as the periodic solution problem.

1. Critical point theory

1.1. *Basic critical point theory*

Let E be a real Hilbert space with an inner product $\langle .,. \rangle$ and Φ a functional in $C^1(E, \mathbb{R})$. Via the Riesz representation theorem we shall identify the Fréchet derivative $\Phi'(x) \in E^*$ with a corresponding element of E, and we shall write $\langle \Phi'(x), y \rangle$ rather than $\Phi'(x)y$. Our goal here is to discuss those methods of critical point theory which will be useful in our applications to Hamiltonian systems. In particular, although most of the results presented here can be easily extended to real Banach spaces, we do not carry out such extension as it will not be needed for our purposes.

Recall that $\{x_j\}$ is said to be a *Palais–Smale sequence* (a (PS)-*sequence* in short) if $\Phi(x_j)$ is bounded and $\Phi'(x_j) \to 0$. The functional Φ satisfies the *Palais–Smale condition* (the (PS)-*condition*) if each (PS)-sequence possesses a convergent subsequence. If $\Phi(x_j) \to c$ and $\Phi'(x_j) \to 0$, we shall sometimes refer to $\{x_j\}$ as a $(\mathrm{PS})_c$-sequence. Φ satisfies the Palais–Smale condition at the level c (the $(\mathrm{PS})_c$-condition) if every $(\mathrm{PS})_c$-sequence has a convergent subsequence.

We shall frequently use the following notation:

$$\Phi^c := \{x \in E\colon\ \Phi(x) \leqslant c\},$$
$$K := \{x \in E\colon\ \Phi'(x) = 0\}, \qquad K_c := \{x \in K\colon\ \Phi(x) = c\}.$$

One of the basic technical tools in critical point theory is the deformation lemma. Below we state a version of it, called the quantitative deformation lemma. It is due to Willem [101], see also [23, Theorem I.3.4] and [102, Lemma 2.3].

A continuous mapping $\eta : A \times [0,1] \to E$, where $A \subset E$, is said to be a *a deformation of A in E* if $\eta(x, 0) = x$ for all $x \in A$. Denote the distance from x to the set B by $d(x, B)$.

LEMMA 1.1. *Suppose $\Phi \in C^1(E, \mathbb{R})$ and let $c \in \mathbb{R}$, $\bar{\varepsilon}, \delta > 0$ and a set $N \subset E$ be given. If*

$$\|\Phi'(x)\| \geqslant \delta \quad \textit{whenever} \quad d(x, E \setminus N) \leqslant \delta \quad \textit{and} \quad |\Phi(x) - c| \leqslant \bar{\varepsilon}, \tag{1.1}$$

then there exists an $\varepsilon \in (0, \bar{\varepsilon})$, depending only on $\bar{\varepsilon}$ and δ, and a deformation $\eta : E \times [0,1] \to E$ such that:

(i) *$\eta(x,t) = x$ whenever $|\Phi(x) - c| \geqslant \bar{\varepsilon}$;*
(ii) *$\eta(\Phi^{c+\varepsilon} \setminus N, 1) \subset \Phi^{c-\varepsilon}$ and $\eta(\Phi^{c+\varepsilon}, 1) \subset \Phi^{c-\varepsilon} \cup N$;*
(iii) *The mapping $t \mapsto \Phi(\eta(x,t))$ is nonincreasing for each $x \in E$.*

PROOF. Since the argument is well known, we omit some details. A complete proof may be found, e.g., in [23, Theorem I.3.4] or [102, Lemma 2.3].

A mapping $V : E \setminus K \to E$ is said to be a *pseudo-gradient vector field* for Φ if V is locally Lipschitz continuous and satisfies

$$\|V(x)\| \leqslant 2\|\Phi'(x)\|, \qquad \langle \Phi'(x), V(x)\rangle \geqslant \|\Phi'(x)\|^2 \tag{1.2}$$

for each $x \in E \setminus K$. It is well known and not difficult to prove that any $\Phi \in C^1(E, \mathbb{R})$ has a pseudo-gradient vector field; see, e.g., [23,80,102].

Let $\chi : E \to [0, 1]$ be a locally Lipschitz continuous function such that

$$\chi(x) = \begin{cases} 0 & \text{if } |\Phi(x) - c| \geqslant \bar{\varepsilon} \text{ or } d(x, E \setminus N) \geqslant \delta, \\ 1 & \text{if } |\Phi(x) - c| \leqslant \bar{\varepsilon}/2 \text{ and } d(x, E \setminus N) \leqslant \delta/2 \end{cases}$$

and consider the Cauchy problem

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = -\frac{1}{2}\delta\chi\big(\eta(x,t)\big)\frac{V(\eta(x,t))}{\|V(\eta(x,t))\|}, \quad \eta(x, 0) = x.$$

Since the vector field above is locally Lipschitz continuous and bounded, $\eta(x,t)$ is uniquely determined and continuous for each $(x, t) \in E \times \mathbb{R}$. It is now easy to see that (i) is satisfied. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi\big(\eta(x,t)\big) = \left\langle \Phi'\big(\eta(x,t)\big), \frac{\mathrm{d}\eta}{\mathrm{d}t} \right\rangle \leqslant -\frac{\delta}{4}\chi\big(\eta(x,t)\big)\big\|\Phi'\big(\eta(x,t)\big)\big\| \tag{1.3}$$

according to (1.2). Hence also (iii) holds.

Let $x \in \Phi^{c+\varepsilon} \setminus N$ and $0 < \varepsilon \leqslant \bar{\varepsilon}/2$. In order to establish the first part of (ii) we must show that $\Phi(\eta(x, 1)) \leqslant c - \varepsilon$. Since

$$\|\eta(x,t) - x\| \leqslant \int_0^t \left\|\frac{\mathrm{d}\eta}{\mathrm{d}s}\right\| \mathrm{d}s \leqslant \frac{1}{2}\delta t,$$

$d(\eta(x,t), E \setminus N) \leqslant \delta/2$ whenever $0 \leqslant t \leqslant 1$. We may assume $\Phi(\eta(x, 1)) \geqslant c - \bar{\varepsilon}/2$ (otherwise we are done). Then, according to (1.3) and the definition of χ,

$$\begin{aligned} \Phi\big(\eta(x,1)\big) &= \Phi(x) + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\Phi\big(\eta(x,t)\big)\,\mathrm{d}t \leqslant \Phi(x) - \frac{\delta}{4}\int_0^1 \big\|\Phi'\big(\eta(x,t)\big)\big\|\,\mathrm{d}t \\ &\leqslant c + \varepsilon - \frac{\delta^2}{4}. \end{aligned}$$

Hence $\Phi(\eta(x, 1)) \leqslant c - \varepsilon$ if we choose $\varepsilon \leqslant \min\{\bar{\varepsilon}/2, \delta^2/8\}$.

In order to prove the second part of (ii) it remains to observe that if $x \in \Phi^{c+\varepsilon}$ and $\eta(x, 1) \notin N$, then $d(\eta(x,t), E \setminus N) \leqslant \delta/2$ and therefore again $\Phi(\eta(x, 1)) \leqslant c - \varepsilon$. □

We emphasize that the constant ε is independent of the functional Φ and the space E as long as Φ satisfies (1.1). We shall make repeated use of this fact.

It is easy to see that if Φ satisfies (PS) and N is a neighbourhood of K_c, then there exist $\bar{\varepsilon}, \delta > 0$ such that (1.1) holds.

Next we introduce the concept of local linking, due to Li and Liu [60]. Let $\Phi \in C^1(E, \mathbb{R})$ and denote the ball of radius r and center at the origin by B_r. The corresponding sphere will be denoted by S_r. The function Φ is said to satisfy the local linking condition at 0 if there exists a subspace $F_0 \subset E$ and $\alpha, \rho > 0$ such that F_0 and $F_0^\perp$ have positive dimension,

$$\Phi \leqslant 0 \quad \text{on } F_0 \cap \bar{B}_\rho, \qquad \Phi \leqslant -\alpha \quad \text{on } F_0 \cap S_\rho \tag{1.4}$$

and

$$\Phi \geqslant 0 \quad \text{on } F_0^\perp \cap \bar{B}_\rho, \qquad \Phi \geqslant \alpha \quad \text{on } F_0^\perp \cap S_\rho. \tag{1.5}$$

We shall denote the inner product of x and y in $\mathbb{R}^m$ by $x \cdot y$ and we set $|x| := (x \cdot x)^{1/2}$. For a symmetric matrix B we denote the Morse index of the quadratic form corresponding to B by $M^-(B)$.

THEOREM 1.2. *Suppose $\Phi \in C^1(\mathbb{R}^m, \mathbb{R})$ satisfies the local linking conditions* (1.4) *and* (1.5) *for some $F_0 \subset \mathbb{R}^m$. Then Φ has a critical point $\bar{x}$ with $|\Phi(\bar{x})| \geqslant \alpha$ in each of the following two cases*:

(i) *There exists $R > 0$ such that $\Phi < 0$ in $\mathbb{R}^m \setminus B_R$;*

(ii) *$\Phi(x) = \frac{1}{2}Bx \cdot x + \psi(x)$, where $\psi'(x) = o(|x|)$ as $|x| \to \infty$, B is a symmetric invertible matrix and $M^-(B) > \dim F_0$.*

PROOF. We first consider case (ii) which is more difficult. If there exists a critical point $\bar{x}$ with $\Phi(\bar{x}) \leqslant -\alpha$, we are done. If there is no such point, then there exists a pseudogradient vector field V whose domain contains $\Phi^{-\alpha}$, and since $|\Phi'(x)|$ is bounded away from 0 as $|x|$ is large, $|\Phi'(x)| \geqslant \delta$ (where $\delta > 0$) whenever $x \in \Phi^{-\alpha}$. Hence the Cauchy problem

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = -V\big(\gamma(x,t)\big), \quad \gamma(x,0) = x$$

has a solution for all $x \in \Phi^{-\alpha}$, $t \geqslant 0$ and

$$\begin{aligned} \Phi\big(\gamma(x,t)\big) &= \Phi(x) + \int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\Phi\big(\gamma(x,s)\big)\,\mathrm{d}s \leqslant -\alpha - \int_0^t \big|\Phi'\big(\gamma(x,s)\big)\big|^2\,\mathrm{d}s \\ &\leqslant -\delta^2 t. \end{aligned} \tag{1.6}$$

Choose $R > 0$ such that $|\psi'(x)| \leqslant \frac{1}{2}\lambda_0|x|$ for all $|x| \geqslant R$, where $\lambda_0 := \{\inf|\lambda_j|\colon \lambda_j$ is an eigenvalue of $B\}$. Let $\mathbb{R}^m = F^+ \oplus F^-$, with $F^\pm$ respectively being the positive and the negative space of B. For $x \in \mathbb{R}^m$ write $x = x^+ + x^-$, $x^\pm \in F^\pm$. By (1.6) and the form of Φ, $|\gamma(x,T)^-| \geqslant R$ for any $x \in F_0 \cap S_\rho$ provided T is large enough.

Let S^n and D^{n+1} be the unit sphere and the unit closed ball in $\mathbb{R}^{n+1}$. Recall that a space X is called l-connected if any mapping from S^n to X, $0 \leqslant n \leqslant l$, can be extended to a mapping from D^{n+1} to X (cf. [84, Section 1.8]). We want to show that the set $\Phi^{-\alpha} \cap$

$\{x \in \mathbb{R}^m\colon |x^-| \geqslant R\}$ is $(k-1)$-connected if $k < M^-(B)$, possibly after choosing a larger R. This will imply in particular that any homeomorphic image of S^{k-1} contained in this set is contractible there. Let $r(x,t) := (1-t)x^+ + x^-$, $0 \leqslant t \leqslant 1$. Then r is a strong deformation retraction of $\Phi^{-\alpha} \cap \{x \in \mathbb{R}^m\colon |x^-| \geqslant R\}$ onto $F^- \setminus B_R$. To see this, we only need to verify that $r(x,t) \in \Phi^{-\alpha}$ for all x, t. Suppose first $Bx^+ \cdot x^+ \leqslant -\frac{1}{2}Bx^- \cdot x^-$. Then $|x^+| \leqslant C|x^-|$ for some C; thus $\psi((1-t)x^+ + x^-) = \mathrm{o}(|x^-|^2)$ as $|x^-| \to \infty$ and

$$\begin{aligned}\Phi\bigl(r(x,t)\bigr) &= \frac{1}{2}(1-t)^2 Bx^+ \cdot x^+ + \frac{1}{2}Bx^- \cdot x^- + \psi\bigl((1-t)x^+ + x^-\bigr)\\ &\leqslant \frac{1}{4}Bx^- \cdot x^- + \psi\bigl((1-t)x^+ + x^-\bigr) \leqslant -\alpha\end{aligned}$$

if R is large enough. Let now $Bx^+ \cdot x^+ \geqslant -\frac{1}{2}Bx^- \cdot x^-$; then $|x^-| \leqslant D|x^+|$ and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Phi\bigl(r(x,t)\bigr) &= -\Phi'\bigl(r(x,t)\bigr) \cdot x^+\\ &= -Bx^+ \cdot x^+ - \psi'\bigl((1-t)x^+ + x^-\bigr) \cdot x^+ \leqslant 0,\end{aligned}$$

again provided R is large enough. Hence in this case $\Phi(r(x,t)) \leqslant \Phi(x) \leqslant -\alpha$. Since $F^- \setminus B_R$ is homeomorphic to $S^{l-1} \times [R, \infty)$, where $l := M^-(B)$, $F^- \setminus B_R$ is $(k-1)$-connected for any $k < l$. It follows that so is the set $\Phi^{-\alpha} \cap \{x \in \mathbb{R}^m\colon |x^-| \geqslant R\}$.

The set $\{\gamma(x,T)\colon x \in F_0 \cap S_\rho\}$ is contained in $\Phi^{-\alpha} \cap \{x \in \mathbb{R}^m\colon |x^-| \geqslant R\}$ and homeomorphic to S^{k-1}, $k < M^-(B)$. Hence it can be contracted to a point x^* in $\Phi^{-\alpha} \cap \{x \in \mathbb{R}^m\colon |x^-| \geqslant R\}$. Denote this contraction by γ_0, let $D_0 := F_0 \cap \bar{B}_\rho$, $D := D_0 \times [0,1]$ and define a mapping $f : \partial D \to \mathbb{R}^m$ by setting

$$f(x_0, s) := \begin{cases} x_0, & s = 0,\ x_0 \in D_0,\\ \gamma(x_0, 2sT), & 0 \leqslant s \leqslant \frac{1}{2},\ x_0 \in \partial D_0 = F_0 \cap S_\rho,\\ \gamma_0\bigl(\gamma(x_0, T), 2s-1\bigr), & \frac{1}{2} \leqslant s \leqslant 1,\ x_0 \in \partial D_0,\\ x^*, & s = 1,\ x_0 \in D_0. \end{cases} \tag{1.7}$$

It is clear from the construction that $\Phi(f(x_0, s)) \leqslant 0$ whenever $(x_0, s) \in \partial D$ and < 0 if $x_0 \in \partial D$. Let

$$\Gamma := \bigl\{g \in C\bigl(D, \mathbb{R}^m\bigr)\colon g|_{\partial D} = f\bigr\} \tag{1.8}$$

and

$$c := \inf_{g \in \Gamma} \max_{(x_0,s) \in D} \Phi\bigl(g(x_0, s)\bigr). \tag{1.9}$$

We shall show that $f(\partial D)$ links $F_0^\perp \cap S_\rho$ in the sense that if $g \in \Gamma$, then $g(x_0, s) \in F_0^\perp \cap S_\rho$ for some $(x_0, s) \in D$. Assuming this, we see that the maximum in (1.9) is always $\geqslant \alpha$, and hence $c \geqslant \alpha$. We claim c is a critical value. Otherwise $K_c = \emptyset$, so (1.1) holds for g with $N = \emptyset$, $\bar{\varepsilon} \in (0, c)$ because Φ satisfies the (PS)-condition as a consequence of (ii). Let $\varepsilon < \bar{\varepsilon}$ and η be as in Lemma 1.1. Let $g \in \Gamma$ be such that $g(x_0, s) \in \Phi^{c+\varepsilon}$ for all $(x_0, s) \in D$. Since $\Phi(f(x_0, s)) \leqslant 0$ if $(x_0, s) \in \partial D$ and $\eta(x, 1) = x$ for $x \in \Phi^0$, the mapping $(x_0, s) \mapsto \eta(g(x_0, s), 1)$ is in Γ. But this is impossible according to the definition of c because $\Phi(\eta(g(x_0, s), 1)) \leqslant c - \varepsilon$ for all $(x_0, s) \in D$.

It remains to show that $f(\partial D)$ links $F_0^{\perp} \cap S_\rho$. Write $x = x_0 + x_0^{\perp} \in F_0 \oplus F_0^{\perp}$, $D_0^{\perp} = F_0^{\perp} \cap \overline{B}_\rho$. For $g \in \Gamma$ we consider the map

$$G : \bigl(D_0 \oplus D_0^{\perp}\bigr) \times [0, 1] \to \mathbb{R}^m, \quad G(x, s) = x_0^{\perp} - g(x_0, s).$$

If there is no linking, then $G(x, s) \neq 0$ for some $g \in \Gamma$ and all $x_0 \in D_0$, $x_0^{\perp} \in \partial D_0^{\perp}$, $0 \leqslant s \leqslant 1$. For $x_0 \in \partial D_0$ we have $g(x_0, s) = f(x_0, s)$, hence $\Phi(g(x_0, s)) < 0$ and $G(x, s) \neq 0$ (because $\Phi(x_0^{\perp}) \geqslant 0$). It follows that $G(x, s) \neq 0$ when $x \in \partial(D_0 \oplus D_0^{\perp})$ and G is an admissible homotopy for Brouwer's degree. Hence

$$\deg\bigl(G(.,0), D_0 \oplus D_0^{\perp}, 0\bigr) = \deg\bigl(G(.,1), D_0 \oplus D_0^{\perp}, 0\bigr).$$

Since $G(x, 0) = x_0^{\perp} - g(x_0, 0) = x_0^{\perp} - f(x_0, 0) = x_0^{\perp} - x_0$, the degree on the left-hand side above is $(-1)^{\dim F_0}$. On the other hand, $f(x_0, 1) = x^*$, where x^* is a point outside $D_0 \oplus D_0^{\perp}$; hence $G(x, 1) \neq 0$ for any $x \in D_0 \oplus D_0^{\perp}$ and the degree is 0. This contradiction completes the proof of (ii).

In case (i) the argument is similar but simpler. Suppose there is no critical point $\bar{x}$ with $\Phi(\bar{x}) \leqslant -\alpha$. Since $\Phi \leqslant 0$ in $\mathbb{R}^m \setminus B_R$, $|\gamma(x_0, T)| \geqslant R$ for some $T > 0$ and all $x_0 \in F_0 \cap S_\rho$. It is obvious that the set $\{\gamma(x_0, T){:}\ x_0 \in F_0 \cap S_\rho\}$ (which is homeomorphic to a sphere of dimension $\leqslant m - 2$) can be contracted to a point in $\mathbb{R}^m \setminus B_R$. Now we can proceed as above. Note only that Φ satisfies the $(\mathrm{PS})_c$-condition because any $(\mathrm{PS})_c$-sequence lies eventually in B_R. $\square$

We shall need the following extension of Theorem 1.2:

THEOREM 1.3. *Suppose* $\Phi \in C^1(\mathbb{R}^m, \mathbb{R})$ *satisfies the local linking conditions* (1.4) *and* (1.5) *for some* $F_0 \subset \mathbb{R}^m$. *If there exist subspaces* $\widetilde{F} \supset F \supset F_0$, $\widetilde{F} \neq F$, *and* $R > 0$ *such that* $\Phi < 0$ *on* $\widetilde{F} \setminus B_R$ *and* $\Phi|_F$ *has no critical point* $x \in \Phi^{-\alpha}$, *then* $\Phi'(\bar{x}) = 0$ *for some* $\bar{x}$ *with* $\alpha \leqslant \Phi(\bar{x}) \leqslant \max_{x \in \widetilde{F} \cap \overline{B}_{R+1}} \Phi(x)$.

PROOF. This time we obtain γ by solving the Cauchy problem

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = -\chi\bigl(\gamma(x, t)\bigr)V\bigl(\gamma(x, t)\bigr), \quad \gamma(x, 0) = x,$$

where $\chi \in C^\infty(\mathbb{R}^m, [0, 1])$ is such that $\chi = 1$ on B_R, $\chi = 0$ on $\mathbb{R}^m \setminus B_{R+1}$ and $V : F \cap \Phi^{-\alpha} \to F$ is a pseudogradient vector field for $\Phi|F$. Now we proceed as in the proof of case (i) above and obtain T such that $\gamma(x_0, T) \in F \cap (B_{R+1} \setminus B_R)$ whenever $x_0 \in F_0 \cap S_\rho$.

This set can be contracted to a point in $\widetilde{F} \cap (\overline{B}_{R+1} \setminus B_R)$, hence we obtain a map $f \in C(\partial D, \widetilde{F} \cap \overline{B}_{R+1})$ as in (1.7). Since there exists $g \in \Gamma \cap C(D, \widetilde{F} \cap \overline{B}_{R+1})$ where Γ is as in (1.8) it follows that

$$c \leqslant \max_{(x_0,s)\in D} \Phi\big(g(x_0,s)\big) \leqslant \max_{x\in\widetilde{F}\cap\overline{B}_{R+1}} \Phi(x). \qquad \square$$

An infinite-dimensional version of the linking theorems (in a setting which corresponds to Theorems 1.2 and 1.3) may be found in [62]. However, we shall only make use of the finite-dimensional versions stated above.

If the functional Φ is invariant with respect to a representation of some symmetry group, then Φ usually has multiple critical points. In order to exploit such symmetries, we introduce index theories.

Let E be a Hilbert space and

$$\Sigma := \{A = C \cap O \subset E\colon\ C \text{ is closed, } O \text{ is open and } -A = A\}. \tag{1.10}$$

Intersections of an open and a closed set (of a topological space) are called *locally closed*. Thus Σ consists of the locally closed symmetric subsets of E. Let $A \in \Sigma$, $A \neq \emptyset$. The *genus* of A, denoted $\gamma(A)$, is the smallest integer k such that there exists an odd mapping $f \in C(A, \mathbb{R}^k \setminus \{0\})$. If such a mapping does not exist for any k, then $\gamma(A) := +\infty$. Finally, $\gamma(\emptyset) = 0$. Equivalently, $\gamma(A) = 1$ if $A \neq \emptyset$ and if there exists an odd map $A \to \{+1, -1\}$; $\gamma(A) \leqslant k$ if A can be covered by k subsets $A_1, \ldots, A_k \in \Sigma$ such that $\gamma(A_j) \leqslant 1$.

PROPOSITION 1.4. *The two definitions of genus given above are equivalent for $A \in \Sigma$.*

PROOF. If $f : A \to \mathbb{R}^k \setminus \{0\}$ is as in the first definition, then the sets $A_j := \{x \in A\colon f_j(x) \neq 0\}$, $j = 1, \ldots, k$, cover A, are open in A and $-A_j = A_j$, hence $A_j \in \Sigma$. The map $f_j/|f_j| : A_j \to \{+1, -1\}$ shows that $\gamma(A_j) \leqslant 1$.

Suppose $\gamma(A) \leqslant k$ in the sense of the second definition. Since $A \cap A_j \in \Sigma$, we may assume $A_j \subset A$, $A = C \cap O$, $A_j = C_j \cap O_j$ and $C_j \subset C$, $O_j \subset O$, where C, C_j, O, O_j are as in the definition of Σ. If $f_j : A_j \to \{+1, -1\}$ is odd, we may extend it to a continuous map $f_j : O_j \to \mathbb{R}$. This is a consequence of Tietze's theorem because A_j is a closed subset of O_j. Replacing $f(x)$ by $\frac{1}{2}(f(x) - f(-x))$ we may assume that the extension is also odd. Let $\pi_j : A \to [0, 1]$, $j = 1, \ldots, k$, be a partition of unity subordinated to the covering $O_1, \ldots, O_k$ of A. Replacing $\pi_j(x)$ by $\frac{1}{2}(\pi_j(x) + \pi_j(-x))$ we may assume that all π_j are even. Now the map

$$f : A \to \mathbb{R}^k, \quad f(x) = \big(\pi_1(x) f_1(x), \ldots, \pi_k(x) f_k(x)\big)$$

is well defined, continuous, odd, and satisfies $f(A) \subset \mathbb{R}^k \setminus \{0\}$. $\square$

The above definitions of genus do not need to coincide for arbitrary subsets $A = -A$ which are not locally closed.

PROPOSITION 1.5. *Let* $A, B \in \Sigma$.

(i) *If there exists an odd mapping* $g \in C(A, B)$, *then* $\gamma(A) \leqslant \gamma(B)$.

(ii) $\gamma(A \cup B) \leqslant \gamma(A) + \gamma(B)$.

(iii) *There exists an open neighbourhood* $N \in \Sigma$ *of* A *such that* $\gamma(A) = \gamma(N)$.

(iv) *If* A *is compact and* $0 \notin A$, *then* $\gamma(A) < \infty$.

(v) *If* $U \in \Sigma$ *is an open bounded neighbourhood of* $0 \in \mathbb{R}^l$, *then* $\gamma(\partial U) = l$. *In particular,* $\gamma(S^{l-1}) = l$, *where* S^{l-1} *is the unit sphere in* $\mathbb{R}^l$.

(vi) *If* X *is a subspace of codimension* m *in* E *and* $\gamma(A) > m$, *then* $A \cap X \neq \emptyset$.

(vii) *If* $0 \notin A$ *and* $i(A) \geqslant 2$, *then* A *is an infinite set.*

A proof of this classical result may be found, e.g., in [80,85] if Σ contains only closed sets. This restriction is however not needed; see Proposition 1.7 below.

Let G be a compact topological group. A *representation* T of G in a Hilbert space E is a family $\{T_g\}_{g \in G}$ of bounded linear operators $T_g : E \to E$ such that $T_e = id$ (where e is the unit element of G and id the identity mapping), $T_{g_1 g_2} = T_{g_1} T_{g_2}$ and the mapping $(g, x) \mapsto T_g x$ is continuous. T is *an isometric* representation if each T_g is an isometry. A set $A \subset E$ is called T*-invariant* if $T_g A = A$ for all $g \in G$. When there is no risk of ambiguity we shall say A is G-invariant or simply invariant. The set

$$\mathcal{O}(x) := \{T_g x\colon\ g \in G\}$$

will be called the *orbit* of x and

$$E^G := \{x \in E\colon\ T_g x = x \text{ for all } g \in G\}$$

the set of *fixed points* of the representation T. Obviously, E^G is a closed subspace of E and $\mathcal{O}(x) = \{x\}$ if and only if $x \in E^G$.

Let

$$\Sigma := \{A \subset E\colon\ A \text{ is locally closed and } T_g A = A \text{ for all } g \in G\}. \tag{1.11}$$

Note that the definition (1.11) of Σ coincides with (1.10) if $G = \mathbb{Z}/2 \equiv \{1, -1\}$ and $T_{\pm 1} x = \pm x$. A mapping $f : E \to \mathbb{R}$ is said to be T*-invariant* (or simply invariant) if $f(T_g x) = x$ for all $g \in G$ and $x \in E$. If T and S are two (possibly different) representations of G in E and F, then a mapping $f : E \to F$ is *equivariant* with respect to T and S (or equivariant) if $f(T_g x) = S_g f(x)$ for all $g \in G$, $x \in E$. Finally, if $f : E \to F$, we set

$$f_G(x) := \int_G S_{g^{-1}} f(T_g x)\, \mathrm{d}g, \tag{1.12}$$

where the integration is performed with respect to the normalized Haar measure. It is easy to see that f_G is equivariant. As a special case, for $G = \mathbb{Z}/2$ acting via the antipodal map on E and F we have $f_G(x) = \frac{1}{2}(f(x) - f(-x))$, so f_G is odd. If G acts trivially on F (i.e., $S_{\pm 1} x = x$) we obtain $f_G(x) = \frac{1}{2}(f(x) + f(-x))$, so f_G is even.

If $\Phi \in C^1(E, \mathbb{R})$ is invariant with respect to an isometric representation T of G, then it is easy to see that $\Phi'(T_g x) = T_g \Phi'(x)$ for all $x \in E$, $g \in G$. Hence x is a critical point of Φ if and only if so are all $y \in \mathcal{O}(x)$. The set $\mathcal{O}(x)$ will be called a *critical orbit* of Φ.

In what follows we restrict our attention to isometric representations of $G = \mathbb{Z}/p$, where $p \geqslant 2$ is a prime, and $G = S^1 = \mathbb{R}/2\pi\mathbb{Z}$. If $G = S^1$, we do not distinguish between $\theta \in \mathbb{R}$ and the corresponding element of G, and we may also identify this element with $\mathrm{e}^{\mathrm{i}\theta}$. The same applies for $G = \mathbb{Z}/p \subset S^1$, where we identify the elements of G with roots of unity, again represented as $\mathrm{e}^{\mathrm{i}\theta}$.

Next we define an index $i : \Sigma \to \mathbb{N}_0 \cup \{\infty\}$ for $G = S^1$ and $G = \mathbb{Z}/p$, p a prime number. In the case $p = 2$ we recover the genus. For $A \in \Sigma$, $A \neq \emptyset$, we define $i(A) = 1$ if there exists a continuous map $f : A \to G \subset \mathbb{C} \setminus \{0\}$ such that $f(T_\theta x) = \mathrm{e}^{\mathrm{i}n\theta} f(x)$ for some $n \in \mathbb{N}$ and all x, θ ($n/p \notin \mathbb{N}$ if $G = \mathbb{Z}/p$). And $i(A) \leqslant k$ if A can be covered by k sets $A_1, \dots, A_k \in \Sigma$ such that $i(A_j) \leqslant 1$. If such a covering does not exist for any k, then $i(A) := +\infty$. Finally, we set $i(\emptyset) := 0$. We have a version of Proposition 1.4 for $G = S^1$.

PROPOSITION 1.6. *If $G = S^1$, then $i(A)$ is the smallest integer k for which there exists a mapping $f \in C(A, \mathbb{C}^k \setminus \{0\})$ such that $f(T_\theta x) = \mathrm{e}^{\mathrm{i}n\theta} f(x)$ for some $n \in \mathbb{N}$ and all x, θ.*

The proof is similar to that of Proposition 1.4. Note only that (1.12) needs to be used and if $f_j(T_\theta x) = \mathrm{e}^{\mathrm{i}n_j\theta} f_j(x)$, then $f(x) = (f_1(x)^{n/n_1}, \dots, f_k(x)^{n/n_k})$ where n is the least common multiple of $n_1, \dots, n_k$. The corresponding version for $G = \mathbb{Z}/p$, p an odd prime, requires spaces lying between $\mathbb{C}^k \setminus \{0\}$ and $\mathbb{C}^{k+1} \setminus \{0\}$; see [10, Proposition 2.9].

The above definition is due to Benci [17,18] in the case $G = S^1$ and to Krasnosel'skii [56] for $G = \mathbb{Z}/p$. Benci used in fact mappings $f \in C(A, \mathbb{C}^k \setminus \{0\})$ as in Proposition 1.6. Let us also remark that a different, cohomological index, has been introduced by Fadell and Rabinowitz [36] for $G = \mathbb{Z}/2$ and $G = S^1$, and by Bartsch [10, Example 4.5] for $G = \mathbb{Z}/p$. While the geometrical indexes of Krasnosel'skii and Benci are much more elementary, the cohomological indexes have some additional properties (which will not be needed here).

Since we only consider isometric representations, it is easy to see that the orthogonal complement $\widetilde{E} := (E^G)^\perp$ is invariant. In order to formulate the properties of the index for $G = S^1$ and $G = \mathbb{Z}/p$ we set

$$d_G := 1 + \dim G = \begin{cases} 1 & \text{for } G = \mathbb{Z}/p, \\ 2 & \text{for } G = S^1. \end{cases}$$

PROPOSITION 1.7. *Suppose $G = S^1$ or $G = \mathbb{Z}/p$, where p is a prime, and let $A, B \in \Sigma$.*

(i) *If there exists an equivariant mapping $g \in C(A, B)$, then $i(A) \leqslant i(B)$.*
(ii) *$i(A \cup B) \leqslant i(A) + i(B)$.*
(iii) *There exists an open neighbourhood $N \in \Sigma$ of A such that $i(A) = i(N)$.*
(iv) *If A is compact and $A \cap E^G = \emptyset$, then $i(A) < \infty$.*
(v) *If U is an open bounded invariant neighbourhood of 0 in a finite-dimensional invariant subspace X of $\widetilde{E}$, then $i(\partial U) = \frac{1}{d_G} \dim X$.*
(vi) *If X is an invariant subspace of $\widetilde{E}$ with finite codimension and if $i(A) > \frac{1}{d_G} \operatorname{codim}_{\widetilde{E}} X$, then $A \cap (E^G \oplus X) \neq \emptyset$.*

(vii) *If* $A \cap E^G \neq \emptyset$, *then* $i(A) = +\infty$. *If* $A \cap E^G = \emptyset$ *and* $i(A) \geqslant 2$, *then* A *contains infinitely many orbits.*

PROOF. (i) Let $i(B) = k < \infty$ (otherwise there is nothing to prove) and $B_1, \dots, B_k$ be a covering of B as in the definition of the index $i(B)$. Then $g^{-1}(B_1), \dots, g^{-1}(B_k)$ is a covering of A as in the definition of $i(A)$, hence $i(A) \leqslant k$.

(ii)–(iv) are obvious.

(v) It follows easily from Proposition 1.4 or 1.6 that $i(\partial U) \leqslant \frac{1}{d_G} \dim X$ if $G = \mathbb{Z}/2$ or $G = S^1$, respectively. In the $\mathbb{Z}/p$-case for $p \geqslant 3$ we may identify X with $\mathbb{C}^l$ and take the covering $A_j := \{z \in \mathbb{C}^l\colon\ z_j \neq 0,\ p \arg(z_j) \neq 0 \bmod 2\pi\}$, $B_j := \{z \in \mathbb{C}^l\colon\ z_j \neq 0,\ p \arg(z_j) \neq \pi \bmod 2\pi\}$, $j = 1, \dots, k$, of $\mathbb{C}^l \setminus \{0\}$ in order to see that $i(\mathbb{C}^l \setminus \{0\}) \leqslant 2l = \dim X$.

The reverse inequality is a consequence of the Borsuk–Ulam theorem. A proof for $G = S^1$ may be found in [73, Theorem 5.4], and for $G = \mathbb{Z}/p$ in [9].

(vi) Let Y be the orthogonal complement of X in $\widetilde{E}$. Then Y is invariant and $\dim Y = \operatorname{codim}_{\widetilde{E}} X$. Suppose $A \cap (E^G \oplus X) = \emptyset$ and let $f(x) = P_Y x$ where P_Y denotes the orthogonal projector onto Y. Then $f : A \to Y \setminus \{0\}$. If $G = \mathbb{Z}/2$ or $G = S^1$ this implies $i(A) \leqslant \frac{1}{d_G} \dim Y$ by Propositions 1.4, 1.6, respectively. If $G = \mathbb{Z}/p$, $p \geqslant 3$, we identify Y with $\mathbb{C}^m$ and write $f = (f_1, \dots, f_m)$. It follows from the Peter–Weyl theorem (see [73, Theorem 5.1], where the case $G = S^1$ is considered) that $f_j(T_\theta x) = \mathrm{e}^{\mathrm{i} n_j \theta} f(x)$ ($n_j \not\equiv 0 \bmod p$). Let $g_j(x) := f_j(x)^{n/n_j}$, where n is the least common multiple of $n_1, \dots, n_k$. Then $g : A \to \mathbb{C}^m \setminus \{0\}$ and $g(T_\theta x) = \mathrm{e}^{\mathrm{i} n\theta} g(x)$, so $i(A) \leqslant i(\mathbb{C}^m \setminus \{0\}) \leqslant 2m = \operatorname{codim}_{\widetilde{E}} X$, a contradiction.

(vii) Suppose $A \cap E^G \neq \emptyset$ and there exists a covering $A_1, \dots, A_k$ of A as in the definition. Then $A_j \cap E^G \neq \emptyset$ for some j. For each $x \in A_j \cap E^G$ we have $f(T_\theta x) = f(x)$. So if $f_j(T_\theta x) = \mathrm{e}^{\mathrm{i} n\theta} f_j(x)$, with n as before, then $f_j(x) = 0$. Thus there is no mapping $f_j : A_j \to G \subset \mathbb{C} \setminus \{0\}$ as required in the definition of index, hence $i(A) = +\infty$. If $A \cap E^G = \emptyset$ and A consists of k orbits $\mathcal{O}(x_1), \dots, \mathcal{O}(x_k)$, then we let $n_j \geqslant 1$ be the largest integer such that $2\pi/n_j \in G$ and $T_{2\pi/n_j} x_j = x_j$. If $G = \mathbb{Z}/p$ then all $n_j = 1$. We define $f : A \to G$ by setting $f(T_\theta x_j) = \mathrm{e}^{\mathrm{i} n\theta}$, where n is the least common multiple of $n_1, \dots, n_k$. □

It is easy to prove an equivariant version of the deformation Lemma 1.1 for invariant functionals $\Phi : E \to \mathbb{R}$. One simply observes that if V is a pseudo-gradient vector field for Φ then

$$V_G(x) := \int_G T_g^{-1} V(T_g x)\,\mathrm{d}g \tag{1.13}$$

is an equivariant pseudo-gradient vector field for Φ. Integrating V_G as in the proof of Lemma 1.1 yields an equivariant deformation η.

THEOREM 1.8. *Suppose* $f \in C^1(S^{n-1}, \mathbb{R})$ *is invariant with respect to a representation of* $\mathbb{Z}/p$ *in* $\mathbb{R}^n$ *without nontrivial fixed points. Then* f *has at least* n $\mathbb{Z}/p$*-orbits of critical points.*

THEOREM 1.9. *Suppose* $f \in C^1(S^{2n-1}, \mathbb{R})$ *is invariant with respect to a representation of* S^1 *in* $\mathbb{R}^{2n}$ *without nontrivial fixed points. Then* f *has at least* n S^1*-orbits of critical points.*

PROOF (outline). The proofs of Theorems 1.8 and 1.9 are standard. Set $M = S^{d_G \cdot n-1}$ and suppose f has finitely many critical orbits $\mathcal{O}_j = \mathcal{O}(x_j)$, $j = 1, \ldots, k$. We may assume that the critical values $c_j = f(x_j)$ are ordered: $c_1 \leqslant \cdots \leqslant c_k$. Then using the properties of the index and an equivariant deformation lemma for functionals defined on manifolds one sees that $i(f^{c_j}) \leqslant j$. The result follows from $i(M) = n$.

We present a simpler proof which works if f' is locally Lipschitz continuous. Let η be the negative gradient flow of f on M and consider the sets

$$A_j := \bigl\{x \in M\colon\ \eta(x,t) \to \mathcal{O}_j \text{ as } t \to \infty\bigr\}$$

and

$$B_j := \bigcup_{i=1}^{j} A_i \quad \text{for } j = 0, \ldots, k.$$

Then $B_0 = \emptyset$, $B_k = M$, and it is not difficult to see that B_{j-1} is an open subset of B_j. Consequently all A_j are locally closed. Using the flow η one constructs an equivariant map $f_j : A_j \to \mathcal{O}_j$. This implies $i(A_j) \leqslant i(\mathcal{O}_j) = 1$, and therefore $n = i(M) = i(B_k) \leqslant k$. This proof does not need the equivariant deformation lemma, and it produces directly a covering of M as in the definition of the index. □

REMARK 1.10. One can define index theories satisfying properties 1.7(i)–(iv), (vii) for arbitrary compact Lie groups G. However, properties 1.7(v), (vi) which are important for applications and computations, cannot be extended in general, except for a very restricted class of groups. This has been investigated in detail in [10]. In certain applications the representation of G in E is of a special form which allows to obtain similar results as above. In order to formulate this, call a finite-dimensional representation space $V \cong \mathbb{R}^n$ of the compact Lie group G admissible if every equivariant map $\overline{\mathcal{O}} \to V^{k-1}$, $\mathcal{O} \subset V^k$ a bounded open and invariant neighbourhood of 0 in V^k, has a zero on $\partial\mathcal{O}$. Clearly the antipodal action of $\mathbb{Z}/2$ on $\mathbb{R}$ is admissible as are the nontrivial representations of $\mathbb{Z}/p$ or S^1 in $\mathbb{R}^2$. Let $E = \bigoplus_{j=1}^{\infty} E_j$ be the Hilbert space sum of the finite-dimensional Hilbert spaces E_j such that each E_j is isomorphic to V as a representation space of G. For instance, $E = L^2(S^1, V)$ with the representation of G given by $(T_g x)(t) = T_g(x(t))$ has this property. The same is true for subspaces like $H^1(S^1, V)$ or $H^{1/2}(S^1, V)$. For an invariant, locally closed set $A \subset E$ let $i(A) = 1$ if $A \neq \emptyset$ and there exists a continuous equivariant map $A \to SV = \{v \in V\colon\ \|v\| = 1\}$. And let $i(A) \leqslant k$ if A can be covered by $A_1, \ldots, A_k \in \Sigma$ with $i(A) \leqslant 1$. Proposition 1.7 can be extended to this index theory. See [6,16] for applications to Hamiltonian systems.

1.2. *Critical point theory for strongly indefinite functionals*

As will be explained in Section 2.1, functionals naturally corresponding to Hamiltonian systems are strongly indefinite. This means that they are of the form $\Phi(z) = \frac{1}{2}\langle Lz, z\rangle -$

$\psi(z)$ where $L: E \to E$ is a selfadjoint Fredholm operator with negative and positive eigenspace both infinite-dimensional, and the same is true for the Hessian $\Phi''(z)$ of a critical point z of Φ. In order to study such functionals it will be convenient to use a variant of the Palais–Smale condition that allows a reduction to the finite-dimensional case and leads to simpler proofs. We shall also present two useful critical point theorems which apply when the Palais–Smale condition does not hold. These will be needed for the existence of homoclinic solutions.

First we introduce certain sequences of finite-dimensional subspaces and replace the Palais–Smale condition by another one which is adapted to these sequences.

Let $\{E_n\}_{n\geqslant 1}$ be a sequence of finite-dimensional subspaces such that $E_n \subset E_{n+1}$ for all n and

$$E = \overline{\bigcup_{n=1}^{\infty} E_n}.$$

Let $P_n : E \to E_n$ denote the orthogonal projection. Then $\{x_j\}$ is called a (PS)*-*sequence* for Φ (with respect to $\{E_n\}$) if $\Phi(x_j)$ is bounded, each $x_j \in E_{n_j}$ for some n_j, $n_j \to \infty$ and $P_{n_j}\Phi'(x_j) \to 0$ as $j \to \infty$. Φ is said to satisfy the (PS)*-*condition* if each (PS)*-sequence has a convergent subsequence. It is easy to see that K_c is compact for each c if (PS)* holds. Indeed, let $x_j \in K_c$, then we can find $n_j \geqslant j$ such that $\|y_j - x_j\| \leqslant 1/j$, $\Phi(y_j) \to c$ and $P_{n_j}\Phi'(y_j) \to 0$, where $y_j := P_{n_j}x_j$. Hence $\{y_j\}$, and therefore also $\{x_j\}$, has a convergent subsequence. We shall repeatedly use the notation

$$\Phi_n := \Phi|_{E_n} \quad \text{and} \quad A_n := A \cap E_n.$$

Observe that $\Phi_n'(x) = P_n\Phi'(x)$ for all $x \in E_n$.

The condition (PS)* (in a slightly different form) has been introduced independently by Bahri and Berestycki [7,8], and Li and Liu [60].

LEMMA 1.11. *If Φ satisfies* (PS)* *and N is a neighbourhood of K_c, then there exist $\bar\varepsilon, \delta > 0$ and $n_0 \geqslant 1$ such that $\|\Phi_n'(x)\| \geqslant \delta$ whenever $d(x, E \setminus N) \leqslant \delta$, $|\Phi(x) - c| \leqslant \bar\varepsilon$ and $n \geqslant n_0$.*

PROOF. If the conclusion is false, then we find a sequence $\{x_j\}$ such that $x_j \in E_{n_j}$ for some $n_j \geqslant j$, $d(x_j, E \setminus N) \to 0$, $\Phi_{n_j}(x_j) \to c$ and $\Phi_{n_j}'(x_j) \to 0$. Hence $\{x_j\}$ is a (PS)*-sequence. Passing to a subsequence, $x_j \to \bar x \in K_c$. However, since K_c is compact, the sequence $\{x_j\}$ is bounded away from K_c and therefore $\bar x \notin K_c$, a contradiction. □

Next we introduce the notion of limit index in order to deal with symmetric functionals. As in Section 1.1 we consider the groups $G = \mathbb{Z}/p$, where p is a prime, or $G = S^1$ and their isometric representations in E. The group $\mathbb{Z}/2$ always acts via the antipodal map (i.e., $T_{\pm 1}x = \pm x$) so that obviously $E^G = \{0\}$. The reason for going beyond the usual index is that we need to distinguish between certain infinite-dimensional sets having $i(A) = \infty$; in particular, we need to compare different spheres of infinite dimension and codimension.

Let $\{E_n\}$ be a sequence of subspaces as above and suppose in addition that each E_n is G-invariant and $E^G \subset E_n$ for some n. Let $\{d_n\}$ be a sequence of integers and

$$\mathcal{E} := \{E_n, d_n\}_{n=1}^{\infty}.$$

The limit index of $A \in \Sigma$ with respect to $\mathcal{E}$, $i_{\mathcal{E}}(A)$, is defined by

$$i_{\mathcal{E}}(A) := \limsup_{n\to\infty} \bigl(i(A_n) - d_n\bigr).$$

Clearly $i_{\mathcal{E}}(A) = \infty$ if $A \cap E^G \neq \emptyset$. The limit index, in a somewhat different form, has been introduced by Y.Q. Li [63], see also [92]. A special case is the limit genus, $\gamma_{\mathcal{E}}(A)$. We note that $i_{\mathcal{E}}(A)$ can take the values $+\infty$ or $-\infty$ and if $E_n = E$, $d_n = 0$ for all n, then $i_{\mathcal{E}}(A) = i(A)$, and similarly for the genus.

REMARK 1.12. The limit index is patterned on the notion of limit relative category introduced by Fournier et al. in [42]. Recall that if Y is a closed subset of X, then a closed set $A \subset X$ is said to be of category k in X relative to Y, denoted $\operatorname{cat}_{X,Y}(A) = k$, if k is the least integer such that there exist closed sets $A_0, \dots, A_k \subset X$, $A_0 \supset Y$, which cover A, all A_j, $1 \leqslant j \leqslant k$, are contractible in X and there exists a deformation $h : A_0 \times [0,1] \to X$ with $h(A_0, 1) \subset Y$ and $h(Y, t) \subset Y$ for all $t \in [0,1]$. If $Y = \emptyset$ (and $A_0 = \emptyset$), then $\operatorname{cat}_X(A) = \operatorname{cat}_{X,\emptyset}(A)$ is the usual Lusternik–Schnirelman category of A in X. For $X \subset E$ and using the above notation for subsets of E, the limit relative category $\operatorname{cat}^{\infty}_{X,Y}(A)$ is by definition equal to $\limsup_{n\to\infty} \operatorname{cat}_{X_n,Y_n}(A_n)$. Note that unlike for the limit index, the limit category is necessarily a nonnegative integer. Note also that if D is the unit closed ball and S its boundary in an infinite-dimensional Hilbert space, then $\operatorname{cat}_{D,S}(D) = \operatorname{cat}_S(S) = 0$ while $\operatorname{cat}^{\infty}_{D,S}(D) = \operatorname{cat}^{\infty}_S(S) = 1$.

Below we formulate some properties of $i_{\mathcal{E}}$ which automatically hold for $\gamma_{\mathcal{E}}$. As before $\widetilde{E}$ is the orthogonal complement of E^G. It follows from the invariance of E_n that the dimension of $\widetilde{E}_n = E_n \cap \widetilde{E}$ is even except when $G = \mathbb{Z}/2$. Recall the notation $d_G = 1 + \dim G$.

PROPOSITION 1.13. *Let $A, B \in \Sigma$.*

(i) *If for almost all n there exists an equivariant mapping $g_n \in C(A_n, B_n)$, then $i_{\mathcal{E}}(A) \leqslant i_{\mathcal{E}}(B)$.*

(ii) $i_{\mathcal{E}}(A \cup B) \leqslant i_{\mathcal{E}}(A) + i(B)$ *if* $i_{\mathcal{E}}(A) \neq -\infty$.

(iii) *Let $l \in \mathbb{Z}$, $R > 0$. If Y is an invariant subspace of $\widetilde{E}$ such that* $\dim Y_n = (d_n + l)d_G$ *for almost all n, then $i_{\mathcal{E}}(Y \cap S_R) = l$.*

(iv) *Let $m \in \mathbb{Z}$. If X is an invariant subspace of $\widetilde{E}$ such that* $\operatorname{codim}_{\widetilde{E}_n} X_n = (d_n + m)d_G$ *for almost all n and if $i_{\mathcal{E}}(A) > m$, then $A \cap (E^G \oplus X) \neq \emptyset$.*

PROOF. (i) It follows from (i) of Proposition 1.7 that $i(A_n) - d_n \leqslant i(B_n) - d_n$. So passing to the limit as $n \to \infty$ we obtain the conclusion.

(ii) $i(A_n \cup B_n) - d_n \leqslant i(A_n) - d_n + i(B_n) \leqslant (i(A_n) - d_n) + i(B)$. Now we can pass to the limit again.

(iii) This follows from (v) of Proposition 1.7.

(iv) There exists a number n such that $E^G \subset E_n$, $\operatorname{codim}_{\widetilde{E}_n} X_n = (d_n + m)d_G$ and $i(A_n) > d_n + m$. So by (vi) of Proposition 1.7, $\emptyset \neq A_n \cap (E^G \oplus X_n) \subset A \cap (E^G \oplus X)$. □

Recall that if x is a critical point of an invariant functional Φ, then so are all $y \in \mathcal{O}(x)$. We have the following results concerning the existence of critical orbits.

THEOREM 1.14. *Suppose that* $\Phi \in C^1(E, \mathbb{R})$ *is* G*-invariant, satisfies* (PS)* *and* $\Phi(0) = 0$. *Moreover, suppose there exist numbers* $\rho > 0$, $\alpha < \beta < 0$, *integers* $m < l$, *and invariant subspaces* $X, Y \subset \widetilde{E}$ *such that*:

(i) $E^G \subset E_n$ *for almost all* n;

(ii) $\operatorname{codim}_{\widetilde{E}_n} X_n = (d_n + m)d_G$ *and* $\dim Y_n = (d_n + l)d_G$ *for almost all* n;

(iii) $\Phi|_{Y \cap S_\rho} \leqslant \beta$;

(iv) $\Phi|_{E^G \oplus X} \geqslant \alpha$ *and* $\Phi|_{E^G} \geqslant 0$.

Then Φ *has at least* $l - m$ *distinct critical orbits* $\mathcal{O}(x_j)$ *such that* $\mathcal{O}(x_j) \cap E^G = \emptyset$. *The corresponding critical values can be characterized as*

$$c_j = \inf_{i_{\mathcal{E}}(A) \geqslant j} \sup_{x \in A} \Phi(x), \qquad m + 1 \leqslant j \leqslant l,$$

and are contained in the interval $[\alpha, \beta]$.

PROOF. It is clear that $\{A \in \Sigma\colon\ i(A) \geqslant j + 1\} \subset \{A \in \Sigma\colon\ i(A) \geqslant j\}$, hence $c_{m+1} \leqslant c_{m+2} \leqslant \cdots \leqslant c_l$. According to (iii) of Proposition 1.13, $i_{\mathcal{E}}(Y \cap S_\rho) \geqslant l$, hence by (iii), $c_l \leqslant \beta$. Suppose $i_{\mathcal{E}}(A) \geqslant m + 1$. Then $A \cap (E^G \oplus X) \neq \emptyset$ by (iv) of Proposition 1.13 and it follows from (iv) that $c_{m+1} \geqslant \alpha$. Moreover, (iv) implies $K_{c_j} \cap E^G = \emptyset$.

Suppose $c := c_j = \cdots = c_{j+p}$ for some $p \geqslant 0$. The proof will be complete if we can show that $i(K_c) \geqslant p + 1$ (because either all c_j are distinct and $K_{c_j} \neq \emptyset$, or $i(K_{c_j}) \geqslant 2$ for some j and K_{c_j} contains infinitely many orbits according to (vii) of Proposition 1.7). By (iii) of Proposition 1.7 there exists a neighbourhood $N \in \Sigma$ such that $i(N) = i(K_c)$, and for this N we may find $\bar{\varepsilon}, \delta > 0$ and $n_0 \geqslant 1$ such that the conclusion of Lemma 1.11 holds. It follows from Lemma 1.1 that we can find an $\varepsilon > 0$ such that for each $n \geqslant n_0$ there exists a deformation $\eta_n : E_n \times [0, 1] \to E_n$ with $\eta_n(\Phi_n^{c+\varepsilon}, 1) \subset \Phi_n^{c-\varepsilon} \cup N_n$. Moreover, using (1.13) we may assume that $\eta_n(., t)$ is equivariant for each t. So by (i) and (ii) of Proposition 1.13 and the definition of c,

$$j + p \leqslant i_{\mathcal{E}}\bigl(\Phi^{c+\varepsilon}\bigr) \leqslant i_{\mathcal{E}}\bigl(\Phi^{c-\varepsilon} \cup N\bigr) \leqslant i_{\mathcal{E}}\bigl(\Phi^{c-\varepsilon}\bigr) + i(N) < j + i(N). \tag{1.14}$$

Hence $i(K_c) = i(N) > p$. □

Applying Theorem 1.14 to $-\Phi$ we immediately obtain the following result which will be more convenient in our applications:

COROLLARY 1.15. *Suppose that* $\Phi \in C^1(E, \mathbb{R})$ *is* G*-invariant, satisfies* (PS)* *and* $\Phi(0) = 0$. *Moreover, suppose there exist numbers* $\rho > 0$, $0 < \alpha < \beta$, *integers* $m < l$, *and invariant subspaces* $X, Y \subset \widetilde{E}$ *such that*:

(i) $E^G \subset E_n$ *for almost all* n;
(ii) $\operatorname{codim}_{\widetilde{E}_n} X_n = (d_n + m)d_G$ *and* $\dim Y_n = (d_n + l)d_G$ *for almost all* n;
(iii) $\Phi|_{Y \cap S_\rho} \geqslant \alpha$;
(iv) $\Phi|_{E^G \oplus X} \leqslant \beta$ *and* $\Phi|_{E^G} \leqslant 0$.
Then Φ *has at least* $l - m$ *distinct critical orbits* $\mathcal{O}(x_j)$ *such that* $\mathcal{O}(x_j) \cap E^G = \emptyset$. *The corresponding critical values can be characterized as*

$$c_j = \sup_{i_{\mathcal{E}}(A) \geqslant j} \inf_{x \in A} \Phi(x), \quad m + 1 \leqslant j \leqslant l,$$

and are contained in the interval $[\alpha, \beta]$.

COROLLARY 1.16. *If the hypotheses of Theorem* 1.14 *or Corollary* 1.15 *are satisfied with* $l \in \mathbb{Z}$ *fixed and* $m \in \mathbb{Z}$ *arbitrarily small, then* Φ *has infinitely many geometrically distinct critical orbits* $\mathcal{O}(x_j)$ *such that* $\mathcal{O}(x_j) \cap E^G = \emptyset$. *Moreover,* $c_j \to -\infty$ *in Theorem* 1.14 *and* $c_j \to \infty$ *in Corollary* 1.15 *as* $j \to -\infty$.

PROOF. It suffices to consider the case of Theorem 1.14. The value c_j is defined for all $j \leqslant l$, $j \in \mathbb{Z}$ and since the sequence $\{c_j\}$ is nondecreasing, either $c_j \to -\infty$ and we are done, or $c_j \to c \in \mathbb{R}$ as $j \to -\infty$. In the second case K_c is nonempty and compact according to (PS)*. Let $N \in \Sigma$ be a neighbourhood of K_c such that $i(N) = i(K_c) < \infty$ and let $\varepsilon > 0$ be as in Lemma 1.1. Since $c + \varepsilon \geqslant c_{j_0}$ for some j_0 and $c - \varepsilon < c_j$ for all $j \leqslant l$, we have (cf. (1.14))

$$j_0 \leqslant i_{\mathcal{E}}\big(\Phi^{c+\varepsilon}\big) \leqslant i_{\mathcal{E}}\big(\Phi^{c-\varepsilon}\big) + i(N) = -\infty,$$

a contraction. □

REMARK 1.17. Proposition 1.13, Theorem 1.14 and Corollaries 1.15 and 1.16 are valid if $G = \mathbb{Z}/2$ and $T_{\pm 1}x = \pm x$ (i.e., Φ is even). For this G, $i(A)$ is just the genus $\gamma(A)$. If $G = \mathbb{Z}/p$ and $p \geqslant 3$, then $l - m$ is necessarily an even integer.

We now state a critical point theorem which needs tools from algebraic topology.

THEOREM 1.18. *Let* M *be a compact differentiable manifold and* $\Phi : E \times M \to \mathbb{R}$ *a* C^1-*functional defined on the product of the Hilbert space* E *and* M. *Suppose* Φ *satisfies* (PS)*, *there exist numbers* $\rho > 0$, $\alpha < \beta \leqslant \gamma$ *and subspaces* W, Y, *where* $E = W \oplus Y$, $W_n \subset W$, $Y_n \subset Y$, $\dim W_n \geqslant 1$, *such that*:
(i) $\Phi|_{(W \cap S_\rho) \times M} \leqslant \alpha$;
(ii) $\Phi|_{Y \times M} \geqslant \beta$;
(iii) $\Phi|_{(W \cap \overline{B}_\rho) \times M} \leqslant \gamma$.
Then Φ *possesses at least* $\operatorname{cupl}(M) + 1$ *critical points.*

Here $\operatorname{cupl}(M)$ denotes the cuplength of M with respect to singular cohomology theory with coefficients in an arbitrary field. For a proof we refer to Fournier et al. [42]. The argument there uses the limit relative category (see Remark 1.12) and is in the spirit of Theorem 1.14. In particular, the numbers c_j are defined by minimaxing over sets $A \supset D$ with $\operatorname{cat}^{\infty}_{C,D}(A) \geqslant j$, where $(C, D) := (E \times M, (W \cap S_\rho) \times M)$. An important role is played by the inequality $\operatorname{cat}^{\infty}_{C,D}((W \cap \overline{B}_\rho) \times M) \geqslant \operatorname{cupl}(M) + 1$. A related result can be found in [90].

For applications to homoclinic solutions one has to deal with functionals where neither the (PS)- nor the (PS)*-condition holds. We present two abstract critical point theorems which are helpful in this case. The proofs involve again a reduction to a finite-dimensional situation.

THEOREM 1.19. *Let E be a separable Hilbert space with the orthogonal decomposition $E = E^+ \oplus E^-$, $z = z^+ + z^-$, and suppose $\Phi \in C^1(E, \mathbb{R})$ satisfies the hypotheses*:

(i) $\Phi(z) = \frac{1}{2}(\|z^+\|^2 - \|z^-\|^2) - \psi(z)$ *where $\psi \in C^1(E, \mathbb{R})$ is bounded below, weakly sequentially lower semicontinuous with $\psi' : E \to E$ weakly sequentially continuous;*

(ii) $\Phi(0) = 0$ *and there are constants $\kappa, \rho > 0$ such that $\Phi(z) > \kappa$ for every $z \in S_\rho \cap E^+$;*

(iii) *there exists $e \in E^+$ with $\|e\| = 1$, and $R > \rho$ such that $\Phi(z) \leqslant 0$ for $z \in \partial M$ where $M = \{z = z^- + \zeta e \colon z^- \in X^-,\ \|z\| \leqslant R,\ \zeta \geqslant 0\}$.*

Then there exists a sequence $\{z_j\}$ in E such that $\Phi'(z_j) \to 0$ and $\Phi(z_j) \to c$ for some $c \in [\kappa, m]$, where $m := \sup \Phi(\overline{M})$.

The theorem is due to Kryszewski and Szulkin [58]. Some compactness is hidden in condition (i) where the weak topology is used. In the applications the concentration-compactness method, see [65], can sometimes be used in order to obtain an actual critical point. Of course, if the Palais–Smale condition holds then there exists a critical point at the level c.

PROOF (outline). Let $P^\pm : E \to E^\pm$ be the orthogonal projections. We choose a Hilbert basis $\{e_k\}_{k\in\mathbb{N}}$ of E^- and define the norm

$$|||u||| := \max\left\{ \|P^+u\|, \sum_{k=1}^{\infty} \frac{|\langle u, e_k\rangle|}{2^k} \right\}.$$

The topology induced on E by this norm will be denoted by τ. On subsets $\{u \in E\colon \|P^-u\| \leqslant R\}$ this topology coincides with the weak × strong product topology $(E^-, w) \times (E^+, \|\cdot\|)$ on E. In particular, for a $\|\cdot\|$-bounded sequence $\{u_j\}$ in E^- we have $u_j \rightharpoonup u$ if and only if $u_j \to u$ with respect to $|||\cdot|||$. Given a finite-dimensional subspace $F \subset E^+$, $\|\cdot\|$-bounded subsets of $E^- \oplus F$ are $|||\cdot|||$-precompact.

We prove the theorem arguing indirectly. Suppose there exists $\alpha > 0$ with $\|\Phi'(u)\| \geqslant \alpha$ for all $u \in \Phi^m_\kappa := \{u \in E\colon \kappa \leqslant \Phi(u) \leqslant m\}$. Then we construct a deformation $h : I \times \Phi^m \to \Phi^m$, $I = [0, 1]$, with the properties:

(h_1) $h : I \times \Phi^m \to \Phi^m$ is continuous with respect to the $\|\cdot\|$-topology on Φ^m, and with respect to the τ-topology;

(h_2) $h(0,u)=u$ for all $u \in \Phi^m$;

(h_3) $\Phi(h(t,u)) \leqslant \Phi(u)$ for all $t \in I$, $u \in \Phi_\kappa^m$;

(h_4) each $(t,u) \in I \times \Phi^m$ has a τ-open neighbourhood W such that the set $\{v - h(s,v)\colon (s,v) \in W\}$ is contained in a finite-dimensional subspace of E;

(h_5) $h(1,\Phi^m) \subset \Phi^\kappa$.

This leads to a contradiction as follows. Since M is τ-compact, by (h_1) and (h_4) there exists a finite-dimensional subspace $F \subset E$ containing the set $\{v-h(s,v)\colon (s,v) \in I \times M\}$, hence $h(I \times (M \cap F)) \subset F$. Since

$$h(I \times \partial M) \subset F \cap \Phi^\kappa \subset F \setminus \big(S_\rho \cap E^+\big)$$

a standard argument using the Brouwer degree yields $h(1, M \cap F) \cap S_\rho \cap E^+ \neq \emptyset$. (The sets $F \cap \partial M$ and $F \cap S_\rho \cap E^+$ link in F.) Now condition (ii) of the theorem implies $h(1, M \cap F) \not\subset \Phi^\kappa$, contradicting ($h_5$).

It remains to construct a deformation h as above. For each $u \in \Phi_\kappa^m$ we choose a pseudo-gradient vector $w(u) \in E$, that is $\|w(u)\| \leqslant 2$ and $\langle \Phi'(u), w(u)\rangle > \|\Phi'(u)\|$ (this definition differs somewhat from (1.2)). By condition (i) of the theorem there exists a τ-open neighbourhood $N(u)$ of u in E such that $\langle \Phi'(v), w(u)\rangle > \|\Phi'(u)\|$ for all $v \in N(u) \cap \Phi_\kappa^m$. If $\Phi(u) < \kappa$ we set $N(u) = \{v \in E\colon \Phi(v) < \kappa\}$. As a consequence of condition (i) of the theorem this set is τ-open. Let $(\pi_j)_{j \in J}$ be a τ-Lipschitz continuous partition of unity of Φ^m subordinated to the covering $N(u)$, $u \in \Phi^m$. This exists because the τ-topology is metric. Clearly, the $\pi_j : E \to [0,1]$ are also $\|\cdot\|$-Lipschitz continuous. For each $j \in J$ there exists $u_j \in \Phi^m$ with $\operatorname{supp} \pi_j \subset N(u_j)$. We set $w_j = w(u_j)$ and define the vector field

$$f : \Phi^m \to E, \quad f(u) := \frac{m-\kappa}{\alpha} \sum_{j \in J} \pi_j(u) w_j .$$

This vector field is locally Lipschitz continuous and τ-locally Lipschitz continuous. It is also τ-locally finite-dimensional. Thus we may integrate it and obtain a flow $\eta : [0,\infty) \times \Phi^m \to \Phi^m$. It is easy to see that the restriction of η to $[0,1] \times \Phi^m$ satisfies the properties (h_1)–(h_5). □

REMARK 1.20. A sequence $\{z_j\}$ is called a *Cerami sequence* if $\Phi(z_j)$ is bounded and $(1 + \|z_j\|)\Phi'(z_j) \to 0$. This definition has been introduced by Cerami in [22]. Note in particular that if $\{z_j\}$ is as above, then $\langle \Phi'(z_j), z_j\rangle \to 0$ which does not need to be the case for an (a priori unbounded) (PS)-sequence. It has been shown in [59] that under the hypotheses of Theorem 1.19 a stronger conclusion holds: there exists a Cerami sequence $\{z_j\}$ such that $\Phi(z_j) \to c \in [\kappa, m]$.

The next result of this section deals with $\mathbb{Z}/p$-invariant functionals $\Phi \in C^1(E,\mathbb{R})$. As a substitute for the (PS)- or (PS)*-condition we introduce the concept of (PS)-attractor. Given an interval $I \subset \mathbb{R}$, we call a set $\mathcal{A} \subset E$ a (PS)$_I$-attractor if for any (PS)$_c$-sequence $\{z_j\}$ with $c \in I$, and any $\varepsilon, \delta > 0$ one has $z_j \in U_\varepsilon(\mathcal{A} \cap \Phi_{c-\delta}^{c+\delta})$ provided j is large enough. Here $U_\varepsilon(F)$ denotes the ε-neighbourhood of F in E.

THEOREM 1.21. *Let E be a separable Hilbert space with an isometric representation of the group $G = \mathbb{Z}/p$, where p is a prime, such that $E^{\mathbb{Z}/p} = \{0\}$. Let $E = E^+ \oplus E^-$, $z = z^+ + z^-$, be an orthogonal decomposition and $E^\pm$ be $\mathbb{Z}/p$-invariant. Let $\Phi \in C^1(E,\mathbb{R})$ be a $\mathbb{Z}/p$-invariant functional satisfying the following conditions*:

(i) $\Phi(z) = \frac{1}{2}(\|z^+\|^2 - \|z^-\|^2) - \psi(z)$ *where* $\psi \in C^1(E,\mathbb{R})$ *is bounded below, weakly sequentially lower semicontinuous with* $\psi' : E \to E$ *weakly sequentially continuous*;

(ii) $\Phi(0) = 0$ *and there exist* $\kappa, \rho > 0$ *such that* $\Phi(z) > \kappa$ *for every* $z \in S_\rho \cap E^+$;

(iii) *there exists a strictly increasing sequence of finite-dimensional* $\mathbb{Z}/p$*-invariant subspaces* $F_n \subset E^+$ *such that* $\sup \Phi(E_n) < \infty$ *where* $E_n := E^- \oplus F_n$, *and an increasing sequence of real numbers* $R_n > 0$ *with* $\sup \Phi(E_n \setminus B_{R_n}) < \inf \Phi(B_\rho)$;

(iv) *for any compact interval* $I \subset (0,\infty)$ *there exists a* $(\mathrm{PS})_I$*-attractor* $\mathcal{A}$ *such that* $\inf\{\|z^+ - w^+\|: z, w \in \mathcal{A},\ z^+ \neq w^+\} > 0$.

Then Φ *has an unbounded sequence* $\{c_j\}$ *of positive critical values.*

PROOF (outline). Let τ be the topology on E introduced in the proof of Theorem 1.19. For $c \in \mathbb{R}$ we consider the set $\mathcal{M}(c)$ of maps $g : \Phi^c \to E$ satisfying:

(P_1) g is τ-continuous and equivariant;

(P_2) $g(\Phi^a) \subset \Phi^a$ for all $a \geqslant \inf \Phi(B_\rho) - 1$ where ρ is from condition (ii);

(P_3) each $u \in \Phi^c$ has a τ-open neigbourhood $W \subset E$ such that the set $(\mathrm{id} - g)(W \cap \Phi^c)$ is contained in a finite-dimensional linear subspace of E.

Let i be the $\mathbb{Z}/p$-index from Section 1.1 and set

$$i_0(c) := \min_{g \in \mathcal{M}(c)} i\big(g(\Phi^c) \cap S_\rho \cap E^+\big) \in \mathbb{N}_0 \cup \{\infty\}.$$

Clearly i_0 is nondecreasing and $i_0(c) = 0$ for $c \leqslant \kappa$ where κ is from (ii). i_0 is a kind of *pseudoindex* in the sense of Benci's paper [18]. Now we define the values

$$c_k := \inf\big\{c > 0:\ i_0(c) \geqslant k\big\}.$$

One can show that $i_0(c)$ is finite for every $c \in \mathbb{R}$ and can only change at a critical level of Φ. In order to see the latter, given an interval $[c,d]$ without critical values one needs to construct maps $g \in \mathcal{M}(d)$ with $g(\Phi^d) \subset \Phi^c$. Such a map can be obtained as time-1-map of a deformation as in the proof of Theorem 1.19. Of course one has to make sure that the deformation is equivariant which is the case if the vector field is equivariant. This can be easily achieved, see (1.13). Given a finite-dimensional subspace $F_n \subset E^+$ from condition (iii) one next proves that $i_0(c) \geqslant \dim F_n$ for any $c \geqslant \Phi(E^- \oplus F_n)$. This is a consequence of the properties of the index stated in Proposition 1.7. No extension of the index to infinite dimensions is needed.

Details of the proof, of a slightly more general result in fact, can be found in [12] for $p = 2$ and in [13] for p an odd prime. □

2. Periodic solutions

2.1. *Variational setting for periodic solutions*

In this section we reformulate the problem of existence of 2π-periodic solutions of the Hamiltonian system

$$\dot{z} = JH_z(z,t) \tag{2.1}$$

in terms of the existence of critical points of a suitable functional and we collect some basic facts about this functional. When looking for periodic solutions of (2.1) we shall always assume that the Hamiltonian $H = H(z,t)$ satisfies the following conditions:

(H_1) $H \in C(\mathbb{R}^{2N} \times \mathbb{R}, \mathbb{R})$, $H_z \in C(\mathbb{R}^{2N} \times \mathbb{R}, \mathbb{R}^{2N})$ and $H(0,t) \equiv 0$;
(H_2) H is 2π-periodic in the t-variable;
(H_3) $|H_z(z,t)| \leqslant c(1 + |z|^{s-1})$ for some $c > 0$ and $s \in (2,\infty)$.

We note that it causes no loss of generality to assume $H(0,t) \equiv 0$. Occasionally we shall need two additional conditions:

(H_4) $H_{zz} \in C(\mathbb{R}^{2N} \times \mathbb{R}, \mathbb{R}^{4N^2})$;
(H_5) $|H_{zz}(z,t)| \leqslant d(1 + |z|^{s-2})$ for some $d > 0$ and $s \in (2,\infty)$.

Clearly, (H_5) implies (H_3).

Let $E := H^{1/2}(S^1, \mathbb{R}^{2N})$ be the Sobolev space of 2π-periodic $\mathbb{R}^{2N}$-valued functions

$$z(t) = a_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt), \quad a_0, a_k, b_k \in \mathbb{R}^{2N} \tag{2.2}$$

such that $\sum_{k=1}^{\infty} k(|a_k|^2 + |b_k|^2) < \infty$. Then E is a Hilbert space with an inner product

$$\langle z, w\rangle := 2\pi a_0 \cdot a_0' + \pi \sum_{k=1}^{\infty} k\big(a_k \cdot a_k' + b_k \cdot b_k'\big), \tag{2.3}$$

where a_k', b_k' are the Fourier coefficients of w. It is well known that the Sobolev embedding $E \hookrightarrow L^q(S^1, \mathbb{R}^{2N})$ is compact for any $q \in [1,\infty)$ (see, e.g., [2]) but $z \in E$ does not imply z is bounded. There is a natural action of $\mathbb{R}$ on $L^q(S^1, \mathbb{R}^{2N})$ and E given by time translation:

$$(T_\theta z)(t) := z(t+\theta) \quad \text{for } \theta, t \in \mathbb{R}.$$

Since the functions z are 2π-periodic in t, T induces an isometric representation of $G = S^1 \equiv \mathbb{R}/2\pi\mathbb{Z}$. In the notation of Section 1.1, we have $\mathcal{O}(z_1) = \mathcal{O}(z_2)$ if and only if $z_2(t) = z_1(t+\theta)$ for some θ and all $t \in \mathbb{R}$. Let

$$\Phi(z) := \frac{1}{2}\int_0^{2\pi} (-J\dot{z} \cdot z)\,\mathrm{d}t - \int_0^{2\pi} H(z,t)\,\mathrm{d}t$$

and

$$\psi(z) := \int_0^{2\pi} H(z,t)\,\mathrm{d}t.$$

PROPOSITION 2.1. *If H satisfies* (H_1)–(H_3)*, then $\Phi \in C^1(E,\mathbb{R})$ and $\Phi'(z) = 0$ if and only if z is a 2π-periodic solution of* (2.1)*. Moreover, ψ' is completely continuous in the sense that $\psi'(z_j) \to \psi'(z)$ whenever $z_j \rightharpoonup z$. If, in addition, H satisfies* (H_4) *and* (H_5)*, then $\Phi \in C^2(E,\mathbb{R})$ and $\psi''(z)$ is a compact linear operator for each z.*

PROOF. We only outline the argument. The details may be found, e.g., in [80, Appendix B] or [102, Appendix A and Lemma 2.16]. Although the results in [102] concern elliptic partial differential equations, the proofs are easy to adapt to our situation.

Let $s' = s/(s-1)$ be the conjugate exponent. By (H_3),

$$H_z : L^s\bigl(S^1,\mathbb{R}^{2N}\bigr) \to L^{s'}\bigl(S^1,\mathbb{R}^{2N}\bigr)$$

is a continuous mapping, and using this one shows that $\psi \in C^1(E,\mathbb{R})$ and

$$\bigl\langle \psi'(z), w\bigr\rangle = \int_0^{2\pi} H_z(z,t)\cdot w\,\mathrm{d}t.$$

Moreover, it follows by the compact embedding of E into $L^s(S^1,\mathbb{R}^{2N})$ that ψ' is completely continuous.

Since $-J\dot{z}\cdot w = \dot{z}\cdot Jw$, the bilinear form $(z,w) \mapsto \int_0^{2\pi}(-J\dot{z}\cdot w)\,\mathrm{d}t$ is (formally) self-adjoint. According to (2.2) and (2.3),

$$\int_0^{2\pi}(-J\dot{z}\cdot w)\,\mathrm{d}t = \pi\sum_{k=1}^{\infty} k\bigl(-Jb_k\cdot a'_k + Ja_k\cdot b'_k\bigr),$$

hence this form is continuous in E and the quadratic form $z \mapsto \int_0^{2\pi}(-J\dot{z}\cdot z)\,\mathrm{d}t$ is of class C^1. Now it is easy to see that $\Phi'(z) = 0$ if and only if z is a 2π-periodic solution of (2.1). Moreover, by elementary regularity theory, $z \in C^1(S^1,\mathbb{R}^{2N})$.

If (H_4) and (H_5) are satisfied, then, referring to the arguments in [80], [102] again, we see that $\psi \in C^2(E,\mathbb{R})$ and

$$\bigl\langle \psi''(z)w, y\bigr\rangle = \int_0^{2\pi} H_{zz}(z,t)w\cdot y\,\mathrm{d}t.$$

Since ψ' is completely continuous, $\psi''(z)$ is a compact linear operator. □

Note that complete continuity of ψ' implies weak continuity of ψ (i.e., $\psi(z_j) \to \psi(z)$ whenever $z_j \rightharpoonup z$).

REMARK 2.2. If system (2.1) is *autonomous*, i.e., $H = H(z)$, then $\Phi(T_\theta z) = \Phi(z)$ for all $\theta \in \mathbb{R}$. Thus Φ is T-invariant. Two 2π-periodic solutions z_1, z_2 of an autonomous system are *geometrically distinct* if and only if $\mathcal{O}(z_1) \neq \mathcal{O}(z_2)$. When $H = H(z)$, we shall write $H'(z)$ instead of $H_z(z)$.

Let $z(t) = a_k \cos kt \pm J a_k \sin kt$. Then

$$\int_0^{2\pi} (-J\dot{z} \cdot z)\,\mathrm{d}t = \pm 2\pi k |a_k|^2 = \pm\|z\|^2.$$

It follows that E has the orthogonal decomposition $E = E^+ \oplus E^0 \oplus E^-$, where

$$E^0 = \{z \in E\colon\ z \equiv a_0 \in \mathbb{R}^{2N}\},$$
$$E^\pm = \left\{z \in E\colon\ z(t) = \sum_{k=1}^\infty a_k \cos kt \pm J a_k \sin kt,\ a_k \in \mathbb{R}^{2N}\right\}$$

and if $z = z^0 + z^+ + z^-$, then

$$\int_0^{2\pi} (-J\dot{z} \cdot z)\,\mathrm{d}t = \|z^+\|^2 - \|z^-\|^2.$$

Hence

$$\Phi(z) = \frac{1}{2}\int_0^{2\pi} (-J\dot{z} \cdot z)\,\mathrm{d}t - \int_0^{2\pi} H(z,t)\,\mathrm{d}t = \frac{1}{2}\|z^+\|^2 - \frac{1}{2}\|z^-\|^2 - \psi(z). \tag{2.4}$$

REMARK 2.3. Φ is called the action functional and the fact that z is a 2π-periodic solution of (2.1) if and only if $\Phi'(z) = 0$ is the least action (or the Euler–Maupertuis) principle. However, although the solutions z are critical points (or extremals) of Φ, they can never be minima (or maxima). Indeed, let $a \in \mathbb{R}^{2N}$ and $z_j = a \cos jt \pm Ja \sin jt$. Then $\Phi(z_j) \to \pm\infty$, so Φ is unbounded below and above. Moreover, using (2.4) and the weak continuity of ψ, it is easy to see that Φ has neither local maxima nor minima.

The first to develop a variational method for finding periodic solutions of a Hamiltonian system as critical points of the action functional was Rabinowitz [78]. Among other he has shown that a star-shaped compact energy surface necessarily carries a closed Hamiltonian orbit (see Section 2.3 for a discussion of this problem).

LEMMA 2.4. *Suppose* (H_1)–(H_3) *are satisfied.*

(i) *If* $H(z,t) = \frac{1}{2}A_0(t)z \cdot z + G_0(z,t)$ *and* $(G_0)_z(z,t) = o(|z|)$ *uniformly in* t *as* $z \to 0$, *then* $\psi_0'(z) = o(\|z\|)$ *and* $\psi_0(z) = o(\|z\|^2)$ *as* $z \to 0$, *where* $\psi_0(z) := \int_0^{2\pi} G_0(z,t)\,\mathrm{d}t$.

(ii) *If* $H(z,t) = \frac{1}{2}A_\infty(t)z \cdot z + G_\infty(z,t)$ *and* $(G_\infty)_z(z,t) = \mathrm{o}(|z|)$ *uniformly in* t *as* $|z| \to \infty$, *then* $\psi'_\infty(z) = \mathrm{o}(\|z\|)$ *and* $\psi_\infty(z) = \mathrm{o}(\|z\|^2)$ *as* $\|z\| \to \infty$, *where* $\psi_\infty(z) := \int_0^{2\pi} G_\infty(z,t)\,\mathrm{d}t$.

PROOF. Since $H(0,t) = 0$, $\psi_0(0) = \psi_\infty(0) = 0$. It follows from ($H_3$) that for each $\varepsilon > 0$ there is a $c_1(\varepsilon) > 0$ such that $|(G_0)_z(z,t)| \leqslant \varepsilon|z| + c_1(\varepsilon)|z|^{s-1}$. Hence by the Sobolev inequality,

$$\bigl|\langle\psi'_0(z), w\rangle\bigr| \leqslant \int_0^{2\pi} \bigl(\varepsilon|z| + c_1(\varepsilon)|z|^{s-1}\bigr)|w|\,\mathrm{d}t \leqslant \bigl(\varepsilon\|z\| + c_2(\varepsilon)\|z\|^{s-1}\bigr)\|w\|.$$

Taking the supremum over $\|w\| \leqslant 1$ and letting $z \to 0$ we see that $\psi'_0(z) = \mathrm{o}(\|z\|)$ as $z \to 0$. Since

$$\psi_0(z) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\psi_0(sz)\,\mathrm{d}s = \int_0^1 \bigl\langle\psi'_0(sz), z\bigr\rangle\,\mathrm{d}s,$$

$\psi_0(z) = \mathrm{o}(\|z\|^2)$.

Similarly, for each $\varepsilon > 0$ there is a $c_3(\varepsilon)$ such that $|(G_\infty)_z(z,t)| \leqslant \varepsilon|z| + c_3(\varepsilon)$. So

$$\bigl|\langle\psi'_\infty(z), w\rangle\bigr| \leqslant \bigl(\varepsilon\|z\| + c_4(\varepsilon)\bigr)\|w\|$$

and $\psi'_\infty(z) = \mathrm{o}(\|z\|)$ as $\|z\| \to \infty$. Since

$$\bigl|\psi_\infty(z)\bigr| = \left|\int_0^1 \bigl\langle\psi'_\infty(sz), z\bigr\rangle\,\mathrm{d}s\right| \leqslant \frac{1}{2}\varepsilon\|z\|^2 + c_4(\varepsilon)\|z\|,$$

$\psi_\infty(z) = \mathrm{o}(\|z\|^2)$ as $\|z\| \to \infty$. □

Let now

$$\widetilde{E}_n := \left\{z \in E\colon\ z(t) = \sum_{k=1}^{n}(a_k\cos kt + b_k\sin kt)\right\} \quad \text{and} \quad E_n := E^0 \oplus \widetilde{E}_n. \tag{2.5}$$

Then $E_n \subset E_{n+1}$ for all n,

$$E = \overline{\bigcup_{n=1}^{\infty} E_n}$$

and

$$E_n^0 = E^0, \quad \widetilde{E}_n^\pm = E_n^\pm = E^\pm \cap E_n. \tag{2.6}$$

Moreover, $E^{S^1} = E^0$ (so $E^{S^1} \subset E_n$ for all n) and each subspace E_n is S^1-invariant (or more precisely, T-invariant). When using limit index theories we shall have $\mathcal{E} = \{E_n, d_n\}$, where

$$d_n \cdot d_G = 2N(1+n) = \dim E^0 + \frac{1}{2} \dim \widetilde{E}_n = \dim E^0 + \dim \widetilde{E}_n^+.$$

The orthogonal projection $E \to E_n$ will be denoted by P_n.

PROPOSITION 2.5. *Suppose* (H$_1$)–(H$_3$) *are satisfied and let* $\{z_j\}$ *be a sequence such that* $z_j \in E_{n_j}$ *for some* n_j, $n_j \to \infty$ *and* $P_{n_j}\Phi'(z_j) \to 0$ *as* $j \to \infty$. *Then* $\{z_j\}$ *has a convergent subsequence in each of the following two cases*:

(i) $H(z,t) = \frac{1}{2}A_\infty(t)z \cdot z + G_\infty(z,t)$, *where* $(G_\infty)_z(z,t) = \mathrm{o}(|z|)$ *as* $|z| \to \infty$ *and* $z = 0$ *is the only* 2π*-periodic solution of the linear system*

$$\dot{z} = JA_\infty(t)z.$$

(ii) $\Phi(z_j)$ *is bounded above and there exist* $\mu > \max\{2, s-1\}$ *and* $R > 0$ *such that*

$$0 < \mu H(z,t) \leqslant z \cdot H_z(z,t) \quad \text{for all } |z| \geqslant R.$$

So in particular, (PS)* *holds if* H *satisfies one of the conditions above.*

It follows upon integration that (ii) implies

$$H(z,t) \geqslant a_1|z|^\mu - a_2 \tag{2.7}$$

for some $a_1, a_2 > 0$. Hence H grows superquadratically and H_z superlinearly as $|z| \to \infty$. Note also that $\mu \leqslant s$ according to (H$_3$).

PROOF. Let $z_j \in E_{n_j}$, $n_j \to \infty$ and $P_{n_j}\Phi'(z_j) \to 0$ as $j \to \infty$. Since $P_{n_j}\Phi'(z_j) = z_j^+ - z_j^- - P_{n_j}\psi'(z_j) \to 0$ and ψ' is completely continuous, it follows that if $\{z_j\}$ is bounded, then $z_j \to z$ after passing to a subsequence. Moreover, $\Phi'(z) = 0$. Hence it remains to show that $\{z_j\}$ must be bounded.

Suppose (i) is satisfied and let

$$\langle B_\infty z, w\rangle := \int_0^{2\pi} A_\infty(t)z \cdot w \,\mathrm{d}t, \qquad L_\infty z := z^+ - z^- - B_\infty z.$$

Then $B_\infty : E \to E$ is a compact linear operator (cf. Proposition 2.1). Since $L_\infty z = 0$ if and only if $\dot{z} = JA_\infty(t)z$, L_∞ is invertible and it follows that $(P_n L_\infty|_{E_n})^{-1}$ is uniformly bounded for large n. Hence

$$(P_{n_j}L_\infty)^{-1}P_{n_j}\Phi'(z_j) = z_j - (P_{n_j}L_\infty)^{-1}P_{n_j}\psi_\infty'(z_j) \to 0$$

and $\{z_j\}$ is bounded because $\psi'_\infty(z) = \mathrm{o}(\|z\|)$ as $\|z\| \to \infty$.

Suppose now (ii) holds. Below $c_1, c_2, \ldots$ will denote different constants whose exact values are insignificant. Since $z_j \in E_{n_j}$,

$$\begin{aligned} c_1\|z_j\| + c_2 &\geqslant \Phi(z_j) - \frac{1}{2}\langle P_{n_j}\Phi'(z_j), z_j\rangle \\ &= \int_0^{2\pi} \left(\frac{1}{2}z_j \cdot H_z(z_j,t) - H(z_j,t)\right) \mathrm{d}t \\ &\geqslant \left(\frac{\mu}{2} - 1\right)\int_0^{2\pi} H(z_j,t)\,\mathrm{d}t \\ &\geqslant c_3\|z_j\|_\mu^\mu - c_4, \end{aligned} \tag{2.8}$$

where the last inequality follows from (2.7). Since E^0 is finite dimensional, $\|z_j^0\| \leqslant c_5\|z_j\|_\mu$ and by (2.8),

$$\|z_j^0\| \leqslant c_6\|z_j\|^{1/\mu} + c_7. \tag{2.9}$$

By the Hölder and Sobolev inequalities,

$$\begin{aligned} \|z_j^+\|^2 &= \langle\Phi'(z_j), z_j^+\rangle + \int_0^{2\pi} H_z(z_j,t)\cdot z_j^+\,\mathrm{d}t \\ &\leqslant c_8\|z_j^+\| + c_9\int_0^{2\pi} |z_j|^{s-1}|z_j^+|\,\mathrm{d}t \\ &\leqslant c_8\|z_j^+\| + c_{10}\|z_j\|_\mu^{s-1}\|z_j^+\| \end{aligned} \tag{2.10}$$

(here we have used that $\mu > s-1$) and a similar inequality holds for z_j^-. Hence

$$\|z_j^\pm\| \leqslant c_8 + c_{10}\|z_j\|_\mu^{s-1}$$

and by (2.8),

$$\|z_j^\pm\| \leqslant c_{11} + c_{12}\|z_j\|^{(s-1)/\mu}.$$

This and (2.9) combined imply $\{z_j\}$ is bounded. □

REMARK 2.6. (a) If $H(z,t) = \frac{1}{2}A(t)z\cdot z + G(z,t)$, where A is a symmetric $2N \times 2N$ matrix with periodic entries and G satisfies the superlinearity condition of Proposition 2.5, it is easy to see by an elementary computation that so does H, possibly with a smaller $\mu > \max\{2, s-1\}$ and a larger R.

(b) In some problems the growth restriction (H_3) may be removed and the condition $\mu > \max\{2, s-1\}$ can be replaced by $\mu > 2$. For this purpose one introduces a modified

Hamiltonian H_K such that $H_K(z,t) = H(z,t)$ for $|z| \leqslant K$ and $H_K(z,t) = C|z|^s$ for $|z| \geqslant K+1$ and some convenient C, s. Then the modified functional satisfies (PS)*, one can apply a suitable variational method to obtain one or more solutions which are uniformly bounded independently of K. Hence these solutions satisfy the original equation for K large. We shall comment on that when appropriate.

Let now H satisfy (H_1)–(H_5). For each fixed z we have $\langle \Phi''(z)w, w\rangle = \|w^+\|^2 - \|w^-\|^2 - \langle \psi''(z)w, w\rangle$, and since $\psi''(z)$ is a compact linear operator, it is easy to see that the quadratic forms $\pm\Phi''(z)$ have infinite Morse index. However, it is possible to define a certain relative index which will always be finite. Let A be a symmetric $2N \times 2N$ constant matrix and

$$\langle Lz, w\rangle := \int_0^{2\pi} (-J\dot{z} - Az) \cdot w \, \mathrm{d}t. \tag{2.11}$$

It follows from (2.2) and (2.3) that

$$\langle Lz, z\rangle = -2\pi A a_0 \cdot a_0 + \pi \sum_{k=1}^{\infty} k\left(\left(-Jb_k - \frac{1}{k}Aa_k\right)\cdot a_k + \left(Ja_k - \frac{1}{k}Ab_k\right)\cdot b_k\right). \tag{2.12}$$

The restriction of this quadratic form to a subspace corresponding to a fixed $k \geqslant 1$ is represented by the $(4N \times 4N)$-matrix $\pi k T_k(A)$, where

$$T_k(A) := \begin{pmatrix} -\frac{1}{k}A & -J \\ J & -\frac{1}{k}A \end{pmatrix}.$$

Let $M^+(\cdot)$ and $M^-(\cdot)$ respectively denote the number of positive and negative eigenvalues of a symmetric matrix (counted with their multiplicities) and let $M^0(\cdot)$ be the dimension of the nullspace of this matrix. Then $M^0(T_k(A)) = 0$ and $M^\pm(T_k(A)) = 2N$ for all k large enough. Indeed, a simple computation shows that the matrix

$$\begin{pmatrix} 0 & -J \\ J & 0 \end{pmatrix}$$

has the eigenvalues ± 1, each of multiplicity $2N$, so by a simple perturbation argument, $M^\pm(T_k(A)) = 2N$ for almost all k (cf. [3, Section 12], [4, Section 2]). Therefore the following numbers are well defined and finite:

$$i^-(A) := M^+(A) - N + \sum_{k=1}^{\infty} \big(M^-\big(T_k(A)\big) - 2N\big), \tag{2.13}$$

$$i^{+}(A) := M^{-}(A) - N + \sum_{k=1}^{\infty} \big(M^{+}\big(T_k(A)\big) - 2N\big) \tag{2.14}$$

and

$$i^{0}(A) := M^{0}(A) + \sum_{k=1}^{\infty} M^{0}\big(T_k(A)\big). \tag{2.15}$$

Clearly, $i^{\pm}(A)$, $i^{0}(A)$ are finite and $i^{-}(A) + i^{+}(A) + i^{0}(A) = 0$. The quantity $i^{-}(A)$ is a relative Morse index in the sense that it provides a measure for the difference between the negative parts of the quadratic forms $z \mapsto \int_0^{2\pi} (-J\dot{z} - Az) \cdot z \,\mathrm{d}t$ and $z \mapsto \int_0^{2\pi} (-J\dot{z} \cdot z) \,\mathrm{d}t$.

It is easy to see that $i^{0}(A) = \dim N(L)$ is the number of linearly independent 2π-periodic solutions of the linear system

$$\dot{z} = JAz. \tag{2.16}$$

Hence in particular $i^{0}(A) \leqslant 2N$. Moreover, $i^{0}(A) = 0$ if and only if $\sigma(JA) \cap \mathrm{i}\mathbb{Z} = \emptyset$. Indeed, it follows from (2.12) (or by substituting (2.2) into (2.16)) that (2.16) has a nontrivial 2π-periodic solution if and only if either A is singular (so $0 \in \sigma(JA)$) or

$$-kJb_k = Aa_k \quad \text{and} \quad kJa_k = Ab_k \tag{2.17}$$

for some $(a_k, b_k) \neq (0, 0)$ and $k \geqslant 1$. (2.17) is equivalent to

$$JA(a_k - \mathrm{i}b_k) = \mathrm{i}k(a_k - \mathrm{i}b_k).$$

Hence $\pm\mathrm{i}k \in \sigma(JA)$ and (2.16) has a nontrivial 2π-periodic solution if and only if $\sigma(JA) \cap \mathrm{i}\mathbb{Z} \neq \emptyset$. We also see that P_n commutes with L, hence if $E = E^{+}(L) \oplus E^{0}(L) \oplus E^{-}(L)$ is the orthogonal decomposition corresponding to the positive, zero and negative part of the spectrum of L and $E_n^{\pm}(L) := E^{\pm}(L) \cap E_n$, $E_n^{0}(L) := E^{0}(L) \cap E_n$, then $E_n = E_n^{+}(L) \oplus E_n^{0}(L) \oplus E_n^{-}(L)$ is an orthogonal decomposition into the positive, zero and negative part of $L_n := P_n L|_{E_n}$. Note that $E^{0}(L) = N(L)$ and $E_n^{0}(L) = E^{0}(L)$ for almost all n.

In one of our applications we shall need a slight extension of Proposition 2.5.

COROLLARY 2.7. *Suppose H is as in* (*a*) *of Remark* 2.6 *and A is a constant matrix. If $\{z_j\}$ is a sequence such that $\Phi(z_j)$ is bounded above, $z_j = w_j^{+} + w_j^{0} + w_j^{-} \in E_{m_j}^{+}(L) \oplus E^{0}(L) \oplus E_{n_j}^{-}(L)$, $m_j, n_j \to \infty$ and $(P_{m_j}^{+} + P^{0} + P_{n_j}^{-})\Phi'(z_j) \to 0$ as $j \to \infty$ ($P_n^{\pm}$ and P^{0} denote the orthogonal projections $E \to E_n^{\pm}(L)$, $E \to E^{0}(L)$), then $\{z_j\}$ has a convergent subsequence.*

The proof follows by inspection of the argument of (ii) in Proposition 2.5. Note in particular that (2.8) holds with $G(z,t) = H(z,t) - \frac{1}{2}Az \cdot z$ replacing H and $E^{0}(L)$ replacing E^{0}. Moreover, since

$$\langle \pm L_n z, z\rangle \geqslant \varepsilon \|z\|^2 \quad \text{for some } \varepsilon > 0 \text{ and all } z \in E_n^{\pm}(L) \tag{2.18}$$

(ε independent of n), also (2.10) can be easily adapted.

REMARK 2.8. (a) If A is a nonconstant matrix with 2π-periodic entries, the definitions of $i^{\pm}(A)$ and $i^{0}(A)$ no longer make sense. Therefore we need some other quantities to measure the size of the positive and the negative part of L. Assume for simplicity that the operator L corresponding to A is invertible and let $M_n^-(L)$ be the Morse index of the quadratic form $z \mapsto \langle Lz, z\rangle$ restricted to E_n. Let

$$j^-(A) := \lim_{n\to\infty} \big(M_n^-(L) - (1+2n)N\big)$$

and $j^+(A) := j^-(-A)$. It can be shown that $j^{\pm}(A)$ are well defined and finite, and $j^{\pm}(A) = i^{\pm}(A)$ whenever A is a constant matrix [57, Sections 7 and 5]. See also the references below.

(b) Morse-type indices for Hamiltonian systems have been introduced by Amann and Zehnder [3,4] and Benci [18]. In [3,4] and [18] computational formulas for these indices are also discussed. Our definitions of $i^{\pm}(A)$, $i^{0}(A)$ follow Li and Liu [61] (more precisely, the indices $i^{\pm}(A)$ as defined here differ from those in [61] by N). The number $j^-(A)$ equals the Conley–Zehnder (or Maslov) index of the fundamental solution $\gamma : [0, 2\pi] \to Sp(2N)$ of the equation $\dot{z}(t) = JA(t)$. Here $Sp(2N)$ denotes the group of symplectic $2N \times 2N$-matrices. Recall that a matrix C is symplectic if $C^t JC = J$. See the books of Abbondandolo [1], Chang [23, Section IV.1] and in particular Long [68,69] for a comprehensive discussion of the Conley–Zehnder index.

2.2. *Periodic solutions near equilibria*

The first existence and multiplicity results for periodic solutions of (2.1) are concerned with solutions near an equilibrium. The classical results of Lyapunov [72], Weinstein [98], and Moser [76] have been very influential for the development of the theory and can be proved using the basic variational methods from Section 1.1.

We consider the autonomous Hamiltonian system

$$\dot{z} = JH'(z) \tag{2.19}$$

where the Hamiltonian $H : \mathbb{R}^{2N} \to \mathbb{R}$ is of class C^2. Since the vector field H' is of class C^1, each initial value problem has a unique solution $z = z(t)$ defined on some maximal interval I. Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}t} H\big(z(t)\big) = H'\big(z(t)\big) \cdot \dot{z}(t) = -J\dot{z}(t) \cdot \dot{z}(t) = 0, \tag{2.20}$$

hence $H(z(t))$ is constant for all $t \in I$. Throughout this section we assume that H has 0 as critical point. We first consider the case where 0 is nondegenerate. The constant function $z \equiv 0$ is then an isolated, hyperbolic stationary solution of (2.19). It is well known that periodic orbits near 0 can only exist if $JH''(0)$ has purely imaginary eigenvalues. The

Lyapunov center theorem states that if $JH''(0)$ has a pair of purely imaginary eigenvalues $\pm \mathrm{i}\omega$ which are simple, and if no integer multiples $\pm \mathrm{i}k\omega$ are eigenvalues of $JH''(0)$ then (2.19) has a one-parameter family of periodic solutions emanating from the equilibrium point. More precisely, let $E(\pm \mathrm{i}\omega) \subset \mathbb{R}^{2N}$ be the two-dimensional eigenspace associated to $\pm \mathrm{i}\omega$ and let $\sigma \in \{\pm 2\}$ be the signature of the quadratic form $Q(z) = \frac{1}{2}H''(0)z \cdot z$ on $E(\pm \mathrm{i}\omega)$. Then for each $\varepsilon > 0$ small enough there exists a periodic solution of (2.19) on the energy surface $H = H(0) + \sigma\varepsilon^2$ with period converging to $2\pi/\omega$ as $\varepsilon \to 0$.

If $\pm \mathrm{i}\omega$ is not simple, or if integer multiples are eigenvalues then there may be no periodic solutions near 0 as elementary examples show; see [76,27]. In order to formulate a sufficient condition we assume that all eigenvalues of $JH''(0)$ which are of the form $\mathrm{i}k\omega$, $k \in \mathbb{Z}$, are semisimple, i.e., their geometric and algebraic multiplicities are equal. Let $E_\omega \subset \mathbb{R}^{2N}$ be the generalized eigenspace of $JH''(0)$ corresponding to the eigenvalues of the form $\pm \mathrm{i}k\omega$, $k \in \mathbb{N}$. Let $Q : \mathbb{R}^{2N} \to \mathbb{R}$, $Q(z) = \frac{1}{2}H''(0)z \cdot z$, be the quadratic part of H at 0 and let $\sigma = \sigma(\omega) \in \mathbb{Z}$ be the signature of the quadratic form $Q|E_\omega$ on E_ω. Observe that σ is automatically an even integer.

THEOREM 2.9. *If $\sigma \neq 0$ then one of the following statements hold.*

(i) *There exists a sequence of nonconstant T_k-periodic solutions z_k of* (2.19) *which lie on the energy surface $H = H(0)$ with $z_k \to 0$ and $T_k \to 2\pi/\omega$ as $k \to \infty$. The period T_k is not necessarily minimal.*

(ii) *For $\varepsilon > 0$ small enough there are at least $|\sigma/2|$ nonconstant periodic solutions z_j^ε, $j = 1, \ldots, |\sigma/2|$, of* (2.19) *on the energy surface $H = H(0) + \sigma\varepsilon^2$ with* (*not necessarily minimal*) *period T_j^ε. These solutions converge towards* 0 *as $\varepsilon \to 0$. Moreover, $T_j^\varepsilon \to 2\pi/\omega$ as $\varepsilon \to 0$.*

This theorem is due to Bartsch [11]. It generalizes the Weinstein–Moser theorem [98, 76] which corresponds to the case where $Q|E_\omega$ is positive or negative definite, hence $|\sigma| = \dim E_\omega$. Observe that the energy surfaces $H = c$ are not necessarily compact for c close to $H(0)$.

PROOF (outline). We may assume $H(0) = 0$. The τ-periodic solutions of (2.19) correspond to 2π-periodic solutions of

$$\dot{z} = \frac{\tau}{2\pi} JH'(z). \tag{2.21}$$

These in turn correspond to critical points of the action functional

$$\mathcal{A}(z) = \int_0^{2\pi} J\dot{z}(t) \cdot z(t)\,\mathrm{d}t$$

restricted to the surface $\{z \in E\colon\ \psi(z) = 2\pi H(0) + \lambda\}$ where $E = H^{1/2}(S^1, \mathbb{R}^{2N})$ and $\psi(z) = \int_0^{2\pi} H(z(t))\,\mathrm{d}t$ are as in Section 2.1. The period appears as Lagrange multiplier in this approach. After performing a Lyapunov–Schmidt reduction of the equation

$$\mathcal{A}'(z) = \frac{\tau}{2\pi}\psi'(z)$$

near $\tau = 2\pi/\omega$ and $z \equiv 0$, one is left with the problem of finding critical points of a function

$$\mathcal{A}_0(v) = \mathcal{A}\bigl(v + \bar{w}(v)\bigr)$$

constrained to the level set $\{v \in V\colon\ \psi_0(v) = \lambda\}$. Here V is the kernel of the linearization

$$E \ni z \mapsto \mathcal{A}''(0)z - \frac{\tau}{2\pi}\psi''(0)z \in E,$$

$\bar{w}\colon V \supset U \to V^{\perp} \subset E$ is defined on a neighborhood U of 0 in V, and $\psi_0(v) := \psi(v + \bar{w}(v))$. Thus $V \cong E_\omega$ and one checks that $\mathcal{A}_0$ and ψ_0 are of class C^1 and that $\psi_0''(0)$ exists. In fact:

$$\bigl\langle \psi_0''(0)v, w\bigr\rangle = \int_0^{2\pi} H''(0)v(t) \cdot w(t)\,\mathrm{d}t,$$

hence we can apply the Morse lemma to ψ_0 near 0. After a change of coordinates ψ_0 looks (in the sense of the Morse lemma) near 0 like the nondegenerate quadratic form

$$q: V \to \mathbb{R}, \quad v \mapsto \frac{1}{2}H''(0)v \cdot v.$$

Therefore the level surfaces $\psi_0^{-1}(\lambda)$ look locally like the level surfaces $q^{-1}(\lambda)$. If q is positive definite, hence $\sigma = \dim V = \dim E_\omega$ (which is just the situation of the Weinstein–Moser theorem), one can conclude the proof easily upon observing that the functionals $\mathcal{A}$, ψ and, hence $\mathcal{A}_0$ and ψ_0 are invariant under the representation of $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ in E induced by the time shifts. Moreover, $\psi_0^{-1}(\lambda) \cong q^{-1}(\lambda)$ is diffeomorphic to the unit sphere SV of V for $\lambda > 0$ small. By Theorem 1.9 any C^1-functional $SV \to \mathbb{R}$ which is invariant under the action of S^1 has at least $\frac{1}{2}\dim V = \frac{1}{2}\dim E_\omega$ S^1-orbits of critical points. If q is negative definite then $\psi_0^{-1}(\lambda)$ is diffeomorphic to the unit sphere SV of V for $\lambda < 0$ close to 0, and one obtains $|\sigma/2|$ critical orbits on these levels.

This elementary argument from S^1-equivariant critical point theory does not work if q is indefinite. Instead one looks at the flow φ_λ on $\Sigma_\lambda := \psi_0^{-1}(\lambda)$ which is essentially induced by the negative gradient of $\mathcal{A}_0|\Sigma_\lambda$. Since $\mathcal{A}_0$ and ψ_0 are only of class C^1 the gradient vector field is of class C^0, so it may not be integrable and has to be replaced by a pseudogradient vector field which leaves Σ_λ invariant for all λ. Next one observes that the hypersurfaces Σ_λ undergo a surgery as λ passes $H(0) = 0$. If $2n^+$ (respectively $2n^-$) is the maximal dimension of a subspace of V on which q is positive (respectively negative) definite then Σ_ε is obtained from $\Sigma_{-\varepsilon}$ upon replacing a handle of type $B^{2n^+} \times S^{2n^- -1}$ by $S^{2n^+} \times B^{2n^-}$. It is this change in the topology of Σ_λ near 0 which forces the existence of stationary orbits of φ near the origin. In order to analyze the influence of this surgery on the flow φ_λ one has to use methods from equivariant Conley index theory and Borel cohomology. The difference $|n^+ - n^-| = |\sigma|/2$ is a lower bound for the number of stationary S^1-orbits of φ_λ on Σ_λ if $\lambda > 0$ is small and $\sigma \cdot (\lambda - H(0)) > 0$. □

It is also possible to parameterize the nontrivial periodic orbits near an equilibrium by their period. The following result is due to Fadell and Rabinowitz [36].

THEOREM 2.10. *If* $\sigma \neq 0$ *then one of the following statements hold.*

(i) *There exists a sequence of nonconstant periodic orbits* $z_k \to 0$ *of* (2.19) *with* (*not necessarily minimal*) *period* $T = 2\pi/\omega$.

(ii) *There exist integers* $k, l \geqslant 0$ *with* $k + l \geqslant |\sigma|/2$, *and there exists* $\varepsilon > 0$ *such that for each* $\tau \in (T - \varepsilon, T)$ (2.19) *has at least k periodic orbits* z_j^τ, $j = 1, \ldots, k$ *with* (*not necessarily minimal*) *period* τ. *And for each* $\tau \in (T, T + \varepsilon)$ (2.19) *has at least l periodic orbits* z_j^τ, $j = k + 1, \ldots, k + l$ *with* (*not necessarily minimal*) *period* τ. *Moreover,* $z_j^\tau \to 0$ *as* $\tau \to T = 2\pi/\omega$.

The proof uses a cohomological index theory. The integers k, l (and thus the direction of the bifurcating solutions with the period as parameter) are not determined by $H''(0)$—unlike case (ii) in Theorem 2.9.

Now we consider the case of a degenerate equilibrium. Suppose first that 0 is an isolated critical point of H, so there are no stationary orbits of (2.19) in a neighbourhood of the origin. Let $\mathrm{i}\omega$ be an eigenvalue of $JH''(0)$ and let $F_\omega \subset \mathbb{R}^{2N}$ be the generalized eigenspace of $JH''(0)$ corresponding to $\pm\mathrm{i}\omega$. Thus $F_\omega \subset E_\omega$ does not contain generalized eigenvectors of $JH''(0)$ corresponding to multiples $\pm\mathrm{i}k\omega$ with $|k| \geqslant 2$. Let $\sigma_1 = \sigma_1(\omega)$ be the signature of the quadratic form $Q|F_\omega$. Since we allow $H''(0)$ to have a nontrivial kernel we also need the critical groups $C^q(H, 0) = \check{H}^q(H^0, H^0 \setminus \{0\})$ associated to $0 \in \mathbb{R}^{2N}$ as a critical point of the Hamiltonian H. Here $\check{H}^*$ denotes the Čech (or Alexander–Spanier) cohomology with coefficients in an arbitrary field.

THEOREM 2.11. *If* $\sigma_1 \neq 0$ *and* $C^q(H, 0) \neq 0$ *for some* $q \in \mathbb{Z}$, *then there exists a sequence* z_k *of nonconstant periodic orbits of* (2.19) *with* (*not necessarily minimal*) *period* T_n *such that* $\|z_n\|_{L^\infty} \to 0$ *and* $T_n \to 2\pi/\omega$.

The result has been proved by Szulkin in [91] using Morse theoretic methods. It is unknown whether the solutions obtained in Theorems 2.9–2.11 lie on connected branches of periodic solutions. Continua of periodic solutions however do exist under stronger hypotheses when degree theoretic methods apply. We state one such result in this direction.

THEOREM 2.12. *If* $\sigma_1 \neq 0$ *and the local degree* $\deg(\nabla H, 0)$ *of* ∇H *at the isolated critical point* 0 *is nontrivial, then there exists a connected branch of periodic solutions of* (2.19) *near* 0.

For a proof see the paper [27] by Dancer and Rybicki. They work in the space $W^1(S^1, \mathbb{R}^{2N})$ and apply a degree for S^1-gradient maps to the bifurcation equations associated to $\dot{y} = \lambda H'(y)$. 2π-periodic solutions $y(s)$ of this equation correspond to $2\pi/\lambda$-periodic solutions of (2.19). The degree theory allows a classical Rabinowitz type argument yielding a global continuum of solutions in $\mathbb{R} \times W^1(S^1, \mathbb{R}^{2N})$ that bifurcates from $(\lambda_0, 0)$ with $\lambda_0 = \omega$.

Observe that $\sigma = \sum_{k=1}^{\infty} \sigma_k$ where $\sigma_k = \sigma_k(\omega) = \sigma_1(k\omega)$ is the signature of $Q|F_k$, F_k the generalized eigenspace of $JH''(0)$ corresponding to $\pm \mathrm{i}k\omega$. Also observe (see [23, Theorem II.3.2]) that the local degree can be expressed in terms of the critical groups as

$$\deg(H', 0) = \sum_{q=0}^{\infty} (-1)^q \dim C^q(H, 0).$$

Thus the hypotheses of Theorem 2.11 are weaker than those of Theorem 2.12. Correspondingly, the conclusion is also weaker.

Since the above results require only local conditions on the Hamiltonian near a stationary point they immediately generalize to Hamiltonian systems on a symplectic manifold (W, Ω). The last result that we state in this section deals with periodic orbits near a manifold M of equilibria. This result can in general not be reduced to the special case $W = \mathbb{R}^{2N}$ with the standard symplectic structure $\Omega = \sum_{k=1}^{N} \mathrm{d}p_i \wedge \mathrm{d}q_i$ because the manifold M need not lie in a symplectic neighbourhood chart. We therefore state it in the general setting.

THEOREM 2.13. *Let (W, Ω) be a symplectic manifold and let $H : W \to \mathbb{R}$ be a smooth Hamiltonian. Suppose there exists a compact symplectic submanifold $M \subset H^{-1}(c) \subset W$ which is a Bott-nondegenerate manifold of minima of H. Then there exists a sequence of nonconstant periodic trajectories of the Hamiltonian flow associated to H which converge to M.*

The result is due to Ginzburg and Kerman [45]. It clearly applies to (2.19) where $W = \mathbb{R}^{2N}$ and Ω is as above. Compared with the Weinstein–Moser theorem where M is a point, Theorem 2.13 does not yield periodic orbits on all energy surfaces close to M, and neither does it yield a multiplicity result. We refer to [45] and the references therein for further results on periodic orbits of Hamiltonian flows near manifolds of equilibria.

2.3. *Fixed energy problem*

Let $H \in C^2(\mathbb{R}^{2N}, \mathbb{R})$ and suppose $D := \{z \in \mathbb{R}^{2N}\colon\ H(z) \leqslant 1\}$ is a compact subset of $\mathbb{R}^{2N}$ such that $H'(z) \neq 0$ for all $z \in S := H^{-1}(1)$. Then S is a compact hypersurface of class C^2 and we may assume without loss of generality that 0 is in the interior of D. We consider the autonomous Hamiltonian system (2.19). If $z(t_0) \in S$ then $z(t) \in S$ for all t because $H(z(t))$ is constant along solutions of (2.19) (see (2.20)). Since S is compact, $z(t)$ exists for all $t \in \mathbb{R}$.

We will be interested in the existence of *closed Hamiltonian orbits* on S, i.e., the sets $\mathrm{Orb}(z) := \{z(t)\colon\ t \in \mathbb{R}\}$, where $z = z(t)$ is a periodic solution of (2.19) with $z(t) \in S$. Here we use the notation $\mathrm{Orb}(z)$ for closed orbits in order to distinguish them from S^1-orbits $\mathcal{O}(z)$ defined in Section 2.1. If $\widetilde{H} \in C^2(\mathbb{R}^{2N}, \mathbb{R})$ is another Hamiltonian such that $S = \widetilde{H}^{-1}(c)$ for some c and $\widetilde{H}'(z) \neq 0$ on S, then $H'(z)$ and $\widetilde{H}'(z)$ are parallel and nowhere zero on S. It follows that two solutions z and $\tilde{z}$ of the corresponding Hamiltonian systems are equivalent up to reparameterization if $z(t_0) = \tilde{z}(\tilde{t}_0)$ for some $t_0, \tilde{t}_0 \in \mathbb{R}$. In particular, the orbits $\mathrm{Orb}(z)$ and $\mathrm{Orb}(\tilde{z})$ coincide (see [80] for a detailed argument). Consequently, closed orbits depend only on S and not on the particular choice of a Hamiltonian having the properties given above. It is also possible to define closed orbits without referring to any Hamiltonian: given a compact surface S of class C^2, one may look for periodic solutions of the system $\dot{z} = JN(z)$, where $N(z)$ is the unit outer normal to S at z.

Throughout this section we assume that S satisfies the following condition:

($\mathcal{S}$) S is a compact hypersurface of class C^2 in $\mathbb{R}^{2N}$, S bounds a starshaped neighbourhood of the origin and all $z \in S$ are transversal to S.

It follows from ($\mathcal{S}$) that for each $z \in \mathbb{R}^{2N} \setminus \{0\}$ there exists a unique $\alpha(z) > 0$ such that $z/\alpha(z) \in S$. Let $\alpha(0) := 0$ and

$$H(z) := \alpha(z)^4. \tag{2.22}$$

Clearly $\alpha(sz) = s^4\alpha(z)$ for all $s \geqslant 0$, hence H is positively homogeneous of degree 4. Moreover, $S = H^{-1}(1)$, $H \in C^2(\mathbb{R}^{2N}, \mathbb{R})$ and (by Euler's identities) $H'(z) \cdot z = 4H(z) \neq 0$ whenever $z \neq 0$. In particular, $H'(z) \neq 0$ on S.

Suppose z is a periodic solution of (2.19) with the Hamiltonian H given by (2.22). If $\mathrm{Orb}(z) \subset H^{-1}(\lambda)$ then $\hat{z}(t) := \lambda^{-1/4} z(t/\sqrt{\lambda})$ is a periodic solution of (2.19) on $H^{-1}(1)$. If z has minimal period 2π then $\hat{z}$ has minimal period $T := 2\pi\sqrt{\lambda}$. On the other hand, given a periodic solution z of (2.19) on $H^{-1}(1)$ with minimal period T then $\tilde{z}(t) := (T/2\pi)^{1/2} z(Tt/2\pi)$ is a periodic solution of (2.19) on $H^{-1}((T/2\pi)^2)$ with minimal period 2π. One easily checks that $\hat{\tilde{z}} = z$ and $\tilde{\hat{z}} = z$. Given two periodic solutions z_1, z_2 of (2.19) on $H^{-1}(1)$ with minimal period T and having the same orbit $\mathrm{Orb}(z_1) = \mathrm{Orb}(z_2) \subset H^{-1}(1)$ then there exists $\theta \in \mathbb{R}$ with $z_2(t) = z_1(t + \theta)$ for all t. The corresponding solutions $\tilde{z}_1, \tilde{z}_2$ with minimal period 2π then satisfy $\tilde{z}_2(t) = \tilde{z}_1(t + \tilde{\theta})$ for some $\tilde{\theta} \in \mathbb{R}$, hence they are not geometrically distinct in the sense of Remark 2.2. If z_1, z_2 have different orbits $\mathrm{Orb}(z_1), \mathrm{Orb}(z_2) \subset H^{-1}(1)$ then $\tilde{z}_1, \tilde{z}_2$ are geometrically distinct.

We summarize the above considerations in the following

THEOREM 2.14. *Let S be a hypersurface satisfying* ($\mathcal{S}$) *and let H be defined by* (2.22). *A periodic solution $z(t)$ of* (2.19) *on $H^{-1}(\lambda)$ yields a periodic solution $\lambda^{-1/4} z(t/\sqrt{\lambda})$ on $S = H^{-1}(1)$. Moreover, there is a one-to-one correspondence between closed orbits on S and geometrically distinct periodic solutions of* (2.19) *with minimal period 2π.*

We emphasize the importance of the assumption on the minimality of the period. If z is a solution of (2.19) with minimal period T and $\mathrm{Orb}(z) \subset S$, then $z(t)$ covers $\mathrm{Orb}(z)$ k times as t goes from 0 to kT. A corresponding solution $\tilde{z}_k(t) := (kT/2\pi)^{1/2} z(kTt/2\pi)$ has minimal period $2\pi/k$, hence $\tilde{z}_k$ and $\tilde{z}_m$ are geometrically distinct if $k \neq m$, yet $\mathrm{Orb}(\hat{\tilde{z}}_k) = \mathrm{Orb}(\hat{\tilde{z}}_m) = \mathrm{Orb}(z)$.

Now we state the first main result of this section. It is due to Rabinowitz [78] and, if S bounds a convex neighbourhood of the origin to Weinstein [99].

THEOREM 2.15. *Let S be a hypersurface satisfying* ($\mathcal{S}$). *Then S contains a closed Hamiltonian orbit.*

PROOF. By Theorem 2.14 it suffices to show that (2.19) with H given by (2.22) has a 2π-periodic solution $z \neq 0$. Let r be the largest and R the smallest number such that

$$r \leqslant |z| \leqslant R \quad \text{for all } z \in S. \tag{2.23}$$

Then

$$\frac{|z|^4}{R^4} \leqslant H(z) \leqslant \frac{|z|^4}{r^4}, \quad \text{for all } z \in \mathbb{R}^{2N}. \tag{2.24}$$

The functional

$$\Phi(z) = \frac{1}{2}\int_0^{2\pi} (-J\dot{z} \cdot z)\,\mathrm{d}t - \int_0^{2\pi} H(z)\,\mathrm{d}t = \frac{1}{2}\|z^+\|^2 - \frac{1}{2}\|z^-\|^2 - \psi(z)$$

is S^1-invariant. Moreover, $H'(z) \cdot z = 4H(z) > 0$, hence Φ satisfies (PS)* according to Proposition 2.5. Let $\widetilde{E}_n$, E_n be given by (2.5) and let $2d_n = 2N(1+n)$, $Y = E^+$ and $X = E_1^+ \oplus E^-$ (cf. (2.6)). Then $E^{S^1} = E^0 \subset E_n$ for all n, $\operatorname{codim}_{\widetilde{E}_n} X_n = 2(d_n - 2N)$ and $\dim Y_n = 2(d_n - N)$. Hence (i) and (ii) of Corollary 1.15 are satisfied, with $l = -N$ and $m = -2N$. By (2.24) and Lemma 2.4, $\Phi|_{Y \cap S_\rho} \geqslant \alpha$ for some $\alpha, \rho > 0$. Since

$$\Phi(z) \leqslant \frac{1}{2}\|z^+\|^2 - \frac{1}{2}\|z^-\|^2 - \frac{1}{R^4}\int_0^{2\pi} |z|^4\,\mathrm{d}t \tag{2.25}$$

and $\dim E_1^+ < \infty$, $\Phi(z) \to -\infty$ as $\|z\| \to \infty$, $z \in X$. Finally, $\Phi|_{E^0} \leqslant 0$ because $H \geqslant 0$. It follows that also (iii) and (iv) of Corollary 1.15 hold. Hence (2.19) has at least N geometrically distinct 2π-periodic solutions $z \neq 0$, and by Theorem 2.14, the hypersurface S carries a closed Hamiltonian orbit. □

It is not necessary to exploit the S^1-symmetry in order to show the existence of one 2π-periodic solution $z \neq 0$. However, the argument presented here will be needed below.

REMARK 2.16. Since the above proof gives no information on the minimal period of the N geometrically distinct 2π-periodic solutions, we do not know whether they correspond to distinct closed orbits on S. System (2.19) has in fact infinitely geometrically distinct 2π-periodic solutions. Indeed, we may replace $X = E_1^+ \oplus E^1$ by $X = E_r^+ \oplus E^-$ for any positive integer r and use Corollary 1.16 (see also Theorem 2.19). On the other hand, if $z = (p_1, \ldots, p_N, q_1, \ldots, q_N)$ and

$$S = \left\{ z \in \mathbb{R}^{2N} \colon \sum_{j=1}^{N} \frac{1}{2}\alpha_j \left(p_j^2 + q_j^2\right) = 1 \right\},$$

where $\alpha_1, \ldots, \alpha_N$ are rationally independent positive numbers, then it is easy to see that S has exactly N distinct closed orbits.

In an answer to a conjecture of Weinstein, Viterbo [96] has generalized Theorem 2.15 to all compact hypersurfaces admitting a so-called contact structure. Subsequently his proof has been simplified (and a more general result obtained) by Hofer and Zehnder [51]. Struwe [86], building upon the work of Hofer and Zehnder, proved that given an interval $[a,b]$ of regular values of H such that the hypersurfaces $S_c := H^{-1}(c) \subset \mathbb{R}^{2N}$, $c \in [a,b]$, are compact, the set $\{c \in [a,b]\colon\ S_c$ carries a closed Hamiltonian orbit$\}$ has full measure $b-a$. It has been shown by counterexamples of Ginzburg [43] and Herman [49] that in general a compact hypersurface may not have any closed Hamiltonian orbit (see also [44] and the references there).

In view of Remark 2.16 it is natural to ask whether each S satisfying $(\mathcal{S})$ must necessarily have N distinct closed Hamiltonian orbits. We shall show that this is indeed the case under an additional geometric condition.

Denote the tangent hyperplane to S at w by $T_w(S)$, suppose S satisfies $(\mathcal{S})$ and let ρ be the largest number such that

$$T_w(S) \cap \left\{z \in \mathbb{R}^{2N}\colon\ |z| < \rho\right\} = \emptyset \quad \text{for all } w \in S. \tag{2.26}$$

Then ρ is the minimum of the distances from $T_w(S)$ to the origin over all $w \in S$. It follows from $(\mathcal{S})$ that ρ is well defined; moreover, if r is as in (2.23) and S bounds a convex set, then $\rho = r$.

THEOREM 2.17. *Let S be a hypersurface satisfying $(\mathcal{S})$ and suppose $R^2 < 2\rho^2$, where R, ρ are as in* (2.23), (2.26). *Then S contains at least N distinct closed Hamiltonian orbits.*

PROOF. It follows from the proof of Theorem 2.15 that (2.19) (with H given by (2.22)) has at least N geometrically distinct 2π-periodic solutions. According to Theorem 2.14 it suffices to show that these solutions have minimal period 2π.

Recall from the proof of Theorem 2.15 that $l = -N$ and $m = -2N$, so invoking Corollary 1.15 we have

$$c_j = \sup_{i_{\mathcal{E}}(A) \geqslant j} \inf_{z \in A} \Phi(z), \qquad -2N+1 \leqslant j \leqslant -N.$$

Since $\operatorname{codim}_{\widetilde{E}_n} X_n = 2(d_n - 2N)$, it follows from (iv) of Proposition 1.13 that if $i_{\mathcal{E}}(A) \geqslant -2N+1$, then $A \cap (E^G \oplus X) = A \cap (E_1^+ \oplus E^0 \oplus E^-) \neq \emptyset$. Hence

$$c_j \leqslant \sup\left\{\Phi(z)\colon\ z \in E_1^+ \oplus E^0 \oplus E^-\right\}. \tag{2.27}$$

Let $z \in E_1^+ \oplus E^0 \oplus E^-$. Using (2.25), the fact that $\|z^+\| = \|z^+\|_2$ for $z^+ \in E_1^+$ and the Hölder inequality, we obtain

$$\begin{aligned}\Phi(z) &\leqslant \frac{1}{2}\|z^+\|^2 - \frac{1}{2}\|z^-\|^2 - \frac{1}{R^4}\|z\|_4^4 \leqslant \frac{1}{2}\|z^+\|_2^2 - \frac{1}{R^4}\|z\|_4^4 \\ &\leqslant \frac{1}{2}\|z\|_2^2 - \frac{1}{R^4}\|z\|_4^4 \leqslant \sqrt{\frac{\pi}{2}}\|z\|_4^2 - \frac{1}{R^4}\|z\|_4^4 \leqslant \frac{\pi R^4}{8}.\end{aligned}$$

This and (2.27) imply

$$c_j \leqslant \frac{\pi R^4}{8}. \tag{2.28}$$

By the definition of ρ and the homogeneity of H,

$$\rho\left|H'(z)\right| \leqslant z \cdot H'(z) = 4H(z) = 4 = 4H(z)^{3/4}$$

whenever $z \in S$. By the homogeneity again,

$$\rho\left|H'(z)\right| \leqslant 4H(z)^{3/4} \quad \text{for all } z \in \mathbb{R}^{2N}. \tag{2.29}$$

Let now $z = z(t)$ be a 2π-periodic solution of (2.19). Since $H(z(t))$ is constant,

$$\begin{aligned} \Phi(z) &= \frac{1}{2}\int_0^{2\pi} (-J\dot{z} \cdot z)\,\mathrm{d}t - \int_0^{2\pi} H(z)\,\mathrm{d}t \\ &= \int_0^{2\pi} \left(\frac{1}{2}H'(z) \cdot z - H(z)\right)\mathrm{d}t = \int_0^{2\pi} H(z)\,\mathrm{d}t = 2\pi H(z). \end{aligned} \tag{2.30}$$

Suppose z has minimal period $2\pi/m$ and write $z = \bar{z} + \tilde{z}$, $\bar{z} \in E^0$, $\tilde{z} \in E^+ \oplus E^-$. By Wirtinger's inequality,

$$\|\tilde{z}\|_2 \leqslant \frac{1}{m}\|\dot{\tilde{z}}\|_2,$$

and it follows using (2.30), (2.29) that

$$\begin{aligned} 2\pi H(z) = \Phi(z) &= \frac{1}{2}\int_0^{2\pi} (-J\dot{\tilde{z}} \cdot z)\,\mathrm{d}t - \int_0^{2\pi} H(z)\,\mathrm{d}t \\ &\leqslant \frac{1}{2}\|\dot{\tilde{z}}\|_2\|\tilde{z}\|_2 - 2\pi H(z) \leqslant \frac{1}{2m}\|\dot{z}\|_2^2 - 2\pi H(z) \\ &= \frac{1}{2m}\int_0^{2\pi} \left|H'(z)\right|^2 \mathrm{d}t - 2\pi H(z) \leqslant \frac{8}{m\rho^2}\int_0^{2\pi} H(z)^{3/2}\,\mathrm{d}t - 2\pi H(z) \\ &= 2\pi\left(\frac{8H(z)^{3/2}}{m\rho^2} - H(z)\right). \end{aligned}$$

Hence

$$H(z) \geqslant \frac{m^2\rho^4}{16}$$

and

$$\Phi(z) = 2\pi H(z) \geqslant \frac{\pi m^2\rho^4}{8}.$$

Since $R^2 < 2\rho^2$,

$$\Phi(z) > \frac{\pi m^2 R^4}{32}.$$

On the other hand, if the solution z corresponds to c_j, $-2N+1 \leqslant j \leqslant -N$, then $\Phi(z) \leqslant \pi R^4/8$ according to (2.28). Hence $m = 1$ and z has minimal period 2π. □

Theorem 2.17 is due to Ekeland and Lasry [35] (see also [33]) in the case of S bounding a compact strictly convex region and $R^2 < 2r^2$, and by Berestycki et al. [20] in the more general case considered here. We would also like to mention a result by Girardi and Matzeu [46] showing that if S satisfies $(\mathcal{S})$, then the condition $R^2 < 2\rho^2$ may be replaced by $R^2 < \sqrt{3}\rho r$ in Theorem 2.17.

It has been a longstanding conjecture (see, e.g., [33, p. 235]) that if S bounds a compact strictly convex set, then the minimal number of distinct closed Hamiltonian orbits such S must carry is N. Ekeland and Lassoued [34] and Szulkin [89] have shown that S carries at least 2 such orbits if its Gaussian curvature is positive everywhere. In a recent work Liu, Long and Zhu [67] have shown that if in addition S is symmetric about the origin, the number of such orbits is at least N, and for general (possibly nonsymmetric) S as above, Long and Zhu [71] have shown the existence of at least $[\frac{N}{2}]+1$ closed orbits ($[a]$ denotes the integer part of a). They also make a new conjecture that $[\frac{N}{2}]+1$ (and not N) is the lower bound for the number of closed Hamiltonian orbits. See also Long's book [69] for a detailed discussion.

In the case $N = 2$, Hofer, Wysocki and Zehnder [50] proved that if S bounds a strictly convex set then there are either two or infinitely many closed Hamiltonian orbits on S. In [32] Ekeland has shown that a generic S bounding a compact convex set and having positive Gaussian curvature carries infinitely many closed Hamiltonian orbits. This result has been partially generalized by Viterbo [97] to hypersurfaces satisfying a condition similar to $(\mathcal{S})$. The question of the existence of infinitely many closed orbits is extensively discussed in [33,69] where many additional references may be found.

2.4. *Superlinear systems*

Throughout this section we assume that H satisfies (H_1)–(H_3), $H(z,t) = \frac{1}{2}Az\cdot z + G(z,t)$, where A is a symmetric $2N \times 2N$ matrix, $G_z(z,t) = o(z)$ uniformly in t as $z \to 0$, and there exist $\mu > \max\{2, s-1\}$ and $R > 0$ such that

$$0 < \mu G(z,t) \leqslant z \cdot G_z(z,t) \quad \text{for all } |z| \geqslant R. \tag{2.31}$$

Recall from (2.7) that the last condition implies G (and hence H) is superquadratic and H_z superlinear.

THEOREM 2.18. *Suppose H satisfies the hypotheses given above and $\sigma(JA) \cap \mathrm{i}\mathbb{Z} = \emptyset$. Then the system* (2.1) *has a 2π-periodic solution $z \neq 0$.*

PROOF. Let

$$\Phi(z) = \frac{1}{2}\int_0^{2\pi} (-J\dot{z} - Az) \cdot z \,\mathrm{d}t - \int_0^{2\pi} G(z,t)\,\mathrm{d}t$$

and let E_n be given by (2.5). Denote the linear operator corresponding to the quadratic part of Φ by L (cf. (2.11)). Since $\sigma(JA) \cap \mathrm{i}\mathbb{Z} = \emptyset$, L is invertible, $E = E^+(L) \oplus E^-(L)$ and $E_n = E_n^+(L) \oplus E_n^-(L)$ (see the discussion and notation preceding Corollary 2.7). It follows using (2.18) and Lemma 2.4 that there exist $\alpha, \rho > 0$ such that

$$\Phi \leqslant 0 \quad \text{on } E_n^-(L) \cap \overline{B}_\rho, \qquad \Phi \leqslant -\alpha \quad \text{on } E_n^-(L) \cap S_\rho \tag{2.32}$$

and

$$\Phi \geqslant 0 \quad \text{on } E_n^+(L) \cap \overline{B}_\rho, \qquad \Phi \geqslant \alpha \quad \text{on } E_n^+(L) \cap S_\rho \tag{2.33}$$

for all n. If Φ has a critical point $z \in \Phi^{-\alpha}$, then $z \neq 0$, so z is a solution of (2.1) we were looking for. Suppose no such z exists. We shall complete the proof by showing that in this case $\Phi'(z) = 0$ for some z with $\Phi(z) \geqslant \alpha$.

We claim that $\Phi_{mn} := \Phi|_{E_m^+(L)\oplus E_n^-(L)}$ has no critical point $z \in \Phi_{mn}^{-\alpha}$ whenever $m, n \geqslant n_0$ and n_0 is large enough. Indeed, otherwise there is a sequence $\{z_j\} \subset \Phi^{-\alpha}$ such that $z_j \in E_{m_j}^+(L) \oplus E_{n_j}^-(L)$, $m_j, n_j \to \infty$ and $\Phi'_{m_j n_j}(z_j) = 0$. By Corollary 2.7, $z_j \to z$ after passing to a subsequence, so $\Phi(z) \leqslant -\alpha$ and $\Phi'(z) = 0$, a contradiction. Hence we may choose n_0 so that $\Phi_{n_0 n}$ has no critical point in $\Phi^{-\alpha}$ for any $n \geqslant n_0$. Let $z = w^+ + w^- \in E_{n_0+1}^+(L) \oplus E_n^-(L)$. Then

$$\Phi(z) = \frac{1}{2}\langle Lw^+, w^+\rangle + \frac{1}{2}\langle Lw^-, w^-\rangle - \int_0^{2\pi} G(z,t)\,\mathrm{d}t, \tag{2.34}$$

and since $G(z,t) \geqslant a_1|z|^\mu - a_2$ according to (2.7), it follows that $\Phi(z) \leqslant 0$ whenever $|z| \geqslant R$. Moreover, since n_0 is fixed, R does not depend on n. If $n \geqslant n_0 + 1$, then by Corollary 1.3 (with E_n corresponding to $\mathbb{R}^m$, $F_0 = E_n^-(L)$, $F = E_{n_0}^+(L) \oplus E_n^-(L)$ and $\widetilde{F} = E_{n_0+1}^+ \oplus E_n^-(L)$) there exists $z_n \in E_n$ such that $\Phi_n'(z_n) = 0$ and $\alpha \leqslant \Phi(z_n) \leqslant \sup_{\bar{B}_{R+1}} \Phi$. Applying Proposition 2.5 to the sequence $\{z_n\}$ we obtain a critical point z with $\Phi(z) \geqslant \alpha$. □

Next we prove that the autonomous system

$$\dot{z} = JH'(z) = J\big(Az + G'(z)\big) \tag{2.35}$$

with superquadratic Hamiltonian has infinitely many geometrically distinct T-periodic solutions for any $T > 0$. Since there is a one-to-one correspondence between T-periodic solutions for the system $\dot{z} = JH'(z)$ and 2π-periodic solutions for $\dot{z} = \lambda JH'(z)$, where $\lambda = T/2\pi$ (this can be easily seen by substituting $\tau = t/\lambda$), we may assume without loss of generality that $T = 2\pi$.

THEOREM 2.19. *Suppose* $H(z) = \frac{1}{2}Az \cdot z + G(z)$ *satisfies* (H_1), (H_3), $H(z) \geqslant 0$ *for all* $z \in \mathbb{R}^{2N}$, G *satisfies* (2.31) *and* $G'(z) \to 0$ *as* $z \to 0$. *Then the system* (2.35) *has a sequence* $\{z_j\}$ *of nonconstant* 2π*-periodic solutions such that* $\|z_j\|_\infty \to \infty$.

If one can show that for each $T > 0$ the system (2.35) has a nonconstant T-periodic solution $z^T \neq 0$, then the number of geometrically distinct nonconstant 2π-periodic solutions is in fact infinite. Indeed, let $z_k = z^{2\pi/k}$, then z_k and z_l may coincide for some $k \neq l$, yet the sequence $\{z_k\}$ will contain infinitely many distinct elements. However, the result stated above shows much more: the solutions z_j have amplitude which goes to infinity with j.

PROOF OF THEOREM 2.19. We verify the hypotheses of Corollary 1.15. By Proposition 2.5, Φ satisfies (PS)*, and obviously, $E^{S^1} = E^0 \subset E_n$. Let $2d_n = 2N(1+n)$, $X = (E_r^+(L) \oplus E^0(L) \oplus E^-(L)) \cap \widetilde{E}$, where r is a positive integer, and $Y = E^+(L) \cap \widetilde{E}$ (we use the notation of the preceding proof). Employing (2.14) and recalling that $M^+(T_k(A)) = 2N$ for large k, we have for n, r large enough ($n \geqslant r$),

$$\begin{aligned}\dim Y_n &= \sum_{k=1}^{n} M^+\bigl(T_k(A)\bigr) = 2nN + \sum_{k=1}^{n}\bigl(M^+\bigl(T_k(A)\bigr) - 2N\bigr)\\ &= 2d_n + i^+(A) - M^-(A) - N =: 2(d_n + l)\end{aligned} \tag{2.36}$$

and

$$\begin{aligned}\operatorname{codim}_{\widetilde{E}_n} X_n &= \sum_{k=r+1}^{n} M^+\bigl(T_k(A)\bigr) = 2Nn - 2Nr = 2d_n - 2N(r+1)\\ &=: 2(d_n + m).\end{aligned}$$

It follows from (2.33) that $\Phi|_{Y \cap S_\rho} \geqslant \alpha$ and from (2.34) with $z = w^+ + w^0 + w^- \in E_r^+(L) \oplus E^0(L) \oplus E^-(L)$ that $\Phi(z) \to -\infty$ whenever $\|z\| \to \infty$, $z \in E^0 \oplus X$. Hence $\Phi|_{E^0 \oplus X} \leqslant \beta$ for some β. Moreover, $\Phi|_{E^0} \leqslant 0$ because $H \geqslant 0$. We have verified the hypotheses of Corollary 1.15 for all r large enough. Since $m \to -\infty$ as $r \to \infty$, we conclude from Corollary 1.16 that (2.35) has a sequence $\{z_j\}$ of nonconstant 2π-periodic solutions such that $\Phi(z_j) \to \infty$.

It remains to show that $\|z_j\|_\infty \to \infty$. By ($H_3$),

$$\begin{aligned}c_j = \Phi(z_j) = \Phi(z_j) - \frac{1}{2}\bigl\langle \Phi'(z_j), z_j \bigr\rangle &= \int_0^{2\pi} \left(\frac{1}{2} z_j \cdot H'(z_j) - H(z_j)\right) \mathrm{d}t\\ &\leqslant \tilde{c} \int_0^{2\pi} \bigl(1 + |z|^s\bigr)\, \mathrm{d}t \leqslant 2\pi \tilde{c}\bigl(1 + \|z\|_\infty^s\bigr),\end{aligned} \tag{2.37}$$

and the conclusion follows because $c_j \to \infty$. □

The assumption $H \geqslant 0$ is not necessary. Below we show how the proof of Theorem 2.19 can be modified in order to remove it.

COROLLARY 2.20. *The conclusion of Theorem* 2.19 *remains valid without the condition* $H \geqslant 0$.

PROOF. Since G satisfies (2.31), so does H according to Remark 2.6; hence H is bounded below. Let

$$\tilde{\alpha} := \max_{z \in E^0} \Phi(z)$$

and let $r_0 < r < n$ be positive integers. Let $X = (E_r^+(L) \oplus E^0(L) \oplus E^-(L)) \cap \widetilde{E}$ as before and $Y = E_{r_0}(L)^\perp \cap E^+(L) \cap \widetilde{E}$. Then $\dim Y_n = 2(d_n + l)$ and we still have $m < l$ if $r - r_0$ is large enough. Define $S = Y \cap \{z \in E\colon \|z\|_s = 1\}$ and note that S is radially homeomorphic to the unit sphere in Y. We claim that $\Phi|_{Y \cap S} \geqslant \alpha > \tilde{\alpha}$ for all large r_0. Assuming this for the moment, we find r_0 such that the condition above is satisfied. Since $\Phi|_{E^0} \leqslant \tilde{\alpha}$, we can easily see by modifying the argument of Theorem 1.14 that the conclusion of Corollary 1.15 (and hence also of Corollary 1.16) holds.

It remains to prove the claim. Arguing by contradiction, we find $r_j \to \infty$ and $z_j \in E_+^{r_j}(L) \cap \widetilde{E}$ such that $\|z_j\|_s = 1$ and $\Phi(z_j) \leqslant \alpha$. Hence

$$\begin{aligned} \alpha \geqslant \Phi(z_j) &= \frac{1}{2}\langle L z_j, z_j \rangle - \int_0^{2\pi} G(z_j)\,\mathrm{d}t \\ &\geqslant \varepsilon \|z_j\|^2 - \tilde{c}\big(\|z_j\|_s^s + 1\big) = \varepsilon \|z_j\|^2 - 2\tilde{c}, \end{aligned}$$

so $\{z_j\}$ is bounded in E. Passing to a subsequence, $z_j \rightharpoonup z$ in E and $z_j \to z$ in $L^s(S^1, \mathbb{R}^{2N})$. It follows that $\|z\|_s = 1$; in particular, $z \neq 0$. On the other hand, $z_j \in E_{r_j}(L)^\perp \cap E^+(L)$ implies $z_j \rightharpoonup 0$, a contradiction. A somewhat different argument will be given in the proof of Theorem 2.25. □

The first result on the existence of a nontrivial periodic solution of (2.1) is due to Rabinowitz [78]. Theorem 2.18 may be found in [62]. The result contained there is in fact more general: the case $\sigma(JA) \cap \mathrm{i}\mathbb{Z} \neq \emptyset$ is allowed if G has constant sign for small $|z|$. Also some Hamiltonians not satisfying the requirement $\mu > s - 1$ are allowed; for this purpose a truncation argument indicated in Remark 2.6 is employed. Other extensions of Theorem 2.18 are due to Felmer [38] and Long and Xu [70]. Corollary 2.19 is due to Rabinowitz [79] (in [79] no growth restriction (H_3) is needed; again, this is achieved by truncation).

An interesting question concerning the autonomous system (2.35) is whether one can find solutions with prescribed minimal period. Results in this direction, mainly for convex Hamiltonians, can be found in Ekeland's book [33, Section IV.5] and in Long's book [69, Chapter 13].

2.5. *Asymptotically linear systems*

In this section we assume that in addition to (H_1)–(H_3) H satisfies the conditions

$$H(z,t) = \frac{1}{2}A_0 z \cdot z + G_0(z,t),$$
$$\text{where } (G_0)_z(z,t) = o(z) \text{ uniformly in } t \text{ as } z \to 0 \tag{2.38}$$

and

$$H(z,t) = \frac{1}{2}A_\infty z \cdot z + G_\infty(z,t),$$
$$\text{where } (G_\infty)_z(z,t) = o(|z|) \text{ uniformly in } t \text{ as } |z| \to \infty. \tag{2.39}$$

Here A_0, A_∞ are $2N \times 2N$ constant matrices. We assume for simplicity that the system (2.1) is *nonresonant* at the origin and at infinity, that is, $\sigma(JA_0) \cap \mathrm{i}\mathbb{Z} = \sigma(JA_\infty) \cap \mathrm{i}\mathbb{Z} = \emptyset$. This terminology is justified by the fact that the systems $\dot{z} = JA_0 z$ and $\dot{z} = JA_\infty z$ have no other 2π-periodic solutions than $z = 0$.

As a first result in this section we give a sufficient condition for the existence of a nontrivial 2π-periodic solution of (2.1).

THEOREM 2.21. *Suppose H satisfies* (H_1)–(H_3), (2.38), (2.39) *and* $\sigma(JA_0) \cap \mathrm{i}\mathbb{Z} = \sigma(JA_\infty) \cap \mathrm{i}\mathbb{Z} = \emptyset$. *If $i^-(A_0) \neq i^-(A_\infty)$, then the system* (2.1) *has a 2π-periodic solution $z \neq 0$.*

PROOF. Suppose $i^-(A_0) < i^-(A_\infty)$. The same argument applied to $-\Phi$ will give the conclusion for $i^-(A_0) > i^-(A_\infty)$. Let L_0 and L_∞ be given by (2.11), with respectively $A = A_0$ and $A = A_\infty$. As in the proof of Theorem 2.18, we see that (2.32) and (2.33) are satisfied, with L_0 replacing L. Suppose Φ has no other critical points than 0. Then $\Phi_n = \Phi|_{E_n}$ has no critical points with $|\Phi_n(z)| \geqslant \alpha$ provided $n \geqslant n_0$ and n_0 is large enough. For otherwise we find $z_j \in E_{n_j}$ such that $n_j \to \infty$ and $\Phi'_{n_j}(z_j) = 0$. According to Proposition 2.5, $z_j \to z$ after passing to a subsequence, hence z is a critical point and $|\Phi(z)| \geqslant \alpha$ which is impossible. Fix $n \geqslant n_0$. Now we invoke Theorem 1.2. It follows from Lemma 2.4 that $\Phi'_n(z) = P_n L_\infty|_{E_n}(z) + o(\|z\|)$ as $\|z\| \to \infty$. If n is large enough, then (with $2d_n = 2N(1+n)$ as in the proof of Theorem 2.19)

$$\dim E_n^-(L_0) = M^+(A_0) + \sum_{k=1}^{n} M^-\bigl(T_k(A_0)\bigr) = 2d_n + i^-(A_0) - N,$$

and similarly,

$$M^-(P_n L_\infty|_{E_n}) = \dim E_n^-(L_\infty) = 2d_n + i^-(A_\infty) - N.$$

Since $i^-(A_0) < i^-(A_\infty)$, Theorem 1.2 with $F_0 = E_n^-(L_0)$ and $B = P_n L_\infty|_{E_n}$ implies that Φ_n has a critical point z such that $|\Phi_n(z)| \geqslant \alpha$. This contradiction completes the proof. □

As in the preceding section, we now turn our attention to the autonomous case.

THEOREM 2.22. *Suppose* $H = H(z)$ *satisfies* (H_1), (H_3), (2.38), (2.39) *and* $\sigma(JA_0) \cap \mathrm{i}\mathbb{Z} = \sigma(JA_\infty) \cap \mathrm{i}\mathbb{Z} = \emptyset$. *If* $H \geqslant 0$ *and* $i^-(A_0) < i^-(A_\infty)$, *then the system* (2.1) *has at least* $\frac{1}{2}(i^-(A_\infty) - i^-(A_0))$ *geometrically distinct nonconstant* 2π*-periodic solutions.*

PROOF. We verify the assumptions of Corollary 1.15. Φ satisfies (PS)*, $E^{S^1} = E_0 \subset E_n$ and $\Phi|_{E^0} \leqslant 0$. Let $X = E^-(L_\infty) \cap \widetilde{E}$ and $Y = E^+(L_0) \cap \widetilde{E}$. Since $H \geqslant 0$ and $\sigma(JA_0) \cap \mathrm{i}\mathbb{Z} = \sigma(JA_\infty) \cap \mathrm{i}\mathbb{Z} = \emptyset$, A_0 and A_∞ are positive definite and $M^+(A_0) = M^+(A_\infty) = 2N$. Therefore (cf. (2.36))

$$\dim Y_n = 2d_n + i^+(A_0) - N =: 2(d_n + l)$$

and

$$\operatorname{codim}_{\widetilde{E}_n} X_n = \sum_{k=1}^{n} M^+\big(T_k(A_\infty)\big) = 2d_n + i^+(A_\infty) - N =: 2(d_n + m).$$

Since $i^+(A_0) = -i^-(A_0)$ and $i^+(A_\infty) = -i^-(A_\infty)$, $\frac{1}{2}(i^-(A_\infty) - i^-(A_0)) = l - m$. Finally, $\Phi|_{Y \cap S_\rho} \geqslant \alpha > 0$ according to (2.33) and since $E^-(L_\infty) = E^0 \oplus X$, it is easy to see that $\Phi|_{E^0 \oplus X} \leqslant \beta$ for an appropriate $\beta > \alpha$. Corollary 1.15 yields at least $l - m$ geometrically distinct 2π-periodic solutions with $\Phi \geqslant \alpha$. Since $\Phi > 0$ these solutions cannot be constant. □

In the next theorem we drop the hypothesis $H \geqslant 0$ and require H to be even in z.

THEOREM 2.23. *Suppose* $H = H(z)$ *satisfies* (H_1), (H_3), (2.38), (2.39) *and* $\sigma(JA_0) \cap \mathrm{i}\mathbb{Z} = \sigma(JA_\infty) \cap \mathrm{i}\mathbb{Z} = \emptyset$. *If* $H(-z) = H(z)$ *for all* $z \in \mathbb{R}^{2N}$, *then the system* (2.1) *has at least* $\frac{1}{2}|i^-(A_\infty) - i^-(A_0)|$ *geometrically distinct* 2π*-periodic solutions* $z \neq 0$.

PROOF. We only sketch the argument. Since Φ is an even functional, in view of Remark 1.17 we may apply Theorem 1.14 if $i^-(A_0) > i^-(A_\infty)$ and Corollary 1.15 if $i^-(A_0) < i^-(A_\infty)$ in order to get $|i^-(A_\infty) - i^-(A_0)|$ pairs of nonzero 2π-periodic solutions (here we use genus instead of index and disregard the S^1-symmetry). For a critical value c the set K_c consists of critical S^1-orbits, some of them may correspond to constant solutions, the other ones are homeomorphic to S^1. So if K_c contains a nonconstant solution, then $\gamma(K_c) \geqslant 2$. On the other hand, if $\gamma(K_c) > 2$, it is easy to see that K_c contains infinitely many geometrically distinct critical orbits. Hence the number of nonzero geometrically distinct critical orbits is at least $\frac{1}{2}|i^-(A_\infty) - i^-(A_0)|$. □

REMARK 2.24. (a) The argument of Theorem 2.23 does not guarantee the existence of nonconstant solutions.

(b) If $H = H(z, t)$ is even in z and satisfies the other assumptions of the above theorem, then the same argument asserts the existence of at least $|i^-(A_\infty) - i^-(A_0)|$ pairs of nontrivial 2π-periodic solutions (see [4] or [92]). Hamiltonian systems with spatial symmetries will be further discussed in Section 2.6 below.

(c) As we have mentioned in Section 2.4, there is a one-to-one correspondence between 2π-periodic solutions of the system $\dot{z} = JH'(z)$ and T-periodic solutions of $\dot{z} = \lambda JH'(z)$, where $\lambda = 2\pi/T$. Hence, in view of the results of that section, there exist nonconstant T-periodic solutions of any period T whenever H is autonomous and superquadratic. Here the situation is different. If $H \geqslant 0$, then $i^-(\lambda A_0) = i^-(\lambda A_\infty) = N$ for all small $\lambda > 0$. Thus Theorem 2.22 gives no nonconstant solutions of small period T. This is not surprising, for if H' is Lipschitz continuous with Lipschitz constant M, then each nonconstant periodic solution must have period $T \geqslant 2\pi/M$ according to Theorem 4.3 in [20]. On the other hand, it is easy to give examples where $i^-(\lambda A_\infty) - i^-(\lambda A_0) \to \infty$ as $\lambda \to \infty$. So the number of geometrically distinct T-periodic solutions will go to infinity with T. However, there may be no solutions of arbitrarily large *minimal* period (see [92], Remark 6.3 for an example).

There is an extensive literature concerning the existence of one or two nontrivial solutions of (2.1) in the framework of Theorem 2.21. Usually the argument is based on an infinite-dimensional Morse theory and it is possible to weaken the nonresonance conditions at zero and infinity. Also, it is not necessary to have constant matrices A_0 and A_∞. The first results related to Theorem 2.21 may be found in Amann and Zehnder [3,4]. For other results and more references, see, e.g., Abbondandolo [1], Chang [23], Guo [48], Izydorek [53], Kryszewski and Szulkin [57], Li and Liu [61], Szulkin and Zou [93]. Theorem 2.22 is due to Amann and Zehnder [4] and Benci [18]. It has been extended by Degiovanni and Olian Fannio [28], see also [92]. While the proof in [28] uses a cohomological index theory (like the one in [36]) and a variant of Benci's pseudoindex [18], the argument in [92] is based on a relative limit index (which is a generalization of the limit index $i_{\mathcal{E}}$). Another extension, using Conley index theory, has been carried out by Izydorek [54]. Some other aspects of the problem (an estimate of the number of T-periodic solutions in terms of the so-called twist number) are discussed in Abbondandolo [1]. However, in all results related to Theorem 2.22 we know of, the assumptions on H are rather restrictive. This is briefly discussed in Remark 6.4 of [92]. Theorem 2.23 is due to Benci [18]. For results about solutions of the autonomous equation (2.1) with prescribed minimal period we refer again to Long's book [69, Section 13.3].

2.6. *Spatially symmetric Hamiltonian systems*

In this section we consider the non-autonomous Hamiltonian system (2.1) when H is invariant with respect to certain group representations in $\mathbb{R}^{2N}$. More precisely, we consider two different kinds of symmetries:

- A compact Lie group G acts on $\mathbb{R}^{2N}$ via an orthogonal and symplectic representation; the standard example is the antipodal action of $\mathbb{Z}/2$ (i.e., H is even in z).
- The infinite group $\mathbb{Z}^k$ acts on $\mathbb{R}^{2N}$ via space translations; the standard example is $\mathbb{Z}^{2N}$ leading to a Hamiltonian system on the torus $T^{2N} := \mathbb{R}^{2N}/\mathbb{Z}^{2N}$. Another example is $\mathbb{Z}^N$ acting via translation of the q-variables which leads to a Hamiltonian system on the cotangent space T^*T^N of the N-dimensional torus.

In the first case we may think of G as a closed subgroup of $\mathrm{O}(2N) \cap \mathrm{Sp}(2N)$. We shall always assume that $H : \mathbb{R}^{2N} \times \mathbb{R} \to \mathbb{R}$ satisfies the hypotheses (H$_1$)–(H$_3$) from Section 2.1.

First we treat the compact group case and require:

(S) The compact group G acts on $\mathbb{R}^{2N}$ via an orthogonal and symplectic representation T such that the action is fixed point free on $\mathbb{R}^{2N} \setminus \{0\}$ (i.e., $(\mathbb{R}^{2N})^G = \{0\}$). H is invariant with respect to T: $H(T_g z, t) = H(z, t)$ for all $g \in G$, $z \in \mathbb{R}^{2N}$, $t \in \mathbb{R}$.

By an orthogonal and symplectic representation we mean that the matrix of T_g is in $\mathrm{O}(2N) \cap \mathrm{Sp}(2N)$ for all $g \in G$. Clearly, if (S) holds and $z(t)$ is a periodic solution of (2.1) then so is $T_g z(t)$ for every $g \in G$. Thus one has to count G-orbits of periodic solutions and not just periodic solutions.

THEOREM 2.25. *Suppose* (H$_1$)–(H$_3$) *and* (S) *hold for G of prime order. If H is superquadratic in the sense of* (2.31) *then the system* (2.1) *has a sequence of 2π-periodic solutions z_j such that $\|z_j\|_\infty \to \infty$.*

If H is invariant with respect to an orthogonal symplectic representation of a more general compact Lie group G then one can apply Theorem 2.25 provided there exists a subgroup $G_1 \subset G$ of prime order having 0 as the only fixed point. This may or may not be the case. It is always the case for $G = S^1$ or more generally, for $G = (S^1)^k$ a torus, acting without nontrivial fixed points. A general existence result in this direction works for *admissible* group actions; see Remark 1.10.

PROOF. We want to apply Corollary 1.16 to the usual action functional

$$\Phi(z) = \frac{1}{2} \int_0^{2\pi} (-J\dot{z} \cdot z)\,\mathrm{d}t - \int_0^{2\pi} H(z, t)\,\mathrm{d}t.$$

Although the proof of Corollary 2.20 could be used here with minor changes, we provide a slightly different argument as we have mentioned earlier. By Proposition 2.5 the (PS)*-condition holds.

Recall the spaces $E_n \subset E$ from (2.5). We choose $k_0 \in \mathbb{N}$ and set $Y := E_{k_0}^{\perp} \cap E^+$. Now we claim that for k_0 large enough, there exists $\rho, \alpha > 0$ so that Φ satisfies condition (iii) from Corollary 1.15, that is,

$$\Phi(z) \geqslant \alpha \quad \text{for } z \in Y \text{ with } \|z\| = \rho. \tag{2.40}$$

In order to see this we first observe that

$$\|z\| \geqslant \sqrt{k_0}\|z\|_2 \quad \text{holds for } z \in Y,$$

and that there exists $c_1 > 0$ with

$$\bigl|H(z, t)\bigr| \leqslant c_1\bigl(|z|^s + 1\bigr)$$

by (H_3). Using the continuous embedding $E \hookrightarrow L^{2s-2}$ we obtain

$$\|z\|_s^s \leqslant \|z\|_2 \cdot \|z\|_{2s-2}^{s-1} \leqslant \frac{c_2}{\sqrt{k_0}}\|z\| \cdot \|z\|^{s-1} \quad \text{for } z \in Y.$$

This implies

$$\Phi(z) \geqslant \frac{1}{2}\|z\|^2 - c_1\big(\|z\|_s^s - 2\pi\big) \geqslant \frac{1}{2}\|z\|^2 - \frac{c_1c_2}{\sqrt{k_0}}\|z\|^s - 2c_1\pi$$

for every $z \in Y$. Setting $\rho = \big(\frac{\sqrt{k_0}}{sc_1c_2}\big)^{1/(s-2)}$ we therefore have for $z \in Y$ with $\|z\| = \rho$:

$$\Phi(z) \geqslant \left(\frac{1}{2} - \frac{1}{s}\right)\left(\frac{\sqrt{k_0}}{sc_1c_2}\right)^{2/(s-2)} - 2c_1\pi > 0$$

provided k_0 is large. Thus we may fix $k_0 \in \mathbb{N}$ so that (2.40) holds.

Next we define $X_k := E^- + E_k$ for $k \in \mathbb{N}$. Then $\sup \Phi(X_k) < \infty$ because $\Phi(z) \to -\infty$ for $z \in X_k$ with $\|z\| \to \infty$ as we have seen earlier. In order to apply Corollary 1.16 with $X = X_k$ it remains to check the dimension condition from Corollary 1.16. Recall that $d_G = 1$ for $G = \mathbb{Z}/p$. Setting $d_n = 2N(1+n)$ we have

$$\dim Y_n = 2N(n - k_0) = d_n + l$$

with $l = -2N(k_0 + 1)$ and

$$\operatorname{codim}_{E_n}(X_k \cap E_n) = 2N(n-k) = d_n + m(k)$$

with $m(k) = -2Nk - 2N$ ($E_n = \widetilde{E}_n$ here because $E^G = \{0\}$). Clearly $l - m(k) \to \infty$ as $k \to \infty$. Now the theorem follows from Corollary 1.16 and (2.37). □

Comparing Theorem 2.25 with Theorems 2.18 and 2.19 we see that from the variational point of view the spatial symmetry condition (S) has the same effect as the S^1-symmetry of the autonomous problem. This is also true for asymptotically linear Hamiltonian systems—as we already observed in Remark 2.24(b). We state one multiplicity result in this setting.

THEOREM 2.26. *Suppose H satisfies* (H_1)–(H_3) *and is asymptotically quadratic in the sense of* (2.38) *and* (2.39). *Suppose moreover that $\sigma(JA_0) \cap \mathrm{i}\mathbb{Z} = \sigma(JA_\infty) \cap \mathrm{i}\mathbb{Z} = \emptyset$, and let $i^-(A_0)$ and $i^-(A_\infty)$ be the Morse indices defined in* (2.13). *If* (S) *holds for G of prime order or for $G \cong S^1$ then the system* (2.1) *has at least $\frac{1}{d_G}|i^-(A_\infty) - i^-(A_0)|$ G-orbits of nontrivial 2π-periodic solutions.*

PROOF. The result follows from Theorem 1.14 or Corollary 1.15; cf. also the proof of Theorem 2.23. □

Theorem 2.26 is also true for $G = (\mathbb{Z}/p)^k$ a p-torus and $d_G = 1$, or $G = (S^1)^k$ a torus and $d_G = 2$, or if G acts freely on $\mathbb{R}^{2N} \setminus \{0\}$ and $d_G = 1 + \dim(G)$. G acting freely means that $T_g z = z$ for some $z \neq 0$ implies that g is the identity. Such actions exist only for a very restricted class of Lie groups. It does not help much if a subgroup G_1 of G acts freely (or without nontrivial fixed points if G_1 is a torus or p-torus), because a G-orbit consists of several G_1-orbits. So the multiple G_1-orbits of periodic solutions may correspond to just one G-orbit of periodic solutions.

There are various extensions of Theorem 2.26 when the linearized equations at 0 or at ∞ have nontrivial 2π-periodic solutions, mostly for even Hamiltonians; see for instance [54] and the references therein.

Now we consider the spatially periodic case. The classical result is due to Conley and Zehnder [24] and deals with the case of $\mathbb{Z}^{2N}$, that is H is periodic in all variables. Then periodic solutions appear in $\mathbb{Z}^{2N}$-orbits.

THEOREM 2.27. *Suppose $H \in C^1(\mathbb{R}^{2N} \times \mathbb{R})$ is 2π-periodic in all variables. Then* (2.1) *has at least $2N + 1$ distinct $\mathbb{Z}^{2N}$-orbits of 2π-periodic solutions.*

PROOF. We consider the decomposition $E = E^+ \oplus E^0 \oplus E^-$ from Section 2.1 and observe that

$$\Phi(z + 2\pi k) = \Phi(z) \quad \text{for every } z \in E,\ k \in \mathbb{Z}^{2N}.$$

Setting $M = T^{2N} = E^0/2\pi\mathbb{Z}^{2N}$ we obtain an induced C^1-functional

$$\Psi : \left(E^+ \oplus E^-\right) \times M \to \mathbb{R}, \quad \Psi\left(z^+, z^-, z^0 + 2\pi\mathbb{Z}^{2N}\right) = \Phi\left(z^+ + z^- + z^0\right).$$

Critical points of Ψ correspond to $\mathbb{Z}^{2N}$-orbits of 2π-periodic solutions of (2.1). The conclusion follows from Theorem 1.18 and the fact that $\operatorname{cupl}(T^{2N}) = 2N + 1$. More precisely, we let $W = E^-$ and $Y = E^+$. Since H is bounded, β as in (ii) of Theorem 1.18 exists and taking ρ large enough, we also find $\alpha < \beta$ and γ. Finally, it is easy to see that $(PS)^*$-sequences are bounded (cf. Proposition 2.5), consequently, the $(PS)^*$-condition is satisfied. □

Using Theorem 1.18 one can also treat more general periodic symmetries, for instance when $H(p, q, t)$ is invariant under $\mathbb{Z}^N$ acting on the q-variables by translations. Then one needs some condition on the behavior of $H(p, q, t)$ as $|p| \to \infty$. Results in this direction have been obtained by a number of authors, see [23,37,42,40,66,90].

If the periodic solutions are non-degenerate then Conley and Zehnder [24] used Morse theoretic arguments to prove:

THEOREM 2.28. *Suppose $H \in C^2(\mathbb{R}^{2N} \times \mathbb{R})$ is 2π-periodic in all variables and all 2π-periodic solutions of* (2.1) *are non-degenerate. Then* (2.1) *has at least 2^{2N} distinct $\mathbb{Z}^{2N}$-orbits of 2π-periodic solutions.*

The Morse theoretic arguments involve in particular the Conley–Zehnder index; see Section 2.1, in particular Remark 2.8. Theorems 2.27 and 2.28 are special cases of the

Arnold conjecture. This states that a Hamiltonian flow on a compact symplectic manifold M has at least $\mathrm{cat}(M)$ periodic solutions. If all periodic solutions are non-degenerate then it has at least $\sum_{i=0}^{\dim M} \dim H_i(M)$ critical points where $H_i(M)$ denotes the ith homology group of M with coefficients in an arbitrary field. Theorems 2.27 and 2.28 correspond to the case $M = T^{2N}$ where $\mathrm{cat}(M) = 2N + 1$ and $\dim H_i(M) = \binom{2N}{i}$, so that $\sum_{i=0}^{\dim M} \dim H_i(M) = 2^{2N}$. The interested reader can find results and many references concerning the Arnold conjecture in the book [52] and in the paper [41].

We conclude this section with a theorem on Hamiltonian systems where the Hamiltonian is both even and spatially periodic in the z-variables:

$$H(z + 2\pi k, t + 2\pi) = H(z, t) = H(-z, t) \quad \text{for all } z \in \mathbb{R}^{2N},\ k \in \mathbb{Z}^{2N},\ t \in \mathbb{R}. \tag{2.41}$$

It follows that $H_z(z, t) = 0$ for all $z \in (\pi\mathbb{Z})^{2N}$, hence modulo the $\mathbb{Z}^{2N}$-action (2.1) has at least 2^{2N} stationary solutions $z(t) \equiv c_i$ with $c_i \in \{0, \pi\}$, $i = 1, \ldots, 2N$. Thus the Arnold conjecture is trivially satisfied if (2.41) holds. A natural question to ask is whether there exist 2π-periodic solutions in addition to the 2^{2N} trivial equilibria.

THEOREM 2.29. *Suppose $H \in C^2(\mathbb{R}^{2N} \times \mathbb{R})$ satisfies* (2.41) *and suppose all 2π-periodic solutions of* (2.1) *are non-degenerate. For $z \in \pi\mathbb{Z}^{2N}$ let $A_z(t) := H_{zz}(z, t)$ and let $j^-(A_z) \in \mathbb{Z}$ be the Conley–Zehnder index. Then* (2.1) *has at least $k = \max |j^-(A_z)| - N$ pairs $\pm z_1, \ldots, \pm z_k$ of 2π-periodic solutions which lie on different $\mathbb{Z}^{2N}$-orbits.*

Theorem 2.29 has been proved in [14], an extension to Hamiltonian systems on T^*T^N, where H is even in z and periodic in the q-variables can be found in [15].

3. Homoclinic solutions

3.1. *Variational setting for homoclinic solutions*

Up to now we have been concerned with periodic solutions of Hamiltonian systems. In this part we turn our attention to homoclinic solutions of the system

$$\dot{z} = J H_z(z, t). \tag{3.1}$$

We assume H satisfies the following hypotheses:

($\widetilde{\mathrm{H}}_1$) $H \in C(\mathbb{R}^{2N} \times \mathbb{R}, \mathbb{R})$, $H_z \in C(\mathbb{R}^{2N} \times \mathbb{R}, \mathbb{R}^{2N})$ and $H(0, t) \equiv 0$;
($\widetilde{\mathrm{H}}_2$) H is 1-periodic in the t-variable;
($\widetilde{\mathrm{H}}_3$) $|H_z(z, t)| \leqslant c(1 + |z|^{s-1})$ for some $c > 0$ and $s \in (2, \infty)$;
($\widetilde{\mathrm{H}}_4$) $H(z, t) = \frac{1}{2} Az \cdot z + G(z, t)$, where A is a constant symmetric $2N \times 2N$-matrix, $\sigma(JA) \cap \mathrm{i}\mathbb{R} = \emptyset$ and $G_z(z, t)/|z| \to 0$ uniformly in t as $z \to 0$.

Note that the hypotheses ($\widetilde{\mathrm{H}}_1$)–($\widetilde{\mathrm{H}}_3$) are the same as (H_1)–(H_3) in Section 2.1 except that the period is normalized to 1 and not 2π (which is slightly more convenient here).

Let z_0 be a 1-periodic solution of (3.1). A solution z is said to be *homoclinic* (or *doubly asymptotic*) to z_0 if $z \not\equiv z_0$ and $|z(t) - z_0(t)| \to 0$ as $|t| \to \infty$. It has been shown by Coti Zelati, Ekeland and Séré in [25] that if H satisfies $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_3)$, then in many cases a symplectic change of variables will reduce the problem of finding homoclinics to z_0 to that of finding solutions homoclinic to 0 for the system (3.1), with a new Hamiltonian H satisfying $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$. Therefore in what follows we only consider solutions which are homoclinic to 0 (i.e., $z \not\equiv 0$ and $z(t) \to 0$ as $|t| \to \infty$), or *homoclinic solutions* for short. Recall that if $(\widetilde{\mathrm{H}}_4)$ holds, then 0 is called a *hyperbolic point*.

Let $E := H^{1/2}(\mathbb{R}, \mathbb{R}^{2N})$ be the Sobolev space of functions $z \in L^2(\mathbb{R}, \mathbb{R}^{2N})$ such that

$$\int_{\mathbb{R}} \left(1 + \xi^2\right)^{1/2} \left|\hat{z}(\xi)\right|^2 \mathrm{d}\xi < \infty,$$

where $\hat{z}$ is the Fourier transform of z. E is a Hilbert space with an inner product

$$(z, w) := \int_{\mathbb{R}} \left(1 + \xi^2\right)^{1/2} \hat{z}(\xi) \cdot \overline{\hat{w}(\xi)}\, \mathrm{d}\xi.$$

The Sobolev embedding $E \hookrightarrow L^q(\mathbb{R}, \mathbb{R}^{2N})$ is continuous for any $q \in [2, \infty)$ (see, e.g., [2] or Section 10 in [87]) but not compact. Indeed, let $z_j(t) := z(t - j)$, where $z \not\equiv 0$; then $z_j \rightharpoonup 0$ in E as $j \to \infty$ but $z_j \not\to 0$ in L^q. However, the embedding $E \hookrightarrow L^q_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^{2N})$ is compact.

Now we introduce a more convenient inner product. It follows from $(\widetilde{\mathrm{H}}_4)$ that $-\mathrm{i}\xi J - A$ is invertible and $(-\mathrm{i}\xi J - A)^{-1}$ is uniformly bounded with respect to $\xi \in \mathbb{R}$. Using this fact, the equality $-J\dot{z} \cdot w = \dot{z} \cdot Jw$ and Plancherel's formula it can be shown that the mapping $L : E \to E$ given by

$$(Lz, w) = \int_{\mathbb{R}} (-J\dot{z} - Az) \cdot w\, \mathrm{d}t$$

is bounded, selfadjoint and invertible (see Section 10 in [87] for the details). It is also shown in [87] that the spectrum $\sigma(-J(\mathrm{d}/\mathrm{d}t) - A)$ is unbounded below and above in $H^1(\mathbb{R}, \mathbb{R}^{2N})$; hence $E = E^+ \oplus E^-$, where $E^\pm$ are L-invariant infinite-dimensional spaces such that the quadratic form (Lz, z) is positive definite on E^+ and negative definite on E^-. Therefore we can define a new equivalent inner product in E by setting

$$\langle z, w\rangle := \left(Lz^+, w^+\right) - \left(Lz^-, w^-\right),$$

where $z^\pm, w^\pm \in E^\pm$. If $\|\cdot\|$ denotes the corresponding norm, we have

$$\int_{\mathbb{R}} (-J\dot{z} - Az) \cdot z\, \mathrm{d}t = (Lz, z) = \left\|z^+\right\|^2 - \left\|z^-\right\|^2. \tag{3.2}$$

Let

$$\Phi(z) := \frac{1}{2} \int_{\mathbb{R}} (-J\dot{z} - Az) \cdot z\, \mathrm{d}t - \int_{\mathbb{R}} G(z, t)\, \mathrm{d}t$$

and

$$\psi(z) := \int_{\mathbb{R}} G(z,t)\,\mathrm{d}t.$$

Then

$$\Phi(z) = \frac{1}{2}\|z^+\|^2 - \frac{1}{2}\|z^-\|^2 - \psi(z). \tag{3.3}$$

PROPOSITION 3.1. *If H satisfies* $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$, *then $\Phi \in C^1(E,\mathbb{R})$ and z is a homoclinic solution of* (3.1) *if and only if $z \neq 0$ and $\Phi'(z) = 0$. Moreover, ψ' and Φ' are weakly sequentially continuous.*

PROOF. We outline the argument. By $(\widetilde{\mathrm{H}}_3)$ and $(\widetilde{\mathrm{H}}_4)$, $|G_z(z,t)| \leqslant c(|z| + |z|^{s-1})$ for some constant $c > 0$. Hence $\psi \in C^1(E,\mathbb{R})$ and

$$\langle \psi'(z), w \rangle = \int_{\mathbb{R}} G_z(z,t) \cdot w\,\mathrm{d}t$$

according to Lemma 3.10 in [102] (although in [102] E is the Sobolev space $H^1(\mathbb{R}^N)$, an inspection of the proof shows that the argument remains valid in our case). Having this, it is easy to see from (3.2) and (3.3) that $\Phi \in C^1(E,\mathbb{R})$ and $\Phi'(z) = 0$ if and only if $z \in E$ and z is a weak solution of (3.1). Since $z \in L^q(\mathbb{R},\mathbb{R}^{2N})$ for all $q \in [2,\infty)$, $G_z(z(.),.) \in L^2(\mathbb{R},\mathbb{R}^{2N})$; hence $z \in H^1(\mathbb{R},\mathbb{R}^{2N})$. In particular, z is continuous by the Sobolev embedding theorem, and consequently, $z \in C^1(\mathbb{R},\mathbb{R}^{2N})$, i.e., z is a classical solution of (3.1).

It is well known that if $z \in H^1(\mathbb{R},\mathbb{R}^{2N})$, then $z(t) \to 0$ as $|t| \to \infty$. For the reader's convenience we include a proof. Let $t \leqslant u \leqslant t+1$. Given $\varepsilon > 0$, there is R such that if $|t| \geqslant R$, then $\|z\|_{L^2((t,t+1),\mathbb{R}^{2N})} < \varepsilon$ and $\|\dot{z}\|_{L^2((t,t+1),\mathbb{R}^{2N})} < \varepsilon$. Hence $|z(u)| < \varepsilon$ for some $u \in [t,t+1]$ and, using Hölder's inequality,

$$\begin{aligned}|z(t)| \leqslant |z(u)| + |z(u) - z(t)| &\leqslant \varepsilon + \int_t^u |\dot{z}(\tau)|\,\mathrm{d}\tau \\ &\leqslant \varepsilon + \|\dot{z}\|_{L^2((t,t+1),\mathbb{R}^{2N})} < 2\varepsilon.\end{aligned}$$

Finally, weak sequential continuity follows from the compact embedding $E \hookrightarrow L^q_{\mathrm{loc}}$, $2 \leqslant q < \infty$. Indeed, if $z_j \rightharpoonup z$ in E, then $z_j \to z$ in $L^2_{\mathrm{loc}}(\mathbb{R},\mathbb{R}^{2N}) \cap L^s_{\mathrm{loc}}(\mathbb{R},\mathbb{R}^{2N})$, and invoking the argument of Lemma 3.10 in [102] again, we see that $\langle \psi'(z_j), w\rangle \to \langle \psi'(z), w\rangle$ for all $w \in E$, i.e., $\psi'(z_j) \rightharpoonup \psi'(z)$. Clearly, $\Phi'(z_j) \rightharpoonup \Phi'(z)$ as well. □

It will be important in what follows that the functional Φ is invariant with respect to the representation of the group $\mathbb{Z}$ of integers given by

$$(T_a z)(t) := z(t+a), \quad a \in \mathbb{Z} \tag{3.4}$$

(this is an immediate consequence of the periodicity of G). Moreover, since the linear operator L is $\mathbb{Z}$-invariant, so are the subspaces $E^{\pm}$.

It follows from the $\mathbb{Z}$-invariance of Φ that Φ' is $\mathbb{Z}$-equivariant; hence if $z = z(t)$ is a homoclinic, so are all $T_a z$, $a \in \mathbb{Z}$. Therefore Φ cannot satisfy the Palais–Smale condition at any critical level $c \neq 0$. Setting $\mathcal{O}(z) := \{T_a z\colon\ a \in \mathbb{Z}\}$ (cf. Section 1.1), we call two homoclinic solutions z_1, z_2 *geometrically distinct* if $\mathcal{O}(z_1) \neq \mathcal{O}(z_2)$.

LEMMA 3.2. *If H satisfies* $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$, *then* $\psi'(z) = \mathrm{o}(\|z\|)$ *and* $\psi(z) = \mathrm{o}(\|z\|^2)$ *as* $z \to 0$.

We omit the argument which is exactly the same as in Lemma 2.4. Next we turn our attention to Palais–Smale sequences.

PROPOSITION 3.3. *If H satisfies* $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$ *and there are* $\mu > \max\{2, s-1\}$ *and* $\delta > 0$ *such that*

$$\delta|z|^{\mu} \leqslant \mu G(z,t) \leqslant z \cdot G_z(z,t) \quad \textit{for all } z, t, \tag{3.5}$$

then each (PS)*-sequence* $\{z_j\}$ *for* Φ *is bounded. Moreover, if* $\Phi(z_j) \to c$, *then* $c \geqslant 0$, *and if* $c = 0$, *then* $z_j \to 0$.

PROOF. As in (2.8), we have

$$c_1\|z_j\| + c_2 \geqslant \Phi(z_j) - \frac{1}{2}\bigl\langle \Phi'(z_j), z_j\bigr\rangle \geqslant \left(\frac{\mu}{2} - 1\right)\int_{\mathbb{R}} G(z_j,t)\,\mathrm{d}t \geqslant c_3\|z_j\|_{\mu}^{\mu}. \tag{3.6}$$

Since for each $\varepsilon > 0$ there is $C(\varepsilon)$ such that $|G_z(z,t)| \leqslant \varepsilon|z| + C(\varepsilon)|z|^{s-1}$ (by $(\widetilde{\mathrm{H}}_3)$ and $(\widetilde{\mathrm{H}}_4)$), we see as in (2.10) that

$$\|z_j^{\pm}\|^2 \leqslant \alpha_j\|z_j^{\pm}\| + c_4\varepsilon\|z_j\|\,\|z_j^{\pm}\| + c_5(\varepsilon)\|z_j\|_{\mu}^{s-1}\|z_j^{\pm}\|, \tag{3.7}$$

where $\alpha_j := \|\Phi'(z_j)\| \to 0$. Hence choosing ε small enough, we have $\|z_j\| \leqslant c_6 + c_7\|z_j\|_{\mu}^{s-1}$, and taking (3.6) into account,

$$\|z_j\| \leqslant c_8 + c_9\|z_j\|^{(s-1)/\mu}.$$

It follows that $\{z_j\}$ is bounded.

Now we obtain from (3.6) that if $\Phi(z_j) \to c$, then

$$c \geqslant c_3 \limsup_{j\to\infty}\|z_j\|^{\mu},$$

so $c \geqslant 0$. If $c = 0$, then $z_j \to 0$ in L^{μ} and passing to a subsequence, $z_j \rightharpoonup z$ in E. Hence $z = 0$. Letting $j \to \infty$ in (3.7) we see that $z_j \to 0$ also in E. □

Condition (3.5) is rather restrictive. Other conditions (always including the inequality $0 < \mu G(z,t) \leqslant z \cdot G_z(z,t)$ for all $z \neq 0$ and some $\mu > 2$) which imply boundedness of (PS)-sequences may be found, e.g., in [6] and [31].

We shall need the following result which is a special case of P.L. Lions' vanishing lemma (see, e.g., [65] or [102]):

LEMMA 3.4. *Let $r > 0$ be given and let $\{z_j\}$ be a bounded sequence in E. If*

$$\lim_{j\to\infty} \sup_{a\in\mathbb{R}} \int_{a-r}^{a+r} |z_j|^2 \,\mathrm{d}t = 0, \tag{3.8}$$

then $z_j \to 0$ in $L^q(\mathbb{R}, \mathbb{R}^{2N})$ for all $q \in (2, \infty)$.

PROOF. In [65,102] the case of $E = H^1(\mathbb{R}^N)$ has been considered. Below we adapt the argument of [102, Lemma 1.21] to our situation.

Let $q \in (2, 4)$. By Hölder's inequality,

$$\|z_j\|^q_{L^q((a-r,a+r),\mathbb{R}^{2N})} \leqslant \|z_j\|^{q-2}_{L^2((a-r,a+r),\mathbb{R}^{2N})} \|z_j\|^2_{L^p((a-r,a+r),\mathbb{R}^{2N})},$$

where p satisfies $(q-2)/2 + 2/p = 1$. Hence

$$\begin{aligned} \|z_j\|^q_{L^q((a-r,a+r),\mathbb{R}^{2N})} &\leqslant \sup_{b\in\mathbb{R}} \bigl(\|z_j\|^{q-2}_{L^2((b-r,b+r),\mathbb{R}^{2N})}\bigr) \|z_j\|^2_{L^p((a-r,a+r),\mathbb{R}^{2N})} \\ &\leqslant C \sup_{b\in\mathbb{R}} \bigl(\|z_j\|^{q-2}_{L^2((b-r,b+r),\mathbb{R}^{2N})}\bigr) \|z_j\|^2_{H^{1/2}((a-r,a+r),\mathbb{R}^{2N})}. \end{aligned} \tag{3.9}$$

Here we may use the norm in $H^{1/2}((a-r, a+r), \mathbb{R}^{2N})$ given by

$$\|z\|^2_{H^{1/2}((a-r,a+r),\mathbb{R}^{2N})} = \|z\|^2_{L^2((a-r,a+r),\mathbb{R}^{2N})} + \int_{a-r}^{a+r}\int_{a-r}^{a+r} \frac{|z(t)-z(s)|^2}{(t-s)^2} \,\mathrm{d}s\,\mathrm{d}t$$

(see [2, Theorem 7.48]). Covering $\mathbb{R}$ by intervals $(a_n - r, a_n + r)$, $n \in \mathbb{Z}$, in such a way that each $t \in \mathbb{R}$ is contained in at most 2 of them and taking the sum with respect to n in (3.9), we obtain

$$\|z_j\|^q_q \leqslant 2C \sup_{a\in\mathbb{R}} \bigl(\|z_j\|^{q-2}_{L^2((a-r,a+r),\mathbb{R}^{2N})}\bigr) |z_j|^2_E,$$

where

$$|z_j|^2_E := \|z_j\|^2_2 + \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|z(t)-z(s)|^2}{(t-s)^2} \,\mathrm{d}s\,\mathrm{d}t.$$

According to Theorem 7.12 in [64], the norms $|\cdot|_E$ and $\|\cdot\|$ are equivalent. Since $\{z_j\}$ is bounded in E, it follows that $z_j \to 0$ in L^q.

If $q \geqslant 4$, we can choose $q_0 \in (2,4)$ and $p > q$. Then by Hölder's inequality, $\|z_j\|_q^q \leqslant \|z_j\|_{q_0}^{(1-\lambda)q}\|z_j\|_p^{\lambda q}$, where $(1-\lambda)q/q_0 + \lambda q/p = 1$, and the conclusion follows because $\|z_j\|_{q_0} \to 0$ and $\|z_j\|_p$ is bounded. ☐

PROPOSITION 3.5. *Suppose* $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$ *are satisfied and* $\{z_j\}$ *is a bounded* (PS)*-sequence such that* $\Phi(z_j) \to c > 0$. *Then* (3.1) *has a homoclinic solution.*

PROOF. Suppose first $\{z_j\}$ is vanishing in the sense that (3.8) is satisfied. It is clear that (3.7) holds for $\mu = s$. Since $\alpha_j = \|\Phi'(z_j)\| \to 0$ and $z_j \to 0$ in L^s according to Lemma 3.4, it follows from (3.7) with ε appropriately small that $z_j \to 0$ in E; hence $\Phi(z_j) \to 0$. This contradiction shows that $\{z_j\}$ cannot be vanishing. Therefore there exist $\delta > 0$ and a_j such that, up to a subsequence,

$$\int_{a_j-r}^{a_j+r} |z_j|^2 \,\mathrm{d}t \geqslant \delta \tag{3.10}$$

for almost all j. Choosing a larger r if necessary we may assume $a_j \in \mathbb{Z}$. Let $\tilde{z}_j(t) := z_j(a_j + t)$. It follows from the $\mathbb{Z}$-invariance of Φ that $\{\tilde{z}_j\}$ is a bounded (PS)-sequence and $\Phi(\tilde{z}_j) \to c$. Hence $\tilde{z}_j \rightharpoonup \tilde{z}$ in E and $\tilde{z}_j \to \tilde{z}$ in L^2_{loc} after passing to a subsequence. Moreover, since

$$\int_{-r}^{r} |\tilde{z}_j|^2 \,\mathrm{d}t = \int_{a_j-r}^{a_j+r} |z_j|^2 \,\mathrm{d}t, \tag{3.11}$$

$\tilde{z} \neq 0$. According to Proposition 3.1, Φ' is weakly sequentially continuous. Thus $\Phi'(\tilde{z}) = 0$ and the conclusion follows. ☐

3.2. *Existence of homoclinics*

Our first result in this section asserts that if G_z is superlinear, then (3.1) has at least 1 homoclinic.

THEOREM 3.6. *Suppose H satisfies* $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$ *and G satisfies* (3.5) *with* $\mu > \max\{2, s-1\}$. *Then* (3.1) *has a homoclinic solution.*

PROOF. According to (3.3), the functional Φ corresponding to (3.1) has the form required in Theorem 1.19. Moreover, $G \geqslant 0$ and therefore $\psi \geqslant 0$. Let $z_j \rightharpoonup z$. Then $z_j \to z$ in $L^2_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^{2N})$ and $z_j \to z$ a.e. in $\mathbb{R}$ after passing to a subsequence, so it follows from Fatou's lemma that ψ is weakly sequentially lower semicontinuous. Moreover, ψ' is weakly sequentially continuous according to Proposition 3.1. Hence (i) of Theorem 1.19 holds, and so does (ii) because $\Phi(z) = \frac{1}{2}\|z\|^2 + \mathrm{o}(\|z\|^2)$ whenever $z \to 0$, $z \in E^+$.

We shall verify (iii). Let $z_0 \in E^+$, $\|z_0\| = 1$. Since there exists a continuous projection from the closure of $\mathbb{R}z_0 \oplus E^-$ in $L^\mu(\mathbb{R}, \mathbb{R}^{2N})$ to $\mathbb{R}z_0$ and $G(z,t) \geqslant \delta\mu^{-1}|z|^\mu$,

$$\begin{aligned}\Phi\bigl(z^- + \zeta z_0\bigr) &\leqslant \frac{\zeta^2}{2} - \frac{1}{2}\bigl\|z^-\bigr\|^2 - \delta\mu^{-1}\bigl\|z^- + \zeta z_0\bigr\|_\mu^\mu \\ &\leqslant \frac{\zeta^2}{2} - \frac{1}{2}\bigl\|z^-\bigr\|^2 - \delta_0\zeta^\mu\|z_0\|_\mu^\mu\end{aligned}$$

for some $\delta_0 > 0$, and it follows that $\Phi(z^- + \zeta z_0) \leqslant 0$ whenever $\|z^- + \zeta z_0\|$ is large enough. Obviously, $\Phi(z^-) \leqslant 0$ for all $z^- \in E^-$.

We have shown that the hypotheses of Theorem 1.19 are satisfied. Hence there exists a (PS)-sequence $\{z_j\}$ such that $\Phi(z_j) \to c > 0$ and it remains to invoke Propositions 3.3 and 3.5. □

Next we turn our attention to asymptotically linear systems. Suppose $\sigma(JA) \cap \mathrm{i}\mathbb{R} = \emptyset$, let λ_1 be the smallest positive and λ_{-1} the largest negative λ such that $\sigma(J(A + \lambda I)) \cap \mathrm{i}\mathbb{R} \neq \emptyset$ and set $\lambda_0 := \min\{\lambda_1, -\lambda_{-1}\}$. Then

$$\begin{aligned}&\lambda_1 = \inf\bigl\{\|z\|^2\colon\ z \in E^+,\ \|z\|_2 = 1\bigr\}, \\ &\lambda_{-1} = -\inf\bigl\{\|z\|^2\colon\ z \in E^-,\ \|z\|_2 = 1\bigr\}, \qquad (3.12)\\ &\|z\|^2 \geqslant \lambda_0\|z\|^2 \quad \text{for all } z \in E\end{aligned}$$

(see Section 10 of [87] for a detailed argument). We shall need the following two additional assumptions on G:

($\widetilde{\mathrm{H}}_5$) $G(z,t) = \frac{1}{2}A_\infty(t)z \cdot z + F(z,t)$, where $A_\infty(t)z \cdot z \geqslant \lambda|z|^2$ for some $\lambda > \lambda_1$ and $F_z(z,t)/|z| \to 0$ uniformly in t as $|z| \to \infty$;

($\widetilde{\mathrm{H}}_6$) $G(z,t) \geqslant 0$ and $\frac{1}{2}G_z(z,t) \cdot z - G(z,t) \geqslant \alpha(|z|)$, where $\alpha(0) = 0$ and $\alpha(|z|)$ is positive and bounded away from 0 whenever z is bounded away from 0.

A simple example of a function G which satisfies ($\widetilde{\mathrm{H}}_5$) and ($\widetilde{\mathrm{H}}_6$) is given by $G(z,t) = a(t)B(|z|)$, where a is 1-periodic, $a(t) \geqslant a_0 > 0$ for all t, $B \in C^1(\mathbb{R}, \mathbb{R})$, $B(0) = B'(0) = 0$, $B'(s)/s$ is strictly increasing, tends to 0 as $s \to 0$ and to $\lambda > \lambda_1/a_0$ as $s \to \infty$. That ($\widetilde{\mathrm{H}}_6$) holds follows from the identity

$$\frac{1}{2}B'(s)s - B(s) = \int_0^s \left(\frac{B'(s)}{s} - \frac{B'(\sigma)}{\sigma}\right)\sigma\,\mathrm{d}\sigma.$$

LEMMA 3.7. *If H satisfies* ($\widetilde{\mathrm{H}}_1$)–($\widetilde{\mathrm{H}}_6$), *then each Cerami sequence* $\{z_j\}$ (*see Remark* 1.20 *for a definition*) *is bounded.*

PROOF. Suppose $\{z_j\}$ is unbounded and let $w_j := z_j/\|z_j\|$. We may assume taking a subsequence that $w_j \rightharpoonup w$. We shall obtain a contradiction by showing that $\{w_j\}$ is neither vanishing (in the sense that (3.8) holds) nor nonvanishing.

Assume first $\{w_j\}$ is nonvanishing. As in the proof of Proposition 3.5, we find $a_j \in \mathbb{Z}$ such that, passing to a subsequence, $\widetilde{w}_j(t) := w_j(a_j + t)$ satisfy

$$\int_{-r}^{r} |\widetilde{w}_j|^2 \,\mathrm{d}t \geqslant \delta > 0$$

for j large. Passing to a subsequence once more, $\widetilde{w}_j \rightharpoonup \widetilde{w}$ in E and $\widetilde{w}_j \to \widetilde{w}$ in L^2_{loc} and a.e. in $\mathbb{R}$. In particular, $\widetilde{w} \neq 0$. Since $\|\Phi'(z_j)\| = \|\Phi'(\tilde{z}_j)\|$, it follows that $\Phi'(\tilde{z}_j)/\|\tilde{z}_j\| \to 0$ and therefore

$$\begin{aligned}\|\tilde{z}_j\|^{-1}\langle \Phi'(\tilde{z}_j), v\rangle = \langle \widetilde{w}_j^+, v\rangle - \langle \widetilde{w}_j^-, v\rangle - \int_{\mathbb{R}} A_\infty(t)\widetilde{w}_j \cdot v \,\mathrm{d}t \\ - \int_{\mathbb{R}} \frac{F_z(\tilde{z}_j, t)\cdot v}{|\tilde{z}_j|} |\widetilde{w}_j| \,\mathrm{d}t \to 0\end{aligned}$$

for all $v \in C_0^\infty(\mathbb{R}, \mathbb{R}^{2N})$. Since $|F_z(z,t)| \leqslant a|z|$ for some $a > 0$, $F_z(z,t) = \mathrm{o}(|z|)$ as $|z| \to \infty$ and $\operatorname{supp} v$ is bounded, we see by the dominated convergence theorem that the last integral on the right-hand side above tends to 0. Consequently, letting $j \to \infty$, we obtain

$$\dot{\widetilde{w}} = J\big(A + A_\infty(t)\big)\widetilde{w}$$

which contradicts the fact that the operator $-J(\mathrm{d}/\mathrm{d}t) - (A + A_\infty(t))$ has no eigenvalues in $L^2(\mathbb{R}, \mathbb{R}^{2N})$ [31, Proposition 2.2].

Suppose now $\{w_j\}$ is vanishing. Then

$$\|z_j\|^{-1}\langle \Phi'(z_j), w_j^+\rangle = \|w_j^+\|^2 - \int_{\mathbb{R}} \frac{G_z(z_j, t)\cdot w_j^+}{|z_j|} |w_j| \,\mathrm{d}t \to 0$$

and

$$\|z_j\|^{-1}\langle \Phi'(z_j), w_j^-\rangle = -\|w_j^-\|^2 - \int_{\mathbb{R}} \frac{G_z(z_j, t)\cdot w_j^-}{|z_j|} |w_j| \,\mathrm{d}t \to 0.$$

Since $\|w_j\| = 1$,

$$\int_{\mathbb{R}} \frac{G_z(z_j, t)\cdot (w_j^+ - w_j^-)}{|z_j|} |w_j| \,\mathrm{d}t \to 1.$$

Let

$$I_j := \big\{t \in \mathbb{R}\colon \big|z_j(t)\big| \leqslant \varepsilon\big\},$$

where $\varepsilon > 0$ has been chosen so that $|G_z(z,t)| \leqslant \frac{1}{2}\lambda_0|z|$ whenever $|z| \leqslant \varepsilon$ (such ε exists according to $(\widetilde{\mathrm{H}}_4)$). Since w_j^+ and w_j^- are orthogonal in L^2, it follows from (3.12) that

$$\int_{I_j} \frac{G_z(z_j,t)\cdot(w_j^+ - w_j^-)}{|z_j|}|w_j|\,\mathrm{d}t \leqslant \frac{1}{2}\lambda_0\|w_j\|_2^2 \leqslant \frac{1}{2}$$

and therefore, since $|G_z(z,t)| \leqslant a_0|z|$ for some $a_0 > 0$,

$$\begin{aligned}\frac{1}{4} &\leqslant \int_{\mathbb{R}\setminus I_j} \frac{G_z(z_j,t)\cdot(w_j^+ - w_j^-)}{|z_j|}|w_j|\,\mathrm{d}t \leqslant 2a_0\int_{\mathbb{R}\setminus I_j}|w_j|^2\,\mathrm{d}t\\ &\leqslant a_0\,\mathrm{meas}\,(\mathbb{R}\setminus I_j)^{(p-2)/p}\|w_j\|_p^{2/p}\end{aligned}$$

for almost all j ($p > 2$ arbitrary but fixed). As $\{w_j\}$ is vanishing, $w_j \to 0$ in $L^p(\mathbb{R},\mathbb{R}^{2N})$ according to Lemma 3.4 and consequently, $\mathrm{meas}\,(\mathbb{R}\setminus I_j) \to \infty$. Let $\alpha_0 := \inf_{|z|>\varepsilon}\alpha(|z|)$. Then $\alpha_0 > 0$ and $(\widetilde{\mathrm{H}}_6)$ implies

$$\begin{aligned}\Phi(z_j) - \frac{1}{2}\big\langle\Phi'(z_j),z_j\big\rangle &= \int_{\mathbb{R}}\left(\frac{1}{2}G_z(z_j,t) - G(z_j,t)\right)\mathrm{d}t\\ &\geqslant \int_{\mathbb{R}\setminus I_j}\left(\frac{1}{2}G_z(z_j,t) - G(z_j,t)\right)\mathrm{d}t \geqslant \int_{\mathbb{R}\setminus I_j}\alpha_0\,\mathrm{d}t \to \infty.\end{aligned}$$

However, $\Phi(z_j)$ is bounded and since $\{z_j\}$ is a Cerami sequence, $\langle\Phi'(z_j),z_j\rangle \to 0$. Therefore the left-hand side above must be bounded, a contradiction. □

The idea of showing boundedness of $\{z_j\}$ by excluding both vanishing and nonvanishing of $\{w_j\}$ goes back to Jeanjean [55].

THEOREM 3.8. *Suppose H satisfies* $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_6)$. *Then* (3.1) *has a homoclinic solution.*

PROOF. We shall use Theorem 1.19 again, this time together with Remark 1.20. Clearly, (i) and (ii) still hold. So if we can show that also (iii) is satisfied, then in view of Remark 1.20 there exists a Cerami sequence $\{z_j\}$ with $\Phi(z_j) \to c > 0$. By Lemma 3.7, $\{z_j\}$ is bounded, hence it is a (PS)-sequence as well and we can invoke Propositions 3.3 and 3.5 in the same way as before.

It remains to verify (iii) of Theorem 1.19. According to (3.12) and since $\lambda > \lambda_1$ there exists $z_0 \in E^+$, $\|z_0\| = 1$, such that

$$1 = \|z_0\|^2 < \lambda\|z_0\|_2^2. \tag{3.13}$$

Since $\Phi|_{E^-} \leqslant 0$, it suffices to show that

$$\Phi\big(z^- + \zeta z_0\big) = \frac{\zeta^2}{2} - \frac{1}{2}\|z^-\|^2 - \int_{\mathbb{R}} G\big(z^- + \zeta z_0,t\big)\,\mathrm{d}t \leqslant 0 \tag{3.14}$$

whenever $\|z^- + \zeta z_0\|$ is large enough. Assuming the contrary, we find z_j^- and ζ_j such that $z_j = z_j^- + \zeta_j z_0$ satisfies $\|z_j\| \to \infty$ and the reverse inequality holds in (3.14). Setting $w_j := z_j/\|z_j\| = (z_j^- + \zeta_j z_0)/\|z_j\| = w_j^- + \eta_j z_0$, we obtain

$$\frac{\eta_j^2}{2} - \frac{1}{2}\|w_j^-\|^2 - \int_{\mathbb{R}} \frac{G(z_j, t)}{\|z_j\|^2}\, \mathrm{d}t \geqslant 0,$$

and since $G \geqslant 0$,

$$\frac{\eta_j^2}{2} - \frac{1}{2}\|w_j^-\|^2 - \int_I \frac{G(z_j, t)}{\|z_j\|^2}\, \mathrm{d}t \geqslant 0, \tag{3.15}$$

where I is a bounded interval (to be specified). Passing to a subsequence, $\eta_j \to \eta \in [0, 1]$, $w_j^- \rightharpoonup w^-$ in E and $w_j^- \to w^-$ in L^2_{loc} and a.e. in $\mathbb{R}$. It follows from (3.15) that $\eta_j \geqslant \|w_j^-\|$, hence $\eta > 0$ because $\eta_j^2 + \|w_j^-\|^2 = 1$. In view of (3.13) and since z_0 and w^- are orthogonal in L^2,

$$\begin{aligned}
\eta^2 - \|w^-\|^2 - \int_{\mathbb{R}} A_\infty(t) w \cdot w\, \mathrm{d}t &\leqslant \eta^2 - \|w^-\|^2 - \lambda\|w\|_2^2 \\
&= \eta^2\bigl(1 - \lambda\|z_0\|_2^2\bigr) - \|w^-\|^2 - \lambda\|w^-\|_2^2 \\
&< 0,
\end{aligned}$$

hence there exists a bounded interval I such that

$$\eta^2 - \|w^-\|^2 - \int_I A_\infty(t) w \cdot w\, \mathrm{d}t < 0. \tag{3.16}$$

On the other hand,

$$\int_I \frac{G(z_j, t)}{\|z_j\|^2}\, \mathrm{d}t = \frac{1}{2}\int_I A_\infty(t) w_j \cdot w_j\, \mathrm{d}t + \int_I \frac{F(z_j, t)}{|z_j|^2} |w_j|^2\, \mathrm{d}t,$$

and since I is bounded, $|F(z, t)| \leqslant a|z|^2$ for some $a > 0$ and $F(z, t) = o(|z|^2)$ as $|z| \to \infty$, it follows from the dominated convergence theorem that the second integral on the right-hand side above tends to 0. Consequently, passing to the limit in (3.15) we obtain

$$\eta^2 - \|w^-\|^2 - \int_I A_\infty(t) w \cdot w\, \mathrm{d}t \geqslant 0,$$

a contradiction to (3.16). □

While several results concerning the existence of homoclinic solutions for second order systems (e.g., of Newtonian or Lagrangian type) may be found in the literature, much less seems to be known about (3.1) under conditions similar to $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$. The first paper to use modern variational methods for finding homoclinic solutions seems to be [25] of Coti Zelati, Ekeland and Séré. It has been shown there that if H satisfies $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$, $G = G(z,t)$ is superquadratic in an appropriate sense and convex in z, then (3.1) has at least 2 geometrically distinct homoclinics. If $G = G(z)$, there is at least 1 homoclinic. The convexity of G was used in order to reformulate the problem in terms of a dual functional which is better behaved than Φ (see also a comment in Section 3.4). A result comparable to our Theorem 3.6, but stronger in the sense that there is no growth restriction on G, is due to Tanaka [95]. His proof is rather different from ours: he shows using a linking argument and fine estimates that there is a sequence of $2\pi k_j$-periodic solutions z_j of (3.1) which tend to a homoclinic as $j \to \infty$. A somewhat different result has been obtained by Ding and Willem [31]. Their function G is also superquadratic but they allow the matrix A to be time-dependent (and 1-periodic) and moreover, they allow 0 to be the left endpoint of a gap of the spectrum of $-J(\mathrm{d}/\mathrm{d}t) - A(t)$ (more precisely, $\sigma(-J(\mathrm{d}/\mathrm{d}t) - A(t)) \cap (0,\alpha) = \emptyset$ for some $\alpha > 0$). See also Xu [103], where the superlinearity condition has been weakened with the aid of a truncation argument. Theorem 3.8 is due to Szulkin and Zou [94]; however, the argument presented here is simpler.

Finally we would like to mention that if A is time-dependent and $|A(t)| \to \infty$ in an appropriate sense as $|t| \to \infty$, then it can be shown that in many cases Φ satisfies $(\mathrm{PS})^*$ and methods similar to those developed in Sections 2.4–2.6 become available, see, e.g., Ding [29]. Results concerning bifurcation of homoclinics may be found in Stuart [87] and Secchi and Stuart [82].

3.3. *Multiple homoclinic solutions*

In Section 2.4 we have seen that in the autonomous case if H is superquadratic, then the Hamiltonian system has infinitely many periodic solutions whose amplitude tends to infinity. The proof relied in a crucial way on the S^1-invariance of the corresponding functional. For homoclinics the situation is very different. Let $z = (p,q) \in \mathbb{R}^2$ and

$$H(z) = \frac{1}{2}p^2 - \frac{1}{2}q^2 + \frac{1}{4}q^4. \tag{3.17}$$

The corresponding Hamiltonian system reduces to a second order equation $-\ddot{q} = V'(q)$, where the potential $V(q) = \frac{1}{4}q^4 - \frac{1}{2}q^2$. It is easy to see that there exists a homoclinic z_0 and

$$S := \left\{\pm z_0(t-a)\colon\ a \in \mathbb{R}\right\} \tag{3.18}$$

is the set of all homoclinics. So although S consists of infinitely (in fact uncountably) many geometrically distinct $\mathbb{Z}$-orbits (in the sense of Section 3.1), it contains only two homoclinics which are really distinct, the reason for this being that Φ is invariant with respect to the representation (3.4) of $\mathbb{R}$ rather than $\mathbb{Z}$. In particular, there are no homoclinics of large amplitude.

Below we shall assume that, in addition to the $\mathbb{Z}$-invariance, Φ is also invariant with respect to a representation of $\mathbb{Z}/p$ (p a prime) in $\mathbb{R}^{2N}$ and (S) of Section 2.6 is satisfied. We also recall from Section 2.6 that if H is even in z, then (S) holds with $\mathcal{G} = \mathbb{Z}/2$ (here we denote groups by $\mathcal{G}$ in order to distinguish them from functions $G = G(z,t)$). Our aim is to show that there are infinitely many geometrically distinct homoclinics provided H is superquadratic. Note that since $\mathbb{Z}/p$ is finite, (3.1) has infinitely many homoclinics which are geometrically distinct when both the representations of $\mathbb{Z}$ and $\mathbb{Z}/p$ are taken into account if and only if it has infinitely many geometrically distinct homoclinics with respect to the representation of $\mathbb{Z}$ only.

THEOREM 3.9. *Suppose H and G satisfy* ($\widetilde{\mathrm{H}}_1$)–($\widetilde{\mathrm{H}}_4$) *and* (3.5) *with* $\mu > \max\{2, s-1\}$, *there exist* $c_0, \varepsilon_0 > 0$ *such that*

$$\left|G_z(z+w,t) - G_z(z,t)\right| \leqslant c_0|w|\left(1+|z|^{s-1}\right) \quad \textit{whenever} \quad |w| \leqslant \varepsilon_0 \tag{3.19}$$

and (S) *of Section* 2.6 *holds for* $\mathcal{G} = \mathbb{Z}/p$, p *a prime. Then* (3.1) *has infinitely many geometrically distinct homoclinic solutions.*

Clearly, (3.19) is satisfied if H_{zz} is continuous and $|H_{zz}(z,t)| \leqslant c(1+|z|^{s-1})$ for some $c > 0$.

According to our comments in Section 2.6, if H is invariant with respect to an orthogonal and symplectic representation of a group $\mathcal{G}$, if $\mathbb{Z}/p \subset \mathcal{G}$ and $(\mathbb{R}^{2N})^{\mathbb{Z}/p} = \{0\}$, then the conclusion of the above theorem remains valid. However, if the system is autonomous or $\mathcal{G}$ is infinite, then already the existence of one homoclinic (which follows from Theorem 3.6) implies that there are infinitely many homoclinics which are geometrically distinct in the $\mathbb{Z} \times \mathbb{Z}/p$-sense but not in the sense of a representation of the larger group.

An important step in the proof of Theorem 3.9 is the following:

PROPOSITION 3.10. *Suppose* ($\widetilde{\mathrm{H}}_1$)–($\widetilde{\mathrm{H}}_4$), (3.5) *and* (3.19) *are satisfied and let* $\{z_j\}$ *be a* $(\mathrm{PS})_c$*-sequence with* $c > 0$. *Then there exist* (*not necessarily distinct*) *homoclinics* $w_1, \dots, w_k$ *and sequences* $\{b_j^m\}$ $(1 \leqslant m \leqslant k)$ *of integers such that, passing to a subsequence if necessary,*

$$\left\| z_j - \sum_{m=1}^{k} T_{b_j^m} w_m \right\| \to 0 \quad \textit{and} \quad \sum_{m=1}^{k} \Phi(w_m) = c.$$

PROOF (outline). We shall only very briefly sketch the argument which is exactly the same as in [58] (where a Schrödinger equation has been considered) or [30], see also [26]. By Proposition 3.3, $\{z_j\}$ is bounded, and it is nonvanishing by the argument of Proposition 3.5. Hence (3.10) is satisfied, and so is (3.11), where $\tilde{z}_j = T_{a_j^1} z_j$. It follows that $\tilde{z}_j \rightharpoonup w_1 \neq 0$ after passing to a subsequence and w_1 is a homoclinic. Let $v_j^1 := \tilde{z}_j - w_1$. Then one shows that $\{v_j^1\}$ is a (PS)-sequence such that $\Phi(v_j^1) \to c - \Phi(w_1)$. This argument is rather technical, and it is here (in the proof that $\Phi'(v_j^1) \to 0$ to be more precise) that the condition (3.19) plays a role. Moreover, there exists $\alpha > 0$ such that $\Phi(w) \geqslant \alpha$ for all critical points $w \neq 0$.

Indeed, otherwise we find a sequence of critical points $w_n \neq 0$ with $\Phi(w_n) \to 0$. But then, according to Proposition 3.3, $w_n \to 0$ which is impossible because (3.3) and the fact that $\psi'(z) = \mathrm{o}(\|z\|)$ as $z \to 0$ imply $w = 0$ is the only critical point in some neighbourhood of 0.

Now we can repeat the same argument for v_j^1 and obtain after passing to a subsequence again that $\tilde{v}_j^1 := T_{a_j^2} v_j^1 \rightharpoonup w_2$ and $\Phi(v_j^2) \to c - \Phi(w_1) - \Phi(w_2)$, where $v_j^2 = \tilde{v}_j^1 - w_2$. Since $\alpha > 0$, after a finite number of steps

$$\Phi\bigl(v_j^k\bigr) \to \beta := c - \sum_{m=1}^{k} \Phi(w_m) \leqslant 0.$$

But then $\beta = 0$ and $v_j^k \to 0$. Since (up to a subsequence) $T_{b_j^k} v_j^k = z_j - \sum_{m=1}^{k} T_{b_j^m} w_m$, where $b_j^m = -(a_j^1 + \cdots + a_j^m)$, the conclusion follows. □

PROOF OF THEOREM 3.9. Assuming that (3.1) has finitely many geometrically distinct homoclinics, we shall show that the hypotheses of Theorem 1.21 are satisfied thereby obtaining a contradiction.

Since $T_g \in \mathrm{O}(2N) \cap \mathrm{Sp}(2N)$ for $g \in \mathcal{G}$, the quadratic form (3.2) is $\mathcal{G}$-invariant, hence Φ is $\mathcal{G}$-invariant. It has been shown in the proof of Theorem 3.6 that ψ is weakly sequentially lower semicontinuous. This and Proposition 3.1 imply (i). We already know that (ii) holds. In order to verify (iii), we first note that there is an increasing sequence of $\mathcal{G}$-invariant subspaces $F_n \subset E^+$ such that $\dim F_n = n$ if $p = 2$ (in this case all subspaces are invariant) and $\dim F_n = 2n$ if $p > 2$. Since there exists a continuous projection $L^\mu(\mathbb{R}, \mathbb{R}^{2N}) \to F_n$, we obtain

$$\Phi(z) \leqslant \frac{1}{2}\bigl\|z^+\bigr\|^2 - \frac{1}{2}\bigl\|z^-\bigr\|^2 - \delta\mu^{-1}\|z\|_\mu^\mu \leqslant \frac{1}{2}\bigl\|z^+\bigr\|^2 - \delta_0\bigl\|z^+\bigr\|_\mu^\mu - \frac{1}{2}\bigl\|z^-\bigr\|^2$$

for some $\delta_0 > 0$ and all $z \in E_n$. The right-hand side above tends to $-\infty$ as $\|z\| \to \infty$ because $\dim F_n < \infty$.

It remains to verify (iv). Choose a unique point in each $\mathbb{Z}$-orbit of homoclinics and denote the set of all such points by $\mathcal{F}$. According to our assumption, $\mathcal{F}$ is finite. Changing the b_j^m:s if necessary we may assume $w_m \in \mathcal{F}$ in Proposition 3.10. For a given positive integer l, let

$$[\mathcal{F}, l] := \left\{ \sum_{m=1}^{k} T_{a_m} w_m \colon\ 1 \leqslant k \leqslant l,\ a_m \in \mathbb{Z},\ w_m \in \mathcal{F} \right\}.$$

If $I \subset (0, \infty)$ is a compact interval and l is large enough, then $[\mathcal{F}, l]$ is a $(\mathrm{PS})_I$-attractor according to Proposition 3.10. Finally, that

$$\inf\bigl\{\bigl\|z^+ - w^+\bigr\|\colon\ z, w \in [\mathcal{F}, l],\ z^+ \neq w^+\bigr\} > 0 \tag{3.20}$$

is a consequence of the result below. □

PROPOSITION 3.11. *Let $\mathcal{F}$ be a finite set of points in E. Then* (3.20) *holds.*

Here it is not assumed that $\mathcal{F}$ is a set whose points have any special property. The proof is straightforward though rather tedious and may be found in [26, Proposition 1.55]. In [26] this result is proved for z and w (and not z^+ and w^+), however, the argument is exactly the same in our case.

Theorem 3.9, for $p = 2$ (even Hamiltonian), is due to Ding and Girardi [30]. They have allowed A to be time-dependent and 0 to be the left endpoint of a gap of the spectrum of $-J(\mathrm{d}/\mathrm{d}t) - A(t)$. A result similar to Theorem 3.9 but allowing much more general (also infinite) groups has been obtained by Arioli and Szulkin [6]. Subsequently the superquadraticity condition in [6] has been weakened by Xu [103] by means of a truncation argument mentioned in the preceding section.

3.4. *Multibump solutions and relation to the Bernoulli shift*

It is known by Melnikov's theory that certain integrable Hamiltonian systems having 0 as a hyperbolic point can be perturbed in such a way that the stable and unstable manifold at 0 intersect transversally. This in turn implies that there exists a compact set, invariant with respect to the Poincaré mapping and conjugate to the Bernoulli shift (these notions will be defined later), see Palmer [77], or [47], [100] for a more comprehensive account of the subject. However, in general it is not an easy task to decide whether the intersection is transversal. In this section we shall see that sometimes under conditions which are weaker than transversality it is still possible to show the existence of an invariant set which is *semiconjugate* to the Bernoulli shift.

Consider the Hamiltonian system (3.1), with H satisfying $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$ and suppose that z_0 is a homoclinic solution. Let $\chi \in C^\infty(\mathbb{R}, [0,1])$ be a function such that $\chi(t) = 1$ for $|t| \leqslant 1/8$ and $\operatorname{supp}\chi \subset (-1/4, 1/4)$. If $\chi_j(t) := \chi(t/j)$, one easily verifies that $\{\chi_j z_0\}$ is a (PS)-sequence, $\Phi(\chi_j z_0) \to \Phi(z_0) = c$ and $\operatorname{supp}(\chi_j z_0) \subset (-j/4, j/4)$. Let $w_j(t) := \chi_j(t)z_0(t) + \chi_j(t-j)z_0(t-j)$. Since $\chi_j z_0$ and $\chi_j(\cdot - j)z_0(\cdot - j)$ have disjoint supports, it follows that $\{w_j\}$ is a (PS)-sequence such that $\Phi(z_j) \to 2c$. One can therefore expect that under suitable conditions there is a large j and a homoclinic solution

$$z(t) = \chi_j(t)z_0(t) + \chi_j(t-j)z_0(t-j) + \tilde{v}(t) = z_0(t) + z_0(t-j) + v(t)$$

such that $\|\tilde{v}\|_\infty$ (and hence also $\|v\|_\infty$) is small compared to $\|z_0\|_\infty$. We shall call this z a 2-bump solution. In a similar way one can look for k-bump solutions with $k > 2$.

Suppose now (3.1) has a homoclinic solution z_0 and for some ε reasonably small, say $\varepsilon \leqslant \frac{1}{2}\|z_0\|_\infty$, there exists $M > 0$ such that for any $k \in \mathbb{N}$ and any sequence of integers $a_1 < a_2 < \cdots < a_k$ satisfying $a_j - a_{j-1} \geqslant M$ for all j, there is a homoclinic solution

$$z(t) = \sum_{j=1}^{k} z_0(t - a_j) + v(t), \tag{3.21}$$

where $\|v\|_\infty \leqslant \varepsilon$ (so z is a k-bump solution). We emphasize that M is independent of k here. Existence of k-bump solutions for a class of superquadratic second order Hamiltonian systems has been shown by Coti Zelati and Rabinowitz [26]. However, in [26] M may depend on k. In [83] Séré has shown that under appropriate conditions on H, there is $M = M(\varepsilon)$ (independent of k) such that homoclinic solutions of the form (3.21) exist. More precisely, he has assumed that $H \in C^2(\mathbb{R}^{2N}, \mathbb{R})$ satisfies $(\widetilde{\mathrm{H}}_1)$–$(\widetilde{\mathrm{H}}_4)$, G is convex in z and

$$\delta_1|z|^s \leqslant G(z,t) \leqslant \delta_2|z|^s, \qquad sG(z,t) \leqslant z \cdot G_z(z,t)$$

for all z, t and some $\delta_1, \delta_2 > 0$, $s > 2$. Let $s' = s/(s-1)$. Since G is convex, one can use Clarke's duality principle in order to construct a dual functional $\Psi \in C^1(L^{s'}(\mathbb{R}, \mathbb{R}^{2N}), \mathbb{R})$ such that there is a one-to-one correspondence between critical points $u \neq 0$ of Ψ and homoclinic solutions z of (3.1) ([25], see also [33]). The functional Ψ is better behaved than Φ; in particular, it is not strongly indefinite and under the conditions specified above it has the mountain pass geometry near 0. This fact has been employed in [25] in order to obtain a homoclinic.

Let c be the mountain pass level for Ψ, or more precisely, let

$$c := \inf_{h \in \Gamma} \max_{\tau \in [0,1]} \Psi\bigl(h(\tau)\bigr),$$

where

$$\Gamma := \bigl\{h \in C\bigl([0,1], L^{s'}\bigl(\mathbb{R}, \mathbb{R}^{2N}\bigr)\bigr) \colon\ h(0) = 0,\ \Psi\bigl(h(1)\bigr) < 0\bigr\}.$$

Since 0 is a strict local minimum of Ψ and Ψ is unbounded below, $c > 0$.

THEOREM 3.12 (Séré [83]). *Suppose that H, G satisfy the hypotheses above and there is $c' > c$ such that the set $\mathcal{K}_{c'} := \{u \in \Psi^{c'} \colon\ \Psi'(u) = 0\}$ is countable. Then for each $\varepsilon > 0$ there exists $M = M(\varepsilon)$ such that to every choice of integers $a_1 < a_2 < \cdots < a_k$ satisfying $a_j - a_{j-1} \geqslant M$ there corresponds a homoclinic solution z of* (3.1) *given by* (3.21).

The countability of $\mathcal{K}_{c'}$ is a sort of nondegeneracy condition. A similar (but stronger) condition has also been employed in [26]. Consider the Hamiltonian system with H given by (3.17). If z_0 is a homoclinic, u_0 corresponds to z_0 and $c = \Psi(u_0)$, then $\mathcal{K}_{c'}$ is uncountable (see (3.18)). On the other hand, there are no multibump solutions in this case. Therefore in general it is necessary to assume some kind of nondegeneracy. Note also that since autonomous systems are invariant with respect to time-translations by a for any $a \in \mathbb{R}$, the countability condition can never be satisfied in this case. However, as has been shown by Bolotin and Rabinowitz [21], autonomous systems may have multibumps.

The proof of Theorem 3.12, in particular the construction of M independent of k, is lengthy and very technical. Therefore we omit it and refer the reader to [83]. To our knowledge no multibump results are known for first order Hamiltonian systems with nonconvex G (except when a reduction to a second order system like in [26] can be made).

From now on we assume that (3.1) has k-bump solutions of the form (3.21) for any k and M is independent of k. Choose $a \geqslant M$ and let

$$z_j(t) := z_0(t - aj), \quad j \in \mathbb{Z}.$$

We claim that, given any sequence $\{s_j\}$ of 0's and 1's, there exists a solution

$$z(t) = \sum_{j \in \mathbb{Z}} s_j z_j(t) + v(t)$$

of (3.1) such that $\|v\|_\infty \leqslant \varepsilon$. Note that if $s_j = 1$ for infinitely many $j's$, then z has infinitely many bumps and is *not* a homoclinic. By our assumption, for any positive integer m we can find a solution

$$z^m(t) = \sum_{j=-m}^{m} s_j z_j(t) + v^m(t)$$

with $\|v^m\|_\infty \leqslant \varepsilon$. Since 0 is a hyperbolic point and $z_0(t) \to 0$ as $|t| \to \infty$, z_0 decays to 0 exponentially (this follows by exponential dichotomy [77], see also [25] or [31]). Therefore $\|z^m\|_\infty$ is bounded uniformly in m, and by (3.1), the same is true of $\|\dot{z}^m\|_\infty$. Hence z^m, and a posteriori also v^m, are uniformly bounded in $H^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^{2N})$. Since the embedding $H^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^{2N}) \hookrightarrow L^\infty_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^{2N})$ is compact and $\|v^m\|_\infty \leqslant \varepsilon$, $v^m \to v$ in $L^\infty_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^{2N})$ for some v with $\|v\|_\infty \leqslant \varepsilon$. It follows that the corresponding function z is a weak solution of (3.1) and $z \neq 0$. Moreover, $z \in H^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^{2N}) \cap C^2(\mathbb{R}, \mathbb{R}^{2N})$ (recall H is of class C^2).

Let now

$$\Sigma_2 := \{0,1\}^{\mathbb{Z}} = \bigl\{s = \{s_j\}_{j \in \mathbb{Z}}\colon\ s_j \in \{0,1\}\bigr\}$$

be the set of doubly infinite sequences of 0's and 1's, endowed with the metric

$$d(s, \tilde{s}) := \sum_{j \in \mathbb{Z}} 2^{-|j|} |s_j - \tilde{s}_j|.$$

The space (Σ_2, d) is easily seen to be compact, totally disconnected and perfect (it is in fact homeomorphic to the Cantor set). The mapping $\sigma \in C(\Sigma_2, \Sigma_2)$ given by

$$\bigl(\sigma(s)\bigr)_j = s_{j+1}$$

is called the *Bernoulli shift on two symbols*. It is often considered as a prototype of a chaotic map. In particular, it has a countable infinity of periodic orbits, an uncountable infinity of nonperiodic orbits, a dense orbit, and it exhibits sensitive dependence on initial conditions. The details may be found, e.g., in Wiggins [100, Chapter 2].

Let

$$Z := \Bigl\{ z \in L^\infty\bigl(\mathbb{R}, \mathbb{R}^{2N}\bigr)\colon\ z(t) = \sum_{j \in \mathbb{Z}} s_j z_j(t) + v(t),\ s_j \in \{0,1\},\ \|v\|_\infty \leqslant \varepsilon \Bigr\}$$

and $I_j := [a(j-\frac{1}{2}), a(j+\frac{1}{2})]$. In Z we introduce a metric d by setting

$$d(z,0) = \sum_{j\in\mathbb{Z}} 2^{-|j|}\bigl(s_j\|z_j\|_\infty + \|v\|_{L^\infty(I_j,\mathbb{R}^{2N})}\bigr). \tag{3.22}$$

Since z_0 decays exponentially, the topology induced by d coincides with the L^∞_{loc}-topology on Z as one readily verifies. Using this, the compactness of Σ_2 and of the embedding $H^1_{\text{loc}}(\mathbb{R},\mathbb{R}^{2N}) \hookrightarrow L^\infty_{\text{loc}}(\mathbb{R},\mathbb{R}^{2N})$, it follows that the set

$$X := \bigl\{z \in Z\colon\ z \text{ is a solution of (3.1)}\bigr\}$$

is compact.

As before, let $T_a\colon X \to X$ be the mapping given by $(T_a z)(t) = z(t+a)$ and let $f_a\colon\mathbb{R}^{2N} \to \mathbb{R}^{2N}$ be the Poincaré (or time-a) mapping defined by $f_a(z_0) = z(a)$ where $z = z(t)$ is the unique solution of (3.1) satisfying the initial condition $z(0) = z_0$. Finally, let $Ev\colon X \to \mathbb{R}^{2N}$ be the evaluation mapping, $Ev(z) := z(0)$, and let $I := Ev(X)$. Since Ev is continuous and injective, it is a homeomorphism between X and I, and it is easy to verify that the diagram

$$\begin{tikzcd} I \arrow[r, "f_a"] & I \\ X \arrow[u, "Ev"] \arrow[r, "T_a"] & X \arrow[u, "Ev"'] \end{tikzcd} \tag{3.23}$$

is commutative. Note in particular that the set I is invariant with respect to the Poincaré mapping f_a.

THEOREM 3.13. *There exists a continuous surjective mapping $g\colon I \to \Sigma_2$ such that the diagram*

$$\begin{tikzcd} I \arrow[r, "f_a"] \arrow[d, "g"'] & I \arrow[d, "g"] \\ \Sigma_2 \arrow[r, "\sigma"] & \Sigma_2 \end{tikzcd}$$

is commutative.

If a continuous surjective mapping g as above exists, we shall say that $f_a\colon I \to I$ is *semiconjugate* to $\sigma\colon\Sigma_2 \to \Sigma_2$, and f_a will be called *conjugate* to σ if g is a homeomorphism.

PROOF. For $z = \sum_{j \in \mathbb{Z}} s_j z_j + v \in X$ we define $\varphi(z) = s = \{s_j\}_{j \in \mathbb{Z}}$. Then φ is a continuous mapping from X onto Σ_2 (cf. (3.22)) and the diagram

$$\begin{tikzcd} X \arrow[r, "T_a"] \arrow[d, "\varphi"'] & X \arrow[d, "\varphi"] \\ \Sigma_2 \arrow[r, "\sigma"] & \Sigma_2 \end{tikzcd} \tag{3.24}$$

commutes. Now the conclusion follows from (3.23) and (3.24) upon setting $g := \varphi \circ (Ev)^{-1}$. $\square$

REMARK 3.14. It follows from the definition of topological entropy $h(.)$ (see, e.g., [47, Definition 5.8.3] or [88, Definition 5.8.4]) and the uniform continuity of g that $h(f_a) \geqslant h(\sigma)$ (cf. [88, Exercise 5.8.1.B]). It is well known that $h(\sigma) > 0$. Hence $f_a|_I$, and therefore also the time-1-mapping f_1, has positive entropy. More precisely, $h(\sigma) = \log 2$ and $h(f_1) \geqslant (\log 2)/a$ according to [88, Example 5.8.1 and Theorem 5.8.4]. The same conclusion about the entropy may also be found in [83], where a different (in a sense, dual) argument has been used.

The approach presented in this section is taken from a work in progress by W. Zou and the second author. They study a certain second order system and hope to show that, in addition to the result of Theorem 3.13, to each m-periodic sequence $s \in \Sigma_2$ there corresponds a $z \in X$ which has period ma.

References

[1] A. Abbondandolo, *Morse Theory for Hamiltonian Systems*, Research Notes in Mathematics, Vol. 425, Chapman & Hall/CRC Press, Boca Raton, FL (2001).

[2] R.A. Adams, *Sobolev Spaces*, Academic Press, New York (1975).

[3] H. Amann and E. Zehnder, *Nontrivial solutions for a class of nonresonance problems and applications to nonlinear differential equations*, Ann. Scuola Norm. Sup. Pisa (4) **7** (1980), 539–603.

[4] H. Amann and E. Zehnder, *Periodic solutions of asymptotically linear Hamiltonian systems*, Manuscripta Math. **32** (1980), 149–189.

[5] A. Ambrosetti and V. Coti Zelati, *Periodic Solutions of Singular Lagrangian Systems*, Birkhäuser (1993).

[6] G. Arioli and A. Szulkin, *Homoclinic solutions of Hamiltonian systems with symmetry*, J. Differential Equations **158** (1999), 291–313.

[7] A. Bahri and H. Berestycki, *Existence of forced oscillations for some nonlinear differential equations*, Comm. Pure Appl. Math. **37** (1984), 403–442.

[8] A. Bahri and H. Berestycki, *Forced vibrations of superquadratic Hamiltonian systems*, Acta Math. **152** (1984), 143–197.

[9] T. Bartsch, *A simple proof of the degree formula for $\mathbb{Z}/p$-equivariant maps*, Math. Z. **212** (1993), 285–292.

[10] T. Bartsch, *Topological Methods for Variational Problems with Symmetries*, Springer-Verlag, Berlin (1993).

[11] T. Bartsch, *A generalization of the Weinstein–Moser theorems on periodic orbits of a Hamiltonian system near an equilibrium*, Ann. Inst. H. Poincaré Anal. Non Linéaire **14** (1997), 691–718.

[12] T. Bartsch and Y.H. Ding, *Homoclinic solutions of an infinite-dimensional Hamiltonian system*, Math. Z. **240** (2002), 289–310.

[13] T. Bartsch and Y.H. Ding, *Deformation theorems on non-metrizable vector spaces and applications to critical point theory*, Preprint.

[14] T. Bartsch and Z.-Q. Wang, *Periodic solutions of even Hamiltonian systems on the torus*, Math. Z. **224** (1997), 65–76.

[15] T. Bartsch and Z.-Q. Wang, *Periodic solutions of spatially periodic, even Hamiltonian systems*, J. Differential Equations (1997), 103–128.

[16] T. Bartsch and M. Willem, *Periodic solutions of non-autonomous Hamiltonian systems with symmetries*, J. Reine Angew. Math. **451** (1994), 149–159.

[17] V. Benci, *A geometrical index for the group S^1 and some applications to the study of periodic solutions of ordinary differential equations*, Comm. Pure Appl. Math. **34** (1981), 393–432.

[18] V. Benci, *On critical point theory of indefinite functionals in the presence of symmetries*, Trans. Amer. Math. Soc. **274** (1982), 533–572.

[19] V. Benci and P.H. Rabinowitz, *Critical point theorems for indefinite functionals*, Invent. Math. **52** (1979), 241–273.

[20] H. Berestycki, J.M. Lasry, G. Mancini and B. Ruf, *Existence of multiple periodic orbits on star-shaped Hamiltonian surfaces*, Comm. Pure Appl. Math. **38** (1985), 253–289.

[21] S.V. Bolotin and P.H. Rabinowitz, *A variational construction of chaotic trajectories for a Hamiltonian system on a torus*, Boll. Unione Mat. Ital. Ser. B (8) **1** (1998), 541–570.

[22] G. Cerami, *Un criterio di esistenza per i punti critici su varietà illimitate*, Istit. Lombardo Accad. Sci. Lett. Rend. A **112** (1978), 332–336.

[23] K.C. Chang, *Infinite Dimensional Morse Theory and Multiple Solution Problem*, Birkhäuser, Boston (1993).

[24] C. Conley and E. Zehnder, *The Birkhoff–Lewis fixed point theorem and a conjecture of V. Arnold*, Invent. Math. **73** (1983), 33–49.

[25] V. Coti Zelati, I. Ekeland and E. Séré, *A variational approach to homoclinic orbits in Hamiltonian systems*, Math. Ann. **288** (1990), 133–160.

[26] V. Coti Zelati and P.H. Rabinowitz, *Homoclinic orbits for second order Hamiltonian systems possessing superquadratic potentials*, J. Amer. Math. Soc. **4** (1991), 693–727.

[27] N. Dancer and S. Rybicki, *A note on periodic solutions of autonomous Hamiltonian systems emanating from degenerate stationary solutions*, Differential Integral Equations **12** (1999), 147–160.

[28] M. Degiovanni and L. Olian Fannio, *Multiple periodic solutions of asymptotically linear Hamiltonian systems*, Nonlinear Anal. **26** (1996), 1437–1446.

[29] Y.H. Ding, *Infinitely many homoclinic orbits for a class of Hamiltonian systems with symmetry*, Chinese Ann. Math. Ser. B **19** (1998), 167–178.

[30] Y.H. Ding and M. Girardi, *Infinitely many homoclinic orbits of a Hamiltonian system with symmetry*, Nonlinear Anal. **38** (1999), 391–415.

[31] Y.H. Ding and M. Willem, *Homoclinic orbits of a Hamiltonian system*, Z. Angew. Math. Phys. **50** (1999), 759–778.

[32] I. Ekeland, *Une théorie de Morse pour les systèmes hamiltoniens convexes*, Ann. Inst. H. Poincaré Anal. Non Linéaire **1** (1984), 19–78.

[33] I. Ekeland, *Convexity Methods in Hamiltonian Mechanics*, Springer-Verlag, Berlin (1990).

[34] I. Ekeland and L. Lassoued, *Multiplicité des trajectoires fermées de systèmes hamiltoniens convexes*, Ann. Inst. H. Poincaré Anal. Non Linéaire **4** (1987), 307–335.

[35] I. Ekeland and J.M. Lasry, *On the number of periodic trajectories for a Hamiltonian flow on a convex energy surface*, Ann. Math. **112** (1980), 283–319.

[36] E.R. Fadell and P.H. Rabinowitz, *Generalized cohomological index theories for Lie group actions with an application to bifurcation questions for Hamiltonian systems*, Invent. Math. **45** (1978), 139–174.

[37] P. Felmer, *Periodic solutions of spatially periodic Hamiltonian systems*, J. Differential Equations **98** (1992), 143–168.

[38] P. Felmer, *Periodic solutions of "superquadratic" Hamiltonian systems*, J. Differential Equations **102** (1993), 188–207.

[39] P. Felmer and Z.-Q. Wang, *Multiplicity for symmetric indefinite functionals and application to Hamiltonian and elliptic systems*, Topol. Methods Nonlinear Anal. **12** (1998), 207–226.

[40] A. Fonda and J. Mawhin, *Multiple periodic solutions of conservative systems with periodic nonlinearity*, Differential Equations and Applications, Vols. I and II, Ohio Univ. Press, Athens, OH (1989).

[41] K. Fukaya and K. Ono, *Arnold conjecture and Gromov–Witten invariants*, Topology **38** (1999), 933–1048.

[42] G. Fournier, D. Lupo, M. Ramos and M. Willem, *Limit relative category and critical point theory*, Dynamics Reported (New Series), Vol. 3, C.K.R.T. Jones, U. Kirchgraber and H.O. Walther, eds., Springer-Verlag, Berlin (1994), pp. 1–24.

[43] V.L. Ginzburg, *An embedding $S^{2n-1} \to \mathbb{R}^{2n}$, $2n-1 \geqslant 7$, whose Hamiltonian flow has no periodic trajectories*, Internat. Math. Res. Notices 1995, 83–98.

[44] V.L. Ginzburg and B.Z. Gürel, *A C^2-smooth counterexample to the Hamiltonian Seifert conjecture in $\mathbb{R}^4$*, Ann. Math. **158** (2003), 953–976.

[45] V.L. Ginzburg and E. Kerman, *Periodic orbits of Hamiltonian flows near symplectic extrema*, Pacific J. Math. **206** (2002), 69–91.

[46] M. Girardi and M. Matzeu, *Solutions of minimal period for a class of nonconvex Hamiltonian systems and applications to the fixed energy problem*, Nonlinear Anal. **10** (1986), 371–382.

[47] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York (1983).

[48] Y.X. Guo, *Nontrivial periodic solutions for asymptotically linear Hamiltonian systems with resonance*, J. Differential Equations **175** (2001), 71–87.

[49] M. Herman, *Examples of compact hypersurfaces in $\mathbb{R}^{2p}$, $2p \geqslant 6$*, Preprint (1994).

[50] H. Hofer, K. Wysocki and E. Zehnder, *The dynamics on three-dimensional strictly convex energy surfaces*, Ann. Math. **148** (1998), 197–289.

[51] H. Hofer and E. Zehnder, *Periodic solutions on hypersurfaces and a result by C. Viterbo*, Invent. Math. **90** (1987), 1–9.

[52] H. Hofer and E. Zehnder, *Symplectic Invariants and Hamiltonian Dynamics*, Birkhäuser, Basel (1994).

[53] M. Izydorek, *A cohomological Conley index in Hilbert spaces and applications to strongly indefinite problems*, J. Differential Equations **170** (2001), 22–50.

[54] M. Izydorek, *Equivariant Conley index in Hilbert spaces and applications to strongly indefinite problems*, Nonlinear Anal. **51** (2002), 33–66.

[55] L. Jeanjean, *On the existence of bounded Palais–Smale sequences and application to a Landesman–Lazer type problem set on $\mathbf{R}^N$*, Proc. Roy. Soc. Edinburgh Sect. A **129** (1999), 787–809.

[56] M.A. Krasnosel'skii, *Topological Methods in the Theory of Nonlinear Integral Equations*, Pergamon Press, Oxford (1964).

[57] W. Kryszewski and A. Szulkin, *An infinite dimensional Morse theory with applications*, Trans. Amer. Math. Soc. **349** (1987), 3181–3234.

[58] W. Kryszewski and A. Szulkin, *Generalized linking theorem with an application to a semilinear Schrödinger equation*, Adv. Differential Equations **3** (1998), 441–472.

[59] G. Li and A. Szulkin, *An asymptotically periodic Schrödinger equation with indefinite linear part*, Commun. Contemp. Math. **4** (2002), 763–776.

[60] S.J. Li and J.Q. Liu, *Some existence theorems on multiple critical points and their applications*, Kexue Tongbao **17** (1984), 1025–1027 (in Chinese).

[61] S.J. Li and J.Q. Liu, *Morse theory and asymptotic linear Hamiltonian system*, J. Differential Equations **78** (1989), 53–73.

[62] S.J. Li and A. Szulkin, *Periodic solutions for a class of nonautonomous Hamiltonian systems*, J. Differential Equations **112** (1994), 226–238.

[63] Y.Q. Li, *A limit index theory and its applications*, Nonlinear Anal. **25** (1995), 1371–1389.

[64] E.H. Lieb and M. Loss, *Analysis*, Amer. Math. Soc., Providence, RI (1997).

[65] P.L. Lions, *The concentration compactness principle in the calculus of variations. The locally compact case. Part II*, Ann. Inst. H. Poincaré Anal. Non Linéaire **1** (1984), 223–283.

[66] J.Q. Liu, *A generalized saddle point theorem*, J. Differential Equations **82** (1989), 372–385.

[67] C.G. Liu, Y. Long and C. Zhu, *Multiplicity of closed characteristics on symmetric convex hypersurfaces in $\mathbb{R}^{2n}$*, Math. Ann. **323** (2002), 201–215.

[68] Y. Long, *Maslov-type index, degenerate critical points, and asymptotically linear Hamiltonian systems*, Sci. China Ser. A **33** (1990), 1409–1419.

[69] Y. Long, *Index Theory for Symplectic Paths with Applications*, Birkhäuser, Basel (2002).

[70] Y. Long and X. Xu, *Periodic solutions for a class of nonautonomous Hamiltonian systems*, Nonlinear Anal. 41 (2000), 455–463.

[71] Y. Long and C. Zhu, *Closed characteristics on compact convex hypersurfaces in* $\mathbf{R}^{2n}$, Ann. Math. 155 (2002), 317–368.

[72] A.M. Lyapunov, *Problème général de la stabilité du mouvement*, Ann. Fac. Sci. Toulouse Math. (6) **2** (1907), 203–474.

[73] J. Mawhin and M. Willem, *Critical Point Theory and Hamiltonian Systems*, Springer-Verlag, New York (1989).

[74] D. McDuff and D. Salamon, *Introduction to Symplectic Topology*, Oxford University Press, New York (1998).

[75] K.R. Meyer and G.R. Hall, *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, Springer-Verlag (1992).

[76] J. Moser, *Periodic orbits near an equilibrium and a theorem by A. Weinstein*, Comm. Pure Appl. Math. **29** (1976), 727–747.

[77] K.J. Palmer, *Exponential dichotomies and transversal homoclinic points*, J. Differential Equations **55** (1984), 225–256.

[78] P.H. Rabinowitz, *Periodic solutions of Hamiltonian systems*, Comm. Pure Appl. Math. **31** (1978), 157–184.

[79] P.H. Rabinowitz, *Periodic solutions of large norm of Hamiltonian systems*, J. Differential Equations **50** (1983), 33–48.

[80] P.H. Rabinowitz, *Minimax Methods in Critical Point Theory with Applications to Differential Equations*, CBMS Reg. Conf. Ser. Math., Vol. 65, Amer. Math. Soc., Providence, RI (1986).

[81] P.H. Rabinowitz, *Variational methods for Hamiltonian systems*, Handbook of Dynamical Systems, Vol. 1A, B. Hasselblatt and A. Katok, eds., Elsevier (2002), 1091–1127.

[82] S. Secchi and C.A. Stuart, *Global bifurcation of homoclinic solutions of Hamiltonian systems*, Discrete Contin. Dynam. Systems **9** (2003), 1493–1518.

[83] E. Séré, *Looking for the Bernoulli shift*, Ann. Inst. H. Poincaré Anal. Non Linéaire **10** (1993), 561–590.

[84] E.H. Spanier, *Algebraic Topology*, Springer-Verlag, New York (1966).

[85] M. Struwe, *Variational Methods*, Springer-Verlag, Berlin (1990).

[86] M. Struwe, *Existence of periodic solutions of Hamiltonian systems on almost every energy surface*, Bol. Soc. Brasil Mat. (N.S.) **20** (1990), 49–58.

[87] C.A. Stuart, *Bifurcation into spectral gaps*, Bull. Belg. Math. Soc., Supplement, 1995.

[88] W. Szlenk, *An Introduction to the Theory of Smooth Dynamical Systems*, Polish Scientific Publishers, Warsaw; Wiley, New York (1984).

[89] A. Szulkin, *Morse theory and existence of periodic solutions of convex Hamiltonian systems*, Bull. Soc. Math. France **116** (1988), 171–197.

[90] A. Szulkin, *A relative category and applications to critical point theory for strongly indefinite functionals*, Nonlinear Anal. **15** (1990), 725–739.

[91] A. Szulkin, *Bifurcation for strongly indefinite functionals and a Liapunov type theorem for Hamiltonian systems*, Differential Integral Equations **7** (1994), 217–234.

[92] A. Szulkin, *Index theories for indefinite functionals and applications*, Topological and Variational Methods for Nonlinear Boundary Value Problems, P. Drábek, ed., Pitman Research Notes in Mathematics Series, Vol. 365, Harlow, Essex (1997), 89–121.

[93] A. Szulkin and W. Zou, *Infinite dimensional cohomology groups and periodic solutions of asymptotically linear Hamiltonian systems*, J. Differential Equations **174** (2001), 369–391.

[94] A. Szulkin and W. Zou, *Homoclinic orbits for asymptotically linear Hamiltonian systems*, J. Funct. Anal. **187** (2001), 25–41.

[95] K. Tanaka, *Homoclinic orbits in a first order superquadratic Hamiltonian system: convergence of subharmonic orbits*, J. Differential Equations **94** (1991), 315–339.

[96] C. Viterbo, *A proof of the Weinstein conjecture in* $\mathbb{R}^{2n}$, Ann. Inst. H. Poincaré Anal. Non Linéaire **4** (1987), 337–356.

[97] C. Viterbo, *Equivariant Morse theory for starshaped Hamiltonian systems*, Trans. Amer. Math. Soc. **311** (1989), 621–655.
[98] A. Weinstein, *Normal modes for nonlinear Hamiltonian systems*, Invent. Math. **20** (1973), 47–57.
[99] A. Weinstein, *Periodic orbits for convex Hamiltonian systems*, Ann. Math. **108** (1978), 507–518.
[100] S. Wiggins, *Global Bifurcations and Chaos*, Springer-Verlag, New York (1988).
[101] M. Willem, *Lectures on Critical Point Theory*, Trabalho de Mat., Vol. 250, Fundação Univ. Brasília, Brasília (1983).
[102] M. Willem, *Minimax Theorems*, Birkhäuser, Boston (1996).
[103] X. Xu, *Homoclinic orbits for first order Hamiltonian systems possessing super-quadratic potentials*, Nonlinear Anal. **51** (2002), 197–214.

CHAPTER 3

Differential Equations on Closed Sets[*,†]

Ovidiu Cârjă

Faculty of Mathematics, “Al. I. Cuza” University of Iaşi, Romania

Ioan I. Vrabie

Faculty of Mathematics, “Al. I. Cuza” University of Iaşi, Romania and
“O. Mayer” Mathematics Institute of the Romanian Academy, Iaşi, Romania

Contents

1. Introduction . . . 149
   Notations . . . 155
2. Preliminaries . . . 156
   2.1. Brezis–Browder ordering principle . . . 156
   2.2. Projections . . . 158
   2.3. Tangent cones . . . 159
   2.4. The proximal normal cone . . . 161
   2.5. Clarke’s tangent cone . . . 163
3. Problems of viability . . . 166
   3.1. Nagumo’s viability theorem . . . 166
   3.2. Proof of the necessity . . . 167
   3.3. Existence of approximate solutions . . . 167
   3.4. Convergence of approximate solutions . . . 170
   3.5. Existence of noncontinuable solutions . . . 172
   3.6. Viability of the relative closure . . . 174
   3.7. Comparison and viability . . . 175
   3.8. Viable preordered subsets . . . 176
4. Problems of invariance . . . 178
   4.1. Preliminary facts . . . 178
   4.2. Sufficient conditions for local invariance . . . 179
   4.3. Viability and comparison imply exterior tangency . . . 180

---

[*] Partially supported by CNCSIS Grant Code 1345/2003.

[†] The writing of this material was facilitated by numerous and very fruitful discussions with Professor Corneliu Ursescu and by a careful reading of the whole manuscript by Professor Mihai Necula. The authors take this opportunity to express their warmest thanks to both of them.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 2
Edited by A. Cañada, P. Drábek and A. Fonda
© 2005 Elsevier B.V. All rights reserved

  4.4. Sufficient conditions for invariance. Revisited . . . 183
  4.5. When tangency implies exterior tangency? . . . 187
  4.6. Local invariance and monotonicity . . . 189
5. Carathéodory solutions . . . 189
  5.1. A Lebesgue type derivation theorem . . . 189
  5.2. Characterizations of Carathéodory viability . . . 192
  5.3. Sufficient conditions for Carathéodory local invariance . . . 197
  5.4. Sufficient conditions via generalized distance . . . 198
6. Differential inclusions . . . 200
  6.1. Multifunctions . . . 200
  6.2. Viability with respect to a multifunction . . . 201
  6.3. Existence of ε-approximate solutions . . . 203
  6.4. Convergence of the ε-approximate solutions . . . 206
  6.5. Noncontinuable solutions . . . 207
  6.6. Viable preordered subsets . . . 210
  6.7. Local invariance. Sufficient conditions . . . 210
  6.8. Local invariance. Necessary conditions . . . 212
7. Applications . . . 214
  7.1. Invariance with respect to parametrized multifunctions . . . 214
  7.2. Differentiability along trajectories . . . 215
  7.3. Liapunov functions . . . 216
  7.4. Hukuhara's theorem . . . 219
  7.5. Kneser's theorem . . . 221
  7.6. The characteristics method for a first order PDE . . . 222
8. Notes and comments . . . 227
  8.1. The upper semicontinuous case . . . 227
  8.2. The case of Carathéodory mappings . . . 228
  8.3. The lower semicontinuous case . . . 228
  8.4. The semilinear single-valued case . . . 229
  8.5. The semilinear multivalued case . . . 231
  8.6. The nonlinear perturbed single-valued case . . . 233
  8.7. The multivalued perturbed nonlinear case . . . 233
  8.8. Applications . . . 233
Appendix . . . 234
References . . . 235

1. Introduction

One of the most important problems in the Theory of Differential Equations is that of local existence. This consists in checking whether or not, for each initial datum, the Cauchy problem attached to a given differential equation has at least one solution. Unless otherwise specified, in that follows, we will assume that $\mathbb{R}^n$ is endowed with one of its equivalent norms $\|\cdot\|$. Sometimes, we will assume that $\|\cdot\|$ is the Euclidean norm, i.e., that it is defined by means of the usual inner product $\langle\cdot,\cdot\rangle$ as $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$. In all these cases, either we will explicitly specify that, or we assume that it is implicitly notified by the simple use of the symbol $\langle\cdot,\cdot\rangle$. By a *domain* we understand a nonempty assumed and open subset $\mathbb{D} \subseteq \mathbb{R}^n$, and by $\mathbb{I}$ we denote an open interval in $\mathbb{R}$. Further, we denote by $\mathbb{K}$ a nonempty subset in $\mathbb{R}^n$. Let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be a given function. We recall that, excepting some other special mentions to be done in due course, by a *solution* to the differential equation

$$u'(t) = f\bigl(t, u(t)\bigr) \tag{1.1.1}$$

we mean a function $u : \mathbb{J} \to \mathbb{K}$, with $\mathbb{J}$ a non-degenerate interval included in $\mathbb{I}$, and satisfying the equality (1.1.1) for all $t \in \mathbb{J}$. Clearly, whenever f is continuous, all solutions to (1.1.1) are of class C^1. The first significant local existence result is due to Cauchy and refers to the case in which $\mathbb{K} = \mathbb{D}$ and f is a C^1-function. This was extended by Lipschitz to the class of all functions f satisfying the homonymous condition, and by Peano [80] to general continuous functions. More precisely, Peano [80] proved:

THEOREM 1.1.1 (Peano [80]). *If $\mathbb{D}$ is a domain and $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ is a continuous function, then, for every $(\tau, \xi) \in \mathbb{I} \times \mathbb{D}$, there exists $T \in \mathbb{I}$, $T > \tau$, such that the Cauchy problem*

$$\begin{cases} u'(t) = f\bigl(t, u(t)\bigr), \\ u(\tau) = \xi \end{cases} \tag{1.1.2}$$

has at least one solution $u : [\tau, T] \to \mathbb{D}$.

On the other hand, there are situations in which, instead of a domain $\mathbb{D}$, one has to consider a set $\mathbb{K}$ which contains non-interior points, say for instance when the state u of a certain system must evolve within a given closed subset $\mathbb{K}$ in $\mathbb{R}^n$. These considerations lead to the concept of *viability* of a set $\mathbb{K}$ with respect to a given function f. More precisely, we say that the subset $\mathbb{K} \subseteq \mathbb{D}$ is *right viable* with respect to the function $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ if for each $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ there exists $T \in \mathbb{I}$, $T > \tau$, such that the differential equation (1.1.1) has at least one solution $u : [\tau, T] \to \mathbb{K}$ satisfying $u(\tau) = \xi$. In order to be consistent with the usual procedure, i.e. $g(t, u) = (1, f(t, u))$, which reduces the nonautonomous case to the autonomous one, in the latter situation, i.e. when $f : \mathbb{D} \to \mathbb{R}^n$, we say that $\mathbb{K}$ is *right viable* with respect to f if for each $\xi \in \mathbb{K}$ there exists $T > 0$ such that (1.1.1) has at least one solution $u : [0, T] \to \mathbb{K}$ satisfying $u(0) = \xi$. Sometimes, when we will be interested to get solutions defined at the left of τ, we will speak about *left viability*. Moreover, for the sake of simplicity, whenever no confusion may occur, we will use the term *viable* to design right viable. Let us notice that $\mathbb{K}$ is right viable with respect to f if and only if $\mathbb{K}$ is left viable with respect to $-f$. The next simple characterization of right viable subsets with respect to a given—possibly discontinuous—function which does not depend on the t-variable is almost obvious.

PROPOSITION 1.1.1. *Let $\mathbb{D}$ be a domain and $f:\mathbb{D}\to\mathbb{R}^n$. Then $\mathbb{K}\subseteq\mathbb{D}$ is right* (*left*) *viable with respect to f if and only if it is the union of a certain family of right* (*left*) *trajectories of* (1.1.1).

If $\mathbb{K}$ is open, then it is viable with respect to any continuous function f, and this is nothing else that the celebrated Peano's local existence theorem 1.1.1 just mentioned. By contrary, if $\mathbb{K}$ is not open, in general, $\mathbb{K}$ may fail to be viable with respect to any continuous f, as we can see from the simple example below.

EXAMPLE 1.1.1. Let us consider the plane $\mathbb{K}=\{(u_1,u_2,u_3);\ u_3=1\}$ and the function $f:\mathbb{K}\to\mathbb{R}^3$, defined by $f(u_1,u_2,u_3)=(u_2+u_3,-u_1,-u_1)$ for every $(u_1,u_2,u_3)\in\mathbb{K}$. Then, if ξ is the projection of the origin on this plane, i.e. $\xi=(0,0,1)$, the problem (1.1.2) has no local solution. Indeed, assuming by contradiction that there exists such a solution $u:[0,T]\to\mathbb{K}$, we have $\langle u'(t),u(t)\rangle=\langle f(u(t)),u(t)\rangle=0$ and therefore $\|u(t)\|=\|\xi\|=1$ for every $t\in[0,T]$. Hence $u(t)$ lies on the sphere of center 0 and radius 1 which has only one point in common with $\mathbb{K}$, namely ξ. Then, necessarily $u(t)=\xi$ for every $t\in[0,T]$, which is impossible, because, in this case, one should have $u_1(t)=0$ and $u_1'(t)=u_2(t)+u_3(t)=1$ for every $t\in[0,T]$. This contradiction can be eliminated only if (1.1.2) has no local solution.

Thus, whenever $\mathbb{K}$ is not open, one has to compensate the lack of this crucial topological assumption by something else, as less restrictive as possible. In order to understand what extra-condition we have to impose, some comments are needed. More precisely, let us consider for the moment that $\mathbb{D}$ is a domain, $f:\mathbb{D}\to\mathbb{R}^n$ is continuous (and does not depend on t) and $U:\mathbb{D}\to\mathbb{R}$ is a function of class C^1 with $\nabla U(\xi)\neq 0$ on $\mathbb{D}$. Then, if U is a prime integral for the system $u'(t)=f(u(t))$, which amounts of saying that f is parallel to the tangent plane to each surface of constant level $\Sigma_\eta=\{\xi\in\mathbb{D};\ U(\xi)=U(\eta)\}$ at every point of this surface, i.e.

$$\sum_{i=1}^{n} f(\xi)\frac{\partial U}{\partial x_i}(\xi)=0,$$

for each $\xi\in\Sigma_\eta$, then the restriction $f_{|\Sigma_\eta}$, of the function f to any surface Σ_η, has the property that, for every $\xi\in\Sigma_\eta$, the equation $u'(t)=f_{|\Sigma_\eta}(u(t))$ has at least one local solution $u:[0,T]\to\Sigma_\eta$ satisfying $u(0)=\xi$. This condition constitutes a first step through a partial answer to the question: *what extra-conditions must satisfy the set $\mathbb{K}\subseteq\mathbb{R}^n$* (*not necessarily a level set*) *and the continuous function $f:\mathbb{K}\to\mathbb{R}^n$, in order that, for every $\xi\in\mathbb{K}$, to exist at least one function of class C^1, $u:[0,T]\to\mathbb{K}$, such that $u(0)=\xi$ and $u'(t)=f(u(t))$*. However, the conditions offered by the result just mentioned in the case

$\mathbb{K} = \Sigma_\eta$ have three weak points. First, they ask f to be defined on the union of all surfaces Σ_η and not on a single one. Second, f must satisfy the mentioned "tangency condition" on each one of the surfaces of the family. Finally, the set $\mathbb{K}$ is in this case of a very specific type, namely it is a surface of constant level for a function $U : \mathbb{D} \to \mathbb{R}$, of class C^1 and satisfying $\nabla U(\xi) \neq 0$ for all $\xi \in \mathbb{D}$.

The first step in order to get an acceptable result in this direction would be expressed as in Theorem 1.1.2 below. See Vrabie [106, Theorem 6.1.2, p. 212].

THEOREM 1.1.2. *Let $\Sigma \subseteq \mathbb{D}$ be a regular surface and $f : \mathbb{D} \to \mathbb{R}^n$ a continuous function. The necessary and sufficient condition in order that, for every $\xi \in \Sigma$ to exist $T > 0$ and a function of class C^1, $u : [0, T] \to \Sigma$, such that $u(0) = \xi$ and $u'(t) = f(u(t))$ for every $t \in [0, T]$, is that for every $\eta \in \Sigma$, $f(\eta)$ be parallel to the tangent plane to Σ at η.*

As far as we know, the first general necessary and sufficient condition for viability has been discovered by Nagumo [70] in 1942 in the case when $\mathbb{K}$ is a closed, or merely locally closed subset in $\mathbb{R}^n$. We recall that $\mathbb{K}$ is *locally closed* if for each $\xi \in \mathbb{K}$ there exists $\rho > 0$ such that $B(\xi, \rho) \cap \mathbb{K}$ is closed. We notice that Nagumo used the term "*rechts zulässig*", i.e. *right admissible* to design viable, and proved the following fundamental characterization for viability:

THEOREM 1.1.3 (Nagumo [70]). *Let $\mathbb{K} \subseteq \mathbb{D}$ be locally closed and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ be continuous. A necessary and sufficient condition in order that $\mathbb{K}$ be viable with respect to f is that, for each $(t, \xi) \in \mathbb{I} \times \mathbb{K}$, $f(t, \xi)$ be tangent to $\mathbb{K}$ at ξ in the sense of Bouligand–Severi, i.e.*

$$\liminf_{h \downarrow 0} \frac{1}{h} \operatorname{dist}\bigl(\xi + hf(t, \xi); \mathbb{K}\bigr) = 0. \tag{1.1.3}$$

Here and thereafter, $\operatorname{dist}(\eta; \mathbb{K})$ denotes the distance from $\eta \in \mathbb{R}^n$ to the subset $\mathbb{K}$ in $\mathbb{R}^n$, i.e. $\operatorname{dist}(\eta; \mathbb{K}) = \inf\{\|\eta - \mu\|;\ \mu \in \mathbb{K}\}$.

Since (1.1.3) is obviously satisfied at each interior point of $\mathbb{K}$, Peano's theorem 1.1.1 is a direct consequence of Nagumo's theorem 1.1.3. Clearly, (1.1.3) is invariant with respect to equivalent norms on $\mathbb{R}^n$. Moreover, this essentially metric condition can be described, and therefore defined, only by means of linear topological concepts. See, for instance, Ursescu [97].

We notice that Nagumo's result (or variants of it) has been rediscovered independently, in the late sixties and early seventies, by Yorke [108,109], Crandall [38] and Hartman [55] among others. More precisely, Yorke [108] uses viability (*weak positive invariance* in his terminology) in order to prove necessary and sufficient conditions for stability by means of Liapunov functions, as well as to give very simple and elegant proofs for both Hukuhara and Kneser's theorems referring to the solution funnel. We notice that Yorke [108] analyzed the case of subsets $\mathbb{K}$ which are *closed relative to* $\mathbb{D}$, i.e., for which there exists a closed set $\mathbb{C} \subseteq \mathbb{R}^n$ such that $\mathbb{K} = \mathbb{C} \cap \mathbb{D}$. Nevertheless, as Remark 1.1.1 below shows, all his results hold true for locally closed sets as well.

REMARK 1.1.1. If $\mathbb{K} \subseteq \mathbb{D}$ is closed relative to $\mathbb{D}$, then it is locally closed. Conversely, if $\mathbb{K} \subseteq \mathbb{D}$ is locally closed, then there exists an open neighborhood $\mathbb{V} \subseteq \mathbb{D}$ of $\mathbb{K}$ such that $\mathbb{K}$ is closed relative to $\mathbb{V}$.

It should be emphasized that, from another point of view, Yorke [108] considered the strictly more general case in which, instead of a cylindrical set $\mathbb{I} \times \mathbb{D}$, f is defined on a non-cylindrical one $\mathcal{D} \subseteq \mathbb{R} \times \mathbb{R}^n$, situation requiring a more delicate analysis. Crandall [38] considers the case when $\mathbb{D}$ is arbitrary, $\mathbb{K} \subseteq \mathbb{D}$ is locally closed and $f : \mathbb{D} \to \mathbb{R}^n$ is continuous. He shows that a sufficient condition for $\mathbb{K}$ to be viable (*forward invariant* in his terminology) with respect to f is (1.1.3). Hartman [55] proves essentially the same result for $\mathbb{D}$ open, $\mathbb{K}$ closed relative to $\mathbb{D}$ and $f : \mathbb{D} \to \mathbb{R}^n$ continuous, and shows in addition that (1.1.3) is necessary for the viability of $\mathbb{K}$ with respect to f.

An extension of Nagumo's viability theorem to Carathéodory functions f was proved by Ursescu [98] by using Scorza Dragoni's theorem.

At this point, we can easily see that viability is independent of the values of f on $\mathbb{D} \setminus \mathbb{K}$, and therefore, in the study of such kind of problems there is no need for f to be defined "outside" $\mathbb{K}$. So, very often in that follows, we will speak about the viability of a subset $\mathbb{K} \subseteq \mathbb{R}^n$ with respect to a function f defined either on $\mathbb{I} \times \mathbb{D}$, or merely on $\mathbb{I} \times \mathbb{K}$. This is no longer true if we consider the problem of *local invariance* to be defined below. Namely, let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$, where $\mathbb{D}$ is a domain and let $\mathbb{K} \subseteq \mathbb{D}$. The subset $\mathbb{K}$ is *locally right invariant* with respect to f if for each $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ and each solution $u : [\tau, c] \to \mathbb{D}$, $c \in \mathbb{I}$, $c > \tau$, of (1.1.1), satisfying the initial condition $u(\tau) = \xi$, there exists $T \in (\tau, c]$ such that we have $u(t) \in \mathbb{K}$ for each $t \in [\tau, T]$. It is *right invariant* if it satisfies the above condition of local invariance with $T = c$.

As in the case of right viability, whenever we will consider solutions defined at the left of τ we will speak about *local left invariance*. Again, for the sake of simplicity, if no confusion can occur, we will use the term *local invariant* to design local right invariant. We notice that if $\mathbb{K}$ is viable with respect to f, and (1.1.1) has the uniqueness property on $\mathbb{I} \times \mathbb{D}$, then $\mathbb{K}$ is locally invariant with respect to f. Furthermore, if $\mathbb{K}$ is locally invariant with respect to f, and the Cauchy problem (1.1.2) has the local existence property, which happens, for instance, when f is continuous, then $\mathbb{K}$ is viable with respect to f. However, in general, local invariance does not imply viability simply because the local invariance of a given set $\mathbb{K}$ could be a consequence of the lack of local existence for (1.1.2) for some $\xi \in \mathbb{K}$. The following example is instructive in this respect.

EXAMPLE 1.1.2. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(\xi) = \begin{cases} -1 & \text{if } \xi \geqslant 0, \\ 1 & \text{if } \xi < 0. \end{cases}$$

First we notice that the Cauchy problem $u'(t) = f(u(t))$, $u(0) = 0$ has no right solution. See Vrabie [106, Example 2.2.1, p. 57]. Then $\mathbb{K} = \{0\}$ is locally invariant, but not viable, with respect to f.

A completion of Proposition 1.1.1 is:

PROPOSITION 1.1.2. *Let $\mathbb{D}$ be a domain and let $f : \mathbb{D} \to \mathbb{R}^n$. If $\mathbb{K}$ is the union of all right* (*left*) *trajectories of* (1.1.1) *in $\mathbb{D}$, issuing from a given subset $\mathbb{C} \subseteq \mathbb{D}$, then $\mathbb{K}$ is locally right* (*left*) *invariant with respect to f. In particular, the subset $\mathbb{K} \subseteq \mathbb{D}$ is locally right* (*left*) *invariant with respect to the continuous function f if and only if it is the union of all right* (*left*) *trajectories of* (1.1.1) *issuing from $\mathbb{K}$.*

We also notice the following simple characterization of local right (left) invariance of arbitrary subsets with respect to a possibly discontinuous function.

PROPOSITION 1.1.3 (Yorke [108]). *Let $\mathbb{D}$ be a domain and $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. The subset $\mathbb{K} \subseteq \mathbb{D}$ is locally right invariant with respect to f if and only if $\mathbb{D} \setminus \mathbb{K}$ is locally left invariant with respect to f.*

It should be noticed that Proposition 1.1.3 cannot be extended to handle right viability, even if f is continuous, as the simple example below shows.

EXAMPLE 1.1.3. For $n = 1$, $\mathbb{K} = (-1, 1)$ is viable with respect to the function $f : \mathbb{R} \to \mathbb{R}$, $f \equiv 1$, but, nevertheless, $\mathbb{R} \setminus \mathbb{K}$ is neither right nor left viable with respect to f.

Sufficient conditions for local invariance were obtained by using some usual uniqueness assumptions ensuring that once a solution to (1.1.1) lies in $\mathbb{K}$, there is no other one issuing from the same initial point and which leaves $\mathbb{K}$ "immediately". Typical examples of this kind were obtained by Brezis [18], Bony [13], Redheffer [85] and Martin [66]. Brezis [18] analyzes the case when $\mathbb{D}$ is open, $\mathbb{K} \subseteq \mathbb{D}$ is relatively closed and f is locally Lipschitz, and proves that (1.1.3) with "lim" instead of "lim inf" is necessary and sufficient for $\mathbb{K}$ to be local invariant, "*flow invariant*" in his terminology, with respect to f. By means of a general concept of normal to a given set, and using a tangency condition expressed in the terms of this concept, Bony [13] get sufficient conditions for invariance. More precisely, let $\mathbb{K}$ be a given subset in $\mathbb{R}^n$. Let $\xi \in \mathbb{K}$ be such that there exists a sphere $B(\eta, \rho)$ containing ξ on its boundary but whose interior has empty intersection with $\mathbb{K}$. Then, the vector $\nu = \eta - \xi$ is *a metric normal to* $\mathbb{K}$ *at* ξ. A similar concept has been introduced and studied subsequently by Mordukhovich [69]. Coming back to local invariance we have:

THEOREM 1.1.4 (Bony [13]). *If $\mathbb{D}$ is a domain, $\mathbb{K}$ is relatively closed in $\mathbb{D}$, $f : \mathbb{D} \to \mathbb{R}^n$ is Lipschitz, and*

$$\langle \nu, f(\xi) \rangle \leqslant 0 \tag{1.1.4}$$

for each ν which is metric normal to $\mathbb{K}$ at ξ, then $\mathbb{K}$ is locally invariant with respect to f.

We notice that unlike (1.1.3) which is invariant with respect to equivalent Banach norms, (1.1.4) is not invariant with respect to equivalent Euclidean norms. Redheffer [85] extends the main results in both Brezis [18] and Bony [13], by proving:

THEOREM 1.1.5 (Redheffer [85]). *If* $\mathbb{D}$ *is a domain,* $\mathbb{K}$ *is relatively closed in* $\mathbb{D}$*,* $f:\mathbb{D}\to\mathbb{R}^n$ *is continuous and there exists a uniqueness function*[1] $\omega:\mathbb{R}_+\to\mathbb{R}$ *such that*

$$\big\langle \xi-\eta, f(\xi)-f(\eta)\big\rangle \leqslant \|\xi-\eta\|\omega\big(\|\xi-\eta\|\big), \tag{1.1.5}$$

for all $\xi,\eta\in\mathbb{D}$ *and* (1.1.4) *is satisfied, then* $\mathbb{K}$ *is locally invariant with respect to* f*.*

THEOREM 1.1.6 (Redheffer [85]). *If* $\mathbb{D}$ *is a domain,* $\mathbb{K}$ *is relatively closed in* $\mathbb{D}$*,* $f:\mathbb{D}\to\mathbb{R}^n$ *is continuous and there exists a uniqueness function* $\omega:\mathbb{R}_+\to\mathbb{R}$ *such that* (1.1.5) *is satisfied for all* $\xi,\eta\in\mathbb{D}$ *and* (1.1.3) *is satisfied at all points* $\xi\in\mathbb{K}$ *at which there exists at least one metric normal to* $\mathbb{K}$*, then* $\mathbb{K}$ *is locally invariant with respect to* f*.*

It is interesting to notice that, in all Theorems 1.1.4, 1.1.5 and 1.1.6, it may happened that $\mathbb{K}$ fails to have metric normal vectors at many of its points, at which, in spite of the fact that it is by no means evident that some solution could not escape from $\mathbb{K}$ through such a point, no condition is imposed. The explanation of this apparently strange situation consists in that, every point at which $\mathbb{K}$ fails to have metric normal vectors is an interior-like point of $\mathbb{K}$, simply because the Bony's tangent cone to $\mathbb{K}$ at such a point is all of $\mathbb{R}^n$. Finally, Martin [66] analyzes the special case when f is continuous and dissipative. For other invariance results see [3,18,25,27,38,65,66,108]. Necessary conditions, which are expressed in the terms of a tangency condition of the type (1.1.3), require also some uniqueness hypotheses. In a slightly different spirit, more closely related to dynamical systems than to differential equations, the local invariance problem was studied by Ursescu [99,100]. The main idea in [100] was to consider from the very beginning that a given abstract evolution operator which stands for the set of "all solutions" satisfies a certain tangency condition coupled with a uniqueness hypothesis. It should be mentioned that, in this general context, there is no need of a "right-hand side" f of the associated differential equation—if any.

Unlike the above mentioned approaches, Cârjă et al. [24] consider the classical differential equation (1.1.1) and look for general sufficient and even necessary conditions for local invariance expressed only in terms of f, $\mathbb{K}$ and $\mathbb{D}$, but not in the terms of the panel of solutions to (1.1.1). The conditions there obtained, although by means of a comparison function—see definition below—allow (1.1.1) to have multiple solutions in $\mathbb{K}$.

We recall that a function $\omega:\mathbb{I}\times[0,a)\to\mathbb{R}$ is a *comparison function* if $\omega(t,0)=0$ for each $t\in\mathbb{I}$, and for each $[\tau,T)\subseteq\mathbb{I}$, the only continuous solution $x:[\tau,T)\to[0,a)$ to the differential inequality $[D_+x](t)\leqslant\omega(t,x(t))$ for $t\in[\tau,T)$, satisfying $x(\tau)=0$, is the null function.

Cârjă et al. [24] show that, if there exists an open neighborhood $\mathbb{V}\subseteq\mathbb{D}$ of $\mathbb{K}$ such that f satisfies the surprisingly simple "exterior tangency" condition

$$\liminf_{h\downarrow 0}\frac{1}{h}\big[\operatorname{dist}\big(\xi+hf(t,\xi);\mathbb{K}\big)-\operatorname{dist}(\xi;\mathbb{K})\big]\leqslant\omega\big(t,\operatorname{dist}(\xi;\mathbb{K})\big) \tag{1.1.6}$$

---

[1] We recall that $\omega:\mathbb{R}_+\to\mathbb{R}$ is a *uniqueness function* if the only continuous function $\phi:[0,\delta)\to\mathbb{R}_+$ satisfying $[D^-\phi](t)\leqslant\omega(\phi(t))$, $[D^+\phi](t)\leqslant\omega(\phi(t))$ for each $t\in(0,\delta)$ and $\phi(0)=0$ is the null function. Here D^- and D^+ are the upper left and right Dini derivatives.

for each $(t,\xi) \in \mathbb{I} \times \mathbb{V}$, where ω is a certain comparison function, then $\mathbb{K}$ is locally invariant with respect to f. This condition reduces to the classical Nagumo's tangency condition when applied to $\xi \in \mathbb{K}$, and this simply because, at each such point $\xi \in \mathbb{K}$, $\operatorname{dist}(\xi;\mathbb{K}) = 0$. So, we can easily see that, whenever $\mathbb{K}$ is open, (1.1.6) is automatically satisfied for the choice $\mathbb{V} = \mathbb{K}$. More than this, they prove that, in many situations, the condition above is even necessary for local invariance. The philosophy of this result rests on the simple observation that the local invariance is equivalent to the "$(\mathbb{D},\mathbb{K})$-separating uniqueness" property defined below, while (1.1.6) implies $(\mathbb{D},\mathbb{K})$-separating uniqueness and, even viability, if f is continuous. More precisely, we say that (1.1.1) has the $(\mathbb{D},\mathbb{K})$-*separating uniqueness property* if, for each $(\tau,\xi) \in \mathbb{I} \times \mathbb{D}$ and every solutions $u, v : [\tau, T] \to \mathbb{D}$ of (1.1.1), satisfying $u(\tau) = v(\tau) = \xi$, there exists $c \in (\tau, T]$ such that both $u([\tau,c))$ and $v([\tau,c))$ are included either in $\mathbb{D} \setminus \mathbb{K}$, or in $\mathbb{K}$. In fact, if f is continuous, the Nagumo's tangency condition (1.1.3) combined with $(\mathbb{D},\mathbb{K})$-separating uniqueness is nothing else than a simple rephrasing of the local invariance property. Indeed, *if* $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, *with* $\mathbb{K}$ *closed and* $\mathbb{D}$ *open, and* $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ *is continuous, then* $\mathbb{K}$ *is locally invariant with respect to* f *if and only if* (1.1.3) *is satisfied and* (1.1.1) *has the* $(\mathbb{D},\mathbb{K})$-*separating uniqueness property*.

Notations

$B(\xi,\rho)$ — the closed ball in $\mathbb{R}^n$ centered at ξ and of radius $\rho > 0$

$\mathcal{B}_{\mathbb{K}}(\xi)$ — the Bony's tangent cone to $\mathbb{K}$ at ξ

$\mathcal{C}_{\mathbb{K}}(\xi)$ — $\{\eta \in \mathbb{R}^n;\ \lim_{h\downarrow 0,\mu\to\xi} \frac{1}{h}\operatorname{dist}(\mu + h\eta;\mathbb{K}) = 0\}$, i.e. the Clarke's tangent cone to $\mathbb{K}$ at ξ

$\mathbb{D}$ — a domain, i.e. a nonempty and open subset in $\mathbb{R}^n$

$\frac{\mathrm{d}^+}{\mathrm{d}s}$ — the right derivative

$\operatorname{dist}(\xi;\mathbb{K})$ — $\inf\{\|\xi - \eta\|;\ \eta \in \mathbb{K}\}$, i.e. the distance between $\xi \in \mathbb{R}^n$ and $\mathbb{K} \subseteq \mathbb{R}^n$

$\mathcal{F}_{\mathbb{K}}(\xi)$ — $\{\eta \in \mathbb{R}^n;\ \lim_{h\downarrow 0} \frac{1}{h}\operatorname{dist}(\xi + h\eta;\mathbb{K}) = 0\}$, i.e. the Federer's tangent cone to $\mathbb{K}$ at ξ

$F_{\tau,\xi}$ — $\{(s,u(s));\ s \geqslant \tau,\ u \in \mathcal{S}(\tau,\xi)\}$

$F_{\tau,\xi}(t)$ — $\{u(t);\ u \in \mathcal{S}(\tau,\xi)\}$

$\mathbb{I}$ — a nonempty and open interval in $\mathbb{R}$

$\mathbb{J}$ — a nonempty and non-degenerate, i.e. with nonempty interior interval in $\mathbb{R}$

$\mathbb{K}$ — a nonempty subset in $\mathbb{R}^n$

$L^p(\mathbb{S};\mathbb{R}^n)$ — the space of all equivalence classes, with respect to the λ almost everywhere equality, of measurable functions from $\mathbb{S}$ to $\mathbb{R}^n$ whose norms are p-Lebesgue integrable over $\mathbb{S}$

$L^1_{\mathrm{loc}}(\mathbb{I})$ — the space of all equivalence classes, with respect to the λ almost everywhere equality, of measurable functions from $\mathbb{I}$ to $\mathbb{R}$ whose norms are Lebesgue integrable over each compact subset in $\mathbb{I}$

l.s.c. — lower semicontinuous

$\mathbb{N}$ — the set of natural numbers, i.e. $0, 1, \dots n, \dots$

$\mathbb{N}^*$ — the set of strictly positive natural numbers, i.e. $1, 2, \dots n, \dots$

$\mathcal{N}_{\mathbb{K}}(\xi)$ — the cone of normals in the sense of Bony to $\mathbb{K}$ at ξ

$(\mathcal{N}_{\mathbb{K}}(\xi))^*$	— the conjugate of $\mathcal{N}_{\mathbb{K}}(\xi)$, i.e. the set of all $\eta \in \mathbb{R}^n$ satisfying $\langle \eta, \nu \rangle \leqslant 0$ for each $\nu \in \mathcal{N}_{\mathbb{K}}(\xi)$
ν	— a vector which is normal in the sense of Bony to $\mathbb{K}$ at ξ
$\mathbb{R}^n$	— the linear space over $\mathbb{R}$ of all n-tuples of real numbers
$\mathbb{R}$	— the set of real numbers
$\overline{\mathbb{R}}$	— $\mathbb{R} \cup \{-\infty, +\infty\}$, i.e. the extended set of real numbers
$\mathcal{S}(\tau, \xi)$	— the set of all noncontinuable solutions u of $u'(t) = f(t, u(t))$ satisfying $u(\tau) = \xi$
$\mathcal{T}_{\mathbb{K}}(\xi)$	— $\{\eta \in \mathbb{R}^n;\ \liminf_{h \downarrow 0} \frac{1}{h} \operatorname{dist}(\xi + h\eta; \mathbb{K}) = 0\}$, i.e. the Bouligand–Severi's tangent cone to $\mathbb{K}$ at ξ
u.s.c.	— upper semicontinuous
$[x, y]_+$	— the right directional derivative of the norm calculated at x in the direction y

2. Preliminaries

2.1. *Brezis–Browder ordering principle*

The goal of this subsection is to prove a general and very simple principle concerning preorder relations which unifies a number of various results in nonlinear functional analysis, principle due to Brezis and Browder [19]. We notice that this is an ordering principle similar to Zorn's lemma, but based on the *axiom of dependent choice* which, as shown by Feferman [42], turns out to be strictly weaker than *the axiom of choice*. For easy reference we recall:

THE AXIOM OF DEPENDENT CHOICE. *Let $\mathcal{S}$ be a nonempty set and let $\mathcal{R} \subseteq \mathcal{S} \times \mathcal{S}$ be a binary relation with the property that, for each $\xi \in \mathcal{S}$, the set $\{\eta \in \mathcal{S};\ \xi \mathcal{R} \eta\}$ is nonempty. Then, for each $\xi \in \mathcal{S}$, there exists a sequence $(\xi_m)_m$ in $\mathcal{S}$ such that $\xi_0 = \xi$ and $\xi_m \mathcal{R} \xi_{m+1}$ for each $m \in \mathbb{N}$.*

We notice that, in its turn, the axiom of dependent choice implies *the axiom of countable choice* stated below, which is sufficient to prove that a lot of remarkable properties in Real Analysis can be described by means of sequences. We emphasize that the axiom of dependent choice is "far enough" from the axiom of countable choice, as shown recently by Howard and Rubin [57].

THE AXIOM OF COUNTABLE CHOICE. *Let $\mathcal{S}$ be a nonempty set and let $\mathcal{F} = \{\mathbb{F}_m;\ m \in \mathbb{N}\}$ be a countable family of nonempty subsets in $\mathcal{S}$. Then, there exists a sequence $(\xi_m)_m$ with the property that $\xi_m \in \mathbb{F}_m$ for each $m \in \mathbb{N}$.*

Finally, we notice that we preferred this framework simply because the results based on the axiom of dependent choice remain true no matter which initial assumption we make, i.e. no matter if we assume that either the axiom of choice, or its negation, holds true.

Let $\mathcal{S}$ be a nonempty set. A binary relation $\preceq \subseteq \mathcal{S} \times \mathcal{S}$ is a *preorder on $\mathcal{S}$* if it is reflexive, i.e. $\xi \preceq \xi$ for each $\xi \in \mathcal{S}$, and transitive, i.e. $\xi \preceq \eta$ and $\eta \preceq \zeta$ imply $\xi \preceq \zeta$.

DEFINITION 2.1.1. Let $\mathcal{S}$ be a nonempty set, $\preceq$ a preorder on $\mathcal{S}$ and let $\mathcal{M}:\mathcal{S}\to\mathbb{R}\cup\{+\infty\}$ be an increasing function. An *$\mathcal{M}$-maximal element* is an element $\bar{\xi}\in\mathcal{S}$ satisfying $\mathcal{M}(\xi)=\mathcal{M}(\bar{\xi})$, for every $\xi\in\mathcal{S}$ with $\bar{\xi}\preceq\xi$.

THEOREM 2.1.1 (Brezis–Browder ordering principle [19]). *Let $\mathcal{S}$ be a nonempty set, $\preceq$ a preorder on $\mathcal{S}$ and let $\mathcal{M}:\mathcal{S}\to\mathbb{R}\cup\{+\infty\}$ be a function. Suppose that*:

(i) *for any increasing sequence $(\xi_k)_k$ in $\mathcal{S}$, there exists some $\eta\in\mathcal{S}$ such that $\xi_k\preceq\eta$ for all $k\in\mathbb{N}$;*

(ii) *the function $\mathcal{M}$ is increasing.*

Then, for each $\xi_0\in\mathcal{S}$ there exists an $\mathcal{M}$-maximal element $\bar{\xi}\in\mathcal{S}$ satisfying $\xi_0\preceq\bar{\xi}$.

PROOF. Suppose first that the function $\mathcal{M}$ is bounded from above. Let us consider a fixed element $\xi_0\in\mathcal{S}$ and let us construct (inductively) an increasing sequence $(\xi_k)_k$ as follows: if ξ_k is given, let us consider the set $\mathcal{S}_k=\{\xi\in\mathcal{S};\ \xi_k\preceq\xi\}$ and let us denote by $\beta_k=\sup\{\mathcal{M}(\xi);\ \xi\in\mathcal{S}_k\}$. If ξ_k satisfies the conclusion, we have nothing to prove. If not, then $\beta_k>\mathcal{M}(\xi_k)$, and so we get ξ_{k+1} such that $\xi_k\preceq\xi_{k+1}$ and

$$\mathcal{M}(\xi_{k+1})>\beta_k-\frac{\beta_k-\mathcal{M}(\xi_k)}{2}. \tag{2.1.1}$$

We have thus constructed an increasing sequence (here is the point where we have used the axiom of dependent choices) $(\xi_k)_k$ with the property that the sequence $(\mathcal{M}(\xi_k))_k$ is strictly increasing. By the assumption (i), $(\xi_k)_k$ is bounded above in $\mathcal{S}$, i.e., there exists $\bar{\xi}\in\mathcal{S}$ such that $\xi_k\preceq\bar{\xi}$ for every $k\in\mathbb{N}$. We show that this $\bar{\xi}$ satisfies the conclusion. Suppose by contradiction that there exists $\eta\in\mathcal{S}$ such that $\bar{\xi}\preceq\eta$ and $\mathcal{M}(\bar{\xi})<\mathcal{M}(\eta)$. Since $\mathcal{M}$ is bounded from above, the sequence $(\mathcal{M}(\xi_k))_k$ is bounded from above, and thus convergent. Moreover $\lim_{k\to\infty}\mathcal{M}(\xi_k)\leqslant\mathcal{M}(\bar{\xi})$ and $\eta\in\mathcal{S}_k$ for each $k\in\mathbb{N}$. Therefore we have $\beta_k\geqslant\mathcal{M}(\eta)$. From (2.1.1), we deduce

$$2\mathcal{M}(\xi_{k+1})-\mathcal{M}(\xi_k)\geqslant\beta_k\geqslant\mathcal{M}(\eta),$$

for all $k\in\mathbb{N}$. Passing to the limit as $k\to\infty$, we obtain $\mathcal{M}(\bar{\xi})\geqslant\mathcal{M}(\eta)$, which is a contradiction. This achieves the proof under the extra-condition that $\mathcal{M}$ is bounded above.

Consider now the general case, and let us define the auxiliary function $\mathcal{M}_1:\mathcal{S}\to(-\frac{\pi}{2},\frac{\pi}{2}]$ by

$$\mathcal{M}_1(\xi)=\begin{cases}\arctan\bigl(\mathcal{M}(\xi)\bigr) & \text{if } \mathcal{M}(\xi)<+\infty,\\ \dfrac{\pi}{2} & \text{if } \mathcal{M}(\xi)=+\infty.\end{cases}$$

The function $\mathcal{M}_1$ is increasing and bounded from above. Therefore there exists an element $\bar{\xi}\in\mathcal{S}$ which verifies the conclusion with $\mathcal{M}_1$ instead of $\mathcal{M}$. But $\arctan\mathcal{M}(\bar{\xi})=\arctan\mathcal{M}(\xi)$ implies $\mathcal{M}(\bar{\xi})=\mathcal{M}(\xi)$, which completes the proof in the general case. □

In its original formulation of Brezis–Browder ordering principle, it is assumed that $\mathcal{M}$ is bounded from above. In order to handle a larger class of applications this condition has been dropped in Cârjă and Ursescu [25], by obtaining the very slight extension here presented. A simple inspection of the proof shows that the conclusion of Theorem 2.1.1 remains true if (i) is replaced by the weaker condition:

(j) *For any increasing sequence* $(\xi_k)_k$ *in* $\mathcal{S}$ *with the property that the sequence* $(\mathcal{M}(\xi_k))_k$ *is strictly increasing, there exists some* $\eta \in \mathcal{S}$ *such that* $\xi_k \preceq \eta$ *for all* $k \in \mathbb{N}$.

2.2. *Projections*

We begin with:

DEFINITION 2.2.1. A subset $\mathbb{K} \subseteq \mathbb{R}^n$ is *locally closed* if for every $\xi \in \mathbb{K}$ there exists $\rho > 0$ such that $\mathbb{K} \cap B(\xi, \rho)$ is closed.

REMARK 2.2.1. Obviously every closed set is locally closed. Furthermore, if $\mathbb{D}$ is a given open subset in $\mathbb{R}^n$, then every relatively closed subset $\mathbb{K}$ in $\mathbb{D}$ is locally closed too. Indeed, if $\mathbb{K}$ is relatively closed, then $\mathbb{K} = \overline{\mathbb{K}} \cap \mathbb{D}$ and so, for every $\xi \in \mathbb{K}$ there exists $\rho > 0$ such that $B(\xi, \rho) \subseteq \mathbb{D}$. Consequently $\mathbb{K} \cap B(\xi, \rho) = \overline{\mathbb{K}} \cap \mathbb{D} \cap B(\xi, \rho) = \overline{\mathbb{K}} \cap B(\xi, \rho)$ which is closed, and this proves the assertion. Thus, each open subset $\mathbb{K}$ in $\mathbb{R}^n$ is locally closed. There exist however locally closed sets which are neither open, nor closed, nor even closed relatively to $\mathbb{D}$, as for example $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^2$ defined by $\mathbb{D} = \{(x, y) \in \mathbb{R}^2;\ x > 0\}$ and $\mathbb{K} = \{(x, y) \in \mathbb{R}^2;\ 0 < x < 1, y = 0\}$. The set $\mathbb{K}$, which is in fact a line segment in the plane xOy, is locally closed but is neither open, nor closed, nor even closed relatively to $\mathbb{D}$.

We say that $\xi \in \mathbb{R}^n$ *has projection on* $\mathbb{K}$ if there exists $\eta \in \mathbb{K}$ such that $\|\xi - \eta\| = \operatorname{dist}(\xi; \mathbb{K})$. Any $\eta \in \mathbb{K}$ enjoying the above property is called *a projection* of ξ on $\mathbb{K}$, and the set of all projections of ξ on $\mathbb{K}$ is denoted by $\Pi_{\mathbb{K}}(\xi)$.

DEFINITION 2.2.2. An open neighborhood $\mathbb{V}$ of $\mathbb{K}$, with $\Pi_{\mathbb{K}}(\xi) \neq \emptyset$ for each $\xi \in \mathbb{V}$, is called a *proximal neighborhood* of $\mathbb{K}$. If $\mathbb{V}$ is a proximal neighborhood of $\mathbb{K}$, then every single-valued selection, $\pi_{\mathbb{K}} : \mathbb{V} \to \mathbb{K}$, of $\Pi_{\mathbb{K}}$, i.e. $\pi_{\mathbb{K}}(\xi) \in \Pi_{\mathbb{K}}(\xi)$ for each $\xi \in \mathbb{V}$, is a *projection subordinated to* $\mathbb{V}$.

The next lemma, proved in Cârjă and Ursescu [25, Lemma 18], shows that for each locally closed set $\mathbb{K}$, all the points which are sufficiently close to $\mathbb{K}$ do have projections, i.e. each locally closed set $\mathbb{K}$ has one proximal neighborhood.

LEMMA 2.2.1. *Let* $\mathbb{K}$ *be locally closed. Then the set of all* $\xi \in \mathbb{R}^n$ *such that* $\Pi_{\mathbb{K}}(\xi)$ *is nonempty is a neighborhood of* $\mathbb{K}$.

PROOF. Let $\xi \in \mathbb{K}$. Since $\mathbb{K}$ is locally closed, there exists $\rho > 0$ such that $\mathbb{K} \cap B(\xi, \rho)$ is closed. To complete the proof, it suffices to show that, for each $\eta \in \mathbb{K}$ satisfying $\|\xi - \eta\| < \rho/2$, $\Pi_{\mathbb{K}}(\eta)$ is nonempty. Indeed, for each η as above, there exists a sequence $(\zeta_k)_k$ in $\mathbb{K}$ such that $(\|\zeta_k - \eta\|)_k$ converges to $\operatorname{dist}(\eta; \mathbb{K})$. We can suppose, by taking a subsequence if necessary, that $(\zeta_k)_k$ converges to a point $\zeta \in \mathbb{R}^n$. So, we have $\operatorname{dist}(\eta; \mathbb{K}) = \|\zeta - \eta\|$. Clearly, $\|\zeta_k - \xi\| \leqslant \|\zeta_k - \eta\| + \|\eta - \xi\|$ for all $k \in \mathbb{N}$, and consequently $\|\zeta - \xi\| \leqslant \operatorname{dist}(\eta; \mathbb{K}) + \|\eta - \xi\| \leqslant 2\|\eta - \xi\| < \rho$. Finally, $\|\zeta_k - \eta\| < \rho$ for all $k \in \mathbb{N}$ sufficiently large. Therefore $\zeta_k \in \mathbb{K} \cap B(\xi, \rho)$, and since the latter is closed, it follows that $\zeta \in \mathbb{K}$. Thus $\Pi_{\mathbb{K}}(\eta)$ is nonempty, and this completes the proof. □

2.3. *Tangent cones*

The next tangency concept was introduced independently by Bouligand [16] and Severi [89].

DEFINITION 2.3.1. Let $\mathbb{K} \subseteq \mathbb{R}^n$ and $\xi \in \mathbb{K}$. The vector $\eta \in \mathbb{R}^n$ is *tangent in the sense of Bouligand–Severi* to the set $\mathbb{K}$ at the point ξ if

$$\liminf_{h \downarrow 0} \frac{1}{h} \operatorname{dist}(\xi + h\eta; \mathbb{K}) = 0.$$

PROPOSITION 2.3.1. *The set $\mathcal{T}_{\mathbb{K}}(\xi)$ of all vectors which are tangent in the sense of Bouligand–Severi to the set $\mathbb{K}$ at the point ξ is a closed cone.*

PROOF. Let $\xi \in \mathbb{K}$. According to Definition 2.3.1, $\eta \in \mathcal{T}_{\mathbb{K}}(\xi)$ if

$$\liminf_{t \downarrow 0} \frac{1}{t} \operatorname{dist}(\xi + t\eta; \mathbb{K}) = 0.$$

Let $s > 0$ and let us observe that

$$\begin{aligned} \liminf_{t \downarrow 0} \frac{1}{t} \operatorname{dist}(\xi + ts\eta; \mathbb{K}) &= s \liminf_{t \downarrow 0} \frac{1}{ts} \operatorname{dist}(\xi + ts\eta; \mathbb{K}) \\ &= s \liminf_{\tau \downarrow 0} \frac{1}{\tau} \operatorname{dist}(\xi + \tau\eta; \mathbb{K}) = 0. \end{aligned}$$

Hence $s\eta \in \mathcal{T}_{\mathbb{K}}(\xi)$. In order to complete the proof, it remains to show that $\mathcal{T}_{\mathbb{K}}(\xi)$ is a closed set. To this aim let $(\eta_k)_{k \in \mathbb{N}^*}$ be a sequence of elements in $\mathcal{T}_{\mathbb{K}}(\xi)$, convergent to η. We have

$$\begin{aligned} \frac{1}{t} \operatorname{dist}(\xi + t\eta; \mathbb{K}) &\leqslant \frac{1}{t} \left\| t(\eta - \eta_k) \right\| + \frac{1}{t} \operatorname{dist}(\xi + t\eta_k; \mathbb{K}) \\ &= \|\eta - \eta_k\| + \frac{1}{t} \operatorname{dist}(\xi + t\eta_k; \mathbb{K}) \end{aligned}$$

for every $k \in \mathbb{N}^*$. So $\liminf_{t \downarrow 0} \frac{1}{t} \operatorname{dist}(\xi + t\eta; \mathbb{K}) \leqslant \|\eta - \eta_k\|$ for every $k \in \mathbb{N}^*$. Since $\lim_{k \to \infty} \|\eta - \eta_k\| = 0$, it follows that $\liminf_{t \downarrow 0} \frac{1}{t} \operatorname{dist}(\xi + t\eta; \mathbb{K}) = 0$, which achieves the proof. □

Following Bouligand [16], the cone $\mathcal{T}_{\mathbb{K}}(\xi)$ is called *the contingent cone* to the set $\mathbb{K}$ at the point ξ.

PROPOSITION 2.3.2. *A vector $\eta \in \mathbb{R}^n$ belongs to the cone $\mathcal{T}_{\mathbb{K}}(\xi)$ if and only if for every $\varepsilon > 0$ there exist $h \in (0, \varepsilon)$ and $p_h \in B(0, \varepsilon)$ with the property*

$$\xi + h(\eta + p_h) \in \mathbb{K}.$$

PROOF. Obviously $\eta \in \mathcal{T}_{\mathbb{K}}(\xi)$ if and only if, for every $\varepsilon > 0$ there exists $h \in (0, \varepsilon)$ and $z_h \in \mathbb{K}$ such that $\frac{1}{h}\|\xi + h\eta - z_h\| \leqslant \varepsilon$. Now, let us define $p_h = \frac{1}{h}(z_h - \xi - h\eta)$, and let us observe that we have both $\|p_h\| \leqslant \varepsilon$, and $\xi + h(\eta + p_h) = z_h \in \mathbb{K}$, thereby completing the proof. □

A simple but useful consequence is:

COROLLARY 2.3.1. *A vector $\eta \in \mathbb{R}^n$ belongs to the cone $\mathcal{T}_{\mathbb{K}}(\xi)$ if and only if there exist two sequences $(h_m)_m$ in $\mathbb{R}_+$ and $(p_m)_m$ in $\mathbb{R}^n$ with $h_m \downarrow 0$ and $p_m \to 0$ as $m \to \infty$, and such that $\xi + h_m(\eta + p_m) \in \mathbb{K}$ for each $m \in \mathbb{N}$.*

REMARK 2.3.1. We notice that, if ξ is an interior point of the set $\mathbb{K}$, then $\mathcal{T}_{\mathbb{K}}(\xi) = \mathbb{R}^n$. Indeed, in this case there exists $\rho > 0$ such that $B(\xi, \rho) \subset \mathbb{K}$ and, therefore, for $t > 0$ sufficiently small, $\xi + t\eta \in B(\xi, \rho) \subseteq \mathbb{K}$. Obviously, for such numbers $t > 0$, we have $\operatorname{dist}(\xi + t\eta; \mathbb{K}) = 0$, from where it follows the condition in Definition 2.3.1.

An interesting consequence of Corollary 2.3.1 is given below.

THEOREM 2.3.1. *Let $\mathbb{K}_1, \mathbb{K}_2 \subseteq \mathbb{R}^n$ be locally closed. If $\xi \in \mathbb{K}_1 \cap \mathbb{K}_2$ is an interior point of $\mathbb{K}_1 \cup \mathbb{K}_2$, then we have*

$$\mathcal{T}_{\mathbb{K}_1 \cap \mathbb{K}_2}(\xi) = \mathcal{T}_{\mathbb{K}_1}(\xi) \cap \mathcal{T}_{\mathbb{K}_2}(\xi).$$

PROOF. Obviously, for each $\xi \in \mathbb{K}_1 \cap \mathbb{K}_2$, $\mathcal{T}_{\mathbb{K}_1 \cap \mathbb{K}_2}(\xi) \subseteq \mathcal{T}_{\mathbb{K}_1}(\xi) \cap \mathcal{T}_{\mathbb{K}_2}(\xi)$. To prove that, whenever, in addition, ξ is in the interior of $\mathbb{K}_1 \cup \mathbb{K}_2$, the converse inclusion holds true, let $\xi \in \mathbb{K}_1 \cap \mathbb{K}_2$ and let $\eta \in \mathcal{T}_{\mathbb{K}_1}(\xi) \cap \mathcal{T}_{\mathbb{K}_2}(\xi)$. By Corollary 2.3.1, there exist four sequences $(h_m)_m$, $(\tilde{h}_m)_m$ in $\mathbb{R}_+$, $(p_m)_m$ and $(\tilde{p}_m)_m$ in X with $h_m \downarrow 0$, $\tilde{h}_m \downarrow 0$, $p_m \to 0$ and $\tilde{p}_m \to 0$ as $m \to \infty$, and such that $\xi + h_m(\eta + p_m) \in \mathbb{K}_1$ and $\xi + \tilde{h}_m(\eta + \tilde{p}_m) \in \mathbb{K}_2$ for each $m \in \mathbb{K}$. Now, if we assume further that ξ is in the interior of $\mathbb{K}_1 \cup \mathbb{K}_2$, there exists $\rho > 0$ such that $B(\xi, \rho) \subset \mathbb{K}_1 \cup \mathbb{K}_2$. Since $\mathbb{K}_1$ and $\mathbb{K}_2$ are locally closed, diminishing $\rho > 0$ if necessary, we may assume that both $\mathbb{K}_1 \cap B(\xi, \rho)$ and $\mathbb{K}_2 \cap B(\xi, \rho)$ are closed. Let $m_0 \in \mathbb{N}$ be such that, for each $m \geqslant m_0$, we have both $\xi_m = \xi + h_m(\eta + p_m) \in B(\xi, \rho)$ and $\tilde{\xi}_m = \xi + \tilde{h}_m(\eta + \tilde{p}_m) \in B(\xi, \rho)$. As a consequence, if $m \geqslant m_0$, the line segment $[\xi_m, \tilde{\xi}_m] \subseteq B(\xi, \rho) \subseteq \mathbb{K}_1 \cup \mathbb{K}_2$. Since $B(\xi, \rho)$ is connected, while $B(\xi, \rho) \cap \mathbb{K}_1$ and $B(\xi, \rho) \cap \mathbb{K}_2$ are closed, there exists $\eta_m \in [\xi_m, \tilde{\xi}_m] \cap \mathbb{K}_1 \cap \mathbb{K}_2$. Since $\eta_m \in [\xi_m, \tilde{\xi}_m]$, there exists $\theta_m \in [0, 1]$ such that $\eta_m = (1 - \theta_m)\xi_m + \theta_m \tilde{\xi}_m$. So, denoting

$$t_m = (1 - \theta_m)h_m + \theta_m \tilde{h}_m \quad \text{and} \quad q_m = \frac{(1 - \theta_m)h_m}{t_m} p_m + \frac{\theta_m \tilde{h}_m}{t_m} \tilde{p}_m,$$

we have

$$\eta_m = \xi + t_m(\eta + q_m) \in \mathbb{K}_1 \cap \mathbb{K}_2.$$

Finally, observing that $t_m \downarrow 0$ and $q_m \to 0$, and using Corollary 2.3.1, we get the conclusion. □

In the case when $\mathbb{K}_1, \mathbb{K}_2$ are closed subsets in a normed vector space, a similar result was obtained by Quincampoix [84, Corollary 2.3].

The next tangency concept was introduced by Federer [41]. See also Girsanov [49].

DEFINITION 2.3.2. Let $\mathbb{K} \subseteq \mathbb{R}^n$ and $\xi \in \mathbb{K}$. The vector $\eta \in \mathbb{R}^n$ is *tangent in the sense of Federer* to the set $\mathbb{K}$ at the point ξ if

$$\lim_{h \downarrow 0} \frac{1}{h} \operatorname{dist}(\xi + h\eta; \mathbb{K}) = 0.$$

The set of all points $\eta \in \mathbb{R}^n$ which are tangent in the sense of Federer to $\mathbb{K}$ at ξ is denoted by $\mathcal{F}_{\mathbb{K}}(\xi)$. This set is a cone which clearly is included in $\mathcal{T}_{\mathbb{K}}(\xi)$. A not necessarily metrical tangency concept in general topological vector spaces which, in $\mathbb{R}^n$, reduces to that one of Federer [41], is due to Ursescu [95]. As we shall show next, for a continuous function $f : \mathbb{K} \to \mathbb{R}^n$, the following surprising equivalence holds true: $f(\xi) \in \mathcal{T}_{\mathbb{K}}(\xi)$ for each $\xi \in \mathbb{K}$ if and only if $f(\xi) \in \mathcal{F}_{\mathbb{K}}(\xi)$ for each $\xi \in \mathbb{K}$, and this in spite of the fact that $\mathcal{F}_{\mathbb{K}}(\xi) \neq \mathcal{T}_{\mathbb{K}}(\xi)$.

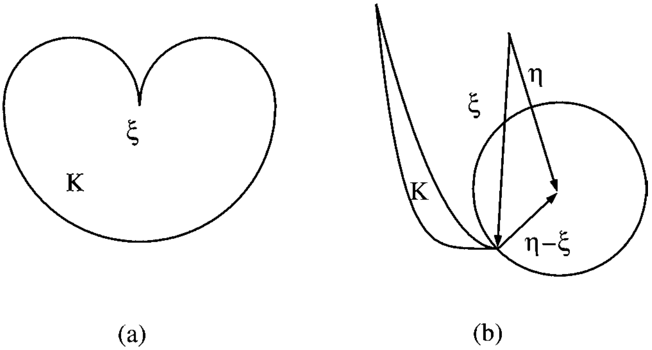
2.4. *The proximal normal cone*

We begin by recalling:

DEFINITION 2.4.1. Let $\xi \in \mathbb{K}$. We say that $\nu \in \mathbb{R}^n$ is *metric normal to* $\mathbb{K}$ *at* ξ, if there exist $\eta \in \mathbb{R}^n$ and $\rho > 0$ such that $B(\eta, \rho)$ contains ξ on its boundary, its interior has empty intersection with $\mathbb{K}$, and $\nu = \eta - \xi$.

In Fig. 1(a) is illustrated a point ξ at which there is no metric normal vector. We notice that we have $\mathcal{T}_{\mathbb{K}}(\xi) = \mathbb{R}^n$. For the case in which there is at least one normal vector to $\mathbb{K}$ at a point ξ, see Fig. 1(b).

DEFINITION 2.4.2. The *proximal normal cone to* $\mathbb{K}$ *at* $\xi \in \mathbb{K}$ is the set of all $\zeta \in \mathbb{R}^n$ of the form $\zeta = \lambda\nu$, where ν is metric normal to $\mathbb{K}$ at ξ and $\lambda \geqslant 0$, whenever such a metric normal ν exists, and $\{0\}$ if there is no metric normal to $\mathbb{K}$ at ξ. We denote this cone by $\mathcal{N}_{\mathbb{K}}(\xi)$.



(a) (b)

Fig. 1.

The use of the term "cone" in Definition 2.4.2, is justified by the simple observation that, for each $\xi \in \mathbb{K}$, $\mathcal{N}_{\mathbb{K}}(\xi)$ is a cone in the usual sense, i.e., for each $\zeta \in \mathcal{N}_{\mathbb{K}}(\xi)$ and $\lambda > 0$, we have $\lambda\zeta \in \mathcal{N}_{\mathbb{K}}(\xi)$.

The next tangency concept was introduced by Ursescu [101]. Let now $\mathbb{S}_{\mathbb{K}}(\xi)$ be the set of all $\eta \in \mathbb{R}^n$ such that $\eta - \xi$ is a metric normal to $\mathbb{K}$ at ξ whenever such a metric normal exists, and $\mathbb{S}_{\mathbb{K}}(\xi) = \{\xi\}$ otherwise. Let $\eta \in \mathbb{S}_{\mathbb{K}}(\xi)$, and let $\mathbb{E}(\xi, \eta) = \{\zeta \in \mathbb{R}^n;\ \|\eta - \zeta\| \geqslant \|\eta - \xi\|\}$. Since $\mathbb{K} \subseteq \mathbb{E}(\xi, \eta)$, for each $\eta \in \mathcal{S}_{\mathbb{K}}(\xi)$, we have

$$\mathcal{T}_{\mathbb{K}}(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi), \tag{2.4.1}$$

where

$$\mathcal{B}_{\mathbb{K}}(\xi) = \bigcap_{\eta \in \mathbb{S}_{\mathbb{K}}(\xi)} \mathcal{T}_{\mathbb{E}(\xi,\eta)}(\xi).$$

One may easily see that $\mathcal{B}_{\mathbb{K}}(\xi)$ is a cone in $\mathbb{R}^n$.

DEFINITION 2.4.3. The set $\mathcal{B}_{\mathbb{K}}(\xi)$, defined as above, is the *Bony tangent cone to* $\mathbb{K}$ *at* ξ, and its elements are the *tangents in the sense of Bony* to $\mathbb{K}$ at $\xi \in \mathbb{K}$.

Let $\|\cdot\|$ be a given norm on $\mathbb{R}^n$. If $x, y \in \mathbb{R}^n$, we denote by $[x, y]_+$ the *right directional derivative of the norm* calculated at x in the direction y, i.e.

$$[x, y]_+ = \lim_{h \downarrow 0} \frac{\|x + hy\| - \|x\|}{h}.$$

If $\|\cdot\|$ is the Euclidean norm, we have

$$[x, y]_+ = \begin{cases} \dfrac{\langle x, y\rangle}{\|x\|} & \text{if } x \neq 0, \\ \|y\| & \text{if } x = 0. \end{cases}$$

REMARK 2.4.1. Taking into account the definitions of both $\mathbb{E}(\xi,\eta)$ and $[\cdot,\cdot]_+$, we easily deduce that $\zeta \in \mathcal{T}_{\mathbb{E}(\xi,\eta)}(\xi)$ if and only if $[\xi-\eta,\zeta]_+ \geqslant 0$. Therefore, $\zeta \in \mathcal{B}_{\mathbb{K}}(\xi)$ if and only if $[-\nu,\zeta]_+ \geqslant 0$ for each ν which is a metric normal to $\mathbb{K}$ at ξ. In particular, when $\mathbb{R}^n$ is endowed with the Euclidean norm, we easily deduce that, for each $\xi \in \mathbb{K}$ and each $\eta \in \mathbb{S}_{\mathbb{K}}(\xi)$, $\mathcal{T}_{\mathbb{E}(\xi,\eta)}(\xi)$ is a closed half-space having the exterior normal $\eta-\xi$. Therefore, in this case, we have

$$\mathcal{B}_{\mathbb{K}}(\xi) = \big(\mathcal{N}_{\mathbb{K}}(\xi)\big)^*,$$

where $(\mathcal{N}_{\mathbb{K}}(\xi))^*$ is the so-called *conjugate cone* of $\mathcal{N}_{\mathbb{K}}(\xi)$, i.e.

$$\big(\mathcal{N}_{\mathbb{K}}(\xi)\big)^* = \big\{\eta \in \mathbb{R}^n;\ \langle \nu,\eta\rangle \leqslant 0,\ \text{for each } \nu \in \mathcal{N}_{\mathbb{K}}(\xi)\big\}.$$

REMARK 2.4.2. If there is no metric normal vector to $\mathbb{K}$ at ξ, we may easily see that $\mathcal{B}_{\mathbb{K}}(\xi) = \mathbb{R}^n$. See also Fig. 1(a).

2.5. *Clarke's tangent cone*

We are now ready to study another useful tangency concept defined by Clarke [32].

DEFINITION 2.5.1. Let $\mathbb{K} \subseteq \mathbb{R}^n$ and $\xi \in \mathbb{K}$. The vector $\eta \in \mathbb{R}^n$ is *tangent in the sense of Clarke* to the set $\mathbb{K}$ at the point ξ if

$$\lim_{\substack{h\downarrow 0\\ \mu\to\xi;\,\mu\in\mathbb{K}}} \frac{1}{h}\,\mathrm{dist}(\mu+h\eta;\mathbb{K}) = 0.$$

We denote by $\mathcal{C}_{\mathbb{K}}(\xi)$ the set of all vectors $\eta \in \mathbb{R}^n$ which are tangent to $\xi \in \mathbb{K}$ in the sense of Clarke. It is not difficult to check out that $\mathcal{C}_{\mathbb{K}}(\xi)$ is a closed convex cone.

REMARK 2.5.1. One may easily see that, for each $\mathbb{K}$ and each $\xi \in \mathbb{K}$, we have

$$\mathcal{C}_{\mathbb{K}}(\xi) \subseteq \mathcal{F}_{\mathbb{K}}(\xi) \subseteq \mathcal{T}_{\mathbb{K}}(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi).$$

We notice that, whenever $\mathbb{K}$ admits a classical tangent space at $\xi \in \mathbb{K}$, all tangent cones previously introduced coincide with that tangent space. This happens for instance if $\mathbb{K}$ is a C^1 curve, or surface or, even if $\mathbb{K}$ is an m-dimensional C^1 manifold. However, if $\mathbb{K}$ is not smooth enough locally around ξ, the inclusions $\mathcal{C}_{\mathbb{K}}(\xi) \subseteq \mathcal{C}_{\mathbb{K}}(\xi) \subseteq \mathcal{T}_{\mathbb{K}}(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi)$ may be strict, as the following example shows.

EXAMPLE 2.5.1 (Necula [73]). Let $\mathbb{K} \subseteq \mathbb{R}^2$ be defined as $\mathbb{K} = \mathbb{K}_1 \cup \mathbb{K}_2$, where

$$\mathbb{K}_1 = \big\{(x,y);\ (x,y) \in \mathbb{R}^2,\ y \leqslant |x|\big\}$$

and

$$\mathbb{K}_2 = \left\{ \left(0, 1/2^m\right);\ m \in \mathbb{N} \right\}$$

and let $\xi = (0, 0) \in \mathbb{K}$. Then, we have

$$\begin{cases} \mathcal{C}_{\mathbb{K}}(\xi) = \{0\}, \\ \mathcal{F}_{\mathbb{K}}(\xi) = \mathbb{K}_1, \\ \mathcal{T}_{\mathbb{K}}(\xi) = \mathbb{K}_1 \cup \left\{(0, y);\ y \geqslant 0\right\}, \\ \mathcal{B}_{\mathbb{K}}(\xi) = \mathbb{R}^2. \end{cases}$$

For a multifunction[2] $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$, we define

$$\liminf_{\substack{\xi \to \xi_0 \\ \xi \in \mathbb{K}}} F(\xi) = \left\{ \eta \in \mathbb{R}^n;\ \lim_{\substack{\xi \to \xi_0 \\ \xi \in \mathbb{K}}} \operatorname{dist}\bigl(\eta, F(\xi)\bigr) = 0 \right\}.$$

LEMMA 2.5.1. *Let us assume that the norm* $\|\cdot\|$ *on* $\mathbb{R}^n$ *satisfies*

$$[x, y]_+ = -[-x, y]_+$$

for each $x, y \in \mathbb{R}^n$.[3] *Let* $\mathbb{K} \subseteq \mathbb{R}^n$ *be locally closed. Then, for each* $\xi_0 \in \mathbb{K}$, *we have*

$$\liminf_{\xi \to \xi_0; \xi \in \mathbb{K}} \mathcal{B}_{\mathbb{K}}(\xi) = \mathcal{C}_{\mathbb{K}}(\xi_0). \tag{2.5.1}$$

PROOF. We begin by showing that

$$\liminf_{\xi \to \xi_0; \xi \in \mathbb{K}} \mathcal{B}_{\mathbb{K}}(\xi) \subseteq \mathcal{C}_{\mathbb{K}}(\xi_0). \tag{2.5.2}$$

Let $\eta \neq 0$, $\eta \in \liminf_{\xi \to \xi_0; \xi \in \mathbb{K}} \mathcal{B}_{\mathbb{K}}(\xi)$. It follows that, for each $\varepsilon > 0$, there exists $\theta > 0$ such that, for each $\phi \in \mathbb{K} \cap B(\xi_0, \theta)$, we have

$$B(\eta, \varepsilon) \cap \mathcal{B}_{\mathbb{K}}(\phi) \neq \emptyset. \tag{2.5.3}$$

Take a sufficiently small θ so that, for all $\xi \in \mathbb{K} \cap B(\xi_0, \frac{\theta}{4})$ and $t \in [0, \frac{\theta}{4\|\eta\|}]$, we have $\Pi_{\mathbb{K}}(\xi + t\eta) \neq \emptyset$. By virtue of Lemma 2.2.1, this is always possible. With ξ and t as above, let us define $g(t) = \operatorname{dist}(\xi + t\eta; \mathbb{K})$. In order to prove that $\eta \in \mathcal{C}_{\mathbb{K}}(\xi_0)$, it suffices to show that $g(t) \leqslant \varepsilon t$ for each $t \in [0, \frac{\theta}{4\|\eta\|}]$. Further, since $g(0) = 0$, it suffices to show that $g'(t) \leqslant \varepsilon$, whenever $g'(t)$ exists. To this aim, take $\phi \in \Pi_{\mathbb{K}}(\xi + t\eta)$, and let us observe that we have $\phi \in \mathbb{K} \cap B(\xi_0, \theta)$. Indeed,

$$\|\phi - \xi_0\| \leqslant \|\xi + t\eta - \phi\| + \|\xi + t\eta - \xi_0\| \leqslant 2\|\xi + t\eta - \xi_0\| \leqslant \theta,$$

---

[2] A function $F : \mathbb{K} \to \mathcal{P}(\mathbb{R}^n)$, where $\mathcal{P}(\mathbb{R}^n)$ is the class of all subsets of $\mathbb{R}^n$, is called *multifunction* and is denoted by $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$.

[3] This happens if $\|\cdot\|$ is Gâteaux differentiable.

as claimed. Now, for a sufficiently small $h > 0$, we obtain

$$g(t+h) - g(t) \leqslant \|\xi + t\eta + h\eta - \phi\| - \|\xi + t\eta - \phi\|.$$

Dividing by h and letting $h \downarrow 0$, we get

$$g'(t) \leqslant [\xi + t\eta - \phi, \eta]_+. \tag{2.5.4}$$

Taking into account that $\phi \in \Pi_{\mathbb{K}}(\xi + t\eta)$, from Definition 2.4.1, we deduce that $\xi + t\eta - \phi$ is metric normal to $\mathbb{K}$ at ϕ. In view of (2.5.3), there exists $w \in \mathcal{B}_{\mathbb{K}}(\phi)$ with $\|w - \eta\| \leqslant \varepsilon$. From Remark 2.4.1, we conclude

$$[-\xi - t\eta + \phi, w]_+ \geqslant 0.$$

According to the hypothesis, we have

$$[\xi + t\eta - \phi, \eta]_+ \leqslant [\xi + t\eta - \phi, \eta - w]_+ + [\xi + t\eta - \phi, w]_+ \leqslant \varepsilon.$$

From this inequality and (2.5.4), we get $g'(t) \leqslant \varepsilon$, as claimed. Thus (2.5.2) holds. Since $\mathcal{T}_{\mathbb{K}}(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi)$, to complete the proof, it suffices to show that

$$\mathcal{C}_{\mathbb{K}}(\xi_0) \subseteq \liminf_{\xi \to \xi_0; \xi \in \mathbb{K}} \mathcal{T}_{\mathbb{K}}(\xi).$$

So, let $\eta \in \mathcal{C}_{\mathbb{K}}(\xi_0)$ and let $\xi_m \in \mathbb{K}$ with $\xi_m \to \xi_0$. The idea is to find $\eta_m \in \mathcal{T}_{\mathbb{K}}(\xi_m)$ with $\eta_m \to \eta$. This would imply that $\eta \in \liminf_{\xi \to \xi_0; \xi \in \mathbb{K}} \mathcal{T}_{\mathbb{K}}(\xi)$. To this end, let us observe that, since $\eta \in \mathcal{C}_{\mathbb{K}}(\xi_0)$, for every $\varepsilon > 0$, there exist $m_\varepsilon \in \mathbb{N}$ and $h_\varepsilon > 0$ such that, for all $0 < h < h_\varepsilon$ and $m \geqslant m_\varepsilon$, we have

$$\operatorname{dist}(\xi_m + h\eta; \mathbb{K}) < h\varepsilon.$$

Fix m as above and take $\mu_m^h \in \mathbb{K}$ with $\|\mu_m^h - \xi_m - h\eta\| < h\varepsilon$. Let us consider

$$\eta_m^h = \frac{1}{h}\big(\mu_m^h - \xi_m\big).$$

Since $\|\eta_m^h - \eta\| \leqslant \varepsilon$, $\{\eta_m^h;\ h < 0 < h_\varepsilon\}$ is bounded and therefore it has a limit point η_m as $h \downarrow 0$. In its turn, η_m satisfies $\|\eta_m - \eta\| \leqslant \varepsilon$. A simple computational argument shows that $\eta_m \in \mathcal{T}_{\mathbb{K}}(\xi_m)$ and this completes the proof. □

REMARK 2.5.2. From Lemma 2.5.1, and Remarks 2.4.1 and 2.5.1, we easily deduce

$$\liminf_{\xi \to \xi_0; \xi \in \mathbb{K}} \mathcal{F}_{\mathbb{K}}(\xi) = \liminf_{\xi \to \xi_0; \xi \in \mathbb{K}} \mathcal{T}_{\mathbb{K}}(\xi) = \liminf_{\xi \to \xi_0; \xi \in \mathbb{K}} \mathcal{B}_{\mathbb{K}}(\xi) = \mathcal{C}_{\mathbb{K}}(\xi_0). \tag{2.5.5}$$

Concerning the equalities (2.5.5), as far as we know, the main and most difficult part, i.e. the equality of the second term to the last one has been proved independently by Ursescu [96] and by Cornet [36] in $\mathbb{R}^n$. See also Treiman [93] for the proof of the inclusion $\liminf_{\xi\to\xi_0;\xi\in\mathbb{K}} \mathcal{T}_{\mathbb{K}}(\xi) \subseteq \mathcal{C}_{\mathbb{K}}(\xi_0)$ in general Banach spaces. We emphasize that there are examples showing that the converse inclusion does not hold in infinite-dimensional Banach spaces. See, for instance, Treiman [93]. We notice that Ursescu [96] proves a characterization of $\mathcal{C}_{\mathbb{K}}(\xi)$ in general Banach spaces, pertaining also an immediate proof of Treiman's main result in [93].

PROPOSITION 2.5.1. *Let $\mathbb{K} \subseteq \mathbb{R}^n$ be locally closed and let $f:\mathbb{K}\to\mathbb{R}^n$ be continuous. Then, the following conditions are equivalent*:

(i) *for each $\xi\in\mathbb{K}$, $f(\xi)\in\mathcal{C}_{\mathbb{K}}(\xi)$,*
(ii) *for each $\xi\in\mathbb{K}$, $f(\xi)\in\mathcal{T}_{\mathbb{K}}(\xi)$,*
(iii) *for each $\xi\in\mathbb{K}$, $f(\xi)\in\mathcal{B}_{\mathbb{K}}(\xi)$.*

In general, if $\mathcal{G}:\mathbb{K}\rightsquigarrow\mathbb{R}^n$ is such that $\mathcal{C}_{\mathbb{K}}(\xi)\subseteq\mathcal{G}(\xi)\subseteq\mathcal{B}_{\mathbb{K}}(\xi)$ for each $\xi\in\mathbb{K}$, then each one of the conditions above is equivalent to:

(iv) *for each $\xi\in\mathbb{K}$, $f(\xi)\in\mathcal{G}(\xi)$.*

PROOF. In view of Remark 2.5.1, it suffices to show that (i) is equivalent to (iii). But this easily follows from Lemma 2.5.1 and this completes the proof. □

We notice that the equivalences in Proposition 2.5.1 were called to our attention by Ursescu [101].

3. Problems of viability

3.1. *Nagumo's viability theorem*

We begin with some background material we will need subsequently.

DEFINITION 3.1.1. Let $\mathbb{K}\subseteq\mathbb{R}^n$ be nonempty and $f:\mathbb{I}\times\mathbb{K}\to\mathbb{R}^n$. The subset $\mathbb{K}$ is *viable* with respect to f if for every $(\tau,\xi)\in\mathbb{I}\times\mathbb{K}$ there exists $T\in\mathbb{I}$, $T>\tau$, such that (1.1.1) has at least one solution, $u:[\tau,T]\to\mathbb{K}$, satisfying $u(\tau)=\xi$.

We can now proceed to the main result in this section.

THEOREM 3.1.1 (Nagumo [70]). *Let $\mathbb{K}\subset\mathbb{R}^n$ be a nonempty and locally closed set and let $f:\mathbb{I}\times\mathbb{K}\to\mathbb{R}^n$ be a continuous function. The necessary and sufficient condition in order that $\mathbb{K}$ be viable with respect to f is that, for every $(t,\xi)\in\mathbb{I}\times\mathbb{K}$, $f(t,\xi)\in\mathcal{T}_{\mathbb{K}}(\xi)$.*

3.2. *Proof of the necessity*

PROOF. In order to prove the necessity let $(t,\xi)\in\mathbb{I}\times\mathbb{K}$. Then, there exist $T\in\mathbb{I}$, $T>t$, and a function $u:[t,T]\to\mathbb{K}$, satisfying $u(t)=\xi$ and $u'(s)=f(s,u(s))$ for every $s\in[t,T]$. Consequently, we have

$$\lim_{h\downarrow 0}\frac{1}{h}\left\|\xi+hf(t,\xi)-u(t+h)\right\|=\lim_{h\downarrow 0}\left\|f\big(t,u(t)\big)-\frac{u(t+h)-u(t)}{h}\right\|=0.$$

But, this relation shows that, for every $(t,\xi)\in\mathbb{I}\times\mathbb{K}$, $f(t,\xi)\in\mathcal{T}_{\mathbb{K}}(\xi)$ and the proof of the necessity is complete.

We notice that, in fact, we have proved much more than claimed by the necessity part of Theorem 3.1.1. Namely, we deduced:

THEOREM 3.2.1. *If $\mathbb{K}$ is viable with respect to $f:\mathbb{I}\times\mathbb{K}\to\mathbb{R}^n$, then, for each $(t,\xi)\in\mathbb{I}\times\mathbb{K}$, we have*

$$\lim_{h\downarrow 0}\frac{1}{h}\operatorname{dist}\big(\xi+hf(t,\xi);\mathbb{K}\big)=0.$$

We notice that in Theorem 3.2.1 neither $\mathbb{K}$ is assumed to be locally closed, nor f to be continuous.

3.3. *Existence of approximate solutions*

PROOF OF THEOREM 3.1.1 (Continued). For the sake of simplicity, we divide the proof of the sufficiency into three steps. In the first one we shall prove the existence of a family of approximate solutions for the Cauchy problem (1.1.2) defined on intervals of the form $[\tau,c]$, with $c\in\mathbb{I}$. In the second step we will show that the problem (1.1.2) admits such approximate solutions, all defined on an interval $[\tau,T]$ independent of the "approximation order". Finally, in the last step, we shall prove the uniform convergence on $[\tau,T]$ of a sequence of such approximate solutions to a solution of the problem (1.1.2).

Let $(\tau,\xi)\in\mathbb{I}\times\mathbb{K}$ be arbitrary and let us choose $\rho>0$, $M>0$ and $T\in\mathbb{I}$, $T>\tau$, such that $B(\xi,\rho)\cap\mathbb{K}$ be closed,

$$\left\|f(t,x)\right\|\leqslant M \tag{3.3.1}$$

for every $t\in[\tau,T]$ and $x\in B(\xi,\rho)\cap\mathbb{K}$, and

$$(T-\tau)(M+1)\leqslant\rho. \tag{3.3.2}$$

The existence of these three numbers is ensured by the fact that $\mathbb{K}$ is locally closed (from where it follows the existence of $\rho>0$), by the continuity of f which implies its boundedness on $[\tau,T]\times B(\xi,\rho)$, and so the existence of $M>0$, and by the fact that $T\in\mathbb{I}$, $T>\tau$, can be chosen as close to τ as we wish. We will show that, once fixed an $\varepsilon \in (0,1)$ and the three numbers above, there exist three functions: $\sigma : [\tau, T] \to [\tau, T]$, nondecreasing, $g : [\tau, T] \to \mathbb{R}^n$, Riemann integrable and $u : [\tau, T] \to \mathbb{R}^n$, continuous, such that:

(i) $t - \varepsilon \leqslant \sigma(t) \leqslant t$ for every $t \in [\tau, T]$;
(ii) $\|g(t)\| \leqslant \varepsilon$ for every $t \in [\tau, T]$;
(iii) $u(\sigma(t)) \in B(\xi, \rho) \cap \mathbb{K}$ for every $t \in [\tau, T]$ and $u(T) \in B(\xi, \rho) \cap \mathbb{K}$;
(iv) u satisfies

$$u(t) = \xi + \int_\tau^t f\big(\sigma(s), u\big(\sigma(s)\big)\big)\,\mathrm{d}s + \int_\tau^t g(s)\,\mathrm{d}s$$

for every $t \in [\tau, T]$.

For the sake of simplicity, in all that follows, we will say that such a triple (σ, g, u) is an *ε-approximate solution* to the Cauchy problem (1.1.2) on the interval $[\tau, T]$. □

The first step. Let $\tau \in \mathbb{I}$, $\xi \in \mathbb{K}$ and let $\rho > 0$, $M > 0$ and $T \in \mathbb{I}$, $T > \tau$, be fixed as above. Let $\varepsilon \in (0,1)$. We begin by showing the existence of an ε-approximate solution on an interval $[\tau, c]$ with $c \in (\tau, T]$. Since for every $(t, \xi) \in \mathbb{I} \times \mathbb{K}$, $f(t, \xi) \in \mathcal{T}_{\mathbb{K}}(\xi)$, from Proposition 2.3.2, it follows that there exists $c \in (\tau, T]$, $c - \tau \leqslant \varepsilon$ and $p \in \mathbb{R}^n$ with $\|p\| \leqslant \varepsilon$ such that $\xi + (c - \tau) f(\tau, \xi) + (c - \tau) p \in \mathbb{K}$. At this point, we can define the functions $\sigma : [\tau, c] \to [\tau, c]$, $g : [\tau, c] \to \mathbb{R}^n$ and $u : [\tau, c] \to \mathbb{R}^n$ by

$$\begin{cases} \sigma(t) = \tau & \text{for } t \in [\tau, c], \\ g(t) = p & \text{for } t \in [\tau, c], \\ u(t) = \xi + (t - \tau) f(\tau, \xi) + (t - \tau) p & \text{for } t \in [\tau, c]. \end{cases}$$

One can readily see that the triple (σ, g, u) is an ε-approximate solution to the Cauchy problem (1.1.2) on the interval $[\tau, c]$. Indeed the conditions (i), (ii) and (iv) are obviously fulfilled, while (iii) follows from (3.3.1), (3.3.2) and (i). Indeed, let us observe that $u(\sigma(t)) = \xi$, and therefore we have $u(\sigma(t)) \in B(\xi, \rho) \cap \mathbb{K}$ for every $t \in [\tau, c]$. Obviously $u(c) \in \mathbb{K}$. Moreover, by (3.3.1) and (3.3.2), we have

$$\big\|u(c) - \xi\big\| \leqslant (c - \tau)\big\|f(\tau, \xi)\big\| + (c - \tau)\|p\| \leqslant (T - \tau)(M + 1) \leqslant \rho$$

for every $t \in [\tau, c]$. Thus (iii) is also satisfied.

The second step. Now, we will prove the existence of an ε-approximate solution defined on the whole interval $[\tau, T]$. To this aim we shall make use of Brezis–Browder Theorem 2.1.1, as follows. Let $\mathcal{S}$ be the set of all ε-approximate solutions to the problem (1.1.2) having the domains of definition of the form $[\tau, c]$ with $c \in (\tau, T]$. On $\mathcal{S}$ we define the relation "$\preceq$" by $(\sigma_1, g_1, u_1) \preceq (\sigma_2, g_2, u_2)$ if the domain of definition $[\tau, c_1]$ of the first triple is included in the domain of definition $[\tau, c_2]$ of the second triple, and the two ε-approximate solutions coincide on the common part of the domains. Obviously "$\preceq$" is a preorder relation on $\mathcal{S}$. Let us show first that each increasing sequence $((\sigma_m, g_m, u_m))_{m \in \mathbb{N}}$ is bounded from above. Indeed, let $((\sigma_m, g_m, u_m))_m$ be an increasing sequence, and let $c^* = \lim_m c_m$, where $[\tau, c_m]$ denotes the domain of definition of (σ_m, g_m, u_m). Clearly, $c^* \in (\tau, T]$. We will show that there exists at least one element, $(\sigma^*, g^*, u^*) \in \mathcal{S}$, defined on $[\tau, c^*]$ and satisfying $(\sigma_m, g_m, u_m) \preceq (\sigma^*, g^*, u^*)$ for each $m \in \mathbb{N}$. In order to do this, we have to prove first that there exists $\lim_m u_m(c_m)$. For each $m, k \in \mathbb{N}$, $m \leqslant k$, we have $u_m(s) = u_k(s)$ for all $s \in [\tau, c_m]$. Taking into account (iii), (iv) and (3.3.1), we deduce

$$\begin{aligned}\left\|u_m(c_m) - u_k(c_k)\right\| &\leqslant \int_{c_m}^{c_k} \left\|f\big(\sigma_k(s), u_k\big(\sigma_k(s)\big)\big)\right\| \mathrm{d}s + \int_{c_m}^{c_k} \left\|g_k(s)\right\| \mathrm{d}s \\ &\leqslant (M + \varepsilon)|c_k - c_m|\end{aligned}$$

for every $m, k \in \mathbb{N}$, which proves that there exists $\lim_{m\to\infty} u_m(c_m)$. Since for every $m \in \mathbb{N}$, $u_m(c_m) \in B(\xi, \rho) \cap \mathbb{K}$, and the latter is closed, it readily follows that $\lim_{m\to\infty} u_m(c_m) \in B(\xi, \rho) \cap \mathbb{K}$. Furthermore, because all the functions in the set $\{\sigma_m;\ m \in \mathbb{N}\}$ are nondecreasing, with values in $[\tau, c^*]$, and satisfy $\sigma_m(c_m) \leqslant \sigma_p(c_p)$ for every $m, p \in \mathbb{N}$ with $m \leqslant p$, there exists $\lim_{m\to\infty} \sigma_m(c_m)$ and this limit belongs to $[\tau, c^*]$. This shows that we can define the triple of functions $(\sigma^*, g^*, u^*) : [\tau, c^*] \to [\tau, c^*] \times \mathbb{R}^n \times \mathbb{R}^n$ by

$$\sigma^*(t) = \begin{cases} \sigma_m(t) & \text{for } t \in [\tau, c_m],\ m \in \mathbb{N}, \\ \lim\limits_{m\to\infty} \sigma_m(c_m) & \text{for } t = c^*, \end{cases}$$

$$g^*(t) = \begin{cases} g_m(t) & \text{for } t \in [\tau, c_m],\ m \in \mathbb{N}, \\ 0 & \text{for } t = c^*, \end{cases}$$

$$u^*(t) = \begin{cases} u_m(t) & \text{for } t \in [\tau, c_m],\ m \in \mathbb{N}, \\ \lim\limits_{m\to\infty} u_m(c_m) & \text{for } t = c^*. \end{cases}$$

One can easily see that (σ^*, g^*, u^*) is an ε-approximate solution which is a majorant for $((\sigma_m, g_m, u_m))_m$. Let us define the function $\mathcal{M} : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$ by $\mathcal{M}((\sigma, g, u)) = c$, where $[\tau, c]$ is the domain of definition of (σ, g, u). Clearly $\mathcal{M}$ satisfies the hypotheses of Brezis–Browder Theorem 2.1.1. Then, $\mathcal{S}$ contains at least one $\mathcal{M}$-maximal element $(\bar\sigma, \bar g, \bar u)$, defined on $[\tau, \bar c]$. In other words, if $(\tilde\sigma, \tilde g, \tilde u) \in \mathcal{S}$, defined on $[\tau, \tilde c]$, satisfies $(\bar\sigma, \bar g, \bar u) \preceq (\tilde\sigma, \tilde g, \tilde u)$, then we necessarily have $\bar c = \tilde c$. We will show next that $\bar c = T$. Indeed, let us assume by contradiction that $\bar c < T$. Then, taking into account the fact that $\bar u(\bar c) \in B(\xi, \rho) \cap \mathbb{K}$, we deduce that

$$\begin{aligned}\left\|\bar u(\bar c) - \xi\right\| &\leqslant \int_\tau^{\bar c} \left\|f\big(\bar\sigma(s), \bar u\big(\bar\sigma(s)\big)\big)\right\| \mathrm{d}s + \int_\tau^{\bar c} \left\|\bar g(s)\right\| \mathrm{d}s \leqslant (\bar c - \tau)(M + \varepsilon) \\ &\leqslant (\bar c - \tau)(M + 1) < (T - \tau)(M + 1) \leqslant \rho.\end{aligned}$$

Then, as $\bar u(\bar c) \in \mathbb{K}$ and $f(\bar c, \bar u(\bar c)) \in \mathcal{T}_{\mathbb{K}}(\bar u(\bar c))$, there exist $\delta \in (0, T - \bar c)$, $\delta \leqslant \varepsilon$ and $p \in \mathbb{R}^n$ such that $\|p\| \leqslant \varepsilon$ and $\bar u(\bar c) + \delta f(\bar c, \bar u(\bar c)) + \delta p \in \mathbb{K}$. From the inequality above, it follows that we can diminish δ if necessary, in order to have $\|\bar{u}(\bar{c})+\delta[f(\bar{c},\bar{u}(\bar{c}))+p]-\xi\|\leqslant\rho$. Let us define the functions $\sigma:[\tau,\bar{c}+\delta]\to[\tau,\bar{c}+\delta]$ and $g:[\tau,\bar{c}+\delta]\to\mathbb{R}^n$ by

$$\sigma(t)=\begin{cases}\bar{\sigma}(t) & \text{for } t\in[\tau,\bar{c}],\\ \bar{c} & \text{for } t\in(\bar{c},\bar{c}+\delta],\end{cases}$$

$$g(t)=\begin{cases}\bar{g}(t) & \text{for } t\in[\tau,\bar{c}],\\ p & \text{for } t\in(\bar{c},\bar{c}+\delta].\end{cases}$$

Clearly, g is Riemann integrable on $[\tau,\bar{c}+\delta]$ and $\|g(t)\|\leqslant\varepsilon$ for every $t\in[\tau,\bar{c}+\delta]$. In addition, for every $t\in[\tau,\bar{c}+\delta]$, $\sigma(t)\in[\tau,\bar{c}]$, and therefore $\bar{u}(\sigma(t))$ is well-defined and belongs to the set $B(\xi,\rho)\cap\mathbb{K}$. Accordingly, we can define $u:[\tau,\bar{c}+\delta]\to\mathbb{R}^n$ by

$$u(t)=\xi+\int_\tau^t f\big(\sigma(s),\bar{u}\big(\sigma(s)\big)\big)\,\mathrm{d}s+\int_\tau^t g(s)\,\mathrm{d}s$$

for every $t\in[\tau,\bar{c}+\delta]$. Clearly u coincides with $\bar{u}$ on $[\tau,\bar{c}]$ and then it readily follows that u, σ and g satisfy all the conditions in (i) and (ii). In order to prove (iii) and (iv), let us observe that

$$u(t)=\begin{cases}\bar{u}(t) & \text{for } t\in[\tau,\bar{c}],\\ u(\bar{c})+(t-\bar{c})f\big(\bar{c},\bar{u}(\bar{c})\big)+(t-\bar{c})p & \text{for } t\in(\bar{c},\bar{c}+\delta].\end{cases}$$

Then u satisfies the equation in (iv). Since

$$u\big(\sigma(t)\big)=\begin{cases}\bar{u}\big(\bar{\sigma}(t)\big) & \text{for } t\in[\tau,\bar{c}],\\ \bar{u}(\bar{c}) & \text{for } t\in[\bar{c},\bar{c}+\delta],\end{cases}$$

it follows that $u(\sigma(t))\in B(\xi,\rho)\cap\mathbb{K}$. Furthermore, from the choice of δ and p, we have both $u(\bar{c}+\delta)=\bar{u}(\bar{c})+\delta f(\bar{c},\bar{u}(\bar{c}))+\delta p\in\mathbb{K}$, and

$$\big\|u(\bar{c}+\delta)-\xi\big\|=\big\|\bar{u}(\bar{c})+\delta f\big(\bar{c},\bar{u}(\bar{c})\big)+\delta p-\xi\big\|\leqslant\rho$$

and consequently u satisfies (iii). Thus $(\sigma,g,u)\in\mathcal{S}$. Furthermore, since $(\bar{\sigma},\bar{g},\bar{u})\preceq(\sigma,g,u)$ and $\bar{c}<\bar{c}+\delta$, it follows that $(\bar{\sigma},\bar{g},\bar{u})$ is not an $\mathcal{M}$-maximal element. But this is absurd. This contradiction can be eliminated only if each maximal element in the set $\mathcal{S}$ is defined on $[\tau,T]$.

3.4. *Convergence of approximate solutions*

The third step. Let $(\varepsilon_k)_{k\in\mathbb{N}}$ be a sequence from $(0,1)$ decreasing to 0 and let $((\sigma_k,g_k,u_k))_{k\in\mathbb{N}}$ be a sequence of ε_k-approximate solutions defined on $[\tau,T]$. From (i) and (ii), it follows that

$$\lim_{k\to\infty}\sigma_k(t)=t\quad\text{and}\quad\lim_{k\to\infty}g_k(t)=0\tag{3.4.1}$$

uniformly on $[\tau, T]$. On the other hand, from (iii), (iv) and (3.3.2), we have

$$\begin{aligned}\|u_k(t)\| &\leqslant \|u_k(t)-\xi\| + \|\xi\| \\ &\leqslant \int_\tau^T \|f(\sigma_k(s), u_k(\sigma_k(s)))\|\,\mathrm{d}s + \int_\tau^T \|g_k(s)\|\,\mathrm{d}s + \|\xi\| \\ &\leqslant (T-\tau)(M+1) + \|\xi\| \leqslant \rho + \|\xi\|\end{aligned}$$

for every $k \in \mathbb{N}$ and every $t \in [\tau, T]$. Hence, the sequence $(u_k)_{k\in\mathbb{N}}$ is uniformly bounded on $[\tau, T]$. Again from (iv), we have

$$\begin{aligned}\|u_k(t)-u_k(s)\| &\leqslant \left|\int_s^t \|f(\sigma_k(\tau), u_k(\sigma_k(\tau)))\|\,\mathrm{d}\tau\right| + \left|\int_s^t \|g_k(\tau)\|\,\mathrm{d}\tau\right| \\ &\leqslant (M+1)|t-s|\end{aligned}$$

for every $t, s \in [\tau, T]$. Consequently the sequence $(u_k)_{k\in\mathbb{N}}$ is equicontinuous on $[\tau, T]$. From Arzelà–Ascoli theorem—see Vrabie [106, Theorem 8.2.1, p. 320]—it follows that, at least on a subsequence, $(u_k)_{k\in\mathbb{N}}$ is uniformly convergent on $[\tau, T]$ to a function $u : [\tau, T] \to \mathbb{R}^n$. Taking into account of (iii), (3.4.1) and of the fact that $B(\xi, \rho) \cap \mathbb{K}$ is closed, we deduce that $u(t) \in B(\xi, \rho) \cap \mathbb{K}$ for every $t \in [\tau, T]$. Passing to the limit in the equation

$$u_k(t) = \xi + \int_\tau^t f(\sigma_k(\tau), u_k(\sigma_k(s)))\,\mathrm{d}s + \int_\tau^t g_k(s)\,\mathrm{d}s$$

and taking into account (3.4.1), we deduce that

$$u(t) = \xi + \int_\tau^t f(s, u(s))\,\mathrm{d}s$$

for every $t \in [\tau, T]$, which achieves the proof of the theorem. □

From Remark 2.3.1 combined with Theorem 3.1.1, we deduce Peano's local existence theorem 1.1.1. We mention that Theorem 1.1.2 too is a direct consequence of Theorem 3.1.1 combined with the observation below.

REMARK 3.4.1. Let $\mathbb{D}$ be a nonempty and open subset in $\mathbb{R}^n$ and let $U : \mathbb{D} \to \mathbb{R}$ be a function of class C^1 with $\nabla U(\xi) \neq 0$ on $\mathbb{D}$. Let $c \in \mathbb{R}$ be such that $\mathbb{K} = \{\xi \in \Omega;\ U(\xi) = c\}$ is nonempty. Then $\eta \in \mathbb{R}^n$ is tangent to $\mathbb{K}$ at the point $\xi \in \mathbb{K}$ if and only if $\langle \eta, \nabla U(\xi)\rangle = 0$. In other words, in this case, $\mathcal{T}_{\mathbb{K}}(\xi)$ coincides with the set of vectors in the tangent plane to $\mathbb{K}$ at ξ. Indeed, let us observe that a vector $\eta \in \mathcal{T}_{\mathbb{K}}(\xi)$ if and only if there exists a function $u : [0, 1] \to \mathbb{K}$ with $u(0) = \xi$, differentiable at $t = 0$, with $u'(0) = \eta$, and such that

$$\lim_{t\downarrow 0} \frac{1}{t} \|\xi + t\eta - u(t)\| = 0.$$

But, in the particular case of the set $\mathbb{K}$ considered, this relation holds if and only if $\langle \eta, \nabla U(\xi)\rangle = 0$, which achieves the proof.

Summing up, and using Proposition 2.5.1, we deduce:

THEOREM 3.4.1. *Let* $\mathbb{K}$ *be a nonempty and locally closed subset in* $\mathbb{R}^n$ *and let* $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ *be continuous. Then the following conditions are equivalent*:

(i) *for every* $(t, \xi) \in \mathbb{I} \times \mathbb{K}$, $f(t, \xi) \in \mathcal{C}_{\mathbb{K}}(\xi)$,
(ii) *for every* $(t, \xi) \in \mathbb{I} \times \mathbb{K}$, $f(t, \xi) \in \mathcal{T}_{\mathbb{K}}(\xi)$,
(iii) *for every* $(t, \xi) \in \mathbb{I} \times \mathbb{K}$, $f(t, \xi) \in \mathcal{B}_{\mathbb{K}}(\xi)$,
(iv) *the set* $\mathbb{K}$ *is viable with respect to* f.

In general, if $\mathcal{G} : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ *is such that* $\mathcal{C}_{\mathbb{K}}(\xi) \subseteq \mathcal{G}(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi)$ *for each* $\xi \in \mathbb{K}$, *then each one of the conditions above is equivalent to*

(v) *for each* $\xi \in \mathbb{K}$, $f(t, \xi) \in \mathcal{G}(\xi)$.

3.5. *Existence of noncontinuable solutions*

In this section we will prove some results concerning the existence of noncontinuable, or even global solutions to (1.1.1). We recall that a solution $u : [\tau, T) \to \mathbb{K}$ to (1.1.1) is called *noncontinuable*, if there is no other solution $v : [\tau, \widetilde{T}) \to \mathbb{K}$ of the same equation, with $T < \widetilde{T}$ and satisfying $u(t) = v(t)$ for all $t \in [\tau, T)$. The solution u is called *global* if $T = \sup \mathbb{I}$. The next theorem follows from Brezis–Browder Theorem 2.1.1.

THEOREM 3.5.1. *Let* $\mathbb{K} \subseteq \mathbb{R}^n$ *be nonempty and let* $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ *be a possibly discontinuous function. Then, the following conditions are equivalent*:

(i) $\mathbb{K}$ *is viable with respect to* f,
(ii) *for each* $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ *there exists at least one noncontinuable solution* $u : [\tau, T) \to \mathbb{K}$ *of* (1.1.1), *satisfying* $u(\tau) = \xi$.

PROOF. Clearly (ii) implies (i). To prove that (i) implies (ii) it suffices to show that every solution u can be continued up to a noncontinuable one. To this aim, we will make use of Brezis–Browder Theorem 2.1.1. Let $\mathcal{S}$ be the set of all solutions to (1.1.1). On $\mathcal{S}$ which, by virtue of (i), is nonempty, we define the binary relation "$\preceq$" by $u \preceq v$ if the domain $[\tau, T_v)$ of v is larger that the domain $[\tau, T_u)$ of u, i.e. $T_u \leqslant T_v$, and $u(t) = v(t)$ for all $t \in [\tau, T_u)$. Clearly "$\preceq$" is a preorder on $\mathcal{S}$. Next, let $(u_m)_m$ be an increasing sequence in $\mathcal{S}$, and let us denote by $[\tau, T_m)$ the domain of definition of u_m. Let $T^* = \lim_{m\to\infty} T_m$, which can be finite, or not, and let us define $u^* : [\tau, T^*) \to \mathbb{K}$ by $u^*(t) = u_m(t)$ for each $t \in [\tau, T_m)$. Since $(T_m)_m$ is increasing and $u_m(t) = u_k(t)$ for each $m \leqslant k$ and each $t \in [\tau, T_m)$, u^* is well-defined and belongs to $\mathcal{S}$. Moreover, u^* is a majorant of $(u_m)_m$. Thus each increasing sequence in $\mathcal{S}$ is bounded from above. Moreover, the function $\mathcal{M} : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$, defined by $\mathcal{M}(v) = T_v$, for each $v \in \mathcal{S}$, is monotone, and therefore we are in the hypotheses of Theorem 2.1.1. Accordingly, for $u \in \mathcal{S}$, there exists at least one element $\bar{u} \in \mathcal{S}$ with $u \preceq \bar{u}$ and, in addition, $\bar{u} \preceq \tilde{u}$ implies $T_{\tilde{u}} = T_{\bar{u}}$. But this means that $\bar{u}$ is noncontinuable, and, of course, that it extends u. The proof is complete. □

THEOREM 3.5.2. *Let $\mathbb{K}$ be a nonempty and locally closed subset in $\mathbb{R}^n$, and let $f: \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be a continuous function. Then a necessary and sufficient condition in order that for each $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ there exists at least one noncontinuable solution to* (1.1.1) *satisfying $u(\tau) = \xi$ is one of the five equivalent conditions in Theorem* 3.4.1.

PROOF. In view of Theorem 3.4.1, each one of the first five conditions is equivalent to the viability of $\mathbb{K}$ with respect to f. The conclusion follows from Theorem 3.5.1, and the proof is complete. □

In order to obtain global existence, some extra growth conditions on f are needed. We discuss below a very natural one, introduced in a more general framework in Vrabie [102, Definition 3.2.1, p. 95].

DEFINITION 3.5.1. A function $f: \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ is called *positively sublinear* if there exists a norm $\|\cdot\|$ on $\mathbb{R}^n$ such that, for each $T \in \mathbb{I}$, there exist $a > 0$, $b \in \mathbb{R}$ and $c > 0$ satisfying

$$\big\|f(t,\xi)\big\| \leqslant a\|\xi\| + b$$

for each $(t, \xi) \in \mathbb{K}_+^c(f)$, where

$$\mathbb{K}_+^c(f) = \big\{(t,\xi) \in \mathbb{I} \times \mathbb{K};\ t \leqslant T,\ \|\xi\| > c \text{ and } \big[\xi, f(t,\xi)\big]_+ > 0\big\}.$$

As concerns the existence of global solutions we have:

THEOREM 3.5.3. *Let $\mathbb{K}$ be a nonempty and closed subset in $\mathbb{R}^n$, and let $f: \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be a continuous and positively sublinear function. Then a necessary and sufficient condition in order that for each $(t, \xi) \in \mathbb{I} \times \mathbb{K}$ there exists at least one global solution to* (1.1.1) *satisfying $u(\tau) = \xi$ is each one of the five equivalent conditions in Theorem* 3.4.1.

PROOF. Clearly, each one of the five conditions is necessary. To complete the proof it suffices to show that, whenever $\mathbb{K}$ is viable with respect to f, then, for each $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$, there exists at least one global solution $u: [\tau, T) \to \mathbb{K}$ to (1.1.1) satisfying $u(\tau) = \xi$. To this aim, let $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ and let $u: [\tau, T) \to \mathbb{K}$ be a noncontinuable solution to (1.1.1) satisfying $u(\tau) = \xi$. We will show that $T = \sup \mathbb{I}$. To this aim, let us assume the contrary, i.e., that $T < \sup \mathbb{I}$. In particular, this means that $T < +\infty$. Since $u'(s) = f(s, (u(s))$ for all $s \in [\tau, T)$, we deduce

$$\big[u(s), u'(s)\big]_+ = \big[u(s), f\big(s, u(s)\big)\big]_+.$$

Since $[u(s), u'(s)]_+ = \frac{\mathrm{d}^+}{\mathrm{d}s}(\|u(s)\|)$ for $s \in [\tau, T)$, where $\|\cdot\|$ is the norm whose existence is ensured by Definition 3.5.1, integrating from τ to t the last equality, we get successively

$$\big\|u(t)\big\| = \|\xi\| + \int_\tau^t \big[u(s), f\big(s, u(s)\big)\big]_+ \mathrm{d}s$$

$$
\begin{aligned}
&= \|\xi\| + \int_{\{\tau \leqslant s \leqslant t;\, \|u(s)\| \leqslant c\}} \big[u(s), f\big(s, u(s)\big)\big]_+ \, \mathrm{d}s \\
&\quad + \int_{\{\tau \leqslant s \leqslant t;\, \|u(s)\| > c\}} \big[u(s), f\big(s, u(s)\big)\big]_+ \, \mathrm{d}s \\
&\leqslant \|\xi\| + c\lambda\big(\big\{s \in [\tau, t];\ \big\|u(s)\big\| \leqslant c\big\}\big) \\
&\quad + \int_{\{\tau \leqslant s \leqslant t;\, \|u(s)\| > c\}} \big[a\big\|u(s)\big\| + b\big] \, \mathrm{d}s \\
&\leqslant \|\xi\| + (b + c)(T - \tau) + a \int_\tau^t \big\|u(s)\big\| \, \mathrm{d}s.
\end{aligned}
$$

Here, as usual, λ denotes the Lebesgue measure on $\mathbb{R}$. Thanks to Gronwall's inequality—see Vrabie [106, Lemma 1.5.2, p. 46]—u is bounded on $[\tau, T)$ and, since $T < +\infty$, it follows that $\{f(t, u(t));\ t \in [\tau, T)\}$ is bounded. Therefore u is globally Lipschitz on $[\tau, T)$ and accordingly there exists $\lim_{t \uparrow T} u(t) = u^*$. Since $\mathbb{K}$ is closed and $T < \sup \mathbb{I}$, it follows that $(T, u^*) \in \mathbb{I} \times \mathbb{K}$. Using this observation and recalling that $\mathbb{K}$ is viable with respect to f, we conclude that u can be continued to the right of T. But this is absurd, because u is noncontinuable. This contradiction can be eliminated only if $T = \sup \mathbb{I}$, and this achieves the proof. □

As, whenever $\mathbb{K}$ is compact, each continuous function $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ is positively sublinear, from Theorem 3.5.3 it readily follows:

COROLLARY 3.5.1. *Let $\mathbb{K}$ be a nonempty and compact subset of $\mathbb{R}^n$, and let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be continuous. Then a necessary and sufficient condition in order that for each $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ there exists at least one solution, $u : [\tau, T) \to \mathbb{K}$ to* (1.1.1), *satisfying $T = \sup \mathbb{I}$ and $u(\tau) = \xi$, is any one of the five equivalent conditions in Theorem* 3.4.1.

3.6. *Viability of the relative closure*

PROPOSITION 3.6.1 (Roxin [86]). *Let $\mathbb{D} \subseteq \mathbb{R}^n$ be open, let $\mathbb{K} \subseteq \mathbb{D}$ and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ be continuous. If $\mathbb{K}$ is viable with respect to f, then its closure relative to $\mathbb{D}$, $\overline{\mathbb{K}}^{\mathbb{D}}$, is also viable with respect to f.*

PROOF. Let $\tau \in \mathbb{I}$ and let $(\xi_k)_k$ be a sequence in $\mathbb{K}$ convergent to $\xi \in \overline{\mathbb{K}}^{\mathbb{D}}$. Since $\mathbb{K}$ is viable with respect to f, there exists a sequence $(u(\cdot, \xi_k))_k$ of $\mathbb{K}$-valued noncontinuable solutions to (1.1.1) satisfying $u(\tau, \xi_k) = \xi_k$, for $k = 1, 2, \ldots$. It is well known that the intersection of the domains of this sequence contains a nontrivial interval $[\tau, T]$. See, for instance, Vrabie [106, Lemma 3.2.1, p. 107]. Moreover, diminishing T if necessary, we may assume that there exists $\rho > 0$ such that $u(t, \xi_k) \in B(\xi, \rho) \subseteq \mathbb{D}$, for all $k \in \mathbb{N}$, and $t \in [\tau, T]$. By a compactness argument involving Arzelà–Ascoli theorem, we conclude that, on a subsequence at least, we have $\lim_{k \to \infty} u(\cdot, \xi_k) = u(\cdot, \xi)$ uniformly on $[\tau, T]$. Thus $u(t, \xi) \in \overline{\mathbb{K}}^{\mathbb{D}}$ for all $t \in [\tau, T]$, and this completes the proof. □

3.7. *Comparison and viability*

The next theorem, called to our attention by Ursescu [101], gives a characterization of the viability of an epigraph of a certain function in the terms of a differential inequality. Similar results can be found in Clarke et al. [33, p. 266]. Throughout, we denote by $[D_+x](t)$ the *right lower Dini derivative* of the function x at t, i.e.

$$[D_+x](t) = \liminf_{h\downarrow 0} \frac{x(t+h)-x(t)}{h}.$$

THEOREM 3.7.1. *Let* $\omega : \mathbb{I}\times\mathbb{R}_+ \to \mathbb{R}$ *and* $v : [\tau, T) \to \mathbb{R}_+$ *be continuous, with* $[\tau, T) \subseteq \mathbb{I}$. *Then*

$$\operatorname{epi}(v) = \bigl\{(t,\eta);\ v(t) \leqslant \eta,\ t \in [\tau, T)\bigr\}$$

is viable with respect to $(t, y) \mapsto (1, \omega(t, y))$ *if and only if* v *satisfies*

$$[D_+v](t) \leqslant \omega\bigl(t, v(t)\bigr) \tag{3.7.1}$$

for each $t \in [\tau, T)$.

PROOF. *Sufficiency*. It suffices to show that the set $\{(t, v(t));\ t \in [\tau, T)\}$, included in the boundary $\partial\operatorname{epi}(v)$ of $\operatorname{epi}(v)$, satisfies the Nagumo's tangency condition (1.1.3). From (3.7.1) it follows that

$$\left[D_+\left(v(\cdot) - \int_\tau^{\cdot} \omega\bigl(s, v(s)\bigr)\,\mathrm{d}s\right)\right](t) \leqslant 0$$

for each $t \in [\tau, T)$. Thus, in view of a classical result in Hobson [56, p. 365], we necessarily have that $t \mapsto v(t) - \int_\tau^t \omega(s, v(s))\,\mathrm{d}s$ is non-increasing on $[\tau, T]$. So, for each $t \in [\tau, T)$ and $h > 0$ such that $t + h < T$, we have

$$\left(t+h, v(t) + \int_t^{t+h} \omega\bigl(s, v(s)\bigr)\,\mathrm{d}s\right) \in \operatorname{epi}(v),$$

and therefore

$$\begin{aligned}
&\operatorname{dist}\bigl(\bigl(t, v(t)\bigr) + h\bigl(1, \omega\bigl(t, v(t)\bigr)\bigr); \operatorname{epi}(v)\bigr)\\
&\quad\leqslant \left\|\bigl(t, v(t)\bigr) + h\bigl(1, \omega\bigl(t, v(t)\bigr)\bigr) - \left(t+h, v(t) + \int_t^{t+h} \omega\bigl(s, v(s)\bigr)\,\mathrm{d}s\right)\right\|\\
&\quad= \left|h\omega\bigl(t, v(t)\bigr) - \int_t^{t+h} \omega\bigl(s, v(s)\bigr)\,\mathrm{d}s\right|.
\end{aligned}$$

Dividing by $h > 0$ and passing to $\liminf$ for $h \downarrow 0$ we get (1.1.3) and this completes the proof of the sufficiency.

Necessity. Let us assume that $\operatorname{epi}(v)$ is viable with respect to the function $(t, y) \mapsto (1, \omega(t, y))$, let $t \in [\tau, T)$, and let (s, x) be a solution to $s' = 1$, $x'(s) = \omega(s, x(s))$, satisfying the initial conditions $s(0) = t$ and $x(0) = v(t)$, and which remains in $\operatorname{epi}(v)$. We have

$$\frac{v(t+h) - v(t)}{h} \leqslant \frac{x(h) - x(0)}{h}.$$

Accordingly

$$[D_+ v](t) \leqslant \omega\big(s(0), x(0)\big) = \omega\big(t, v(t)\big),$$

and this achieves the proof of the necessity. □

DEFINITION 3.7.1. A function $\omega : \mathbb{I} \times [0, a) \to \mathbb{R}$, $0 < a \leqslant +\infty$, is a *comparison* function if $\omega(t, 0) = 0$ for each $t \in [0, a)$, and for each $[\tau, T) \subseteq \mathbb{I}$, the only continuous function $x : [\tau, T) \to [0, a)$, satisfying

$$\begin{cases} [D_+ x](t) \leqslant \omega\big(t, x(t)\big) & \text{for all } t \in [\tau, T), \\ x(\tau) = 0, \end{cases}$$

is the null function.

COROLLARY 3.7.1. *Let $\omega : \mathbb{I} \times \mathbb{R}_+ \to \mathbb{R}$ be continuous and such that, for each $\tau \in \mathbb{I}$, the Cauchy problem $y'(t) = \omega(t, y(t))$, $y(\tau) = 0$ has only the null solution. Then ω is a comparison function.*

PROOF. Let $v : [\tau, T] \to \mathbb{R}_+$ be any solution to (3.7.1). By Theorem 3.7.1, $\operatorname{epi}(v)$ is viable with respect to $(t, y) \mapsto (1, \omega(t, y))$. So, the unique solution $y : [\tau, T) \to \mathbb{R}_+$ of the Cauchy problem $y'(t) = \omega(t, y(t))$, $y(\tau) = 0$ satisfies $0 \leqslant v(t) \leqslant y(t) = 0$. □

3.8. *Viable preordered subsets*

Let us assume now that $\mathbb{K}$ is a nonempty subset in $\mathbb{R}^n$, and let "$\preceq$" be a *preorder* on $\mathbb{K}$, i.e. a reflexive and transitive binary relation. For our later purposes, it is convenient to identify "$\preceq$" with the multifunction $\mathcal{P} : \mathbb{K} \rightsquigarrow \mathbb{K}$, defined by

$$\mathcal{P}(\xi) = \{\eta \in \mathbb{K};\ \xi \preceq \eta\}$$

for each $\xi \in \mathbb{K}$, and called also a preorder. We say that a preorder "$\preceq$", or $\mathcal{P}$ is *closed* if "$\preceq$" is a closed subset in $\mathbb{R}^n \times \mathbb{R}^n$. Let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$, and let us consider the differential equation (1.1.1). We say that "$\preceq$", or $\mathcal{P}$, is *viable with respect to* f if for each $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$, there exist $[\tau, T] \subseteq \mathbb{I}$ and a solution $u : [\tau, T] \to \mathbb{R}^n$ of (1.1.1) satisfying $u(\tau) = \xi$, $u(t) \in \mathbb{K}$ for each $t \in [\tau, T]$ and u is "$\preceq$"-monotone on $[\tau, T]$, i.e., for each $\tau \leqslant s \leqslant t \leqslant T$, we have $u(s) \preceq u(t)$. The next lemma in Cârjă and Ursescu [25] is the main tool in our forthcoming analysis.

LEMMA 3.8.1. *Let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be continuous and let $\mathcal{P}$ be a preorder on $\mathbb{K}$. If $\mathcal{P}$ is viable with respect to f then, for each $\xi \in \mathbb{K}$, $\mathcal{P}(\xi)$ is viable with respect to f. If $\mathcal{P}$ is closed in $\mathbb{R}^n \times \mathbb{R}^n$ and, for each $\xi \in \mathbb{K}$, $\mathcal{P}(\xi)$ is viable with respect to f, then $\mathcal{P}$ is viable with respect to f.*

PROOF. Clearly, if $\mathcal{P}$ is viable with respect to f, then, for all $\xi \in \mathbb{K}$, $\mathcal{P}(\xi)$ is viable with respect to f.

Now, if $\mathcal{P}$ is closed, then, for each $\xi \in \mathbb{K}$, $\mathcal{P}(\xi)$ is a fortiori closed. Let us assume that, for each $\xi \in \mathbb{K}$, $\mathcal{P}(\xi)$ is viable with respect to f. Let $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$. We shall show that there exist $[\tau, T] \subseteq \mathbb{I}$ and at least one solution $u : [\tau, T] \to \mathbb{K}$ of (1.1.1), with $u(\tau) = \xi$, and such that $u([s, T]) \subseteq \mathcal{P}(u(s))$ for each $s \in [\tau, T]$. To this aim, we proceed in several steps.

In the first step, we note that, by standard qualitative arguments—see, for instance, Vrabie [106, Lemma 3.2.1, p. 107]—one can show that there exists $T > \tau$, $T \in \mathbb{I}$, such that for every noncontinuable solution $u : [\tau, T_m) \to \mathbb{K}$ to (1.1.1) with $u(\tau) = \xi$ we have $T < T_m$. Since $\mathcal{P}(\xi)$ is viable with respect to f, there exists a solution $u : [\tau, T] \to \mathbb{K}$ of (1.1.1) with $u(\tau) = \xi$ and $u([\tau, T]) \subseteq \mathcal{P}(\xi)$.

In the second step, we remark that, for every solution $v : [\tau, T] \to \mathbb{K}$ to (1.1.1), with $v(\tau) = \xi$ and $v([\tau, T]) \subseteq \mathcal{P}(\xi)$, and for every $\nu \in [\tau, T)$, there exists a solution $w : [\tau, T] \to \mathbb{K}$ to (1.1.1) such that w equals v on $[\tau, \nu]$ and $w([\nu, T]) \subseteq \mathcal{P}(w(\nu))$.

In the third step, we observe that, thanks to the first two steps, for every nonempty and finite subset S of $[\tau, T)$, with $\tau \in S$, there exists a solution $u : [\tau, T] \to \mathbb{K}$ of (1.1.1) satisfying both $u(\tau) = \xi$ and $u([s, T]) \subseteq \mathcal{P}(u(s))$ for all $s \in S$.

In the fourth step, we consider a sequence $(S_k)_{k \in \mathbb{N}}$ of nonempty finite subsets of $[\tau, T)$ such that $\tau \in S_k$, $S_k \subseteq S_{k+1}$ for each $n \in \mathbb{N}$, and the set $S = \cup_{k \in \mathbb{N}} S_k$ is dense in $[\tau, T]$. For example, we can take

$$S_k = \left\{ \tau + \frac{i}{2^k}(T - \tau);\ i \in \left\{0, 1, \ldots, 2^k - 1\right\} \right\}.$$

Further, we shall make use of the third step to get a sequence of solutions $(u_k : [\tau, T] \to \mathbb{K})_k$ to (1.1.1), satisfying $u_k(\tau) = \xi$ and such that $u_k([s, T]) \subseteq \mathcal{P}(u_k(s))$ for each $k \in \mathbb{N}$ and each $s \in S_k$. Now, by virtue of Arzelà–Ascoli theorem, we can assume, taking a subsequence if necessary, that the sequence $(u_k)_k$ converges uniformly on $[\tau, T]$ to a solution $u : [\tau, T] \to \mathbb{K}$ of (1.1.1). Clearly $u(\tau) = \xi$.

In the fifth step, we show that $u([s, T]) \subset \mathcal{P}(u(s))$ for all $s \in S$. Indeed, given s as above, there exists $k \in \mathbb{N}$ such that $s \in S_k$. Then $s \in S_m$ and $u_m([s, T]) \subseteq \mathcal{P}(u_m(s))$ for all $m \in \mathbb{N}$ with $k \leqslant m$. At this point, the closedness of the graph of $\mathcal{P}$ shows that $u([s, T]) \subset \mathcal{P}(u(s))$.

In the sixth and final step, taking into account that S is dense in $[\tau, T]$, u is continuous on $[\tau, T]$ and the graph of $\mathcal{P}$ is closed, we conclude that the preceding relation holds for every $s \in [\tau, T]$, and this completes the proof. □

THEOREM 3.8.1. *Let $\mathcal{P}$ be a closed preorder on $\mathbb{K}$ and let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be continuous. Then a necessary and sufficient condition in order that $\mathcal{P}$ be viable with respect to f is the tangency condition below*:

$$f(t,\xi) \in \mathcal{T}_{\mathcal{P}(\xi)}(\xi)$$

for each $(t,\xi) \in \mathbb{I} \times \mathbb{K}$.

PROOF. The proof follows immediately from Lemma 3.8.1. □

4. Problems of invariance

This section follows very closely Cârjă et al. [24].

4.1. *Preliminary facts*

Let $\mathbb{D}$ be a domain in $\mathbb{R}^n$, $\mathbb{K} \subseteq \mathbb{D}$ a locally closed subset, and let us consider the ordinary differential equation (1.1.1), where $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ is a given function.

DEFINITION 4.1.1. The subset $\mathbb{K}$ is *locally invariant* with respect to f if for each $(\tau,\xi) \in \mathbb{I} \times \mathbb{K}$ and each solution $u : [\tau, c] \to \mathbb{D}$, $c \in \mathbb{I}$, $c > \tau$, of (1.1.1), satisfying the initial condition $u(\tau) = \xi$, there exists $T \in (\tau, c]$ such that $u(t) \in \mathbb{K}$ for each $t \in [\tau, T]$. It is *invariant* if it satisfies the above condition of local invariance with $T = c$.

The relationship between viability and local invariance is clarified in

REMARK 4.1.1. If f is continuous on $\mathbb{I} \times \mathbb{D}$ and $\mathbb{K}$ is locally invariant with respect to f, then $\mathbb{K}$ is viable with respect to f. The converse of this assertion is no longer true, as we can see from the following example.

EXAMPLE 4.1.1. Let $\mathbb{D} = \mathbb{R}$, $\mathbb{K} = \{0\}$ and let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(u) = 3\sqrt[3]{u^2}$ for every $u \in \mathbb{R}$. Then $\mathbb{K}$ is viable with respect to f but $\mathbb{K}$ is not locally invariant with respect to f, because the differential equation $u'(t) = f(u(t))$ has at least two solutions which satisfy $u(0) = 0$, i.e. $u \equiv 0$ and $v(t) = t^3$.

A simple necessary and sufficient condition of invariance is stated below.

THEOREM 4.1.1. *Let $\mathbb{D}$ be a domain, $\mathbb{K} \subseteq \mathbb{D}$ a nonempty and locally closed subset and $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ a continuous function with the property that the associated Cauchy problem has the uniqueness property. Then, a necessary and sufficient condition in order that $\mathbb{K}$ be invariant with respect to f is that, for every $(t,\xi) \in \mathbb{I} \times \mathbb{K}$, $f(t,\xi) \in \mathcal{T}_{\mathbb{K}}(\xi)$.*

PROOF. The conclusion follows from Theorem 3.1.1 and Remark 4.1.1. □

Theorem 4.1.1 says that, in general, if $\mathbb{K}$ is viable with respect to $f_{|\mathbb{I}\times\mathbb{K}}$ and (1.1.1) has the uniqueness property, then $\mathbb{K}$ is locally invariant with respect to f. The preceding example shows that this is no longer true if we assume that $\mathbb{K}$ is viable with respect to f and merely $u'(t) = f_{|\mathbb{I}\times\mathbb{K}}(t, u(t))$ has the uniqueness property.

REMARK 4.1.2. Moreover, if $f:\mathbb{I}\times\mathbb{D}\to\mathbb{R}^n$ is continuous and there exists one and only one point $\xi\in\mathbb{D}$ such that the differential equation (1.1.1) has at least two solutions u and v satisfying $u(\tau)=v(\tau)=\xi$, then, $\mathbb{K}=\{u(t);\ t\in[\tau,T]\}$ is viable with respect to f but it is not locally invariant with respect to f.

The next example reveals another interesting fact about local invariance. It shows that the local invariance of $\mathbb{K}$ with respect to f can take place even if $u'(t)=f_{|\mathbb{I}\times\mathbb{K}}(t,u(t))$ has not the uniqueness property.

EXAMPLE 4.1.2. Let $\mathbb{K}=\{(x,y)\in\mathbb{R}^2;\ y\geqslant 0\}$ and let $f:\mathbb{R}^2\to\mathbb{R}^2$ be defined by

$$f\big((x,y)\big)=\begin{cases}(1,0) & \text{if } (x,y)\in\mathbb{R}^2\setminus\mathbb{K},\\ \big(1,3\sqrt[3]{y^2}\big) & \text{if } (x,y)\in\mathbb{K}.\end{cases}$$

Obviously $\mathbb{K}$ is locally invariant with respect to $f_{|\mathbb{K}}$ but $u'(t)=f_{|\mathbb{K}}(u(t))$ has not the uniqueness property. The latter assertion follows from the remark that, from each point, $(x,0)$ (on the boundary of $\mathbb{K}$), we have at least two solutions to $u'(t)=f(u(t))$, $u(t)=(t+x,0)$ and $v(t)=(t+x,t^3)$ satisfying $u(0)=v(0)=(x,0)$.

4.2. *Sufficient conditions for local invariance*

Our first sufficient condition for local invariance says that, whenever there exists an open neighborhood $\mathbb{V}\subseteq\mathbb{D}$ of $\mathbb{K}$ such that f satisfies the "exterior tangency" condition

$$\liminf_{h\downarrow 0}\frac{1}{h}\big[\operatorname{dist}\big(\xi+hf(t,\xi);\mathbb{K}\big)-\operatorname{dist}(\xi;\mathbb{K})\big]\leqslant\omega\big(t,\operatorname{dist}(\xi;\mathbb{K})\big)\tag{4.2.1}$$

for each $(t,\xi)\in\mathbb{I}\times\mathbb{V}$, where ω is a comparison function in the sense of Definition 3.7.1, then $\mathbb{K}$ is locally invariant with respect to f. More precisely, we have:

THEOREM 4.2.1. *Let $\mathbb{K}\subseteq\mathbb{D}\subseteq\mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $f:\mathbb{I}\times\mathbb{D}\to\mathbb{R}^n$. If* (4.2.1) *is satisfied, then $\mathbb{K}$ is locally invariant with respect to f.*

PROOF. Let $\mathbb{V}\subseteq\mathbb{D}$ be the open neighborhood of $\mathbb{K}$ whose existence is ensured by (4.2.1), and let $\omega:\mathbb{I}\times[0,a)\to\mathbb{R}$ the corresponding comparison function. Let $\xi\in\mathbb{K}$ and let $u:[\tau,c]\to\mathbb{V}$ be any solution to (1.1.1) satisfying $u(\tau)=\xi$. Diminishing c if necessary, we may assume that there exists $\rho>0$ such that $B(\xi,\rho)\cap\mathbb{K}$ is closed, $u(t)\in B(\xi,\rho/2)$ and, in addition, $\operatorname{dist}(u(t);\mathbb{K})<a$ for each $t\in[\tau,c]$. Let $g:[\tau,c]\to\mathbb{R}_+$ be defined by $g(t)=\operatorname{dist}(u(t);\mathbb{K})$ for each $t\in[\tau,c]$. Let $t\in[\tau,c)$ and $h>0$ with $t+h\in[\tau,c]$. We have

$$\begin{aligned}g(t+h)&=\operatorname{dist}\big(u(t+h);\mathbb{K}\big)\\&\leqslant h\left\|\frac{u(t+h)-u(t)}{h}-u'(t)\right\|+\operatorname{dist}\big(u(t)+hu'(t);\mathbb{K}\big).\end{aligned}$$

Therefore

$$\frac{g(t+h)-g(t)}{h} \leqslant \alpha(h) + \frac{\operatorname{dist}(u(t)+hu'(t);\mathbb{K}) - \operatorname{dist}(u(t);\mathbb{K})}{h},$$

where

$$\alpha(h) = \left\| \frac{u(t+h)-u(t)}{h} - u'(t) \right\|.$$

Since $u'(t) = f(t, u(t))$ and $\lim_{h\downarrow 0} \alpha(h) = 0$, passing to the inf-limit in the inequality above for $h \downarrow 0$, and taking into account that $\mathbb{V}$, $\mathbb{K}$, and f satisfy (4.2.1), we get

$$[D_+ g](t) \leqslant \omega\big(t, g(t)\big)$$

for each $t \in [\tau, c)$. So, $g(t) \equiv 0$ which means that $u(t) \in \overline{\mathbb{K}} \cap B(\xi, \rho/2)$. But $\overline{\mathbb{K}} \cap B(\xi, \rho/2) \subseteq \mathbb{K} \cap B(\xi, \rho)$, and this achieves the proof. □

REMARK 4.2.1. Clearly, (4.2.1) is satisfied with $\omega = \omega_f$, where the function $\omega_f : \mathbb{I} \times [0, a) \to \overline{\mathbb{R}}$, $a = \sup_{\xi\in\mathbb{V}} \operatorname{dist}(\xi;\mathbb{K})$ is defined by

$$\omega_f(t,x) = \sup_{\substack{\xi\in\mathbb{V} \\ \operatorname{dist}(\xi;\mathbb{K})=x}} \liminf_{h\downarrow 0} \frac{1}{h}\big[\operatorname{dist}\big(\xi + hf(t,\xi);\mathbb{K}\big) - \operatorname{dist}(\xi;\mathbb{K})\big] \tag{4.2.2}$$

for each $(t, x) \in \mathbb{I} \times [0, a)$.

So, Theorem 4.2.1 can be reformulated as:

THEOREM 4.2.2. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. If there exists an open neighborhood $\mathbb{V}$ of $\mathbb{K}$ with $\mathbb{V} \subseteq \mathbb{D}$ such that ω_f defined by* (4.2.2) *is a comparison function, then $\mathbb{K}$ is locally invariant with respect to f.*

4.3. *Viability and comparison imply exterior tangency*

DEFINITION 4.3.1. Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$. We say that a function $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ has *the comparison property with respect to* $(\mathbb{D}, \mathbb{K})$ if there exist a proximal neighborhood $\mathbb{V} \subseteq \mathbb{D}$ of $\mathbb{K}$, one projection $\pi_{\mathbb{K}} : \mathbb{V} \to \mathbb{K}$ subordinated to $\mathbb{V}$, and one comparison function $\omega : \mathbb{I} \times [0, a) \to \mathbb{R}$, with $a = \sup_{\xi\in\mathbb{V}} \operatorname{dist}(\xi;\mathbb{K})$, such that

$$\big[\xi - \pi_{\mathbb{K}}(\xi), f(t,\xi) - f\big(t, \pi_{\mathbb{K}}(\xi)\big)\big]_+ \leqslant \omega\big(t, \|\xi - \pi_{\mathbb{K}}(\xi)\|\big) \tag{4.3.1}$$

for each $(t, \xi) \in \mathbb{I} \times \mathbb{V}$.

Let us observe that (4.3.1) is automatically satisfied for each $(t,\xi)\in\mathbb{I}\times\mathbb{K}$, and therefore, in Definition 4.3.1, we have only to assume that (4.3.1) holds for each $(t,\xi)\in\mathbb{I}\times[\mathbb{V}\setminus\mathbb{K}]$.

DEFINITION 4.3.2. The function $f:\mathbb{I}\times\mathbb{D}\to\mathbb{R}^n$ is called:

(i) $(\mathbb{D},\mathbb{K})$-*Lipschitz* if there exist a proximal neighborhood $\mathbb{V}\subseteq\mathbb{D}$ of $\mathbb{K}$, a subordinated projection $\pi_{\mathbb{K}}:\mathbb{V}\to\mathbb{K}$, and $L>0$, such that

$$\left\|f(t,\xi)-f\big(t,\pi_{\mathbb{K}}(\xi)\big)\right\|\leqslant L\left\|\xi-\pi_{\mathbb{K}}(\xi)\right\|$$

for each $(t,\xi)\in\mathbb{I}\times[\mathbb{V}\setminus\mathbb{K}]$;

(ii) $(\mathbb{D},\mathbb{K})$-*dissipative* if there exist a proximal neighborhood $\mathbb{V}\subseteq\mathbb{D}$ of $\mathbb{K}$, and a projection, $\pi_{\mathbb{K}}:\mathbb{V}\to\mathbb{K}$, subordinated to $\mathbb{V}$, such that

$$\big[\xi-\pi_{\mathbb{K}}(\xi),\,f(t,\xi)-f\big(t,\pi_{\mathbb{K}}(\xi)\big)\big]_+\leqslant 0$$

for each $(t,\xi)\in\mathbb{I}\times[\mathbb{V}\setminus\mathbb{K}]$.

REMARK 4.3.1. We notice that, if we assume that (4.3.1), or either of the conditions (i), or (ii) in Definition 4.3.2 is satisfied for ξ replaced by ξ_1 and $\pi_{\mathbb{K}}(\xi)$ replaced by ξ_2 with $\xi_1,\xi_2\in\mathbb{V}$, as considered in Kenmochi and Takahashi [60], then, for each $[\tau,T]\subseteq\mathbb{I}$ and $\xi\in\mathbb{K}$, there exists at most one solution $u:[\tau,T]\to\mathbb{K}$ to (1.1.1) satisfying $u(\tau)=\xi$. On contrary, in this more general frame, it may happen that, for certain (or for all) $[\tau,T]\subseteq\mathbb{I}$ and $\xi\in\mathbb{K}$, (1.1.1) have at least two solutions $u,v:[\tau,T]\to\mathbb{K}$ satisfying $u(\tau)=v(\tau)=\xi$.

Let $\mathbb{V}$ be a proximal neighborhood of $\mathbb{K}$, and let $\pi_{\mathbb{K}}:\mathbb{V}\to\mathbb{K}$ be a projection subordinated to $\mathbb{V}$. If $f:\mathbb{I}\times\mathbb{V}\to\mathbb{K}$ is a function with the property that, for each $t\in\mathbb{I}$ and $\eta\in\mathbb{K}$, the restriction of $f(t,\cdot)$ to the "segment"

$$\mathbb{V}_\eta=\big\{\xi\in\mathbb{V}\setminus\mathbb{K};\ \pi_{\mathbb{K}}(\xi)=\eta\big\}$$

is dissipative, then f is $(\mathbb{D},\mathbb{K})$-dissipative.

It is easy to see that if f is either $(\mathbb{D},\mathbb{K})$-Lipschitz, or $(\mathbb{D},\mathbb{K})$-dissipative, then it has the comparison property with respect to $(\mathbb{D},\mathbb{K})$. We notice that there are examples showing that there exist functions f which, although neither $(\mathbb{D},\mathbb{K})$-Lipschitz, nor $(\mathbb{D},\mathbb{K})$-dissipative, do have the comparison property. Moreover, there exist functions which, although $(\mathbb{D},\mathbb{K})$-Lipschitz, are not Lipschitz on $\mathbb{D}$, as well as, functions which although $(\mathbb{D},\mathbb{K})$-dissipative, are not dissipative on $\mathbb{D}$. In fact, these two properties describe merely the local behavior of f at the interface between $\mathbb{K}$ and $\mathbb{D}\setminus\mathbb{K}$. We include below two examples: the first one of an $(\mathbb{D},\mathbb{K})$-Lipschitz function which is not locally Lipschitz, and the second one of a function which, although non-dissipative, is $(\mathbb{D},\mathbb{K})$-dissipative. We notice that both examples refer to the autonomous case.

EXAMPLE 4.3.1. The graph of an $(\mathbb{D},\mathbb{K})$-Lipschitz function $f:\mathbb{R}\to\mathbb{R}$ which is not Lipschitz is illustrated in Fig. 2. Here $\mathbb{K}=[a,b]$ and $\mathbb{D}$ is any open subset in $\mathbb{R}$ including $\mathbb{K}$.

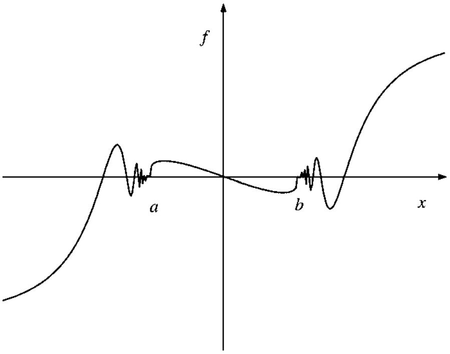


Fig. 2.

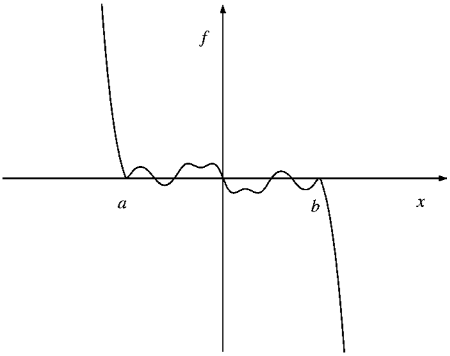


Fig. 3.

EXAMPLE 4.3.2. The graph of a function $f:\mathbb{R}\to\mathbb{R}$ which is $(\mathbb{D},\mathbb{K})$-dissipative but not dissipative is illustrated in Fig. 3. This time, $\mathbb{K}$ is either $(-\infty,\beta]$, or $[\alpha,+\infty)$, or $[\alpha,\beta]$ with $\alpha\leqslant a\leqslant b\leqslant\beta$, and $\mathbb{D}$ is any open subset in $\mathbb{R}$ including $\mathbb{K}$.

We begin with:

THEOREM 4.3.1. *Let* $\mathbb{K}\subseteq\mathbb{D}\subseteq\mathbb{R}^n$, *with* $\mathbb{K}$ *locally closed and* $\mathbb{D}$ *open, and let* $f:\mathbb{I}\times\mathbb{D}\to\mathbb{R}^n$. *If* f *has the comparison property with respect to* $(\mathbb{D},\mathbb{K})$, *and* (1.1.3) *is satisfied, then* (4.2.1) *holds true.*

PROOF. Let $\mathbb{V}\subseteq\mathbb{D}$ be the open neighborhood of $\mathbb{K}$ as in Definition 4.3.1, let $\xi\in\mathbb{V}$ and $[t,T)\subseteq\mathbb{I}$. Let $\rho>0$ and let $\pi_{\mathbb{K}}$ be the selection of $\Pi_{\mathbb{K}}$ as in Definition 4.3.1. Let $h>0$ with $t+h\in[t,T]$. Taking into account that $\|\xi-\pi_{\mathbb{K}}(\xi)\|=\operatorname{dist}(\xi;\mathbb{K})$, we have

$$\begin{aligned}&\operatorname{dist}\bigl(\xi+hf(t,\xi);\mathbb{K}\bigr)-\operatorname{dist}(\xi;\mathbb{K})\\&\quad\leqslant\bigl\|\xi-\pi_{\mathbb{K}}(\xi)+h\bigl[f(t,\xi)-f\bigl(t,\pi_{\mathbb{K}}(\xi)\bigr)\bigr]\bigr\|\end{aligned}$$

$$-\left\|\xi-\pi_{\mathbb{K}}(\xi)\right\|+\operatorname{dist}\bigl(\pi_{\mathbb{K}}(\xi)+hf\bigl(t,\pi_{\mathbb{K}}(\xi)\bigr);\mathbb{K}\bigr).$$

Dividing by h, passing to the inf-limit for $h\downarrow 0$, and using (1.1.3), we get

$$\begin{aligned}&\liminf_{h\downarrow 0}\frac{1}{h}\bigl[\operatorname{dist}\bigl(\xi+hf(t,\xi);\mathbb{K}\bigr)-\operatorname{dist}(\xi;\mathbb{K})\bigr]\\&\quad\leqslant\bigl[\xi-\pi_{\mathbb{K}}(\xi),f(t,\xi)-f\bigl(t,\pi_{\mathbb{K}}(\xi)\bigr)\bigr]_{+}\\&\quad\leqslant\omega\bigl(t,\left\|\xi-\pi_{\mathbb{K}}(\xi)\right\|\bigr).\end{aligned}$$

But this inequality shows that (4.2.1) holds, and this completes the proof. □

In the specific case in which f is continuous, we have:

THEOREM 4.3.2. *Let $\mathbb{K}\subseteq\mathbb{D}\subseteq\mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $f:\mathbb{I}\times\mathbb{D}\to\mathbb{R}^n$ be continuous. Let us assume that f has the comparison property with respect to* $(\mathbb{D},\mathbb{K})$*, and one of the four conditions below is satisfied*:

(i) *for every $(t,\xi)\in\mathbb{I}\times\mathbb{K}$, $f(t,\xi)\in\mathcal{C}_{\mathbb{K}}(\xi)$,*
(ii) *for every $(t,\xi)\in\mathbb{I}\times\mathbb{K}$, $f(t,\xi)\in\mathcal{T}_{\mathbb{K}}(\xi)$,*
(iii) *for every $(t,\xi)\in\mathbb{I}\times\mathbb{K}$, $f(t,\xi)\in\mathcal{B}_{\mathbb{K}}(\xi)$,*
(iv) *the set $\mathbb{K}$ is viable with respect to f.*

Then (4.2.1) *is also satisfied.*

In general, if $\mathcal{G}:\mathbb{K}\rightsquigarrow\mathbb{R}^n$ satisfies $\mathcal{C}_{\mathbb{K}}(\xi)\subseteq\mathcal{G}(\xi)\subseteq\mathcal{B}_{\mathbb{K}}(\xi)$ for each $\xi\in\mathbb{K}$ and:

(v) *for every $(t,\xi)\in\mathbb{I}\times\mathbb{K}$, $f(t,\xi)\in\mathcal{G}(\xi)$,*

then (4.2.1) *is satisfied too.*

PROOF. The conclusion follows from Theorem 3.4.1. □

4.4. *Sufficient conditions for invariance. Revisited*

The next sufficient condition for invariance follows from Theorems 4.2.1 and 4.3.2.

THEOREM 4.4.1. *Let $\mathbb{K}\subseteq\mathbb{D}\subseteq\mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $f:\mathbb{I}\times\mathbb{D}\to\mathbb{R}^n$. If f has the comparison property with respect to* $(\mathbb{D},\mathbb{K})$*, and satisfies one of the conditions* (i), (ii), (iii), *or* (v) *in Theorem* 4.3.2, *then $\mathbb{K}$ is locally invariant with respect to f.*

Now, let $\mathbb{K}\subseteq\mathbb{R}^n$ and let $\mathbb{V}$ be an open neighborhood of $\mathbb{K}$.

DEFINITION 4.4.1. A function $g:\mathbb{V}\to\mathbb{K}$ is *a proximal generalized distance* if:

(i) g is Lipschitz continuous on bounded subsets in $\mathbb{V}$,
(ii) $g(\xi)=0$ if and only if $\xi\in\mathbb{K}$.

If $\mathbb{K}$ is closed, a typical example of proximal generalized distance is offered by $g(\xi) = \alpha(\operatorname{dist}(\xi;\mathbb{K}))$, where $\alpha : [0,+\infty) \to [0,+\infty)$ is Lipschitz on bounded subsets, $\alpha(0) = 0$ and $\alpha(r) \neq 0$ if $r \neq 0$, while $\operatorname{dist}(\xi;\mathbb{K})$ is the usual distance from ξ to $\mathbb{K}$. We notice that if there exists a proximal generalized distance $g : \mathbb{V} \to [0,+\infty)$, then $\mathbb{K}$ is locally closed. Indeed, since $\mathbb{K} = \{\xi \in \mathbb{V};\ g(\xi) = 0\}$ and g is continuous, $\mathbb{K}$ is relatively closed in $\mathbb{V}$. But $\mathbb{V}$ is open and thus $\mathbb{K}$ is locally closed, as claimed.

THEOREM 4.4.2. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. If there exist an open neighborhood $\mathbb{V}$ of $\mathbb{K}$, with $\mathbb{V} \subseteq \mathbb{D}$, a proximal generalized distance $g : \mathbb{V} \to \mathbb{R}_+$ and a comparison function $\omega : \mathbb{I} \times [0,a) \to \mathbb{R}$ such that*

$$\liminf_{h \downarrow 0} \frac{1}{h}\big[g\big(\xi + hf(t,\xi)\big) - g(\xi)\big] \leqslant \omega\big(t, g(\xi)\big) \tag{4.4.1}$$

for each $(t,\xi) \in \mathbb{I} \times \mathbb{V}$, then $\mathbb{K}$ is locally invariant with respect to f.

PROOF. The proof follows closely that one of Theorem 4.2.1, with the mention that here one has to use the obvious inequality $g(\lambda) \leqslant g(\eta) + L\|\lambda - \eta\|$ for each $\lambda, \eta \in B(\xi,\rho) \cap \mathbb{V}$, where $L > 0$ is the Lipschitz constant of g on $B(\xi,\rho) \cap \mathbb{V}$. □

In order to obtain a simple, but useful, extension of Theorem 4.2.1, some observations are needed. Namely, if $g : \mathbb{V} \to [0,+\infty)$ is a generalized distance, we may consider the generalized tangency condition

$$\liminf_{h \downarrow 0} \frac{1}{h} g\big(\xi + hf(t,\xi)\big) = 0 \tag{4.4.2}$$

for each $(t,\xi) \in \mathbb{I} \times \mathbb{K}$, and one may ask whether this implies viability, whenever, of course f is continuous. The answer to this question is in the negative as the simple example below shows.

EXAMPLE 4.4.1. Let $\mathbb{K}$ be locally closed, let $\mathbb{V}$ be any open neighborhood of $\mathbb{K}$ and let $g : \mathbb{V} \to [0,+\infty)$ be defined as $g(\xi) = \operatorname{dist}^2(\xi;\mathbb{K})$ for each $\xi \in \mathbb{V}$. Further, let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be a continuous function such that $\mathbb{K}$ is not viable with respect to f. We can always find such a function whenever $\mathbb{K}$ is not open. Now, since

$$g\big(\xi + hf(t,\xi)\big) \leqslant \big\|\xi + hf(t,\xi) - \xi\big\|^2 \leqslant h^2\big\|f(t,\xi)\big\|^2$$

for each $(t,\xi) \in \mathbb{I} \times \mathbb{K}$, (4.4.2) is trivially satisfied. So, the generalized tangency condition (4.4.2) does not imply the viability of $\mathbb{K}$ with respect to f.

This example shows that if g is a proximal generalized distance, and g^2 satisfies (4.4.2), it may happen that g does not satisfy (4.4.2). Therefore it justifies why, in the next result, we assume explicitly that (4.4.2) holds true, even though it is automatically satisfied by g^2.

THEOREM 4.4.3. *Let* $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, *and let* $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. *If there exist an open neighborhood* $\mathbb{V}$ *of* $\mathbb{K}$, *with* $\mathbb{V} \subseteq \mathbb{D}$, *and a proximal generalized distance* $g : \mathbb{V} \to \mathbb{R}_+$ *satisfying* (4.4.2) *and such that*

$$\liminf_{h \downarrow 0} \frac{1}{2h} \big[g^2\big(\xi + hf(t, \xi)\big) - g^2(\xi) \big] \leqslant g(\xi)\omega\big(t, g(\xi)\big) \tag{4.4.3}$$

for each $(t, \xi) \in \mathbb{I} \times \mathbb{V}$, *then* $\mathbb{K}$ *is locally invariant with respect to* f.

PROOF. We have only to observe that, in the presence of (4.4.2), (4.4.3) and (4.4.1) are equivalent, and to apply Theorem 4.4.2. □

Using Theorem 4.4.3, we will prove next some other sufficient conditions for invariance expressed in the terms of a generalized Lipschitz projection. Namely, a subset $\mathbb{K}$ is a *Lipschitz retract* if there exist an open neighborhood $\mathbb{V}$ of $\mathbb{K}$ and a Lipschitz continuous map, $r : \mathbb{V} \to \mathbb{K}$, with $r(\xi) = \xi$ if and only if $\xi \in \mathbb{K}$. The function r as above is a *generalized Lipschitz projection*. For each Lipschitz retract $\mathbb{K}$, one can define a proximal generalized distance, $g : \mathbb{V} \to [0, +\infty)$, by $g(\xi) = \|r(\xi) - \xi\|$ for all $\xi \in \mathbb{V}$. Consequently, each Lipschitz retract is locally closed. Moreover, each open subset $\mathbb{K}$ is Lipschitz retract (take $\mathbb{V} = \mathbb{K}$ and r the identity). Another simple example of a Lipschitz retract is given by a closed subset $\mathbb{K}$ which has an open neighborhood $\mathbb{V}$ for which there exists a single-valued continuous projection $\pi_{\mathbb{K}} : \mathbb{V} \to \mathbb{K}$, i.e. $\operatorname{dist}(\xi; \mathbb{K}) = \|\pi_{\mathbb{K}}(\xi) - \xi\|$ for each $\xi \in \mathbb{V}$. In the latter case we say that $\mathbb{K}$ is *a proximate retract*. It should be noticed that the class of Lipschitz retract subsets is strictly larger than that of proximate retracts as the simple example below shows.

EXAMPLE 4.4.2. Take $\mathbb{R}^2$, endowed with the usual Hilbert structure and let us observe that the set

$$\mathbb{K} = \big\{(x, y) \in \mathbb{R}^2;\ y \leqslant |x|\big\},$$

although Lipschitz retract, is not a proximate retract. Indeed, let $\mathbb{V} = \mathbb{R}^2$, and let $r((x, y))$ be defined, either as (x, y) if $(x, y) \in \mathbb{K}$, or as $(x, |x|)$ if $(x, y) \in \mathbb{V} \setminus \mathbb{K}$. It is easy to verify that r is a generalized Lipschitz projection with Lipschitz constant $\sqrt{2}$, and thus $\mathbb{K}$ is a Lipschitz retract. However, $\mathbb{K}$ is not a proximate retract since any selection $\pi_{\mathbb{K}}$ of the projection $\Pi_{\mathbb{K}}$ is discontinuous at each point $(0, y)$, with $y > 0$.

We emphasize that all the results which will follow can be reformulated to hold also for locally Lipschitz retracts, i.e. for those subsets $\mathbb{K}$ satisfying: *for each* $\xi \in \mathbb{K}$ *there exists* $\rho > 0$ *such that* $B(\xi; \rho) \cap \mathbb{K}$ *is Lipschitz retract*, but for the sake of simplicity we confined ourselves to the simpler case of Lipschitz retracts. First, let $\mathbb{K}$ be Lipschitz retract with the corresponding generalized Lipschitz projection $r : \mathbb{V} \to \mathbb{K}$. In the next two results, we will assume that the norm on $\mathbb{R}^n$ is defined by means of an inner product, i.e. $\|x\|^2 = \langle x, x\rangle$ for each $x \in \mathbb{R}^n$.

THEOREM 4.4.4. *Let $\mathbb{D} \subseteq \mathbb{R}^n$ be open and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. Let us assume that $\mathbb{K} \subseteq \mathbb{D}$ is Lipschitz retract with generalized Lipschitz projection $r : \mathbb{V} \to \mathbb{K}$ satisfying*

$$\liminf_{h\downarrow 0} \frac{1}{h} \big\| r\big(\xi + hf(t,\xi)\big) - \xi - hf(t,\xi)\big\| = 0 \tag{4.4.4}$$

for each $(t,\xi) \in \mathbb{I} \times \mathbb{K}$. Assume, in addition, that there exists a comparison function $\omega : \mathbb{I} \times [0,a) \to \mathbb{R}$, with $a = \sup_{\xi\in\mathbb{V}} \|r(\xi) - \xi\|$ such that

$$\begin{aligned}&\liminf_{h\downarrow 0} \frac{1}{h} \big\langle r\big(\xi + hf(t,\xi)\big) - r(\xi) - hf(t,\xi), r(\xi) - \xi\big\rangle \\ &\quad\leqslant \big\|r(\xi) - \xi\big\| \omega\big(t, \big\|r(\xi) - \xi\big\|\big)\end{aligned} \tag{4.4.5}$$

for each $(t,\xi) \in \mathbb{I} \times \mathbb{V}$. Then $\mathbb{K}$ is locally invariant with respect to f.

PROOF. Let us define $g(\xi) = \|r(\xi) - \xi\|$ for each $\xi \in \mathbb{V}$, and let $L > 0$ be the Lipschitz constant of r. Let us observe that

$$\begin{aligned}&g^2(\xi + h\eta) - g^2(\xi) \\ &\quad= \big\langle r(\xi + h\eta) - (\xi + h\eta) - \big(r(\xi) - \xi\big), r(\xi + h\eta) - (\xi + h\eta) + \big(r(\xi) - \xi\big)\big\rangle \\ &\quad= \big\langle r(\xi + h\eta) - r(\xi), r(\xi + h\eta) + r(\xi) - 2\xi\big\rangle \\ &\qquad - h\big\langle \eta, 2r(\xi + h\eta) - 2\xi\big\rangle + h^2\|\eta\|^2 \\ &\quad= \big\|r(\xi + h\eta) - r(\xi)\big\|^2 + 2\big\langle r(\xi + h\eta) - r(\xi), r(\xi) - \xi\big\rangle \\ &\qquad - h\big\langle \eta, 2r(\xi + h\eta) - 2\xi\big\rangle + h^2\|\eta\|^2 \\ &\quad\leqslant \big(L^2 + 1\big)h^2\|\eta\|^2 + 2\big\langle r(\xi + h\eta) - r(\xi), r(\xi) - \xi\big\rangle - 2h\big\langle \eta, r(\xi + h\eta) - \xi\big\rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}&\liminf_{h\downarrow 0} \frac{1}{2h} \big[g^2(\xi + h\eta) - g^2(\xi)\big] \\ &\quad\leqslant \liminf_{h\downarrow 0} \frac{1}{h} \big\langle r(\xi + h\eta) - r(\xi), r(\xi) - \xi\big\rangle - \big\langle \eta, r(\xi) - \xi\big\rangle.\end{aligned}$$

Since, by (4.4.4), g satisfies (4.4.2), taking $\eta = f(t,\xi)$ and using (4.4.5) and Theorem 4.4.3, we get the conclusion. □

A consequence of Theorem 4.4.4 is stated below.

THEOREM 4.4.5. *Let $\mathbb{K} \subseteq \mathbb{D}$, with $\mathbb{D}$ open, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. Let us assume that $\mathbb{K}$ is a Lipschitz retract with the generalized Lipschitz projection $r : \mathbb{V} \to \mathbb{K}$ satisfying* (4.4.4).

Let us assume in addition that, for each $t \in \mathbb{I}$ and $\xi \in \mathbb{V}$, there exists the directional derivative, $r'(\xi)[f(t,\xi)]$, of r, at ξ in the direction $f(t,\xi)$, and

$$\big\langle r'(\xi)\big[f(t,\xi)\big] - f(t,\xi), r(\xi) - \xi\big\rangle \leqslant \big\|r(\xi) - \xi\big\|\omega\big(t, \big\|r(\xi) - \xi\big\|\big), \tag{4.4.6}$$

where $\omega : \mathbb{I} \times [0, a) \to \mathbb{R}$ is a comparison function, and $a = \sup_{\xi \in \mathbb{V}} \|r(\xi) - \xi\|$. Then $\mathbb{K}$ is locally invariant with respect to f.

PROOF. It is easy to see that, in this specific case, (4.4.6) is equivalent to (4.4.5) and this completes the proof. □

REMARK 4.4.1. Let $\mathbb{K}$ be a Lipschitz retract subset and let $r : \mathbb{V} \to \mathbb{K}$ be the corresponding generalized Lipschitz projection. Let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be a continuous function, let $a = \sup_{\xi \in \mathbb{V}} \|r(\xi) - \xi\|$, and let us define the function $\omega : \mathbb{I} \times [0, a) \to \overline{\mathbb{R}}_+$ by $\omega(t, 0) = 0$ and

$$\omega(t,x) = \sup_{\substack{\xi \in \mathbb{V} \\ \|r(\xi)-\xi\|=x}} \frac{\langle r'(\xi)[f(t,\xi)] - f(t,\xi), r(\xi) - \xi\rangle}{\|r(\xi) - \xi\|} \tag{4.4.7}$$

for each $(t,x) \in \mathbb{I} \times (0,a)$. From Theorem 4.4.5, it follows that $\mathbb{K}$ is invariant with respect to f if (4.4.4) is satisfied and ω, defined by (4.4.7), is a comparison function. Furthermore, if $\mathbb{K} \subseteq \mathbb{R}^n$ is a closed linear subspace in $\mathbb{R}^n$, and r is the projection of $\mathbb{R}^n$ on $\mathbb{K}$, then r is linear, and $r'(\xi)[\eta] = r(\eta)$ for each $\xi, \eta \in \mathbb{R}^n$. So, in this case, the condition (4.4.4) is equivalent to $f(\mathbb{I} \times \mathbb{K}) \subseteq \mathbb{K}$. Take $\mathbb{V} = \{\xi \in \mathbb{R}^n;\ \operatorname{dist}(\xi; \mathbb{K}) < \rho\}$, for some fixed $\rho > 0$, and let us observe that the function ω, defined by (4.4.7), is given by $\omega(t,0) = 0$ and

$$\omega(t,x) = \sup_{\substack{\xi \in \mathbb{V} \\ \|r(\xi)-\xi\|=x}} \frac{\langle r(f(t,\xi)) - f(t,\xi), r(\xi) - \xi\rangle}{\|r(\xi) - \xi\|} \tag{4.4.8}$$

for each $(t,x) \in \mathbb{I} \times (0,\rho)$. Hence, if $f(\mathbb{I} \times \mathbb{K}) \subseteq \mathbb{K}$, and ω defined by (4.4.8) is a comparison function, then $\mathbb{K}$ is invariant with respect to (1.1.1).

4.5. *When tangency implies exterior tangency?*

Next, we will prove that, in special circumstances, the tangency condition (1.1.3) for a function $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}$ comes from the exterior tangency condition (4.2.1) for a suitably defined extension $\tilde{f} : \mathbb{I} \times \mathbb{D} \to \mathbb{R}$ of f. More precisely, we have:

THEOREM 4.5.1. *Let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be a given function satisfying* (1.1.3). *If $\mathbb{V} \subseteq \mathbb{R}^n$ is a proximal neighborhood of $\mathbb{K}$ and $r : \mathbb{V} \to \mathbb{K}$ is a projection subordinated to $\mathbb{V}$, then $\tilde{f} : \mathbb{I} \times \mathbb{V} \to \mathbb{R}^n$, defined by $\tilde{f}(t,\cdot) = f(t, r(\cdot))$ satisfies* (4.2.1).

PROOF. Let $\xi \in \mathbb{V}$ and $h > 0$. We have

$$\begin{aligned}&\operatorname{dist}\big(\xi + h\tilde{f}(t,\xi);\mathbb{K}\big) - \operatorname{dist}(\xi;\mathbb{K})\\&\quad\leqslant \big\|\xi - r(\xi)\big\| + \operatorname{dist}\big(r(\xi) + hf\big(t, r(\xi)\big);\mathbb{K}\big) - \big\|\xi - r(\xi)\big\|\\&\quad= \operatorname{dist}\big(r(\xi) + hf\big(t, r(\xi)\big);\mathbb{K}\big).\end{aligned}$$

Dividing by $h > 0$ and passing to lim inf for $h \downarrow 0$, we get

$$\liminf_{h\downarrow 0}\frac{1}{h}\big[\operatorname{dist}\big(\xi + h\tilde{f}(t,\xi);\mathbb{K}\big) - \operatorname{dist}(\xi;\mathbb{K})\big] \leqslant 0.$$

So, (4.2.1) holds true with $\omega \equiv 0$, and the proof is complete. □

It should be noticed that, the conclusion of Theorem 4.5.1 is no longer true if we are looking for a continuous extension $\tilde{f}$ of a continuous function f satisfying (1.1.3), as the next example shows.

EXAMPLE 4.5.1. This example is adapted from Aubin and Cellina [3, p. 203]. Let $\mathbb{K}_1 = \{(x, 3\sqrt[3]{x^2});\ x \in \mathbb{R}_+\}$, $\mathbb{K}_2 = \{(x, 3\sqrt[3]{x^2});\ x \in \mathbb{R}_-\}$ and let $\mathbb{K} = \mathbb{K}_1 \cup \mathbb{K}_2$. If $\xi \in \mathbb{K}_1$, we define $f(\xi)$ as the unit clockwise oriented tangent vector to $\mathbb{K}_1$ at ξ, and if $\xi \in \mathbb{K}_2$, we define $f(\xi)$ as the unit counterclockwise oriented tangent vector to $\mathbb{K}_1$ at ξ. Of course, $f((0,0)) = (0,1)$. Thus $f : \mathbb{K} \to \mathbb{R}^2$ is continuous and $f(\xi) \in \mathcal{T}_{\mathbb{K}}(\xi)$ for each $\xi \in \mathbb{K}$. By virtue of Theorem 1.1.3, $\mathbb{K}$ is viable with respect to f. Let $\tilde{f}$ be any continuous extension of f to an open neighborhood $\mathbb{V}$ of the origin. We may assume that for each $v \in \mathbb{V}$, $\tilde{f}_2(v) \geqslant \frac{1}{2}$.

In fact, the equation $u'(t) = f(u(t))$ subjected to $u(0) = (0,0)$ has two local solutions $u, v : [0,\delta] \to \mathbb{K}$, with $u([0,\delta]) \subseteq \mathbb{K}_1$ and $v([0,\delta]) \subseteq \mathbb{K}_2$. Diminishing $\delta > 0$, we may assume that no solution to $u'(t) = \tilde{f}(u(t))$, $u(0) = (0,0)$, can escape from $\mathbb{V}$. Now, if we assume that $\mathbb{K}$ is invariant with respect to $\tilde{f}$, we have

$$F_{0,(0,0)}(\delta) = \big\{u(\delta);\ u'(t) = \tilde{f}\big(u(t)\big),\ \text{for all } t \in [0,\delta],\ u(0) = (0,0)\big\} \subseteq \mathbb{K},$$

and by virtue of a classical result due to Kneser—see Theorem 7.5.1—we know that $F_{0,(0,0)}(\delta)$ is connected, and therefore, we conclude that there exists at least one solution $w : [0,\delta] \to \mathbb{K}$ of $u'(t) = \tilde{f}(u(t))$, $u(0) = (0,0)$ with $w(\delta) = (0,0)$. But this is impossible, because $w_2(\delta) \geqslant \frac{1}{2}\delta$.

However, if f is continuous and $\mathbb{K}$ is smooth enough, by the very same proof we deduce:

THEOREM 4.5.2. *Let $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ be a continuous function satisfying* (1.1.3). *If there exists a proximal neighborhood $\mathbb{V} \subseteq \mathbb{R}^n$ of $\mathbb{K}$, and a continuous projection $r : \mathbb{V} \to \mathbb{K}$ subordinated to $\mathbb{V}$, then f can be extended to a continuous function $\tilde{f} : \mathbb{I} \times \mathbb{V} \to \mathbb{R}^n$ satisfying* (4.2.1).

4.6. *Local invariance and monotonicity*

We say that the preorder "$\preceq$" is *locally invariant with respect to* f if for each $(\tau,\xi)\in\mathbb{I}\times\mathbb{K}$, each solution $u:[\tau,c]\to\mathbb{D}$, $c\in(\tau,b]$, of (1.1.1) satisfying $u(\tau)=\xi$, there exists $T\in(\tau,c]$ such that $u(t)\in\mathbb{K}$ for each $t\in[\tau,T]$ and u is "$\preceq$"-monotone on $[\tau,T]$, i.e., for each $\tau\leqslant s\leqslant t\leqslant T$, we have $u(s)\preceq u(t)$. We recall that, for each $\xi\in\mathbb{K}$, $\mathcal{P}(\xi)=\{\eta\in\mathbb{K};\ \xi\preceq\eta\}$.

REMARK 4.6.1. The preorder "$\preceq$" is locally invariant with respect to f if and only if for each $(\tau,\xi)\in\mathbb{I}\times\mathbb{K}$, each solution $u:[\tau,c]\to\mathbb{D}$, $c\in(\tau,b]$, of (1.1.1) satisfying $u(\tau)=\xi$, there exists $T\in(\tau,c]$ such that, for each $s\in[\tau,T]$ and $t\in[s,T]$, we have $u(t)\in\mathcal{P}(u(s))$.

In contrast with Lemma 3.8.1, the next lemma is almost obvious.

LEMMA 4.6.1. *The preorder "$\preceq$" is locally invariant with respect to f if and only if for each $\xi\in\mathbb{K}$, $\mathcal{P}(\xi)$ is locally invariant with respect to f.*

COROLLARY 4.6.1. *If for each $\xi\in\mathbb{K}$ there exists an open neighborhood $\mathbb{V}\subseteq\mathbb{D}$ of ξ and a comparison function $\omega:\mathbb{I}\times[0,a)\to\mathbb{R}$, with $a=\sup_{\eta\in\mathbb{V}}\operatorname{dist}(\eta;\mathcal{P}(\xi))$, and such that*

$$\liminf_{h\downarrow 0}\frac{1}{h}\big[\operatorname{dist}\big(\eta+hf(t,\eta);\mathcal{P}(\xi)\big)-\operatorname{dist}\big(\eta;\mathcal{P}(\xi)\big)\big]\leqslant\omega\big(t,\operatorname{dist}\big(\eta;\mathcal{P}(\xi)\big)\big)$$

for each $(t,\eta)\in\mathbb{I}\times\mathbb{V}$, then "$\preceq$" is locally invariant with respect to f.

COROLLARY 4.6.2. *If, for each $\xi\in\mathbb{K}$, f has the comparison property with respect to $(\mathbb{D},\mathcal{P}(\xi))$ and, for each $(t,\eta)\in\mathbb{I}\times\mathcal{P}(\xi)$, we have*

$$\liminf_{h\downarrow 0}\frac{1}{h}\operatorname{dist}\big(\eta+hf(t,\eta);\mathcal{P}(\xi)\big)=0,$$

then "$\preceq$" is locally invariant with respect to f.

5. Carathéodory solutions

5.1. *A Lebesgue type derivation theorem*

In this section we extend some of the previously established results to a more general case allowing the function f to be discontinuous with respect to the time variable. This case is very important in the study of some control problems to be analyzed in the sequel, when even starting with a continuous function $(t,v,u)\mapsto f(t,v,u)$, due to the discontinuities of the optimal control $t\mapsto v^*(t)$, we have to consider a very irregular right-hand side $(t,u)\mapsto f(t,v^*(t),u)$ which is only measurable with respect to t. First we recall:

DEFINITION 5.1.1. A function $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ is *of Carathéodory type*, if it satisfies the three conditions below.

(H$_1$) For every $\xi \in \mathbb{K}$ the function $f(\cdot, \xi)$ is measurable on $\mathbb{I}$.

(H$_2$) For almost every $t \in \mathbb{I}$ the function $f(t, \cdot)$ is continuous on $\mathbb{K}$.

(H$_3$) For every $m > 0$, there exists a function $\ell_m \in L^1_{\mathrm{loc}}(\mathbb{I})$ such that $\|f(t,u)\| \leqslant \ell_m(t)$ for almost every $t \in \mathbb{I}$ and for all $u \in B(0,m) \cap \mathbb{K}$.

DEFINITION 5.1.2. A function $u : [\tau, T] \to \mathbb{K}$ is a *Carathéodory solution* to the differential equation (1.1.1) if u is absolutely continuous on $[\tau, T]$ and satisfies $u'(t) = f(t, u(t))$ a.e. for $t \in [\tau, T]$.

DEFINITION 5.1.3. We say that $\mathbb{K}$ is *Carathéodory viable* with respect to $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ if for each $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ there exist $T \in \mathbb{I}$, $T > \tau$ and a Carathéodory solution $u : [\tau, T] \to \mathbb{K}$ to (1.1.1), satisfying $u(\tau) = \xi$.

Clearly, if $\mathbb{K}$ is viable with respect to f, then it is Carathéodory viable, but not conversely. In order to prove some necessary and sufficient conditions for Carathédory viability, we recall first a Lebesgue type derivation theorem due to Scorza Dragoni [88]. For more general results see Frankowska et al. [45] and Cârjă and Monteiro Marques [21].

THEOREM 5.1.1. *Assume* (H$_1$), (H$_2$) *and* (H$_3$). *Then there exists a negligible subset* $\mathbb{Z}$ *of* $\mathbb{I}$ *such that, for every* $t \in \mathbb{I} \setminus \mathbb{Z}$ *and every* $\xi \in \mathbb{K}$, *one has*

$$\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} f\big(s, u(s)\big)\,\mathrm{d}s = f(t, \xi) \tag{5.1.1}$$

for all continuous functions $u : \mathbb{I} \to \mathbb{K}$ *with* $u(t) = \xi$.

In order to prove Theorem 5.1.1, we recall for easy reference a specific form of a Lusin type continuity result due to Scorza Dragoni [87]. For more general results see Berliocchi and Lasry [10] and Kucia [63].

Here and thereafter, λ denotes the usual Lebesgue measure on $\mathbb{R}$.

THEOREM 5.1.2 (Scorza Dragoni [87]). *Let* $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ *be a function that satisfies* (H$_1$) *and* (H$_2$). *Then, for each* $\varepsilon > 0$, *there exists a closed set* $\mathbb{A} \subseteq \mathbb{I}$ *such that* $\lambda(\mathbb{I} \setminus \mathbb{A}) < \varepsilon$ *and the restriction of* f *to* $\mathbb{A} \times \mathbb{K}$ *is continuous.*

We note that the conclusion of Theorem 5.1.2 holds also true if $\mathbb{I}$ is replaced by any Lebesgue measurable subset in $\mathbb{R}$, while $\mathbb{K}$ and $\mathbb{R}^n$ are replaced by two separable metric spaces X and respectively Y. We can now proceed to the proof of Theorem 5.1.1.

PROOF OF THEOREM 5.1.1. Since $\mathbb{I}$ is a countable union of finite length intervals, it suffices to consider the case when $\mathbb{I}$ is of finite length. For each $\gamma > 0$, we shall obtain a set $\mathbb{L}_\gamma \subset \mathbb{I}$, with $\lambda(\mathbb{I} \setminus \mathbb{L}_\gamma) < \gamma$, and such that (5.1.1) holds for all $t \in \mathbb{L}_\gamma$. Finally, since $\lambda(\mathbb{I} \setminus \mathbb{L}_\gamma) < \gamma$, it will suffice to consider $\mathbb{Z} = \cap_m (\mathbb{I} \setminus \mathbb{L}_{1/m})$.

Let $\gamma > 0$ and let us observe that, by virtue of Theorem 5.1.2, it follows that there exists a compact set $\mathbb{A}_\gamma \subseteq \mathbb{I}$ such that $\lambda(\mathbb{I} \setminus \mathbb{A}_\gamma) < \gamma$, and the restriction of f to $\mathbb{A}_\gamma \times \mathbb{K}$ is continuous.

We define $\mathbb{L}_\gamma \subseteq \mathbb{A}_\gamma$ as the set of density points of $\mathbb{A}_\gamma$ which are also Lebesgue points of the functions $\tilde{\ell}_m : \mathbb{I} \to \mathbb{R}$, given by $\tilde{\ell}_m(t) = \ell_m(t)\chi_{\mathbb{I}\setminus\mathbb{A}_\gamma}(t)$, where ℓ_m is given by (H3), $m = 1, 2, \ldots$. It is known that $\lambda(\mathbb{L}_\gamma) = \lambda(\mathbb{A}_\gamma)$ and, by definition, for $t \in \mathbb{L}_\gamma$ we have

$$\begin{aligned} &\lim_{t\in\mathbb{J};\lambda(\mathbb{J})\to 0} \frac{\lambda(\mathbb{A}_\gamma \cap \mathbb{J})}{\lambda(\mathbb{J})} = 1, \\ &\lim_{t\in\mathbb{J};\lambda(\mathbb{J})\to 0} \frac{1}{\lambda(\mathbb{J})} \int_{\mathbb{J}} \left|\tilde{\ell}_m(s) - \tilde{\ell}_m(t)\right| \mathrm{d}s = 0, \end{aligned} \tag{5.1.2}$$

where $\mathbb{J}$ denotes intervals of positive length.

Let $t \in \mathbb{L}_\gamma$. Consider a continuous function $u : \mathbb{I} \to \mathbb{K}$, and denote by $\xi = u(t)$. Then, there is $m \geqslant 1$ such that $\|u(\theta)\| < m$ for all $\theta \in [t, t+\delta]$, where $\delta > 0$ is sufficiently small. Let $\varepsilon > 0$ be arbitrary. We can further assume that, for all $\theta \in \mathbb{A}_\gamma \cap [t, t+\delta]$,

$$\left\| f\big(\theta, u(\theta)\big) - f(t, \xi)\right\| \leqslant \frac{\varepsilon}{3}. \tag{5.1.3}$$

By taking a smaller δ if necessary, in view of (5.1.2), we can also ensure that

$$\frac{1}{s}\int_{[t,t+s]\setminus\mathbb{A}_\gamma} \ell_m(\theta)\,\mathrm{d}\theta \leqslant \frac{\varepsilon}{3}, \tag{5.1.4}$$

for every $s \in (0, \delta)$, and also

$$\frac{\lambda([t, t+s] \setminus \mathbb{A}_\gamma)}{s} \left\| f(t, \xi)\right\| \leqslant \frac{\varepsilon}{3} \tag{5.1.5}$$

for every $s \in (0, \delta)$. Then, by (5.1.3), for $s \in (0, \delta)$, we have

$$\frac{1}{s}\int_{[t,t+s]\cap\mathbb{A}_\gamma} \left\| f\big(\theta, u(\theta)\big) - f(t, \xi)\right\| \mathrm{d}\theta \leqslant \frac{\varepsilon}{3}\frac{\lambda([t, t+s] \cap \mathbb{A}_\gamma)}{s} \leqslant \frac{\varepsilon}{3},$$

while by (5.1.4) and (5.1.5) we have

$$\begin{aligned} &\frac{1}{s}\int_{[t,t+s]\setminus\mathbb{A}_\gamma} \left\| f\big(\theta, u(\theta)\big) - f(t, \xi)\right\| \mathrm{d}\theta \\ &\quad\leqslant \frac{1}{s}\int_{[t,t+s]\setminus\mathbb{A}_\gamma} \big(\ell_m(\theta) + \left\| f(t, \xi)\right\|\big)\,\mathrm{d}\theta \\ &\quad\leqslant \frac{1}{s}\int_{[t,t+s]\setminus\mathbb{A}_\gamma} \ell_m(\theta)\,\mathrm{d}\theta + \frac{\lambda([t, t+s] \setminus \mathbb{A}_\gamma)}{s}\left\| f(t, \xi)\right\| \leqslant 2\frac{\varepsilon}{3}. \end{aligned}$$

Finally, we have

$$\left\| \frac{1}{s} \int_t^{t+s} f\big(\theta, u(\theta)\big)\, \mathrm{d}\theta - f(t,\xi) \right\| \leqslant \varepsilon,$$

for all $s \in (0, \delta)$ and this completes the proof. □

Using similar arguments we can prove:

THEOREM 5.1.3. *Assume* (H$_1$), (H$_2$) *and* (H$_3$). *Then there exists a negligible subset* $\mathbb{Z}$ *of* $\mathbb{I}$ *such that*, *for every* $t \in \mathbb{I} \setminus \mathbb{Z}$ *and every* $\xi \in \mathbb{K}$, *one has*

$$\lim_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} f\big(s, v(h)\big)\, \mathrm{d}s = f(t,\xi)$$

for all functions $v : [0, +\infty) \to \mathbb{K}$ *satisfying* $\lim_{h\downarrow 0} v(h) = \xi$.

PROOF. The proof proceeds similarly with that one of Theorem 5.1.1 up to and including the sentence: "Let $t \in \mathbb{L}_\gamma$." We continue as follows. There is $m \geqslant 1$ such that $\|v(h)\| < m$ for $h \in (0, \delta)$ where $\delta > 0$ is sufficiently small. Let $\varepsilon > 0$ be arbitrary. We can assume that, for all $\theta \in \mathbb{A}_\gamma \cap [t, t+\delta]$ and $h \in (0, \delta)$,

$$\big\| f\big(\theta, v(h)\big) - f(t,\xi) \big\| \leqslant \frac{\varepsilon}{3}.$$

By taking a smaller δ if necessary, in view of (5.1.2), we can also ensure that (5.1.4) and (5.1.5) hold. From now on, the proof follows the very same way as that one of Theorem 5.1.1, with the observation that, here, $f(\theta, u(\theta))$ should be replaced by $f(\theta, v(h))$, wherever it appears. □

5.2. *Characterizations of Carathéodory viability*

We are now ready to prove the main characterizations of Carathéodory viability.

THEOREM 5.2.1. *Suppose that* $\mathbb{K}$ *is locally closed and* $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ *is of Carathéodory type. Then the following conditions are equivalent*:

(i) *there exists a negligible set* $\mathbb{Z} \subset \mathbb{I}$ *such that for every* $t \in \mathbb{I} \setminus \mathbb{Z}$ *and for every* $\xi \in \mathbb{K}$, $f(t,\xi) \in \mathcal{C}_{\mathbb{K}}(\xi)$,

(ii) *there exists a negligible set* $\mathbb{Z} \subset \mathbb{I}$ *such that for every* $t \in \mathbb{I} \setminus \mathbb{Z}$ *and for every* $\xi \in \mathbb{K}$, $f(t,\xi) \in \mathcal{T}_{\mathbb{K}}(\xi)$,

(iii) *there exists a negligible set* $\mathbb{Z} \subseteq \mathbb{I}$ *such that for every* $t \in \mathbb{I} \setminus \mathbb{Z}$ *and for every* $\xi \in \mathbb{K}$, $f(t,\xi) \in \mathcal{B}_{\mathbb{K}}(\xi)$,

(iv) *for each* $\xi \in \mathbb{K}$ *there exists a negligible set* $\mathbb{Z}_\xi \subseteq \mathbb{I}$ *such that for every* $t \in \mathbb{I} \setminus \mathbb{Z}_\xi$, $f(t,\xi) \in \mathcal{C}_{\mathbb{K}}(\xi)$,

(v) *for each* $\xi \in \mathbb{K}$ *there exists a negligible set* $\mathbb{Z}_\xi \subseteq \mathbb{I}$ *such that for every* $t \in \mathbb{I} \setminus \mathbb{Z}_\xi$, $f(t,\xi) \in \mathcal{T}_{\mathbb{K}}(\xi)$,

(vi) *for each* $\xi \in \mathbb{K}$ *there exists a negligible set* $\mathbb{Z}_\xi \subseteq \mathbb{I}$ *such that for every* $t \in \mathbb{I} \setminus \mathbb{Z}_\xi$, $f(t,\xi) \in \mathcal{B}_{\mathbb{K}}(\xi)$,

(vii) $\mathbb{K}$ *is Carathéodory viable with respect to* f.

In general, if $\mathcal{G} : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ *is such that* $\mathcal{C}_{\mathbb{K}}(\xi) \subseteq \mathcal{G}(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi)$ *for each* $\xi \in \mathbb{K}$, *then each one of the conditions below*:

(viii) *there exists a negligible set* $\mathbb{Z} \subseteq \mathbb{I}$ *such that for every* $t \in \mathbb{I} \setminus \mathbb{Z}$ *and for every* $\xi \in \mathbb{K}$, $f(t,\xi) \in \mathcal{G}(\xi)$,

(ix) *for each* $\xi \in \mathbb{K}$ *there exists a negligible set* $\mathbb{Z}_\xi \subseteq \mathbb{I}$ *such that for every* $t \in \mathbb{I} \setminus \mathbb{Z}_\xi$, $f(t,\xi) \in \mathcal{G}(\xi)$,

is equivalent to each one of the seven conditions above.

In order to prove Theorem 5.2.1, we need two auxiliary lemmas.

LEMMA 5.2.1. *Let* $\mathbb{K} \subseteq \mathbb{R}^n$ *be locally closed, let* $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ *be a function of Carathéodory type, let* $[\tau, T] \subseteq \mathbb{I}$, $\xi \in \mathbb{K}$, *and let* $\rho > 0$ *be such that* $\mathbb{K} \cap B(\xi,\rho)$ *is closed and there exists* $\ell \in L^1(\tau,T)$ *such that* $\|f(s,\eta)\| \leqslant \ell(s)$ *a.e. for* $s \in [\tau,T]$ *and for all* $\eta \in \mathbb{K} \cap B(\xi,\rho)$. *Then the family*

$$\mathcal{F} = \left\{ f(\cdot,\eta);\ \eta \in \mathbb{K} \cap B(\xi,\rho) \right\}$$

is compact in $L^1(\tau,T;\mathbb{R}^n)$.

PROOF. Let $(f(\cdot,\eta_m))_m$ be a sequence in $\mathcal{F}$. Since for all $m \in \mathbb{N}$, we have $\eta_m \in \mathbb{K} \cap B(\xi,\rho)$ and the latter is compact, there exist a subsequence of $(\eta_m)_m$, denoted for simplicity again by $(\eta_m)_m$, and $\eta \in \mathbb{K} \cap B(\xi,\rho)$, such that $\lim_m \eta_m = \eta$. It readily follows that $\lim_m f(s,\eta_m) = f(s,\eta)$ a.e. for $s \in [\tau,T]$. Since $\|f(s,\eta_m)\| \leqslant \ell(s)$ a.e. for $s \in [\tau,T]$ and for all $m \in \mathbb{N}$, by virtue of Lebesgue dominated convergence theorem, we deduce that

$$\lim_m \int_\tau^T \left\| f(s,\eta_m) - f(s,\eta) \right\| \mathrm{d}s = 0,$$

and this achieves the proof. □

LEMMA 5.2.2. *Let* $\mathbb{K} \subseteq \mathbb{R}^n$ *be locally closed, let* $f : \mathbb{I} \times \mathbb{K} \to \mathbb{R}^n$ *be a function of Carathéodory type, let* $[\tau, T] \subseteq \mathbb{I}$ *with* $T < \sup \mathbb{I}$, *let* $h > 0$ *with* $T + h < \sup \mathbb{I}$ *and let* $f_h : [\tau,T] \times \mathbb{K} \to \mathbb{R}^n$ *be defined by*

$$f_h(t,\xi) = \frac{1}{h} \int_t^{t+h} f(s,\xi)\,\mathrm{d}s$$

for each $(t,\xi) \in [\tau,T] \times \mathbb{K}$. *Then* f_h *is continuous on* $[\tau,T] \times \mathbb{K}$.

PROOF. Let $(t,\xi) \in [\tau,T] \times \mathbb{K}$. Since $\mathbb{K}$ is locally closed and f satisfies (H_3), there exist $\rho > 0$ and $\ell \in L^1(\tau,T)$ such that $K \cap B(\xi,\rho)$ is closed and $\|f(s,\eta)\| \leqslant \ell(s)$ a.e. for

$s \in [\tau, T]$ and for all $\eta \in K \cap B(\xi, \rho)$. Then, for each θ with $[t+\theta, t+h+\theta] \subset \mathbb{I}$, and each $\eta \in K \cap B(\xi, \rho)$, we have

$$\begin{aligned}
&\big\| f_h(t+\theta, \eta) - f_h(t, \xi) \big\| \\
&\quad = \left\| \frac{1}{h} \int_{t+\theta}^{t+\theta+h} f(s, \eta)\, \mathrm{d}s - \frac{1}{h} \int_{t}^{t+h} f(s, \xi)\, \mathrm{d}s \right\| \\
&\quad \leqslant \frac{1}{h} \int_{t}^{t+h} \big\| f(s+\theta, \eta) - f(s, \eta) \big\|\, \mathrm{d}s + \frac{1}{h} \int_{t}^{t+h} \big\| f(s, \eta) - f(s, \xi) \big\|\, \mathrm{d}s.
\end{aligned}$$

In view of Lemma 5.2.1 and Theorem A.1.1, we have

$$\lim_{\theta \to 0} \int_{t}^{t+h} \big\| f(s+\theta, \eta) - f(s, \eta) \big\|\, \mathrm{d}s = 0$$

uniformly for $\eta \in \mathbb{K} \cap B(\xi, \rho)$. Next, using condition (H_2) in Definition 5.1.1, the fact that $\| f(s, \eta) \| \leqslant \ell(s)$ a.e. for $s \in [t, t+h]$ and for all $\eta \in \mathbb{K} \cap B(\xi, \rho)$, by virtue of Lebesgue dominated convergence theorem, we deduce that

$$\lim_{\eta \to \xi} \int_{t}^{t+h} \big\| f(s, \eta) - f(s, \xi) \big\|\, \mathrm{d}s = 0,$$

relation which, along with the preceding one, shows that

$$\lim_{(\theta, \eta) \to (0, \xi)} \big\| f_h(t+\theta, \eta) - f_h(t, \xi) \big\| = 0,$$

and this completes the proof. □

We can now proceed to the proof of Theorem 5.2.1.

PROOF OF THEOREM 5.2.1. In view of Theorem 3.4.1, (i) (ii) and (iii) are equivalent. Furthermore (i) implies (iv), (iv) implies (v) and (v) implies (vi). In order to show that the first seven conditions are equivalent, it remains to prove that (vi) implies (vii) and (vii) implies (ii). Let us prove that (vi) implies (vii). To this aim, let $\mathbb{V}$ be an open neighborhood of $\mathbb{K}$ such that $\Pi_{\mathbb{K}}(\xi) \neq \emptyset$ for each $\xi \in \mathbb{V}$. The existence of the set $\mathbb{V}$, enjoying the specified properties, is ensured by Lemma 2.2.1. Let $\ell \in L^1_{\mathrm{loc}}(\mathbb{I})$ and let $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ be arbitrary. We claim that we can choose $\rho > 0$ and $T \in \mathbb{I}$ with $\tau < T < \sup \mathbb{I}$, such that $B(\xi, \rho) \cap \mathbb{K}$ is closed, $B(\xi, \rho) \subseteq \mathbb{V}$,

$$\big\| f(t, u) \big\| \leqslant \ell(t) \tag{5.2.1}$$

a.e. for $t \in [\tau, T]$ and for each $u \in B(\xi, \rho)$, and

$$\int_{\tau}^{T} \ell(s)\, \mathrm{d}s \leqslant \frac{\rho}{2}. \tag{5.2.2}$$

Indeed, since $\mathbb{V}$ is open and $\mathbb{K}$ is locally closed, we can get a $\rho > 0$ with $B(\xi,\rho)\cap\mathbb{K}$ closed and $B(\xi,\rho)\subseteq\mathbb{V}$. Further, as f satisfies (H$_3$), it follows that there exists $\ell\in L^1_{\mathrm{loc}}(\mathbb{I})$ satisfying (5.2.1). Finally, diminishing T if necessary, we get (5.2.2) and $T<\sup\mathbb{I}$. Next, let $h>0$ with $T+h<\sup\mathbb{I}$, and let us define $f_h:[\tau,T]\to\mathbb{R}^n$ by

$$f_h(t,\xi)=\frac{1}{h}\int_t^{t+h} f(s,\xi)\,\mathrm{d}s.$$

From Lemma 5.2.2, we know that f_h is continuous. On the other hand, since (vi) holds, we have $\langle f_h(t,\xi),\zeta\rangle\leqslant 0$, for all $t\in\mathbb{I}$ and $\zeta\in\mathcal{N}_{\mathbb{K}}(\xi)$, and thus the function f_h satisfies the condition (v) in Theorem 3.4.1 with $\mathcal{G}=\overline{\mathrm{co}}\,\mathcal{T}_{\mathbb{K}}(\xi)$. Therefore, for each $h>0$, with $T+h<\sup\mathbb{I}$, $\mathbb{K}$ is viable with respect to f_h. This implies that, for any $h>0$, $T+h<\sup\mathbb{I}$, there exists $T_h\in(\tau,T]$ and a solution $u_h:[\tau,T_h]\to\mathbb{K}$ of the differential equation $u'(t)=f_h(t,u(t))$ with $u_h(\tau)=\xi$. From Theorem 3.5.1, using (5.2.1) and (5.2.2), we conclude that we may choose $T_h=T$ for each $h>0$, $T+h<\sup\mathbb{I}$. Therefore, for each sufficiently small $h>0$, $u'(t)=f_h(t,u(t))$ has at least one solution $u_h:[\tau,T]\to B(\xi,\rho)\cap\mathbb{K}$ satisfying $u_h(\tau)=\xi$.

Take $h_m\downarrow 0$, and let $u_m=u_{h_m}$ be a solution to $u'(t)=f_m(t,u(t))$ with $f_m=f_{h_m}$ on $[\tau,T]$ and $u_m(\tau)=\xi$. We will prove that, for a certain function $u:[\tau,T]\to\mathbb{K}$, $\lim_m u_m(t)=u(t)$ uniformly for $t\in[\tau,T]$, and u is a Carathéodory solution to (1.1.1) satisfying $u(\tau)=\xi$. In order to do this, we will show first that $\{u_m;\ m\in\mathbb{N}\}$ satisfies the hypotheses of Arzelà–Ascoli theorem. Let $\mathbb{E}\subset[\tau,T]$ be a Lebesgue measurable set and let us observe that

$$\begin{aligned}
&\int_{\mathbb{E}}\left\|u_m'(t)\right\|\mathrm{d}t\\
&\quad=\int_{\mathbb{E}}\left\|f_{h_m}\big(t,u_m(t)\big)\right\|\mathrm{d}t=\int_{\mathbb{E}}\left\|\frac{1}{h_m}\int_t^{t+h_m} f\big(s,u_m(t)\big)\,\mathrm{d}s\right\|\mathrm{d}t\\
&\quad=\int_{\mathbb{E}}\left\|\frac{1}{h_m}\int_0^{h_m} f\big(t+s,u_m(t)\big)\,\mathrm{d}s\right\|\mathrm{d}t\\
&\quad\leqslant\frac{1}{h_m}\int_0^{h_m}\int_{\mathbb{E}}\left\|f\big(t+s,u_m(t)\big)\right\|\mathrm{d}t\,\mathrm{d}s\leqslant\frac{1}{h_m}\int_0^{h_m}\int_{\mathbb{E}}\ell(t+s)\,\mathrm{d}t\,\mathrm{d}s.
\end{aligned}$$

Since $\ell\in L^1_{\mathrm{loc}}(\mathbb{I})$, for each $\varepsilon>0$ there exists $\delta(\varepsilon)>0$ such that, for each $\mathbb{E}\subset[\tau,T+h_1]$ with $\lambda(\mathbb{E})\leqslant\delta(\varepsilon)$, we have

$$\int_{\mathbb{E}}\ell(t)\,\mathrm{d}t\leqslant\varepsilon.$$

Since the Lebesgue measure is translation invariant, it follows that, for each $\mathbb{E}\subseteq[\tau,T]$ with $\lambda(\mathbb{E})\leqslant\delta(\varepsilon)$, we have

$$\int_{\mathbb{E}}\ell(t+s)\,\mathrm{d}t\leqslant\varepsilon$$

for each $s \in [0, h_1]$. Summing up, we conclude that

$$\int_{\mathbb{E}} \|u_m'(t)\| \,\mathrm{d}t \leqslant \frac{1}{h_m} \int_0^{h_m} \int_{\mathbb{E}} \ell(t+s) \,\mathrm{d}t \,\mathrm{d}s \leqslant \varepsilon$$

for each $\mathbb{E} \subseteq [\tau, T]$ with $\lambda(\mathbb{E}) \leqslant \delta(\varepsilon)$. So $\{u_m';\ m \in \mathbb{N}\}$ is uniformly integrable. This implies that $\{u_m;\ m \in \mathbb{N}\}$ is equicontinuous. Moreover, since $u_m(t) \in \mathbb{K} \cap B(\xi, \rho)$ for each $t \in [\tau, T]$, if readily follows that $\{u_m;\ m \in \mathbb{N}\}$ is uniformly bounded. Therefore, there exists $u : [\tau, T] \to \mathbb{K} \cap B(\xi, \rho)$ such that, at least on a subsequence, denoted for simplicity again by $(u_m)_m$, we have $\lim_m u_m(t) = u(t)$ uniformly for $t \in [\tau, T]$. Clearly u is continuous. On the other hand, by Theorem 5.1.3, we have

$$\lim_{m\to\infty} \frac{1}{h_m} \int_s^{s+h_m} f\big(\theta, u_m(s)\big) \,\mathrm{d}\theta = f\big(s, u(s)\big)$$

a.e. for $s \in [\tau, T]$. Since $\{u_m';\ m \in \mathbb{N}\}$ is uniformly integrable, by virtue of Vitali's Theorem A.1.2, we conclude that

$$\begin{aligned}\lim_{m\to\infty} u_m(t) &= \lim_{m\to\infty} \left(\xi + \int_\tau^t \frac{1}{h_m} \int_s^{s+h_m} f\big(\theta, u_m(s)\big) \,\mathrm{d}\theta \,\mathrm{d}s \right) \\ &= \xi + \int_\tau^t f\big(s, u(s)\big) \,\mathrm{d}s,\end{aligned}$$

which shows that u is a Carathéodory solution to (1.1.1) satisfying $u(\tau) = \xi$. Hence $\mathbb{K}$ is Carathéodory viable with respect to f, and this shows that (vi) implies (vii).

Next, to prove that (vii) implies (ii), let $\mathbb{Z} \subset \mathbb{I}$ be a negligible set as in Theorem 5.1.1. Let $t \in \mathbb{I} \setminus \mathbb{Z}$, $\xi \in \mathbb{K}$ and let $u(\cdot)$ be a Carathéodory solution to $u'(s) = f(s, u(s))$ on $[t, t+T]$ with $u(t) = \xi$ and $u(s) \in \mathbb{K}$ for $s \in [t, t+h]$. We have

$$\frac{u(t+h) - u(t)}{h} = \frac{1}{h} \int_t^{t+h} f\big(s, u(s)\big) \,\mathrm{d}s.$$

On the other hand, by (5.1.1), the right-hand side approaches $f(t, \xi)$ as $h \downarrow 0$. Accordingly

$$\lim_{h\downarrow 0} \frac{1}{h} \operatorname{dist}\big(\xi + hf(t, \xi); \mathbb{K}\big) \leqslant \lim_{h\downarrow 0} \left\| f(t, \xi) - \frac{u(t+h) - u(t)}{h} \right\| = 0$$

and so, $f(t, \xi) \in \mathcal{F}_{\mathbb{K}}(\xi) \subseteq \mathcal{T}_{\mathbb{K}}(\xi)$, which shows that (ii) holds.

Now, (viii) implies (iii) which implies (vii), and (vii) implies (i) which implies (viii). Finally, (ix) implies (vi) which implies (vii), and (vii) implies (iv) which implies (ix). The proof is complete. □

5.3. *Sufficient conditions for Carathéodory local invariance*

The aim of this section is to extend the results concerning local invariance to the more general setting allowing the function f to be of Carathéodory type. Unlike the case of necessary and sufficient conditions for Carathéodory viability whose proofs differ essentially from those referring to viability, all the results which will follow are simple copies of their "C^1-solutions" counterparts we already presented when we dealt with local invariance. Therefore, we do not enter into details.

Let $\mathbb{D}$ be a domain in $\mathbb{R}^n$, $\mathbb{K} \subseteq \mathbb{D}$ a locally closed subset, and let us consider the ordinary differential equation (1.1.1), where $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ is a given function.

DEFINITION 5.3.1. The subset $\mathbb{K}$ is *Carathéodory locally invariant* with respect to f if for each $(\tau, \xi) \in \mathbb{I} \times \mathbb{K}$ and each Carathéodory solution $u : [\tau, c] \to \mathbb{D}$, $c \in \mathbb{I}$, $c > \tau$, to (1.1.1), satisfying the initial condition $u(\tau) = \xi$, there exists $T \in (\tau, c]$ such that we have $u(t) \in \mathbb{K}$ for each $t \in [\tau, T]$. It is *Carathéodory invariant* if it satisfies the above condition of Carathéodory local invariance with $T = c$.

REMARK 5.3.1. If $\mathbb{K}$ is Carathéodory locally invariant with respect to f, then $\mathbb{K}$ is locally invariant with respect to f. We notice that here we do have any extra-condition upon f as of being of Carathéodory type. It should be interesting to know if, under the latter assumption on f, the local invariance implies the Carathéodory local invariance.

The relationship between Carathéodory viability and Carathéodory local invariance is clarified in:

REMARK 5.3.2. If f is of Carathéodory type on $\mathbb{I} \times \mathbb{D}$ and $\mathbb{K}$ is Carathéodory locally invariant with respect to f, then $\mathbb{K}$ is Carathéodory viable with respect to f. The converse of this assertion is no longer true, as we already have seen in Example 4.1.1.

As in the case of local invariance, we have the following simple necessary and sufficient condition of Carathéodory local invariance.

THEOREM 5.3.1. *Let $\mathbb{D}$ be a domain, $\mathbb{K} \subseteq \mathbb{D}$ a nonempty and locally closed subset and $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ a Carathéodory type function with the property that* (1.1.1) *has the uniqueness property. Then, a necessary and sufficient condition in order that $\mathbb{K}$ be invariant with respect to f is each one of the nine conditions* (i)–(ix) *in Theorem* 5.2.1.

PROOF. The conclusion follows from the Carathéodory Local Existence Theorem combined with Remark 5.3.2 and the equivalence between (i) and (ix) in Theorem 5.2.1. □

As in the case of local invariance, the Carathéodory local invariance of $\mathbb{K}$ with respect to f can take place even if $u'(t) = f_{|\mathbb{I}\times\mathbb{K}}(t, u(t))$ has not the uniqueness property. See Example 4.1.2.

We say that $\omega : \mathbb{I} \times [0, \rho) \to \mathbb{R}$ is a *Carathéodory comparison function* if, for each $[\tau, T) \subseteq \mathbb{I}$, the only absolutely continuous solution $x : [\tau, T) \to [0, \rho)$, of $[D_+x](t) \leqslant \omega(t, x(t))$ a.e. for $t \in [\tau, T)$, satisfying $x(\tau) = 0$, is the null function.

Our first sufficient condition for Carathéodory local invariance is expressed by means of the "exterior tangency" condition: there exist a negligible subset $\mathbb{Z}$ in $\mathbb{I}$ and an open neighborhood $\mathbb{V} \subseteq \mathbb{D}$ of $\mathbb{K}$ such that

$$\liminf_{h \downarrow 0} \frac{1}{h}\bigl[\operatorname{dist}\bigl(\xi + hf(t,\xi); \mathbb{K}\bigr) - \operatorname{dist}(\xi; \mathbb{K})\bigr] \leqslant \omega\bigl(t, \operatorname{dist}(\xi; \mathbb{K})\bigr) \tag{5.3.1}$$

for each $(t, \xi) \in (\mathbb{I} \setminus \mathbb{Z}) \times \mathbb{V}$, where ω is a certain Carathéodory comparison function.

The next result is a "Carathéodory" counterpart of Theorem 4.2.1.

THEOREM 5.3.2. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. If* (5.3.1) *is satisfied, then $\mathbb{K}$ is Carathéodory locally invariant with respect to f.*

Theorem 5.3.2 can be reformulated as:

THEOREM 5.3.3. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. If there exists an open neighborhood V of $\mathbb{K}$ with $\mathbb{V} \subseteq \mathbb{D}$ such that ω_f defined by* (4.2.2) *is a Carathéodory comparison function, then $\mathbb{K}$ is Carathéodory locally invariant with respect to f.*

DEFINITION 5.3.2. Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$. We say that a function $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ has the *Carathéodory comparison property with respect to* $(\mathbb{D}, \mathbb{K})$ if there exist a negligible subset $\mathbb{Z}$ of $\mathbb{I}$, a proximal neighborhood $\mathbb{V} \subseteq \mathbb{D}$ of $\mathbb{K}$, one projection $\pi_{\mathbb{K}} : \mathbb{V} \to \mathbb{K}$ subordinated to $\mathbb{V}$, and a Carathéodory comparison function $\omega : \mathbb{I} \times [0, a) \to \mathbb{R}$, with $a = \sup_{\xi \in \mathbb{V}} \operatorname{dist}(\xi; \mathbb{K})$, such that

$$\bigl[\xi - \pi_{\mathbb{K}}(\xi), f(t,\xi) - f\bigl(t, \pi_{\mathbb{K}}(\xi)\bigr)\bigr]_+ \leqslant \omega\bigl(t, \bigl\|\xi - \pi_{\mathbb{K}}(\xi)\bigr\|\bigr)$$

for each $(t, \xi) \in (\mathbb{I} \setminus \mathbb{Z}) \times \mathbb{V}$.

THEOREM 5.3.4. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. If f has the Carathéodory comparison property with respect to* $(\mathbb{D}, \mathbb{K})$, *and $\mathbb{K}$ is Carathéodory viable with respect f, then* (5.3.1) *holds true.*

Other sufficient conditions for invariance obtained from Theorem 5.2.1 are:

THEOREM 5.3.5. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ be of Carathéodory type. Let us assume that f has the Carathéodory comparison property with respect to* $(\mathbb{D}, \mathbb{K})$, *and satisfies one of the nine equivalent tangency conditions* (i)–(ix) *in Theorem* 5.2.1. *Then $\mathbb{K}$ is Carathéodory locally invariant with respect to f.*

5.4. *Sufficient conditions via generalized distance*

Here we reconsider the sufficient conditions for local invariance obtained by means of a proximal generalized distance in the context of Carathéodory type functions.

THEOREM 5.4.1. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. If there exist an open neighborhood $\mathbb{V}$ of $\mathbb{K}$, with $\mathbb{V} \subseteq \mathbb{D}$, a proximal generalized distance $g : \mathbb{V} \to \mathbb{R}_+$, a negligible subset $\mathbb{Z}$ in $\mathbb{I}$ and a Carathéodory comparison function $\omega : \mathbb{I} \times [0, a) \to \mathbb{R}$ such that*

$$\liminf_{h \downarrow 0} \frac{1}{h}\bigl[g\bigl(\xi + hf(t,\xi)\bigr) - g(\xi)\bigr] \leqslant \omega\bigl(t, g(\xi)\bigr)$$

for each $(t, \xi) \in (\mathbb{I} \setminus \mathbb{Z}) \times \mathbb{V}$, then $\mathbb{K}$ is Carathéodory locally invariant with respect to f.

Let $g : \mathbb{V} \to [0, +\infty)$ be a proximal generalized distance, and let us introduce the generalized Carathéodory tangency condition: *there exists a negligible subset $\mathbb{Z}$ of $\mathbb{I}$ such that*

$$\liminf_{h \downarrow 0} \frac{1}{h} g\bigl(\xi + hf(t,\xi)\bigr) = 0 \tag{5.4.1}$$

for each $(t, \xi) \in (\mathbb{I} \setminus \mathbb{Z}) \times \mathbb{K}$.

THEOREM 5.4.2. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. If there exist an open neighborhood $\mathbb{V}$ of $\mathbb{K}$, with $\mathbb{V} \subseteq \mathbb{D}$, a proximal generalized distance $g : \mathbb{V} \to \mathbb{R}_+$, a negligible subset $\mathbb{Z}$ in $\mathbb{I}$ and a Carathéodory comparison function $\omega : \mathbb{I} \times [0, a) \to \mathbb{R}$ satisfying* (5.4.1) *and such that*

$$\liminf_{h \downarrow 0} \frac{1}{2h}\bigl[g^2\bigl(\xi + hf(t,\xi)\bigr) - g^2(\xi)\bigr] \leqslant g(\xi)\omega\bigl(t, g(\xi)\bigr)$$

for each $(t, \xi) \in (\mathbb{I} \setminus \mathbb{Z}) \times \mathbb{V}$, then $\mathbb{K}$ is Carathéodory locally invariant with respect to f.

THEOREM 5.4.3. *Let $\mathbb{D} \subseteq \mathbb{R}^n$ be open and let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. Let us assume that $\mathbb{K} \subseteq \mathbb{D}$ is a Lipschitz retract with the generalized Lipschitz projection $r : \mathbb{V} \to \mathbb{K}$, and there exists a negligible subset $\mathbb{Z}$ of $\mathbb{I}$ satisfying*

$$\liminf_{h \downarrow 0} \frac{1}{h}\bigl\|r\bigl(\xi + hf(t,\xi)\bigr) - \xi - hf(t,\xi)\bigr\| = 0 \tag{5.4.2}$$

for each $(t, \xi) \in (\mathbb{I} \setminus \mathbb{Z}) \times \mathbb{K}$. Assume further that there exists a Carathéodory comparison function $\omega : \mathbb{I} \times [0, a) \to \mathbb{R}$, with $a = \sup_{\xi \in \mathbb{V}} \|r(\xi) - \xi\|$, such that

$$\begin{aligned}&\liminf_{h \downarrow 0} \frac{1}{h}\bigl\langle r\bigl(\xi + hf(t,\xi)\bigr) - r(\xi) - hf(t,\xi), r(\xi) - \xi\bigr\rangle \\ &\quad \leqslant \bigl\|r(\xi) - \xi\bigr\|\omega\bigl(t, \bigl\|r(\xi) - \xi\bigr\|\bigr)\end{aligned}$$

for each $(t, \xi) \in (\mathbb{I} \setminus \mathbb{Z}) \times \mathbb{V}$. Then, $\mathbb{K}$ is Carathéodory locally invariant with respect to f.

A consequence of Theorem 5.4.3 is stated below.

THEOREM 5.4.4. *Let* $\mathbb{K} \subseteq \mathbb{D}$ *with* $\mathbb{D}$ *open and let* $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$. *Let us assume that* $\mathbb{K}$ *is a Lipschitz retract with the generalized Lipschitz projection* $r : \mathbb{V} \to \mathbb{K}$ *and there exists a negligible subset* $\mathbb{Z}$ *of* $\mathbb{I}$ *such that* (5.4.2) *is satisfied. Let us assume, in addition, that, for each* $t \in \mathbb{I} \setminus \mathbb{Z}$ *and* $\xi \in \mathbb{V}$, *there exists the directional derivative,* $r'(\xi)[f(t,\xi)]$, *of* r, *at* ξ *in the direction* $f(t,\xi)$, *and*

$$\big\langle r'(\xi)\big[f(t,\xi)\big] - f(t,\xi), r(\xi) - \xi\big\rangle \leqslant \big\|r(\xi) - \xi\big\|\omega\big(t, \big\|r(\xi) - \xi\big\|\big),$$

where $\omega : \mathbb{I} \times [0,a) \to \mathbb{R}$ *is a Carathéodory comparison function, with*

$$a = \sup_{\xi \in \mathbb{V}} \big\|r(\xi) - \xi\big\|.$$

Then, $\mathbb{K}$ *is Carathéodory locally invariant with respect to* f.

6. Differential inclusions

6.1. *Multifunctions*

In this section we extend some of the previously established results to a more general case allowing the function f to be multivalued. First we introduce two classes of multifunctions. Let $\mathbb{K}$ be a nonempty subset in $\mathbb{R}^n$ and let $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ a given mapping with nonempty values.

DEFINITION 6.1.1. The multifunction $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ is *upper semicontinuous* (*u.s.c.*) *at* $\xi \in \mathbb{K}$ if for every open neighborhood $\mathbb{V}$ of $F(\xi)$ there exists an open neighborhood $\mathbb{U}$ of ξ such that $F(\eta) \subseteq \mathbb{V}$ for each $\eta \in \mathbb{U} \cap \mathbb{K}$. We say that F is *upper semicontinuous* (*u.s.c.*) *on* $\mathbb{K}$ if it is u.s.c. at each $\xi \in \mathbb{K}$.

DEFINITION 6.1.2. The multifunction $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ is *lower semicontinuous* (*l.s.c.*) *at* $\xi \in \mathbb{K}$ if for every open set $\mathbb{V}$ in $\mathbb{R}^n$ with $F(\xi) \cap \mathbb{V} \neq \emptyset$ there exists an open neighborhood $\mathbb{U}$ of ξ such that $F(\eta) \cap \mathbb{V} \neq \emptyset$ for each $\eta \in \mathbb{U} \cap \mathbb{K}$. We say that F is *lower semicontinuous* (*l.s.c.*) *on* $\mathbb{K}$ if it is l.s.c. at each $\xi \in \mathbb{K}$.

The next two lemmas will prove useful later.

LEMMA 6.1.1. *If* $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ *is a nonempty and compact valued u.s.c. multifunction, then, for each compact subset* $\mathbb{C}$ *of* $\mathbb{K}$, $\bigcup_{\xi \in \mathbb{C}} F(\xi)$ *is compact. In particular, for each compact subset* $\mathbb{C}$ *of* $\mathbb{K}$, *there exists* $M > 0$ *such that* $\|\eta\| \leqslant M$ *for each* $\xi \in \mathbb{C}$ *and each* $\eta \in F(\xi)$.

PROOF. Let $\mathbb{C}$ be a compact subset in $\mathbb{K}$ and let $\{\mathbb{D}_\sigma;\ \sigma \in \Gamma\}$ be an arbitrary open covering of $\bigcup_{\xi \in \mathbb{C}} F(\xi)$. Since F is compact valued, for each $\xi \in \mathbb{C}$ there exists $n(\xi) \in \mathbb{N}$ such that

$$F(\xi) \subseteq \bigcup_{1 \leqslant k \leqslant n(\xi)} \mathbb{D}_{\sigma_k}.$$

But F is u.s.c. and therefore there exists an open neighborhood $\mathbb{U}(\xi)$ of ξ such that

$$F\big(\mathbb{U}(\xi)\cap\mathbb{K}\big)\subseteq\bigcup_{1\leqslant k\leqslant n(\xi)}\mathbb{D}_{\sigma_k}.$$

The family $\{\mathbb{U}(\xi);\ \xi\in\mathbb{C}\}$ is an open covering of $\mathbb{C}$. As $\mathbb{C}$ is compact, there exists a finite family $\{\xi_1,\xi_2,\dots,\xi_p\}$ in $\mathbb{C}$ such that

$$F(\mathbb{C})\subseteq\bigcup_{1\leqslant j\leqslant p}F\big(\mathbb{U}(\xi_j)\cap\mathbb{K}\big)\subseteq\bigcup_{1\leqslant j\leqslant p}\ \bigcup_{1\leqslant k\leqslant n(\xi_j)}\mathbb{D}_{\sigma_k},$$

and this completes the proof. □

LEMMA 6.1.2. *Let $F:\mathbb{K}\rightsquigarrow\mathbb{R}^n$ be a u.s.c., nonempty, convex, compact valued multifunction, and let $u_m:[0,T]\to\mathbb{K}$ and $f_m\in L^1(0,T;\mathbb{R}^n)$ be such that $f_m(t)\in F(u_m(t))$ for each $m\in\mathbb{N}$ and a.e. for $t\in[0,T]$. If $\lim_m u_m(t)=u(t)$ a.e. for $t\in[0,T]$ and $\lim_m f_m=f$ weakly in $L^1(0,T;\mathbb{R}^n)$, then $f(t)\in F(u(t))$ a.e. for $t\in[0,T]$.*

PROOF. By Mazur's theorem—Dunford–Schwartz [40, Theorem 6, p. 416]—there exists a sequence $(g_m)_m$ of convex combinations of $\{f_k;\ k\geqslant m\}$, i.e. $g_m\in\operatorname{co}\{f_m,f_{m+1},\dots\}$ for each $m\in\mathbb{N}$, which converges strongly in $L^1(0,T;\mathbb{R}^n)$ to f. By a classical result due to Lebesgue, we know that there exists a subsequence (g_{m_p}) of (g_m) which converges almost everywhere on $[0,T]$ to f. Denote by $\mathbb{T}$ the set of all $s\in[0,T]$ such that both $(g_{m_p}(s))_p$ and $(u_m(s))_m$ are convergent to $f(s)$ and to $u(s)$, respectively, and, in addition, $f_m(s)\in F(u_m(s))$ for each $m\in\mathbb{N}$. Clearly $[0,T]\setminus\mathbb{T}$ has null measure. Let $s\in\mathbb{T}$ and let $\mathbb{E}$ be an open half-space in $\mathbb{R}^n$ including $F(u(s))$. Since F is upper semicontinuous at $u(s)$ and $(u_m(s))_m$ converges to $u(s)$, there exists $m(\mathbb{E})$ belonging to $\mathbb{N}$, such that $F(u_m(s))\subseteq\mathbb{E}$ for each $m\geqslant m(\mathbb{E})$. From the relation above, taking into account that $f_m(s)\in F(u_m(s))$ for each $m\in\mathbb{N}$ and a.e. for $s\in[0,T]$, we easily conclude

$$g_{m_p}(s)\in\overline{\operatorname{co}}\bigg(\bigcup_{m\geqslant m(\mathbb{E})}F\big(u_m(s)\big)\bigg)$$

for each $p\in\mathbb{N}$ with $m_p\geq m(\mathbb{E})$. Passing to the limit for $p\to+\infty$ in the relation above we deduce $f(s)\in\overline{\mathbb{E}}$. Since $F(u(s))$ is closed and convex, it is the intersection of all closed half-spaces which include it. So, inasmuch as $\mathbb{E}$ was arbitrary, we, finally, get $f(s)\in F(u(s))$ for each $s\in\mathbb{T}$ and this completes the proof. □

6.2. *Viability with respect to a multifunction*

Let us consider the autonomous differential inclusion

$$u'(t)\in F\big(u(t)\big)\tag{6.2.1}$$

where $F:\mathbb{K}\rightsquigarrow\mathbb{R}^n$ is a given multifunction. By a *solution* of (6.2.1) we mean an absolutely continuous function $u:\mathbb{J}\to\mathbb{K}$, with $\mathbb{J}$ a proper interval, satisfying $u'(t)\in F(u(t))$ a.e. for $t\in\mathbb{J}$.

DEFINITION 6.2.1. We say that $\mathbb{K}$ is *right viable with respect to* F if for each $\xi\in\mathbb{K}$, there exists $T>0$ such that (6.2.1) has at least one solution $u:[0,T]\to\mathbb{K}$ satisfying $u(0)=\xi$.

THEOREM 6.2.1. *Let $\mathbb{K}$ be a locally closed subset in $\mathbb{R}^n$ and $F:\mathbb{K}\rightsquigarrow\mathbb{R}^n$ a u.s.c. nonempty, closed, convex and bounded valued mapping. Then a necessary and sufficient condition in order that $\mathbb{K}$ be right viable with respect to F is the tangency condition*

$$F(\xi)\cap\mathcal{T}_{\mathbb{K}}(\xi)\neq\emptyset \tag{6.2.2}$$

for each $\xi\in\mathbb{K}$.

PROOF. *Necessity*. Since F is upper semicontinuous and has nonempty, closed, convex and bounded values, by virtue of Lemma 6.1.1, we conclude that F is locally bounded. Let $\xi\in\mathbb{K}$ and let $u:[0,T]\to\mathbb{K}$ be a solution to (6.2.1) satisfying $u(0)=\xi$. Let $h\in(0,T]$ and let us define

$$u_h=\frac{1}{h}\int_0^h u'(s)\,\mathrm{d}s.$$

Since F is locally bounded, by diminishing $T>0$ if necessary, we may assume that $\{u'(s);\ s\in[0,T]\}$ is bounded a.e. In other words, there exists $M>0$ such that $\|u'(s)\|\leqslant M$ a.e. for $s\in[0,T]$. This means that the set $\{\|u_h\|;\ h\in(0,T]\}$ is bounded, also by M, and therefore there exists $h_m\downarrow 0$ such that $\eta=\lim_{m\to\infty}u_{h_m}$. Since F is upper semicontinuous at ξ and has convex compact values, we infer that $\eta\in F(\xi)$. Indeed, for each open neighborhood $\mathbb{V}$ of $F(\xi)$ there exists an open neighborhood $\mathbb{U}$ of ξ such that $F(\zeta)\subseteq\mathbb{V}$ for each $\zeta\in\mathbb{K}\cap\mathbb{U}$. Therefore, if $h\in(0,T]$ is sufficiently small, we have $u'(s)\in\mathbb{V}$ a.e. for $s\in(0,h]$. So, $u_h\in\overline{\text{co}}\,\mathbb{V}$. Accordingly $\eta\in\bigcap_{\mathbb{V}\in\mathcal{V}(F(\xi))}\overline{\text{co}}\,\mathbb{V}$, where $\mathcal{V}(F(\xi))$ stands for the set of open neighborhoods of $F(\xi)$. But $F(\xi)$ is convex, and hence $\bigcap_{\mathbb{V}\in\mathcal{V}(F(\xi))}\overline{\text{co}}\,\mathbb{V}=F(\xi)$. To complete the proof of the necessity, it suffices to show that $\eta\in\mathcal{T}_{\mathbb{K}}(\xi)$. To this aim, let us observe that $\xi+hu_h=u(h)\in\mathbb{K}$, and so $\text{dist}(\xi+hu_h;\mathbb{K})=0$. Thus

$$\text{dist}(\xi+h\eta;\mathbb{K})\leqslant\text{dist}(\xi+hu_h;\mathbb{K})+\big\|\xi+hu_h-(\xi+h\eta)\big\|\leqslant h\|u_h-\eta\|.$$

Since η is a limit point of u_h as $h\downarrow 0$, we deduce

$$\liminf_{h\downarrow 0}\frac{1}{h}\,\text{dist}(\xi+h\eta;\mathbb{K})=0$$

which shows that $\eta\in\mathcal{T}_{\mathbb{K}}(\xi)$. The proof of the necessity is complete. □

REMARK 6.2.1. On may ask whether the equivalence between (i) and (ii) in Proposition 2.5.1 can be extended to this framework. Namely, one may ask on the equivalence of the two conditions below:

(j) for each $\xi \in \mathbb{K}$, $F(\xi) \cap \mathcal{C}_{\mathbb{K}}(\xi) \neq \emptyset$,
(jj) for each $\xi \in \mathbb{K}$, $F(\xi) \cap \mathcal{T}_{\mathbb{K}}(\xi) \neq \emptyset$.

The next simple example shows that (j) and (jj) are not equivalent. Namely, take $n = 2$, $\mathbb{K} = \{(x_1, x_2);\ |x_1| \geqslant x_2\}$ and $\mathbb{V} = \{(x_1, 1);\ |x_1| \leqslant 1\}$, and let us define $F : \mathbb{K} \rightsquigarrow \mathbb{R}^2$ by $F(\xi) = \mathbb{V}$, for each $\xi \in \mathbb{K}$. Clearly, F is both u.s.c. and l.s.c., with nonempty, convex, compact values and (jj) holds true. Nevertheless $F(0,0) \cap \mathcal{C}_{\mathbb{K}}((0,0)) = \emptyset$.

We recall that, whenever $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ is l.s.c. at $\xi \in \mathbb{K}$, for each $\eta \in F(\xi)$, and each $(\xi_m)_m$ with $\lim_m \xi_m = \xi$, there exist $(\eta_m)_m$ with $\eta_m \in F(\xi_m)$ for each $m \in \mathbb{N}$, and such that $\lim_m \eta_m = \eta$. See, for instance, Cârjă [20, Theorem 2.4, p. 25]. So, from Remark 2.5.1 and Lemma 2.5.1, we deduce:

PROPOSITION 6.2.1. *If $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ is l.s.c., then the following conditions are equivalent*:

(i) *for each $\xi \in \mathbb{K}$, $F(\xi) \subseteq \mathcal{C}_{\mathbb{K}}(\xi)$,*
(ii) *for each $\xi \in \mathbb{K}$, $F(\xi) \subseteq \mathcal{T}_{\mathbb{K}}(\xi)$,*
(iii) *for each $\xi \in \mathbb{K}$, $F(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi)$,*

In general, if $\mathcal{G} : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ is such that $\mathcal{C}_{\mathbb{K}}(\xi) \subseteq \mathcal{G}(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi)$ for each $\xi \in \mathbb{K}$, then each one of the conditions above is equivalent to:

(iv) *for each $\xi \in \mathbb{K}$, $F(\xi) \subseteq \mathcal{G}(\xi)$.*

REMARK 6.2.2. The next simple example shows that the equivalence between (i) and (ii) in Proposition 6.2.1 is no longer true if F is u.s.c. Namely, take $n = 2$, $\mathbb{K} = \{(x_1, x_2);\ |x_1| \geqslant x_2\}$ and let us define $F : \mathbb{K} \rightsquigarrow \mathbb{R}^2$ by

$$F(\xi) = \begin{cases} \{(0,0)\} & \text{if } \xi \in \mathbb{K} \setminus \{(0,0)\}, \\ \mathbb{K} & \text{if } \xi = (0,0). \end{cases}$$

Clearly, F is u.s.c. with nonempty, closed values and (ii) holds true. Nevertheless (i) does not hold.

6.3. *Existence of ε-approximate solutions*

The proof of the sufficiency consists in showing that (6.2.2) along with Brezis–Browder Theorem 2.1.1 imply that, for each $\xi \in \mathbb{K}$, there exists at least one sequence of "approximate" solutions to (6.2.1), defined on the same interval, $u_m : [0, T] \to \mathbb{R}^n$, satisfying $u_m(0) = \xi$ for each $m \in \mathbb{N}$, and such that $(u_m)_m$ converges in some sense to a solution u of (6.2.1) satisfying $u(0) = \xi$.

The next lemma represents an existence result concerning "approximate solutions" of (6.2.1) satisfying $u_m(0) = \xi$.

LEMMA 6.3.1. *Let $\mathbb{K}$ be a nonempty and locally closed subset in $\mathbb{R}^n$, and let $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ be a nonempty valued mapping which is locally bounded. If $\mathbb{K}$ and F satisfy* (6.2.2) *then for* *each* $\xi \in \mathbb{K}$ *there exist* $\rho > 0$, $T > 0$, $M > 0$ *such that* $\mathbb{K} \cap B(\xi, \rho)$ *is closed, and for each* $\varepsilon \in (0, 1]$ *there exist four functions*: $\sigma : [0, T] \to [0, T]$ *nondecreasing*, $f : [0, T] \to \mathbb{R}^n$ *and* $g : [0, T] \to \mathbb{R}^n$ *measurable, and* $u : [0, T] \to \mathbb{R}^n$ *continuous, satisfying*

(i) $s - \varepsilon \leqslant \sigma(s) \leqslant s$, $u(\sigma(s)) \in \mathbb{K} \cap B(\xi, \rho)$ *and* $f(s) \in F(u(\sigma(s)))$, *a.e. for* $s \in [0, T]$,
(ii) $\|f(s)\| \leqslant M$ *a.e. for* $s \in [0, T]$,
(iii) $u(T) \in \mathbb{K} \cap B(\xi, \rho)$,
(iv) $\|g(s)\| \leqslant \varepsilon$ *a.e. for* $s \in [0, T]$

and

$$u(t) = \xi + \int_0^t f(s)\,\mathrm{d}s + \int_0^t g(s)\,\mathrm{d}s \tag{6.3.1}$$

for each $t \in [0, T]$.

PROOF. Let $\xi \in \mathbb{K}$ be arbitrary, and choose $\rho > 0$, $T > 0$ and $M > 0$ such that $\mathbb{K} \cap B(\xi, \rho)$ is closed,

$$\|\eta\| \leqslant M \tag{6.3.2}$$

for each $x \in \mathbb{K} \cap B(\xi, \rho)$ and $\eta \in F(x)$, and

$$T(M + 1) \leqslant \rho. \tag{6.3.3}$$

This is always possible since $\mathbb{K}$ is locally closed and F is locally bounded. Let $\varepsilon \in (0, 1]$. We start by showing how to define the functions f, g, σ, and u on a sufficiently small interval $[0, \delta]$, $\delta \leqslant T$, and then we will show how to extend them to the whole interval $[0, T]$. We recall that, in view of Proposition 2.3.2, there exist $\delta \in (0, \varepsilon]$, $\eta \in F(\xi)$ and $p \in \mathbb{R}^n$, all depending on ε, with $\|p\| \leqslant \varepsilon$, and satisfying $\xi + \delta(\eta + p) \in \mathbb{K}$. Diminishing δ if necessary, and using (6.3.2), we may also assume that $\|\delta(\eta + p)\| \leqslant \rho$. Let us define $u : [0, \delta] \to \mathbb{R}^n$ by

$$u(t) = \xi + t\eta + tp \tag{6.3.4}$$

for each $t \in [0, \delta]$. By virtue of (6.3.2), u, η and p satisfy

(j) $\eta \in F(u(0))$,
(jj) $\|\eta\| \leqslant M$,
(jjj) $u(\delta) \in \mathbb{K} \cap B(\xi, \rho)$,
(jv) $\|p\| \leqslant \varepsilon$.

Setting $f(s) = \eta$, $g(s) = p$, $\sigma(s) = 0$ for $s \in [0, \delta]$ and u defined by (6.3.4), from (j)–(jv), the fact that $\delta \in (0, \varepsilon]$ and (6.3.4), we can easily see that (σ, f, g, u) satisfies (i)–(iv) and (6.3.1) with T substituted by δ.

Next, we are going to show that, for each $\varepsilon \in (0, 1]$, there exists at least one 4-tuple (σ, f, g, u), whose domain is $[0, T]$, satisfying (i)–(iv) and (6.3.1). To this aim we shall use Brezis–Browder Theorem 2.1.1 as follows. Let $\mathcal{S}$ be the set of all 4-tuples (σ, f, g, u) defined on $[0, a]$ with $a \in (0, T]$ and satisfying (i)–(iv) and (6.3.1) on $[0, a]$. This set is clearly nonempty because (σ, f, g, u), defined as above, belongs to $\mathcal{S}$. On $\mathcal{S}$ we define a binary relation "$\preceq$" as follows. We say that (σ, f, g, u) defined on $[0, a]$ and $(\tilde{\sigma}, \tilde{f}, \tilde{g}, \tilde{u})$ defined on $[0, b]$ satisfy

$$(\sigma, f, g, u) \preceq \left(\tilde{\sigma}, \tilde{f}, \tilde{g}, \tilde{u}\right)$$

if $a \leqslant b$, $\sigma(s) = \tilde{\sigma}(s)$, $f(s) = \tilde{f}(s)$ and $g(s) = \tilde{g}(s)$ a.e. for $s \in [0, a]$. Endowed with "$\preceq$", $\mathcal{S}$ is obviously a preordered set. Let $((\sigma_m, f_m, g_m, u_m))_m$ be a monotone sequence, and let us denote by $[0, a_m]$ the domain of $(\sigma_m, f_m, g_m, u_m)$. Let $a^* = \lim_m a_m$, where $[0, a_m]$ denotes the domain of definition of $(\sigma_m, f_m, g_m, u_m)$. Clearly $a^* \in (0, T]$. We will show that there exists $(\sigma^*, f^*, g^*, u^*) \in \mathcal{S}$, defined on $[0, a^*]$ and satisfying

$$(\sigma_m, f_m, g_m, u_m) \preceq (\sigma^*, f^*, g^*, u^*)$$

for each $m \in \mathbb{N}$. We will prove first that there exists $\lim_m u_m(a_m)$. For each $m, k \in \mathbb{N}$, $m \leqslant k$, we have $u_m(s) = u_k(s)$ for all $s \in [0, a_m]$, and therefore, taking into account (iii), (iv) and (6.3.1), we deduce

$$\begin{aligned}\left\|u_m(a_m) - u_k(a_k)\right\| &\leqslant \int_{a_m}^{a_k} \left\|f_k(s)\right\| \mathrm{d}s + \int_{a_m}^{a_k} \left\|g_k(s)\right\| \mathrm{d}s \\ &\leqslant (M + \varepsilon)|a_k - a_m|\end{aligned}$$

for all $m, k \in \mathbb{N}$, $m \leqslant k$, which proves that there exists $\lim_{m\to\infty} u_m(a_m)$. Since for every $m \in \mathbb{N}$, $u_m(a_m) \in B(\xi, \rho) \cap \mathbb{K}$, and the latter is closed, it readily follows that $\lim_{m\to\infty} u_m(a_m) \in B(\xi, \rho) \cap \mathbb{K}$. In addition, because all the functions in the set $\{\sigma_m;\ m \in \mathbb{N}\}$ are nondecreasing, with values in $[0, a^*]$ and satisfy $\sigma_m(a_m) \leqslant \sigma_p(a_p)$ for every $m, p \in \mathbb{N}$ with $m \leqslant p$, there exists $\lim_{m\to\infty} \sigma_m(a_m)$ and this limit belongs to $[0, a^*]$. This allows us to define the 4-tuple $(\sigma^*, f^*, g^*, u^*) : [0, a^*] \to [0, a^*] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ by

$$\sigma^*(t) = \begin{cases} \sigma_m(t) & \text{for } t \in [0, a_m],\ m \in \mathbb{N}, \\ \lim\limits_{m\to\infty} \sigma_m(a_m) & \text{for } t = a^*, \end{cases}$$

$$f^*(t) = f_m(t) \quad \text{a.e. for } t \in [0, a_m],\ m \in \mathbb{N},$$

$$g^*(t) = \begin{cases} g_m(t) & \text{for } t \in [0, a_m],\ m \in \mathbb{N}, \\ 0 & \text{for } t = a^*, \end{cases}$$

$$u^*(t) = \begin{cases} u_m(t) & \text{for } t \in [0, a_m],\ m \in \mathbb{N}, \\ \lim\limits_{m\to\infty} u_m(a_m) & \text{for } t = a^*. \end{cases}$$

One can easily see that $(\sigma^*, f^*, g^*, u^*)$ is an ε-approximate solution which is a majorant for $((\sigma_m, f_m, g_m, u_m))_m$. Let us define $\mathcal{M} : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$ by $\mathcal{M}((\sigma, f, g, u)) = c$, where $[0, c]$ is the domain of definition of (σ, f, g, u). Clearly $\mathcal{M}$ satisfies the hypotheses of Brezis–Browder Theorem 2.1.1. Then, there exists an $\mathcal{M}$-maximal element

$(\bar{\sigma}, \bar{f}, \bar{g}, \bar{u}) \in \mathcal{S}$, defined on $[0, \bar{c}]$. This means that if $(\tilde{\sigma}, \tilde{f}, \tilde{g}, \tilde{u}) \in \mathcal{S}$ is defined on $[0, \tilde{c}]$ and $(\bar{\sigma}, \bar{f}, \bar{g}, \bar{u}) \preceq (\tilde{\sigma}, \tilde{f}, \tilde{g}, \tilde{u})$, we have $\mathcal{M}((\bar{\sigma}, \bar{f}, \bar{g}, \bar{u})) = \mathcal{M}((\tilde{\sigma}, \tilde{f}, \tilde{g}, \tilde{u}))$. We will show next that $\bar{c} = T$. Indeed, let us assume by contradiction that $\bar{c} < T$. We have

$$\begin{aligned}\left\|\bar{u}(\bar{c}) - \xi\right\| &\leqslant \int_0^{\bar{c}} \left\|\bar{f}(s)\right\| \mathrm{d}s + \int_0^{\bar{c}} \left\|\bar{g}(s)\right\| \mathrm{d}s \leqslant \bar{c}(M + \varepsilon)\\ &\leqslant \bar{c}(M + 1) < \rho.\end{aligned}$$

Then, as $\bar{u}(\bar{c}) \in \mathbb{K}$, by virtue of (6.2.2), there exists $f_{\bar{c}} \in F(\bar{u}(\bar{c})) \cap \mathcal{T}_{\mathbb{K}}(\bar{u}(\bar{c}))$. This means that, there exist $\delta \in (0, T - \bar{c})$, $\delta \leqslant \varepsilon$ and $p \in \mathbb{R}^n$ such that $\|p\| \leqslant \varepsilon$ and $\bar{u}(\bar{c}) + \delta f_{\bar{c}} + \delta p \in \mathbb{K}$. From the inequality above, it follows that we can diminish δ if necessary, in order to have $\|\bar{u}(\bar{c}) + \delta(f_{\bar{c}} + p) - \xi\| \leqslant \rho$. Let us define the functions $\sigma : [0, \bar{c} + \delta] \to [0, \bar{c} + \delta]$, $f : [0, \bar{c} + \delta] \to \mathbb{R}^n$ and $g : [0, \bar{c} + \delta] \to \mathbb{R}^n$ by

$$\sigma(t) = \begin{cases} \bar{\sigma}(t) & \text{for } t \in [0, \bar{c}],\\ \bar{c} & \text{for } t \in (\bar{c}, \bar{c} + \delta],\end{cases}$$

$$f(t) = \begin{cases} \bar{f}(t) & \text{a.e. for } t \in [0, \bar{c}],\\ f_{\bar{c}} & \text{a.e. for } t \in (\bar{c}, \bar{c} + \delta],\end{cases}$$

$$g(t) = \begin{cases} \bar{g}(t) & \text{for } t \in [0, \bar{c}],\\ p & \text{for } t \in (\bar{c}, \bar{c} + \delta].\end{cases}$$

It is not difficult to see that σ is nondecreasing, f and g are measurable on $[0, \bar{c} + \delta]$ and $\|g(t)\| \leqslant \varepsilon$ for every $t \in [0, \bar{c} + \delta]$. We define $u : [0, \bar{c} + \delta] \to \mathbb{R}^n$ by

$$u(t) = \xi + \int_0^t f(s)\,\mathrm{d}s + \int_0^t g(s)\,\mathrm{d}s$$

for every $t \in [0, \bar{c} + \delta]$. Clearly u coincides with $\bar{u}$ on $[0, \bar{c}]$ and then it readily follows that σ, f, g and u satisfy all the conditions in (i) and (ii). In order to prove (iii) and (iv), let us observe that

$$u(t) = \begin{cases} \bar{u}(t) & \text{for } t \in [0, \bar{c}],\\ \bar{u}(\bar{c}) + (t - \bar{c})f_{\bar{c}} + (t - \bar{c})p & \text{for } t \in [\bar{c}, \bar{c} + \delta].\end{cases}$$

As $u(\bar{c} + \delta) = \bar{u}(\bar{c}) + \delta f_{\bar{c}} + \delta p \in \mathbb{K}$, from the choice of δ, we also have $u(\bar{c} + \delta) \in \mathbb{K} \cap B(\xi, \rho)$. So, although $(\bar{\sigma}, \bar{f}, \bar{g}, \bar{u})$ is $\mathcal{M}$-maximal, we have both $(\bar{\sigma}, \bar{f}, \bar{g}, \bar{u}) \preceq (\sigma, f, g, u)$, and $\mathcal{M}((\bar{\sigma}, \bar{f}, \bar{g}, \bar{u})) < \mathcal{M}((\sigma, f, g, u))$, which is absurd. This contradiction can be eliminated only if each maximal element in the set $\mathcal{S}$ is defined on $[0, T]$, and this completes the proof of Lemma 6.3.1. □

6.4. *Convergence of the ε-approximate solutions*

DEFINITION 6.4.1. Let $\xi \in \mathbb{K}$ and $\varepsilon \in (0, 1]$. A 4-tuple (σ, f, g, u) satisfying (i)–(iv) and (6.3.1) is called an *ε-approximate solution* of (6.2.1).

We are now prepared to prove the sufficiency of Theorem 6.2.1.

PROOF. Let $\varepsilon_m \downarrow 0$ and, for each $m \in \mathbb{N}$, let us fix an ε_m-approximate solution $(\sigma_m, f_m, g_m, u_m)$ of (6.2.1) defined on $[0, T]$.

Since, by (iv), $\|g_m(t)\| \leqslant \varepsilon_m$ for each $m \in \mathbb{N}$ and $t \in [0, T]$, we have

$$\lim_{m\to\infty} \int_0^t g_m(s)\, \mathrm{d}s = 0$$

uniformly for $t \in [0, T]$. Moreover, as for each $m \in \mathbb{N}$, f_m satisfies (ii), we may assume with no loss of generality (by extracting a subsequence if necessary) that there exists $f \in L^1(0, T; \mathbb{R}^n)$ such that

$$\lim_{m\to\infty} f_m = f$$

weakly in $L^1(0, T; \mathbb{R}^n)$. As a consequence, from (6.3.1), we infer that there exists $u : [0, T] \to \mathbb{R}^n$ such that

$$\lim_{m\to\infty} u_m(t) = u(t)$$

uniformly for $t \in [0, T]$. Also from (6.3.1) and the last three relations, we easily conclude that

$$u(t) = \xi + \int_0^t f(s)\, \mathrm{d}s$$

for each $t \in [0, T]$. Recalling that, by (i), we have $s - \varepsilon_m \leqslant \sigma_m(s) \leqslant s$ for each $m \in \mathbb{N}$ and a.e. for $s \in [0, T]$, from the remarks above, we obtain

$$\lim_{m\to\infty} u_m\bigl(\sigma_m(s)\bigr) = u(s)$$

a.e. for $s \in [0, T]$. Furthermore, again by (i), $u_m(\sigma_m(s)) \in \mathbb{K} \cap B(\xi, \rho)$ a.e. for $s \in [0, T]$ and since $\mathbb{K} \cap B(\xi, \rho)$ is closed, we have $u(s) \in \mathbb{K}$ a.e. for $s \in [0, T]$. Since u is continuous, using once again the fact that $\mathbb{K} \cap B(\xi, \rho)$ is closed, we deduce that $u(s) \in \mathbb{K} \cap B(\xi, \rho)$ for each $s \in [0, T]$. Finally, since F is upper semicontinuous with nonempty, convex and compact values, by virtue of Lemma 6.1.2, we conclude that $f(s) \in F(u(s))$ a.e. for $s \in [0, T]$, and thus u is a solution to (6.2.1) and $u(0) = \xi$. The proof is complete. □

6.5. *Noncontinuable solutions*

In this section we will prove some results concerning the existence of noncontinuable, or even global solutions to (6.2.1). We recall that a solution $u : [0, T) \to \mathbb{K}$ of (6.2.1) is called *noncontinuable* if there is no other solution $v : [0, \widetilde{T}) \to \mathbb{K}$ of (6.2.1), with $T < \widetilde{T}$, satisfying $u(t) = v(t)$ for all $t \in [0, T)$. The next theorem is a consequence of Brezis–Browder Theorem 2.1.1.

THEOREM 6.5.1. *Let* $\mathbb{K}$ *be a nonempty and locally closed subset in* $\mathbb{R}^n$, *and let* $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ *be a u.s.c., nonempty, closed, convex and bounded valued mapping. Then, a necessary and sufficient condition in order that for each* $\xi \in \mathbb{K}$ *to exist at least one noncontinuable solution to* (6.2.1) *satisfying* $u(0) = \xi$ *is the tangency condition* (6.2.2).

PROOF. The necessity is an easy consequence of Theorem 6.2.1. As concerns the sufficiency, let $\xi \in \mathbb{K}$ and let $u : [0, T) \to \mathbb{K}$ a solution to (6.2.1) satisfying $u(0) = \xi$. The existence of this solution is guaranteed by Theorem 6.2.1. To complete the proof it suffices to show that u can be continued up to a noncontinuable one. To this aim, we will make use of Brezis–Browder Theorem 2.1.1. Let $\mathcal{S}$ be the set of all solutions to (6.2.1), defined at least and coinciding with u on $[0, T)$. On $\mathcal{S}$, we define the binary relation "$\preceq$" by $u \preceq v$ if the domain $[0, T_v)$ of v is larger that the domain $[0, T_u)$ of u, i.e. $T_u \leqslant T_v$, and $u(t) = v(t)$ for all $t \in [0, T_u)$. Clearly "$\preceq$" is a preorder on $\mathcal{S}$. Next, let $(u_m)_m$ be an increasing sequence in $\mathcal{S}$, and let us denote by $[0, T_m)$ the domain of definition of u_m. Let $T^* = \lim_{m\to\infty} T_m$, which can be finite, or not, and let us define $u^* : [0, T^*) \to \mathbb{K}$ by $u^*(t) = u_m(t)$ for each $t \in [0, T_m)$. Since $(T_m)_m$ is increasing and $u_m(t) = u_k(t)$ for each $m \leqslant k$ and each $t \in [0, T_m)$, u^* is well-defined and belongs to $\mathcal{S}$. Moreover, u^* is a majorant of u_m for each $m \in \mathbb{N}$. Thus each increasing sequence in $\mathcal{S}$ is bounded from above. Moreover, the function $\mathcal{M} : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$, defined by $\mathcal{M}(v) = T_v$, for each $v \in \mathcal{S}$, is monotone, and therefore we are in the hypotheses of Theorem 2.1.1. Accordingly, there exists at least one $\mathcal{M}$-maximal element $\bar{u} \in \mathcal{S}$. But this means that $\bar{u}$ is noncontinuable, and, of course, that it extends u. The proof is complete. □

We need next a multivalued counterpart of Definition 3.5.1.

DEFINITION 6.5.1. A mapping $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ is called *positively sublinear* if there exist a norm $\|\cdot\|$ on $\mathbb{R}^n$, $a > 0$, $b \in \mathbb{R}$, and $c > 0$ such that

$$\sup\bigl\{\|\eta\|;\ \eta \in F(\xi)\bigr\} \leqslant a\|\xi\| + b$$

for each $\xi \in \mathbb{K}_+^c(F)$, where

$$\mathbb{K}_+^c(F) = \Bigl\{\xi \in \mathbb{K};\ \|\xi\| > c \text{ and } \sup_{\eta\in F(\xi)} [\xi, \eta]_+ > 0\Bigr\}.$$

As concerns the existence of global solutions to (6.2.1), we have:

THEOREM 6.5.2. *Let* $\mathbb{K} \subseteq \mathbb{R}^n$ *be nonempty and closed, and let* $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ *be a u.s.c., nonempty, closed, convex and bounded valued mapping which is positively sublinear. Then, a necessary and sufficient condition in order that for each* $\xi \in \mathbb{K}$ *there exists at least one global solution to* (6.2.1) *satisfying* $u(0) = \xi$ *is the tangency condition* (6.2.2).

PROOF. Let $\xi \in \mathbb{K}$ and let $u : [0, T) \to \mathbb{K}$ be a noncontinuable solution to (6.2.1) satisfying $u(0) = \xi$. We will show that $T = +\infty$. To this aim, let us assume the contrary, i.e., that $T < +\infty$. Since $u'(s) \in F(u(t))$ a.e. for $s \in [0, T)$, we deduce

$$\big[u(s), u'(s)\big]_+ \leqslant \sup_{\eta \in F(u(s))} \big[u(s), \eta\big]_+.$$

Observing that $[u(s), u'(s)] = \frac{\mathrm{d}}{\mathrm{d}s}(\|u(s)\|)$ a.e. for $s \in [0, T)$, where $\|\cdot\|$ is the norm in Definition 6.5.1, integrating from 0 to t the last equality, we obtain

$$\big\|u(t)\big\| \leqslant \|\xi\| + \int_0^t \sup_{\eta \in F(u(s))} \big[u(s), \eta\big]_+ \,\mathrm{d}s.$$

Denoting by $M = \sup\{\|\eta\|;\ \eta \in F(v),\ \|v\| \leqslant c\}$ and by λ the Lebesgue measure on $\mathbb{R}$, by the positive sublinearity of F and Lemma 6.1.1 we get

$$\begin{aligned}
\big\|u(t)\big\| &\leqslant \|\xi\| + \int_{\{s \leqslant t;\, \|u(s)\| \leqslant c\}} \sup_{\eta \in F(u(s))} \big[u(s), \eta\big]_+ \,\mathrm{d}s \\
&\quad + \int_{\{s \leqslant t;\, \|u(s)\| > c\}} \sup_{\eta \in F(u(s))} \big[u(s), \eta\big]_+ \,\mathrm{d}s \\
&\leqslant \|\xi\| + M\lambda\big(\big\{s \leqslant t;\ \big\|u(s)\big\| \leqslant c\big\}\big) + \int_{\{s \leqslant t;\, \|u(s)\| > c\}} \big[a\big\|u(s)\big\| + b\big] \,\mathrm{d}s \\
&\leqslant \|\xi\| + (b + M)T + a\int_0^t \big\|u(s)\big\| \,\mathrm{d}s.
\end{aligned}$$

So, by Gronwall's Inequality, we deduce that u is bounded on $[0, T)$. Again from Lemma 6.1.1, it follows that the set $\{\eta;\ \eta \in F(u(t)),\ t \in [0, T)\}$ is bounded. Therefore u is globally Lipschitz on $[0, T)$ and accordingly there exists $\lim_{t \uparrow T} u(t) = u^*$. Since $\mathbb{K}$ is closed, it follows that $u^* \in \mathbb{K}$. Using this observation and recalling that (6.2.2) is satisfied, we conclude that u can be continued to the right of T, which is absurd as long as u is noncontinuable. This contradiction can be eliminated only if $T = +\infty$, and this completes the proof. □

Noticing that, in view of Lemma 6.1.1, whenever $\mathbb{K}$ is compact, each upper semicontinuous, compact valued mapping $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ is bounded, and thus positively sublinear, from Theorem 6.5.2, we deduce:

COROLLARY 6.5.1. *Let $\mathbb{K} \subseteq \mathbb{R}^n$ be nonempty and compact, and let $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ be a u.s.c., nonempty, closed, convex and bounded valued mapping. Then, a necessary and sufficient condition in order that for each $\xi \in \mathbb{K}$ to exist at least one global solution, $u : [0, +\infty) \to \mathbb{K}$ of* (6.2.1), *satisfying $u(0) = \xi$ is the tangency condition* (6.2.2).

6.6. *Viable preordered subsets*

As in Section 3.8, let us consider $\mathbb{K}$ a nonempty subset in $\mathbb{R}^n$, and let "$\preceq$" be a preorder on $\mathbb{K}$, identified with the multifunction $\mathcal{P} : \mathbb{K} \rightsquigarrow \mathbb{K}$, defined by

$$\mathcal{P}(\xi) = \{\eta \in \mathbb{K};\ \xi \preceq \eta\}$$

for each $\xi \in \mathbb{K}$, and called also a preorder. Let $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ be a multifunction. We say that "$\preceq$", or $\mathcal{P}$, is *viable with respect to* F if for each $\xi \in \mathbb{K}$, there exist $T > 0$ and a solution $u : [0, T] \to \mathbb{R}^n$ of the differential inclusion (6.2.1) satisfying $u(\tau) = \xi$, $u(t) \in \mathbb{K}$ for each $t \in [0, T]$ and u is "$\preceq$"-monotone on $[\tau, T]$, i.e., for each $\tau \leqslant s \leqslant t \leqslant T$, we have $u(s) \preceq u(t)$. An autonomous multivalued counterpart of Lemma 3.8.1 is proved below. See Cârjă and Ursescu [25].

LEMMA 6.6.1. *Let $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ and let $\mathcal{P}$ be a preorder on $\mathbb{K}$. If $\mathcal{P}$ is viable with respect to F then, for each $\xi \in \mathbb{K}$, $\mathcal{P}(\xi)$ is viable with respect to F. If $\mathcal{P}$ is closed in $\mathbb{R}^n \times \mathbb{R}^n$ and, for each $\xi \in \mathbb{K}$, $\mathcal{P}(\xi)$ is viable with respect to F, then $\mathcal{P}$ is viable with respect to F.*

The proof follows the very same lines as that one of Lemma 3.8.1, and therefore we do not enter into details.

THEOREM 6.6.1. *Let $\mathcal{P}$ be a closed preorder on $\mathbb{K}$ and let $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ be u.s.c. with nonempty, compact and convex values. Then a necessary and sufficient condition in order that $\mathcal{P}$ be viable with respect to F is the tangency condition below*:

$$F(\xi) \cap \mathcal{T}_{\mathcal{P}(\xi)}(\xi) \neq \emptyset$$

for each $\xi \in \mathbb{K}$.

PROOF. The proof follows immediately from Lemma 6.6.1. □

6.7. *Local invariance. Sufficient conditions*

Let $\mathbb{D}$ be a domain in $\mathbb{R}^n$, $\mathbb{K} \subseteq \mathbb{D}$ a locally closed subset, and let us consider the differential inclusion (6.2.1), where $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$ is a given multifunction.

DEFINITION 6.7.1. The subset $\mathbb{K}$ is *locally invariant* with respect to F if for each $\xi \in \mathbb{K}$ and each solution $u : [0, c] \to \mathbb{D}$, $c > 0$, of (6.2.1), satisfying the initial condition $u(0) = \xi$, there exists $T \in (0, c]$ such that we have $u(t) \in \mathbb{K}$ for each $t \in [0, T]$. It is *invariant* if it satisfies the local invariance condition above with $T = c$.

REMARK 6.7.1. If F is upper semicontinuous and nonempty, convex and compact valued on $\mathbb{D}$ and $\mathbb{K}$ is locally invariant with respect to F, then $\mathbb{K}$ is viable with respect to F. The converse of this assertion is no longer true, as we already have seen, even in the single-valued case. See Example 4.1.1.

Our first sufficient condition for local invariance is expressed in terms of the exterior tangency condition below, extending its single-valued counterpart (1.1.6), i.e., there exists an open neighborhood $\mathbb{V}$ of $\mathbb{K}$, with $\mathbb{V} \subseteq \mathbb{D}$, such that

$$\liminf_{h \downarrow 0} \frac{1}{h}\big[\operatorname{dist}(\xi + h\eta; \mathbb{K}) - \operatorname{dist}(\xi; \mathbb{K})\big] \leqslant \omega\big(\operatorname{dist}(\xi; \mathbb{K})\big) \tag{6.7.1}$$

for each $\xi \in \mathbb{V}$ and each $\eta \in F(\xi)$, where ω is an "autonomous" comparison function, i.e. a comparison function which does not depend on t. The main result in this section is:

THEOREM 6.7.1. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$. If* (6.7.1) *is satisfied, then $\mathbb{K}$ is locally invariant with respect to F.*

PROOF. Let $\mathbb{V} \subseteq \mathbb{D}$ be the open neighborhood of $\mathbb{K}$ whose existence is ensured by (6.7.1) and let $\omega : [0, a) \to \mathbb{R}$ the corresponding comparison function. Let $\xi \in \mathbb{K}$ and let $u : [0, c] \to \mathbb{V}$ be any local solution to (6.2.1) satisfying $u(0) = \xi$. Diminishing c if necessary, we may assume that there exists $\rho > 0$ such that $B(\xi, \rho) \cap \mathbb{K}$ is closed, $u(t) \in B(\xi, \rho/2)$ and, in addition, $\operatorname{dist}(u(t); \mathbb{K}) < a$ for each $t \in [0, c]$. Let $g : [0, c] \to \mathbb{R}_+$ be defined by $g(t) = \operatorname{dist}(u(t); \mathbb{K})$ for each $t \in [0, c]$. Let us observe that g is absolutely continuous on $[0, c]$. Let $t \in [0, c)$ be such that both $u'(t)$ and $g'(t)$ exist and $u'(t) \in F(u(t))$, and let $h > 0$ with $t + h \in [0, c]$. We have

$$\begin{aligned} g(t+h) &= \operatorname{dist}\big(u(t+h); \mathbb{K}\big) \\ &\leqslant h\left\|\frac{u(t+h) - u(t)}{h} - u'(t)\right\| + \operatorname{dist}\big(u(t) + hu'(t); \mathbb{K}\big). \end{aligned}$$

Therefore

$$\frac{g(t+h) - g(t)}{h} \leqslant \alpha(h) + \frac{\operatorname{dist}(u(t) + hu'(t); \mathbb{K}) - \operatorname{dist}(u(t); \mathbb{K})}{h},$$

where

$$\alpha(h) = \left\|\frac{u(t+h) - u(t)}{h} - u'(t)\right\|.$$

Since $\lim_{h \downarrow 0} \alpha(h) = 0$, passing to the inf-limit for $h \downarrow 0$ and taking into account that $\mathbb{V}$, $\mathbb{K}$ and F satisfy (6.7.1), we get

$$[D_+ g](t) \leqslant \omega\big(g(t)\big)$$

a.e. for $t \in [0, c)$. So, $g(t) \equiv 0$ which means that $u(t) \in \overline{\mathbb{K}} \cap B(\xi, \rho/2)$ for all $t \in [0, c)$. But $\overline{\mathbb{K}} \cap B(\xi, \rho/2) \subset \mathbb{K} \cap B(\xi, \rho)$ for each $t \in [0, c)$, and this completes the proof. □

REMARK 6.7.2. One may easily see that (6.7.1) is satisfied with $\omega = \omega_F$, where the function $\omega_F : [0,a) \to \overline{\mathbb{R}}$, $a = \sup_{\xi\in\mathbb{V}} \operatorname{dist}(\xi;\mathbb{K})$, is defined by

$$\omega_F(x) = \sup_{\substack{\xi\in\mathbb{V}\\ \operatorname{dist}(\xi;\mathbb{K})=x}} \ \sup_{\eta\in F(\xi)} \liminf_{h\downarrow 0} \frac{1}{h}\bigl[\operatorname{dist}(\xi+h\eta;\mathbb{K}) - \operatorname{dist}(\xi;\mathbb{K})\bigr] \tag{6.7.2}$$

for each $x \in [0,a)$.

So, Theorem 6.7.1 can be reformulated as:

THEOREM 6.7.2. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$. If there exists an open neighborhood V of $\mathbb{K}$ with $\mathbb{V} \subseteq \mathbb{D}$ such that ω_F defined by* (6.7.2) *is a comparison function, then $\mathbb{K}$ is locally invariant with respect to F.*

6.8. *Local invariance. Necessary conditions*

We begin with the following simple necessary condition for local invariance.

THEOREM 6.8.1. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ nonempty and $\mathbb{D}$ open, and let $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$ be l.s.c. with nonempty, convex and compact values. If $\mathbb{K}$ is locally invariant with respect to F, then, for each $\xi \in \mathbb{K}$, $F(\xi) \subseteq \mathcal{C}_{\mathbb{K}}(\xi)$.*

PROOF. Let $\xi \in \mathbb{K}$ and $\eta \in F(\xi)$. Since F is l.s.c. with nonempty, closed and convex values, by the Michael' selection theorem [67,68], there exists a continuous function $f : \mathbb{D} \to \mathbb{R}^n$ such that $f(\xi) = \eta$ and $f(x) \in F(x)$ for each $x \in \mathbb{D}$. As $\mathbb{D}$ is open, by Peano's local existence theorem 1.1.1, there exists at least one solution $u : [0,T] \to \mathbb{D}$ of the equation $u'(t) = f(u(t))$ satisfying $u(0) = \xi$. Clearly $u'(0) = f(u(0)) = \eta$ and, in addition, u is a solution to the differential inclusion (6.2.1). Since $\xi \in \mathbb{K}$ and the latter is locally invariant with respect to F, there exists $0 < a \leqslant T$ such that $u(t) \in \mathbb{K}$ for all $t \in [0,a]$. Now, repeating the same arguments as in the proof of Theorem 3.2.1—Section 3.2—we conclude that $\eta = u'(0) \in \mathcal{F}_{\mathbb{K}}(\xi)$. The conclusion follows from Proposition 6.2.1, and the proof is complete. □

Next, we rephrase some concepts we introduced in the single-valued case.

DEFINITION 6.8.1. Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$. A multifunction $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$ has *the comparison property with respect to* $(\mathbb{D},\mathbb{K})$ if there exist a proximal neighborhood $\mathbb{V} \subseteq \mathbb{D}$ of $\mathbb{K}$, one projection $\pi_{\mathbb{K}} : \mathbb{V} \to \mathbb{K}$ subordinated to $\mathbb{V}$, and one comparison function

$$\omega : [0,a) \to \mathbb{R}, \quad \text{with } a = \sup_{\xi\in\mathbb{V}} \operatorname{dist}(\xi;\mathbb{K}),$$

such that

$$\sup_{\eta\in F(\xi)} \ \inf_{\eta_\pi\in F(\pi_K(\xi))} \bigl[\xi - \pi_{\mathbb{K}}(\xi), \eta - \eta_\pi\bigr]_+ \leqslant \omega\bigl(\bigl\|\xi - \pi_{\mathbb{K}}(\xi)\bigr\|\bigr) \tag{6.8.1}$$

for each $\xi \in \mathbb{V}$.

A condition similar to that one in Definition 6.8.1, with ξ replaced by ξ_1 and $\pi_{\mathbb{K}}(\xi)$ replaced by ξ_2 with $\xi_1, \xi_2 \in \mathbb{V}$, has been used previously by Cârjă and Ursescu [25]. As in the single-valued case, here, (6.8.1) is also automatically satisfied for each $\xi \in \mathbb{K}$, and therefore, in Definition 6.8.1, we have only to assume that (6.8.1) holds for each $\xi \in \mathbb{V} \setminus \mathbb{K}$.

DEFINITION 6.8.2. The multifunction $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$ is called:

(i) $(\mathbb{D}, \mathbb{K})$-*Lipschitz* if there exist a proximal neighborhood $\mathbb{V} \subseteq \mathbb{D}$ of $\mathbb{K}$, a subordinated projection $\pi_{\mathbb{K}} : \mathbb{V} \to \mathbb{K}$, and $L > 0$, such that

$$\sup_{\eta \in F(\xi)} \inf_{\eta_\pi \in F(\pi_K(\xi))} \|\eta - \eta_\pi\| \leqslant L \big\| \xi - \pi_{\mathbb{K}}(\xi) \big\|$$

for each $\xi \in \mathbb{V} \setminus \mathbb{K}$;

(ii) $(\mathbb{D}, \mathbb{K})$-*dissipative* if there exist a proximal neighborhood $\mathbb{V} \subset \mathbb{D}$ of $\mathbb{K}$, and a projection, $\pi_{\mathbb{K}} : \mathbb{V} \to \mathbb{K}$, subordinated to $\mathbb{V}$, such that

$$\sup_{\eta \in F(\xi)} \inf_{\eta_\pi \in F(\pi_K(\xi))} \big[\xi - \pi_{\mathbb{K}}(\xi), \eta - \eta_\pi \big]_+ \leqslant 0$$

for each $\xi \in \mathbb{V} \setminus \mathbb{K}$.

A strictly more restrictive Lipschitz condition, with ξ replaced by ξ_1 and $\pi_K(\xi)$ by ξ_2, with ξ_1, ξ_2 belonging to $\mathbb{D}$, has been considered first by Filippov [43]. In a very same spirit as Filippov [43], Kobayashi [62] has used a dissipative type condition strictly restrictive than that one in Definition 6.8.2. It is easy to see that if F is either $(\mathbb{D}, \mathbb{K})$-Lipschitz, or $(\mathbb{D}, \mathbb{K})$-dissipative, then it has the comparison property with respect to $(\mathbb{D}, \mathbb{K})$. We notice that there are examples showing that there exist multifunctions F which, although neither $(\mathbb{D}, \mathbb{K})$-Lipschitz, nor $(\mathbb{D}, \mathbb{K})$-dissipative, have the comparison property. See, for instance, the “single-valued” Examples 4.3.1 and 4.3.2.

THEOREM 6.8.2. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$ be a multifunction. If F has the comparison property with respect to* $(\mathbb{D}, \mathbb{K})$, *and*

$$\liminf_{h \downarrow 0} \frac{1}{h} \operatorname{dist}(\xi + h\eta; \mathbb{K}) = 0 \tag{6.8.2}$$

for each $\xi \in \mathbb{K}$ and each $\eta \in F(\xi)$, then (6.7.1) *holds true.*

PROOF. Let $\mathbb{V} \subseteq \mathbb{D}$ be the open neighborhood of $\mathbb{K}$ as in Definition 6.8.1, let $\xi \in \mathbb{V}$ and $\eta \in F(\xi)$. Let $\rho > 0$ and let $\pi_{\mathbb{K}}$ be the selection of $\Pi_{\mathbb{K}}$ as in Definition 6.8.1. Let $h \in (0, T]$. Since $\|\xi - \pi_{\mathbb{K}}(\xi)\| = \operatorname{dist}(\xi; \mathbb{K})$, we have

$$\begin{aligned} &\operatorname{dist}(\xi + h\eta; \mathbb{K}) - \operatorname{dist}(\xi; \mathbb{K}) \\ &\quad \leqslant \big\| \xi - \pi_{\mathbb{K}}(\xi) + h[\eta - \zeta] \big\| - \big\| \xi - \pi_{\mathbb{K}}(\xi) \big\| + d\big(\pi_{\mathbb{K}}(\xi) + h\zeta; \mathbb{K}\big), \end{aligned}$$

for each $\zeta \in F(\pi_{\mathbb{K}}(\xi))$.

Dividing by h, passing to the lim inf for $h \downarrow 0$, and using (6.8.2), we get

$$\liminf_{h\downarrow 0} \frac{1}{h}\big[\operatorname{dist}(\xi + h\eta; \mathbb{K}) - \operatorname{dist}(\xi; \mathbb{K})\big] \leqslant \big[\xi - \pi_{\mathbb{K}}(\xi), \eta - \zeta\big]_+.$$

Since $\zeta \in F(\pi_{\mathbb{K}}(\xi))$ is arbitrary, we have

$$\liminf_{h\downarrow 0} \frac{1}{h}\big[\operatorname{dist}(\xi + h\eta; \mathbb{K}) - \operatorname{dist}(\xi; \mathbb{K})\big] \leqslant \inf_{\zeta\in F(\pi_{\mathbb{K}}(\xi))}\big[\xi - \pi_{\mathbb{K}}(\xi), \eta - \zeta\big]_+.$$

Therefore

$$\liminf_{h\downarrow 0} \frac{1}{h}\big[\operatorname{dist}(\xi + h\eta; \mathbb{K}) - \operatorname{dist}(\xi; \mathbb{K})\big] \leqslant \omega\big(\big\|\xi - \pi_{\mathbb{K}}(\xi)\big\|\big).$$

But this inequality shows that (6.7.1) holds, and this completes the proof. □

COROLLARY 6.8.1. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$ be a multifunction. If F has the comparison property with respect to $(\mathbb{D}, \mathbb{K})$ and, for each $\xi \in \mathbb{K}$, $F(\xi) \subseteq \mathcal{T}_{\mathbb{K}}(\xi)$, then $\mathbb{K}$ is local invariant with respect to F.*

Corollary 6.8.1 and Proposition 6.2.1 yield:

COROLLARY 6.8.2. *Let $\mathbb{K} \subseteq \mathbb{D} \subseteq \mathbb{R}^n$, with $\mathbb{K}$ locally closed and $\mathbb{D}$ open, and let $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$ be l.s.c. with nonempty values. If, F has the comparison property with respect to $(\mathbb{D}, \mathbb{K})$ and, for each $\xi \in \mathbb{K}$, $F(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi)$, then $\mathbb{K}$ is local invariant with respect to F.*

7. Applications

7.1. *Invariance with respect to parametrized multifunctions*

Let us consider the parametrized multifunction $F : \mathbb{D} \rightsquigarrow \mathbb{R}^n$,

$$F(\xi) = \big\{f(\xi, v);\ v \in \mathbb{V}\big\},$$

where $\mathbb{D}$ is a nonempty and open subset in $\mathbb{R}^n$, $\mathbb{V}$ is a nonempty and compact subset in $\mathbb{R}^n$ and $f : \mathbb{D} \times \mathbb{R}^n \to \mathbb{R}^n$ is a continuous function. Let $\mathbb{K}$ be a locally closed subset of $\mathbb{D}$ and let us assume that F has convex values. We assume further that, for each measurable function $v : [0, a] \to \mathbb{V}$ and for each $\xi \in \mathbb{D}$, there exists $T \in (0, a]$ such that the equation $u'(t) = f(u(t), v(t))$ has a unique Carathéodory solution $u : [0, T] \to \mathbb{R}^n$ satisfying $u(0) = \xi$.

THEOREM 7.1.1. *Under the above circumstances, the following conditions are equivalent*:

(i) *for each $\xi \in \mathbb{K}$ and $v \in \mathbb{V}$, $f(\xi, v) \in \mathcal{C}_{\mathbb{K}}(\xi)$,*
(ii) *for each $\xi \in \mathbb{K}$ and $v \in \mathbb{V}$, $f(\xi, v) \in \mathcal{T}_{\mathbb{K}}(\xi)$,*

(iii) *for each* $\xi \in \mathbb{K}$ *and* $v \in \mathbb{V}$, $f(\xi, v) \in \mathcal{B}_{\mathbb{K}}(\xi)$,

(iv) $\mathbb{K}$ *is locally invariant with respect to* F.

In general, if $\mathcal{G} : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ *is such that* $\mathcal{C}_{\mathbb{K}}(\xi) \subseteq \mathcal{G}(\xi) \subseteq \mathcal{B}_{\mathbb{K}}(\xi)$ *for each* $\xi \in \mathbb{K}$, *then each one of the conditions above is equivalent to*

(v) *for each* $\xi \in \mathbb{K}$, $f(\xi, v) \in \mathcal{G}(\xi)$.

PROOF. In view of Proposition 2.5.1, it suffices to show that (ii) is equivalent to (iv). Let us assume that (ii) holds true and let $u : [0, T] \to \mathbb{D}$ be a solution to $u'(t) \in F(u(t))$ with $u(0) = \xi \in \mathbb{K}$. We have to show that there exists $\tau \in (0, T]$ such that $u(t) \in \mathbb{K}$ for all $t \in [0, \tau]$. By a classical result of Filippov, there exists a measurable function $v : [0, T] \to \mathbb{V}$ such that

$$u'(t) = f\big(u(t), v(t)\big) \tag{7.1.1}$$

a.e. for $t \in [0, T]$. By the uniqueness property, taking into account (ii) and using Theorem 5.2.1, we deduce that there exists $\tau \in (0, T]$ such that the unique solution to (7.1.1) satisfies $u(t) \in \mathbb{K}$ for all $t \in [0, \tau]$. Thus (ii) implies (iv).

Let us suppose now that (iv) holds true. Fix $\xi \in \mathbb{K}$ and $v \in \mathbb{V}$. The equation $u'(t) = f(u(t), v)$ with $u(0) = \xi$ has a unique solution $u : [0, T] \to \mathbb{D}$. Since $\mathbb{K}$ is locally invariant with respect to F, there exists $\tau \in (0, T]$ such that we have $u(t) \in \mathbb{K}$ for all $t \in [0, \tau]$. As $u'(0) = f(\xi, v)$, it readily follows that $f(\xi, v) \in \mathcal{F}_{\mathbb{K}}(\xi) \subseteq \mathcal{T}_{\mathbb{K}}(\xi)$, and thus (iv) implies (ii). This completes the proof. □

7.2. *Differentiability along trajectories*

In this section we will show how viability can be used in order to obtain sufficient conditions for asymptotic stability via Liapunov functions. Let us consider the autonomous differential equation

$$u'(t) = g\big(u(t)\big), \tag{7.2.1}$$

where $g : \mathbb{R}^n \to \mathbb{R}^n$ is continuous. Let $V : \mathbb{R}^n \to \mathbb{R}$, and let us define

$$V^*(\xi) = \limsup_{h \downarrow 0} \frac{1}{h}\big[V\big(\xi + hg(\xi)\big) - V(\xi)\big]$$

and

$$V_*(\xi) = \liminf_{h \downarrow 0} \frac{1}{h}\big[V\big(\xi + hg(\xi)\big) - V(\xi)\big].$$

If $V^*(\xi) = V_*(\xi)$, we denote this common value by $\dot{V}(\xi)$ and we note that, if V is differentiable, we have $\dot{V}(\xi) = \langle \operatorname{grad} V(\xi), g(\xi)\rangle$. The next result will prove useful in the sequel.

THEOREM 7.2.1 (Yoshizawa [110]). *If* $V:\mathbb{R}^n \to \mathbb{R}$ *is locally Lipschitz,* $u:[0,T)\to\mathbb{R}^n$ *is any solution to* (7.2.1) *and* $t\in[0,T)$*, then*

$$V^*\big(u(t)\big) = \limsup_{h\downarrow 0}\frac{1}{h}\big[V\big(u(t+h)\big) - V\big(u(t)\big)\big].$$

PROOF. We have

$$\begin{aligned} V^*\big(u(t)\big) &= \limsup_{h\downarrow 0}\frac{1}{h}\big[V\big(u(t)\big) + hg\big(u(t)\big) - V\big(u(t)\big)\big] \\ &\leqslant \limsup_{h\downarrow 0}\frac{1}{h}\left[V\left(u(t) + \int_t^{t+h} g\big(u(s)\big)\,\mathrm{d}s\right) - V\big(u(t)\big)\right] \\ &\quad + \limsup_{h\downarrow 0}\frac{1}{h}\left[-V\left(u(t) + \int_t^{t+h} g\big(u(s)\big)\,\mathrm{d}s\right)\right. \\ &\quad \left. + V\big(u(t)\big) + hg\big(u(t)\big)\right]. \end{aligned}$$

Since

$$\begin{aligned} &\left|\limsup_{h\downarrow 0}\frac{1}{h}\left[-V\left(u(t) + \int_t^{t+h} g\big(u(s)\big)\,\mathrm{d}s\right) + V\big(u(t)\big) + hg\big(u(t)\big)\right]\right| \\ &\quad \leqslant L\limsup_{h\downarrow 0}\frac{1}{h}\left\|\int_t^{t+h} g\big(u(s)\big)\,\mathrm{d}s - hg\big(u(t)\big)\right\| = 0, \end{aligned}$$

where $L>0$ is the Lipschitz constant of V on a suitably chosen neighborhood of $u(t)$, we deduce

$$V^*\big(u(t)\big) \leqslant \limsup_{h\downarrow 0}\frac{1}{h}\big[V\big(u(t+h)\big) - V\big(u(t)\big)\big].$$

Similarly, we get

$$V^*\big(u(t)\big) \geqslant \limsup_{h\downarrow 0}\frac{1}{h}\big[V\big(u(t+h)\big) - V\big(u(t)\big)\big]$$

and this completes the proof. □

7.3. *Liapunov functions*

DEFINITION 7.3.1. We say that 0 is *stable* for (7.2.1) if for each $\varepsilon>0$ there exists $\delta(\varepsilon)\in(0,\varepsilon)$ such that for each $\xi\in\mathbb{R}^n$, $\|\xi\|\leqslant\delta(\varepsilon)$, each solution u of (7.2.1), satisfying $u(0)=\xi$, is defined on $[0,+\infty)$, and $\|u(t)\|\leqslant\varepsilon$ for all $t\geqslant 0$.

Clearly, if 0 is stable for (7.2.1), $\{0\}$ is both viable and locally invariant with respect to g, and thus $g(0) = 0$. In other words, if 0 is stable for (7.2.1), then $u \equiv 0$ is necessarily a solution to (7.2.1), and this is the only one issuing from 0.

DEFINITION 7.3.2. We say that V has positive gradient at $v \in \mathbb{R}$ if

$$\liminf_{\substack{\operatorname{dist}(\xi;\mathbb{K})\downarrow 0\\ \xi\notin\mathbb{K}}} \frac{V(\xi)-v}{\operatorname{dist}(\xi;\mathbb{K})} > 0,$$

where $\mathbb{K} = V^{-1}((-\infty, v])$.

If V is of class C^1 and $\|\operatorname{grad} V\|$ is bounded from below on $V^{-1}(\{v\})$ by a constant $c > 0$, then V has positive gradient at v.

PROPOSITION 7.3.1. *Let* $g : \mathbb{R}^n \to \mathbb{R}^n$ *and* $V : \mathbb{R}^n \to \mathbb{R}$ *be continuous, and let* $\mathbb{K} = V^{-1}((-\infty, v])$. *If* V *has positive gradient at* v *and* $V^*(\xi) \leqslant 0$ *whenever* $V(\xi) = v$, *then* $\mathbb{K}$ *is viable with respect to* g.

PROOF. Let us assume by contradiction that $\mathbb{K}$ is not viable with respect to g. In view of Theorem 3.1.1, this means that there exists $\xi \in \partial\mathbb{K}$ such that $g(\xi)$ is not tangent in the sense of Federer to $\mathbb{K}$ at ξ. So, there exist $\gamma > 0$ and a sequence $h_m \downarrow 0$ such that

$$\gamma < \frac{1}{h_m}\operatorname{dist}\bigl(\xi + h_m g(\xi); \mathbb{K}\bigr)$$

for all $m \in \mathbb{N}$. As $\xi \in \partial\mathbb{K}$, it follows that $V(\xi) = v$. Furthermore, since V has positive gradient at v, there exists $\nu > 0$ such that

$$\nu < \frac{V(\xi + h_m g(\xi)) - v}{\operatorname{dist}(\xi + h_m g(\xi); \mathbb{K})} < \frac{1}{h_m\gamma}\bigl[V\bigl(\xi + h_m g(\xi)\bigr) - v\bigr]$$

for each $m \in \mathbb{N}$. Hence

$$0 < \nu \leqslant \frac{1}{\gamma}\limsup_{h\downarrow 0}\frac{V(\xi + hg(\xi)) - V(\xi)}{h} = \frac{1}{\gamma}V^*(\xi),$$

thereby contradicting the hypothesis that $V^*(\xi) \leqslant 0$. This contradiction can be eliminated only if $\mathbb{K}$ is viable with respect to g, and this completes the proof. □

THEOREM 7.3.1 (Yorke [108]). *Assume that* (7.2.1) *has the uniqueness property. Then* 0 *is stable for* (7.2.1) *if and only if there exist a continuous function* $V : \mathbb{R}^n \to [0, +\infty)$ *and a sequence* $v_m \downarrow 0$ *such that* $V(x) = 0$ *if and only if* $x = 0$, *and, for every* $m \in \mathbb{N}$ *for which* $V(x) = v_m$, *we have* $V^*(x) \leqslant 0$, *and, for each* $m \in \mathbb{N}$, V *has positive gradient at* v_m.

PROOF. We denote by $u(\cdot,\xi):[0,+\infty)\to\mathbb{R}^n$ the unique solution of (7.2.1) satisfying $u(0,\xi)=\xi$.

Sufficiency. Let us assume that such a sequence $(v_m)_m$ and function V exist. Let $\mathbb{K}_m = V^{-1}([0,v_m])$. Since (7.2.1) has the uniqueness property, by Proposition 7.3.1, it follows that $\mathbb{K}_m$ is both viable and locally invariant with respect to g. Let $\varepsilon>0$ and let us define both $\mu_*(\varepsilon)=\inf_{\|\xi\|=\varepsilon} V(\xi)$ and $\mu^*(\varepsilon)=\sup_{\|\xi\|=\varepsilon} V(\xi)$. Then, $\mu_*(\varepsilon)>0$. For any $\varepsilon>0$, choose $m=m(\varepsilon)$ such that $v_m<\mu_*(\varepsilon)$, and choose $\delta(\varepsilon)\in(0,\varepsilon)$ such that $\mu^*(\delta)<v_m$ for each $\delta\in(0,\delta(\varepsilon))$. Let $\xi\in\mathbb{R}^n$ with $\|\xi\|\leqslant\delta(\varepsilon)$. Clearly, $\xi\in\mathbb{K}_m$. If there exists $t>0$ such that $\|u(t,\xi)\|>\varepsilon$, then there exists $t_0\in(0,t)$ such that $\|u(t_0,\xi)\|=\varepsilon$. On the other hand, by the choice of m, we have $u(t_0,\xi)\notin\mathbb{K}_m$, thereby contradicting the local invariance of $\mathbb{K}_m$ with respect to g. This contradiction can be eliminated only if $\|u(t,\xi)\|<\varepsilon$ for all $t\geqslant 0$, and this completes the proof of the sufficiency.

Necessity. Suppose that 0 is stable for (7.2.1). Let $\delta(\cdot)$ the function in Definition 7.3.1, choose $\varepsilon_1>0$, and let us define inductively $\varepsilon_{m+1}=\delta(\varepsilon_m)/2$. Clearly, $\varepsilon_m\downarrow 0$. Let us denote by

$$\mathbb{K}_m^0=\bigl\{u(t,\xi);\ \|\xi\|\leqslant\delta(\varepsilon_m),\ \text{and } t\geqslant 0\bigr\}.$$

Since $\mathbb{K}_m^0$ contains only points reached at positive time by solutions starting in $B(0,\delta(\varepsilon_m))$, it follows that $\mathbb{K}_m^0$ is viable with respect to g. By Proposition 3.6.1, we deduce that $\mathbb{K}_m=\overline{\mathbb{K}_m^0}$ is viable with respect to g. Since $u'(t)=g(u(t))$ has the uniqueness property, it follows that $\mathbb{K}_m$ is in fact invariant with respect to g. For each $m=1,2,\ldots,$ let

$$v_m=\sum_{i=m+1}^{\infty}\varepsilon_i.$$

Clearly, $v_m=\frac{1}{2}\sum_{i=m+1}^{\infty}\delta(\varepsilon_{i-1})\leqslant\varepsilon_m$ for $m=1,2,\ldots$. Let $\mathbb{K}_0=\mathbb{R}^n$ and $v_0=+\infty$. We have $\cdots\subseteq\mathbb{K}_2\subseteq\mathbb{K}_1\subseteq\mathbb{K}_0$ and therefore, we may define $N(\xi)=\max\{m;\ \xi\in\mathbb{K}_m\}$. Notice that we want to have $V(\xi)=v_m$ for all $\xi\in\partial\mathbb{K}_m$. Define $V(0)=0$ and

$$V(\xi)=\min\bigl\{v_{N(\xi)},v_{N(\xi)+1}+\operatorname{dist}(\xi;\mathbb{K}_{N(\xi)+1})\bigr\}$$

for $\xi\neq 0$. We begin by proving that V is continuous. First, let us observe that, for each $m\in\mathbb{N}$, the function N is constant on $\mathbb{K}_m\setminus\mathbb{K}_{m+1}$. Accordingly the restriction of V to $\mathbb{K}_m\setminus\mathbb{K}_{m+1}$ is continuous, and so the function V itself is continuous at each interior point of $\mathbb{K}_m\setminus\mathbb{K}_{m+1}$. Next, we show that V is continuous on $\partial\mathbb{K}_m$ for each $m\in\mathbb{N}$. Let ξ be arbitrary in $\partial\mathbb{K}_m$. We have both $\|\xi\|\in[\delta(\varepsilon_m),\varepsilon_m)$ and $\sup_{\eta\in\mathbb{K}_{m+1}}\|\eta\|\leqslant\varepsilon_{m+1}$. Therefore, $N(\xi)=m$ and $\operatorname{dist}(\xi;\mathbb{K}_{m+1})>\delta(\varepsilon_m)-\varepsilon_{m+1}=\delta(\varepsilon_m)/2$. Hence we have $v_{N(\xi)+1}+\operatorname{dist}(\xi;\mathbb{K}_{N(\xi)+1})>v_{N(\xi)}$ and $V(\xi)=v_m$. Choose $\xi_i\notin\mathbb{K}_m$ for $i=1,2,\ldots,$ with $\xi_i\to\xi$. Then, $\lim_{i\to\infty}\operatorname{dist}(\xi_i;\mathbb{K}_m)=0$, and

$$\lim_{i\to\infty}V(\xi_i)=v_m=V(\xi),$$

which shows that V is continuous at ξ. It remains to prove that V is continuous at 0. Take $\xi_i \to 0$. Then $N(\xi_i) \to \infty$ and $V(\xi_i) \leqslant v_{N(\xi_i)}$. But $v_m \to 0$ as $m \to \infty$, so that

$$\lim_{\xi \to 0} V(\xi) = 0 = V(0).$$

Thus V is continuous on $\mathbb{R}^n$.

Finally, we show that $V(\xi) = v_m$ implies $V^*(\xi) \leqslant 0$. Indeed, if for some $m \in \mathbb{N}$ we have $V(\xi) = v_m$, then $\xi \in \mathbb{K}_m$, $\xi + hg(\xi) \in \mathbb{K}_{m-1}$ for $h > 0$ sufficiently small, and

$$V\big(\xi + hg(\xi)\big) \leqslant v_m + \operatorname{dist}\big(\xi + hg(\xi); \mathbb{K}_m\big).$$

Since $\mathbb{K}_m$ is viable with respect to g, by virtue of Theorem 3.2.1, we have $g(\xi) \in \mathcal{T}_{\mathbb{K}_m}(\xi)$, and therefore

$$V^*(\xi) \leqslant \limsup_{h \downarrow 0} \frac{1}{h} \operatorname{dist}\big(\xi + hg(\xi); \mathbb{K}_m\big) = 0.$$

If $0 < \operatorname{dist}(\xi; \mathbb{K}_m) \leqslant v_m - v_{m-1} = \varepsilon_m$, then $V(\xi) = v_m + \operatorname{dist}(\xi; \mathbb{K}_m)$, so that V has positive gradient at each v_m, thereby completing the proof. □

7.4. *Hukuhara's theorem*

In this section, by using viability and invariance techniques, we will prove two celebrated results concerning the funnel of solutions, results due to Hukuhara [47] and Kneser [61]. Let $f : \mathbb{I} \times \mathbb{D} \to \mathbb{R}^n$ be a continuous function, and let us consider the non-autonomous differential equation (1.1.1). Let $(\tau, \xi) \in \mathbb{I} \times \mathbb{D}$, and let us denote by $\mathcal{S}(\tau, \xi)$ the set of all noncontinuable solutions u of (1.1.1) satisfying $u(\tau) = \xi$.

DEFINITION 7.4.1. The *right solution funnel* through $(\tau, \xi) \in \mathbb{I} \times \mathbb{D}$, $F_{\tau,\xi}$, is defined by

$$F_{\tau,\xi} = \big\{\big(s, u(s)\big);\ s \geqslant \tau,\ u \in \mathcal{S}(\tau, \xi)\big\}.$$

If $t \geqslant \tau$, we define the *t-cross section of* $F_{\tau,\xi}$ by

$$F_{\tau,\xi}(t) = \big\{u(t);\ u \in \mathcal{S}(\tau, \xi)\big\}.$$

The next compactness result will prove useful in that follows.

PROPOSITION 7.4.1. *Let $(\tau, \xi) \in \mathbb{I} \times \mathbb{D}$ and let $t > \tau$ be such that, for each $u \in \mathcal{S}(\tau, \xi)$, $u(t)$ is defined. Then*

$$F_{\tau,\xi}\big([\tau, t]\big) = \big\{\big(s, u(s)\big);\ s \in [\tau, t],\ u \in \mathcal{S}(\tau, \xi)\big\}$$

is compact.

PROOF. Let $(t_m, u_m(t_m))_m$ be an arbitrary sequence in $F_{\tau,\xi}([\tau,t])$, with $u_m : (a_m, b_m) \to \mathbb{D}$ for each $m \in \mathbb{N}$. We may assume with no loss of generality that $\lim_m t_m = s$. Obviously, $s \in [\tau, t] \subseteq (a_m, b_m)$ for every $m \in \mathbb{N}$, and therefore, by Hartman [54, Theorem 3.2, p. 26], there exists at least one subsequence of $(u_m)_m$, denoted for simplicity again by $(u_m)_m$, and $u \in \mathcal{S}(\tau,\xi)$, with $\lim_m u_m = u$ uniformly on $[\tau, t]$. But this shows that $\lim_m (t_m, u_m(t_m)) = (s, u(s))$. Since $(s, u(s)) \in F_{\tau,\xi}([\tau, t])$, the proof is complete. □

From Proposition 7.4.1, we deduce:

COROLLARY 7.4.1. *The set $F_{\tau,\xi}$ is locally closed.*

THEOREM 7.4.1 (Hukuhara [47]). *Let $(\tau,\xi) \in \mathbb{I} \times \mathbb{D}$ and let $t > \tau$ be such that, for each $u \in \mathcal{S}(\tau,\xi)$, $u(t)$ is defined. Then, for each $\eta \in \partial F_{\tau,\xi}(t)$ there exists a solution v such that $v(s) \in \partial F_{\tau,\xi}(s)$ for all $s \in [\tau, t]$.*

In order to prove Theorem 7.4.1, we need:

THEOREM 7.4.2 (Yorke [108]). *If $\mathbb{D}$ is a domain, $\mathbb{K}_1, \mathbb{K}_2 \subseteq \mathbb{D}$ are locally closed and viable with respect to f, and if $\mathbb{K}_1 \cup \mathbb{K}_2 = \mathbb{D}$, then $\mathbb{K}_1 \cap \mathbb{K}_2$ is viable with respect to f.*

PROOF. The conclusion is a consequence of Theorem 2.3.1 and Nagumo's viability theorem 3.1.1. □

REMARK 7.4.1. A result similar to Theorem 7.4.2 holds true trivially in the case of local invariance. More precisely, if $\mathbb{D} \subseteq \mathbb{R}^n$ is open, $\mathbb{K}_1, \mathbb{K}_2 \subseteq \mathbb{D}$ are locally closed and locally invariant with respect to f, then $\mathbb{K}_1 \cap \mathbb{K}_2$ is locally invariant with respect to f.

Let $\tau \in \mathbb{I}$ be fixed, let us denote by $\mathcal{D} = \{s \in \mathbb{I};\ s > \tau\} \times \mathbb{D}$, and let us define $\mathcal{F} : \mathcal{D} \to \mathbb{R} \times \mathbb{R}^n$, by $\mathcal{F}(t,\xi) = (1, f(t,\xi))$ for each $(t,\xi) \in \mathcal{D}$. Throughout, we denote by $\partial_{\mathcal{D}} F_{\tau,\xi}$ the boundary of $F_{\tau,\xi}$ relative to $\mathcal{D}$, i.e. $\partial_{\mathcal{D}} F_{\tau,\xi} = \overline{(\mathcal{D} \setminus F_{\tau,\xi})}^{\mathcal{D}} \cap \overline{\mathcal{D} \cap F_{\tau,\xi}}^{\mathcal{D}}$. We will deduce Theorem 7.4.1 from a slightly more general result, i.e. Theorem 7.4.3 below.

THEOREM 7.4.3 (Yorke [108]). *For each $(\tau,\xi) \in \mathbb{I} \times \mathbb{D}$, the set $\partial_{\mathcal{D}} F_{\tau,\xi}$ is left viable with respect to $\mathcal{F}$.*

PROOF. Let us observe that (1.1.1) can be equivalently written as

$$w'(t) = \mathcal{F}\big(w(t)\big),$$

where $\mathcal{F}$ is defined as above, and $w = (s, u)$. By the definition of $F_{\tau,\xi}$, we easily deduce that $\mathcal{D} \cap F_{\tau,\xi}$ is right viable and right locally invariant with respect to $\mathcal{F}$, and hence, by Propositions 1.1.1–1.1.3, it follows that $\mathcal{D} \setminus F_{\tau,\xi}$ is both left viable and left locally invariant with respect to $\mathcal{F}$. So, thanks to Proposition 3.6.1, we conclude that $\mathcal{K}_1 = \overline{\mathcal{D} \setminus F_{\tau,\xi}}^{\mathcal{D}}$ is left viable with respect to $\mathcal{F}$. Further, also by definition, $\mathcal{D} \cap F_{\tau,\xi}$ is left viable with respect to $\mathcal{F}$, and again by Proposition 3.6.1, it follows that $\mathcal{K}_2 = \overline{\mathcal{D} \cap F_{\tau,\xi}}^{\mathcal{D}}$ is left viable with respect to $\mathcal{F}$. Since $\mathcal{K}_1 \cup \mathcal{K}_2 = \mathcal{D}$, by Remark 1.1.1 and Theorem 7.4.2, we conclude that $\mathcal{K}_1 \cap \mathcal{K}_2 = \partial_{\mathcal{D}} F_{\tau,\xi}$ is left viable with respect to $\mathcal{F}$, and this completes the proof. □

We may now proceed to the proof of Theorem 7.4.1.

PROOF OF THEOREM 7.4.1. First, let us observe that thanks to Proposition 7.4.1, it follows that $\partial_{\mathcal{D}} F_{\tau,\xi} \subseteq F_{\tau,\xi}$. Hence, in view of Theorem 7.4.3, we know that, for each $(t, u(t)) \in \partial_{\mathcal{D}} F_{\tau,\xi}$, there exists at least one solution $v : [\theta, t] \to \mathbb{D}$, with $\tau \leqslant \theta < t$, $v(t) = u(t)$ and such that $(s, v(s)) \in \partial_{\mathcal{D}} F_{\tau,\xi}(s)$ for each $s \in [\theta, t]$. But, by virtue of Proposition 7.4.1, $F_{\tau,\xi}([\tau, t])$ is compact, and therefore a simple maximality argument shows that we can always extend such a solution to $[\tau, t]$, and this achieves the proof. □

REMARK 7.4.2. Theorem 7.4.3 implies that, for each $(\tilde{t}, \tilde{y}) \in \partial_{\mathcal{D}} F_{\tau,\xi}$, there exists at least one noncontinuable solution, $v(\cdot) : (\sigma, \tilde{t}] \to \mathbb{R}^n$, of (1.1.1), such that $(s, v(s)) \in \partial_{\mathcal{D}} F_{\tau,\xi}$ for all $s \in (\sigma, \tilde{t}] \cap [\tau, \tilde{t}]$. If $t > \tau$ is chosen as in Theorem 7.4.1 and $\tilde{t} \in [\tau, t]$, then by virtue of both Theorem 7.4.3 and Proposition 7.4.1, it follows that $[\tau, \tilde{t}] \subset (\sigma, \tilde{t}]$.

REMARK 7.4.3. We notice that

$$\partial F_{\tau,\xi}(t) \subseteq (\partial_{\mathcal{D}} F_{\tau,\xi})(t) = \left\{ v \in \mathbb{R}^n;\ (t, v) \in \partial_{\mathcal{D}} F_{\tau,\xi} \right\}$$

and the inclusion can be strict. So, Theorem 7.4.3 is more general that Theorem 7.4.1 because it considers all of $\partial_{\mathcal{D}} F_{\tau,\xi}$.

7.5. *Kneser's theorem*

We conclude this section with the celebrated theorem of Kneser.

THEOREM 7.5.1 (Kneser [61]). *Let $(\tau, \xi) \in \mathbb{I} \times \mathbb{D}$ and let $t > \tau$ be such that, for each $u \in \mathcal{S}(\tau, \xi)$, $u(t)$ is defined. Then $F_{\tau,\xi}(t)$ is connected.*

PROOF. Let us assume by contradiction that $F_{\tau,\xi}(t)$ is not connected. Then there exist two nonempty subsets $\mathbb{C}_1$, $\mathbb{C}_2$ with $F_{\tau,\xi}(t) = \mathbb{C}_1 \cup \mathbb{C}_2$ but $\mathbb{C}_1 \cap \overline{\mathbb{C}_2} = \overline{\mathbb{C}_1} \cap \mathbb{C}_2 = \emptyset$. Let $\mathbb{K}_1$ be the union of all right noncontinuable trajectories[4] of (1.1.1) whose corresponding solutions v, either are not defined at t, or, if defined, satisfy

$$\operatorname{dist}\big(v(t); \mathbb{C}_1\big) \leqslant \operatorname{dist}\big(v(t); \mathbb{C}_2\big). \tag{7.5.1}$$

Similarly, we define $\mathbb{K}_2$ by reversing the inequality (7.5.1). By virtue of Proposition 1.1.1, both $\mathbb{K}_1$ and $\mathbb{K}_2$ are right viable with respect to $\mathcal{F}$. In addition, $\mathbb{K}_1$ and $\mathbb{K}_2$ are locally closed in $\mathbb{I} \times \mathbb{D}$ and $\mathbb{K}_1 \cup \mathbb{K}_2 = \mathbb{I} \times \mathbb{D}$. So, by Proposition 3.6.1, it follows that both $\overline{\mathbb{K}_1}^{\mathcal{D}}$ and $\overline{\mathbb{K}_2}^{\mathcal{D}}$ are right viable with respect to $\mathcal{F}$ and, of course, locally closed. In view of Theorem 7.4.2, we conclude that $\mathbb{K} = \overline{\mathbb{K}}_1^{\mathcal{D}} \cap \overline{\mathbb{K}}_2^{\mathcal{D}}$ is right viable with respect to $\mathcal{F}$. As $(\tau, \xi) \in \mathbb{K}$, there exists a noncontinuable solution v such that $(s, v(s)) \in \mathbb{K}$ for each s in the domain of v. By the choice of t, $v(t)$ is defined and belongs to $\mathbb{C}_1 \cup \mathbb{C}_2$. Since $(t, v(t)) \in \mathbb{K}$, $\operatorname{dist}(v(t); \mathbb{C}_1) = \operatorname{dist}(v(t); \mathbb{C}_2)$ which must be 0. But this is absurd because $v(t)$ would be either in $\mathbb{C}_1 \cap \overline{\mathbb{C}}_2$, or in $\overline{\mathbb{C}}_1 \cap \mathbb{C}_2$. This contradiction can be eliminated only if $F_{\tau,\xi}(t)$ is connected as claimed. □

---

[4] We recall that a trajectory of (1.1.1) is a set of the form $\{v(t);\ t \in \mathbb{I}_v\}$, where $v : \mathbb{I}_v \to \mathbb{D}$ is a solution of (1.1.1).

7.6. *The characteristics method for a first order PDE*

Let Ω be a nonempty open subset of $\mathbb{R}^n$, let $H : \Omega \times \mathbb{R} \rightsquigarrow \mathbb{R}^n \times \mathbb{R}$ be a multifunction with nonempty values, and consider the first order partial differential equation

$$\inf_{(u,v)\in H(x,w(x))} \big(Dw(x)(u) - v\big) = 0, \tag{7.6.1}$$

where D denotes the differentiability concept of Severi [90]. We recall that a function $w : \Omega \to \mathbb{R}$ is *Severi differentiable* at a point $x \in \Omega$ if, for every $u \in \mathbb{R}^n$, there exists the finite limit

$$Dw(x)(u) = \lim_{\substack{s\downarrow 0\\ p\to 0}} \left(\frac{1}{s}\right)\big(w\big(x + s(u+p)\big) - w(x)\big),$$

called the *Severi differential of* w *at* x *in the direction* u. If w is Severi differentiable at x, the function $u \mapsto Dw(x)(u)$ from $\mathbb{R}^n$ to $\mathbb{R}$ is the *Severi differential of* w *at* x. For details see Ursescu [94]. We have

$$\operatorname{graph}\big(Dw(x)\big) = \mathcal{T}_{\operatorname{graph}(w)}\big(x, w(x)\big).$$

For $u \in \mathbb{R}^n$, let us consider the extended real numbers:

$$\underline{D}w(x)(u) = \liminf_{\substack{s\downarrow 0\\ p\to 0}} \left(\frac{1}{s}\right)\big(w\big(x + s(u+p)\big) - w(x)\big),$$

$$\overline{D}w(x)(u) = \limsup_{\substack{s\downarrow 0\\ p\to 0}} \left(\frac{1}{s}\right)\big(w\big(x + s(u+p)\big) - w(x)\big).$$

We have the equalities:

$$\operatorname{epi}\big(\underline{D}w(x)\big) = \mathcal{T}_{\operatorname{epi}(w)}\big(x, w(x)\big),$$

$$\operatorname{hyp}\big(\overline{D}w(x)\big) = \mathcal{T}_{\operatorname{hyp}(w)}\big(x, w(x)\big),$$

where epi stands for the epigraph and hyp for the hypograph. To conclude, w is differentiable at x if and only if $\underline{D}w(x)(u)$ and $\overline{D}w(x)(u)$ are finite and equal to each other for all $u \in \mathbb{R}^n$, while $\overline{D}w(x)(u) \leqslant \underline{D}w(x)(u)$ for all $u \in \mathbb{R}^n$.

The epigraph and hypograph equalities above show that $u \mapsto \underline{D}w(x)(u)$ and $u \mapsto \overline{D}w(x)(u)$ are l.s.c. and u.s.c., respectively.

By a solution to Eq. (7.6.1) we mean a differentiable function $w : \Omega \to \mathbb{R}$ which satisfies equality (7.6.1) for all $x \in \Omega$. Equation (7.6.1) is worth noting since it contains at least two important particular cases: the quasilinear, first order partial differential equations (Goursat [51], Perron [81], Kamke [59], Carathéodory [29], Courant and Hilbert [37]); the Bellman equations (Bellman [7–9], Pontryagin et al. [83], Boltyanskii [11,12], Gonzales [50], Hájek [53], Cesari [30], Clarke and Vinter [35], and many others). Among other particular cases of Eq. (7.6.1) we mention also the eikonal equation (see Ishii [58]). Let us consider further the ordinary differential inclusion

$$\bigl(X'(t), Y'(t)\bigr) \in H\bigl(X(t), Y(t)\bigr), \tag{7.6.2}$$

whose solutions we label as characteristics with regard to Eq. (7.6.1). The theory for the inclusion (7.6.2) is well developed (see Section 6.1 above) and we use it to start developing an existence theory for Eq. (7.6.1). We take here as a model the classical characteristics method and characterize the solutions w of (7.6.1) by means of the behavior of the functions w along solutions (X, Y) of (7.6.2). Since the inclusion (7.6.2) is an autonomous one, we consider only solutions $(X, Y) : [0, T) \to \mathbb{R}^n \times \mathbb{R}$ where $0 < T \leqslant \infty$. In order to characterize the solutions w of (7.6.1), we need the following conditions:

- (C_1) *For every $x \in \Omega$, there exists a solution $(X, Y) : [0, T) \to \mathbb{R}^n \times \mathbb{R}$ of the inclusion* (7.6.2)*, with $(X(0), Y(0)) = (x, w(x))$, such that, for every $s \in (0, T)$, $w(X(s)) \leqslant Y(s)$;*
- (C_2) *For every $x \in \Omega$, for every solution $(X, Y) : [0, T) \to \mathbb{R}^n \times \mathbb{R}$ of the inclusion* (7.6.2)*, with $(X(0), Y(0)) = (x, w(x))$, and for every $s \in (0, T)$, $Y(s) \leqslant w(X(s))$.*

Now we are ready to state a first result concerning Eq. (7.6.1).

THEOREM 7.6.1. *Let $H : \Omega \times \mathbb{R} \rightsquigarrow \mathbb{R}^n \times \mathbb{R}$ be both u.s.c. and l.s.c. with nonempty, compact and convex values, and let $w : \Omega \to \mathbb{R}$ be differentiable and such that H has the comparison property with respect to* $(\Omega \times \mathbb{R}, \operatorname{hyp}(w))$. *Then w is a solution to Eq.* (7.6.1) *if and only if it satisfies conditions* (C_1) *and* (C_2).

As far as we know, the differentiability concept of Severi is the least restrictive for which the characterization above holds true. Every other differentiability concepts used in the literature devoted to particular cases of Eq. (7.6.1) implies the classical Fréchet, referred better as to Stolz–Young–Fréchet–Hadamard, differentiability. And, w is Fréchet differentiable at x if and only if both w is Severi differentiable at x, and the Severi differential $Dw(x)$ is linear on $\mathbb{R}^n$. In addition, Eq. (7.6.1) admits a natural substitute which dispenses with any differentiability restriction, and which can be still characterized by using (C_1)

and (C$_2$). The substitute for Eq. (7.6.1) consists of the couple of first order partial differential inequalities below:

$$\inf_{(u,v)\in H(x,w(x))}\bigl(\underline{D}w(x)(u)-v\bigr)\leqslant 0, \tag{7.6.3}$$

$$0\leqslant \inf_{(u,v)\in H(x,w(x))}\bigl(\overline{D}w(x)(u)-v\bigr), \tag{7.6.4}$$

called *generalized Bellman equation*. A *solution to inequality* (7.6.3) (or (7.6.4)) is a function $w:\Omega\to\mathbb{R}$ which satisfies inequality (7.6.3) (or (7.6.4)) for all $x\in\Omega$. Clearly a function w is a solution to (7.6.1) if and only if it is a differentiable solution to the couple (7.6.3) and (7.6.4). Since differentiability at a point implies continuity at that point, we conclude that Theorem 7.6.1 above is a natural corollary of Theorem 7.6.2 below.

THEOREM 7.6.2. *Let $H:\Omega\times\mathbb{R}\rightsquigarrow\mathbb{R}^n\times\mathbb{R}$ be both u.s.c. and l.s.c. with nonempty, compact and convex values, and let $w:\Omega\to\mathbb{R}$ be continuous such that H has the comparison property with respect to $(\Omega\times\mathbb{R},\operatorname{hyp}(w))$. Then, w is a solution to* (7.6.3) *and* (7.6.4) *if and only if it satisfies conditions* (C$_1$) *and* (C$_2$).

We mention that all solutions to every variational problem satisfy both (C$_1$) and (C$_2$) (the Bellman "principle") with a suitable chosen H. Hence, under rather common hypotheses upon the components of a variational problem, its continuous solution satisfies inequalities (7.6.3) and (7.6.4) (the generalized Bellman "equation"). A typical example is the time optimal control problem associated to a control system and a target. More precisely, consider the multifunction $F:\mathbb{R}^n\rightsquigarrow\mathbb{R}^n$, the differential inclusion

$$X'(t)\in F\bigl(X(t)\bigr), \tag{7.6.5}$$

and fix a *target set* $\mathcal{T}$ (nonempty and closed). Let $\mathcal{R}$ be the *reachable set*, that is, the set of all initial points which can be transferred to $\mathcal{T}$ by trajectories of (7.6.5). For $x\in\mathcal{R}$ define $T(x)$ as the infimum of the transition times. The well-known Bellman equation for the optimal time problem is

$$-1+\sup_{u\in F(x)}-DT(x)(u)=0,\qquad x\in\mathcal{R}\setminus\mathcal{T},$$

which is a particular case of (7.6.1), for the choice $H(x,y)=(F(x),\{-1\})$.

In its turn, Theorem 7.6.2 follows from the following anatomized variant of itself.

THEOREM 7.6.3. (a) *Let $H:\Omega\times\mathbb{R}\rightsquigarrow\mathbb{R}^n\times\mathbb{R}$ be u.s.c. with nonempty, compact and convex values. A continuous function $w:\Omega\to\mathbb{R}$ is a solution to inequality* (7.6.3) *if and only if satisfies condition* (C$_1$).

(b) *Let $H:\Omega\times\mathbb{R}\rightsquigarrow\mathbb{R}^n\times\mathbb{R}$ be l.s.c. with closed and convex values. Let $w:\Omega\to\mathbb{R}$ be a continuous function such that H has the comparison property with respect to $(\Omega\times\mathbb{R},\operatorname{hyp}(w))$. Then w is a solution to inequality* (7.6.4) *if and only if it satisfies condition* (C$_2$).

PROOF. (a) Since $H(x, w(x))$ is compact and since $\underline{D}w(x)$ is l.s.c. on $\mathbb{R}^n$, it follows that "inf" can be replaced by "min" in (7.6.3), hence inequality (7.6.3) states that

$$\phi \neq H\big(x, w(x)\big) \cap \mathcal{T}_{\mathrm{epi}(w)}\big(x, w(x)\big). \tag{7.6.6}$$

Since w is continuous, (7.6.6) is equivalent to (7.6.7) below

$$\phi \neq H(x, t) \cap \mathcal{T}_{\mathrm{epi}(w)}(x, t) \tag{7.6.7}$$

for $t \geqslant w(x)$, and this simply because $\mathcal{T}_{\mathrm{epi}(w)}(x, t) = \mathbb{R}^n \times \mathbb{R}$ whenever $w(x) < t$. On the other hand, since w is continuous, condition (C_1) is equivalent to the viability of epi(w) with respect to H. But the set epi(w) is closed in $\Omega \times \mathbb{R}$, and therefore the conclusion is an immediate consequence of Theorem 6.2.1.

(b) The inequality in (7.6.4) states that $H(x, w(x)) \subseteq \mathcal{T}_{\mathrm{hyp}(w)}(x, w(x))$. Since w is continuous, we have $H(x, t) \subseteq \mathcal{T}_{\mathrm{hyp}(w)}(x, t)$ in case $w(x) > t$. On the other hand, condition (C_2) states that the set hyp(w) is invariant with respect to H. Since the set hyp(w) is closed in $\Omega \times \mathbb{R}$, the conclusion follows from Theorem 6.8.1 and Corollary 6.8.1. □

A natural question is whether we can weaken the continuity property of the function w in Theorem 7.6.2. This question arises from the fact that epi(w) is closed even in the case when w is l.s.c. and hyp(w) is closed in case w is u.s.c. The answer is in the negative as the following examples in Cârjă and Ursescu [25] show. Consider first the inequality

$$\underline{D}w(x)\big(w(x)\big) - 1 \leqslant 0.$$

Here $H(x, y) = \{(y, 1)\}$, for all $(x, y) \in \mathbb{R} \times \mathbb{R}$ and the characteristic system $X'(s) = Y(s)$, $Y'(s) = 1$ has the solution $X(s) = X(0) + sY(0) + s^2/2$, $Y(s) = Y(0) + s$. The l.s.c. function $w : \mathbb{R} \to \mathbb{R}$ given by $w(x) = 0$ for $x = 0$ and by $w(x) = 1$ for $x \neq 0$ is a solution to the preceding inequality but does not satisfy the condition: *for every* $x \in \mathbb{R}$ *and for every* $s \in (0, +\infty)$, $w(x + sw(x) + s^2/2) \leqslant w(x) + s$ (take $x = 0$ and $s \in (0, 1)$ and observe $w(0 + sw(0) + s^2/2) = 1 > s = w(0) + s$). A second example is given by the differential inequality

$$0 \leqslant \overline{D}w(x)\big(w(x)\big) + 1.$$

Here $H(x, y) = \{(y, -1)\}$ for all $(x, y) \in \mathbb{R} \times \mathbb{R}$ and the characteristic system $X'(s) = Y(s)$, $Y'(s) = -1$ has the solutions $X(s) = X(0) + sy(0) - s^2/2$, $Y(s) = Y(0) - s$. The u.s.c. function $w : \mathbb{R} \to \mathbb{R}$ given by $w(x) = 0$ for $x = 0$ and by $w(x) = -1$ for $x \neq 0$ is a solution to the preceding inequality but does not satisfy the condition: *for every* $x \in \mathbb{R}$ *and for every* $s \in (0, +\infty)$, $w(x) - s \leqslant w(x + sw(x) - s^2/2)$ (take $x = 0$ and $s \in (0, 1)$ and observe $w(0) - s = -s > -1 = w(0) + sw(0) - s^2/2)$.

A condition which assures that (C_1) in part (a) of Theorem 7.6.3 holds for a l.s.c. function w, while (C_2) in part (b) holds for a u.s.c. function w is given below:

THEOREM 7.6.4. *Let $H : \Omega \times \mathbb{R} \rightsquigarrow \mathbb{R}^n \times \mathbb{R}$ be a nonempty and convex valued multifunction satisfying*

$$H(x, y_1) \subseteq H(x, y_2) \tag{7.6.8}$$

for each $x \in \Omega$ and $y_1, y_2 \in \mathbb{R}$, with $y_1 \leqslant y_2$. (a) *Assume that H is u.s.c. with compact values. Then, a l.s.c. function $w : \Omega \to \mathbb{R}$ is a solution to the inequality* (7.6.3) *if and only if it satisfies* (C_1).

(b) *Assume that H is l.s.c. with closed values. Let $w : \Omega \to \mathbb{R}$ a u.s.c. function such that H has the comparison property with respect to $(\Omega \times \mathbb{R}, \mathrm{hyp}(w))$. Then, $w : \Omega \to \mathbb{R}$ is a solution to the inequality* (7.6.4) *if and only if it satisfies* (C_2).

PROOF. The proof goes in the same spirit as that of Theorem 7.6.3. Indeed, in order to prove (a), since $\mathcal{T}_{\mathrm{epi}(w)}(x, w(x)) \subseteq \mathcal{T}_{\mathrm{epi}(w)}(x, t)$ if $w(x) \leqslant t$, (7.6.8) implies that (7.6.6) is equivalent to (7.6.7). In its turn, (7.6.6) is equivalent to (7.6.3). On the other hand, since (7.6.8) is satisfied, condition (C_1) is equivalent to the viability of epi(w) with respect to H. Indeed, if $w(x) < t$, and $(X, Y) : [0, T] \to \mathbb{R}^n \times \mathbb{R}$ is a solution to (7.6.2) with $(X(0), Y(0)) = (x, w(x))$, and $w(X(s)) \leqslant Y(s)$ for all $s \in [0, T]$, then the function $s \mapsto (X(s), Y(s) + t - w(x)) = (\overline{X}(s), \overline{Y}(s))$ is a solution to (7.6.2) with $(\overline{X}(0), \overline{Y}(0)) = (x, t)$ and satisfies $w(\overline{X}(s)) \leqslant \overline{Y}(s)$ for all $s \in [0, T]$. This completes the proof of (a). The proof of (b) goes in the very same spirit, and therefore we do not give details. □

REMARK 7.6.1. The property (C_1) is related to the so-called *weakly decreasing systems* discussed in Clarke et al. [34, p. 211], which in turn are related to the Liapunov theory of stabilization. See also [34, p. 208]. The "monotone variant" of (C_1) labelled below as (C_3) is connected to the so-called *strongly decreasing systems* discussed also in [34, p. 217].

(C_3) *For each $x \in \Omega$ there exists a solution $(X, Y) : [0, T] \to \mathbb{R}^n \times \mathbb{R}$ of the differential inclusion* (7.6.2) *with $(X(0), Y(0)) = (x, w(x))$, such that the function $s \mapsto w(X(s)) - Y(s)$ is decreasing.*

We have:

THEOREM 7.6.5. *Let $H : \Omega \times \mathbb{R} \rightsquigarrow \mathbb{R}^n \times \mathbb{R}$ be a u.s.c., nonempty and convex valued multifunction. Let, us assume that* (7.6.8) *is satisfied. Then, a continuous function w satisfies* (C_1) *if and only if it satisfies* (C_3).

PROOF. On epi(w) we define a relation $\mathcal{P}$ by

$$\mathcal{P}(x, t) = \big\{(y, s);\ w(y) - s \leqslant w(x) - t\big\}.$$

Obviously, $\mathcal{P}$ is a preorder on epi(w) and condition (C_3) is equivalent to the viability of the preorder $\mathcal{P}$ with respect to H. Moreover, condition (C_1) is equivalent to the viability of $\mathcal{P}(x, t)$ with respect to H for each $(x, t) \in \mathrm{epi}(w)$. The conclusion follows from Lemma 6.6.1. □

8. Notes and comments

8.1. *The upper semicontinuous case*

Since of its birth in 1942, viability theory emerged in several directions we will discuss sequentially below. The first viability result for the multivalued case is due to Bebernes and Schuur [6]. We mention that, in 1936, Zaremba [111] proved that if $F:\mathbb{D}\rightsquigarrow\mathbb{R}^n$ is u.s.c. with nonempty compact and convex values and $\mathbb{D}$ is open then, for each $\xi\in\mathbb{D}$, the differential inclusion (6.2.1) has at least one solution $u:[0,T]\to\mathbb{D}$ satisfying $u(0)=\xi$. It should be noticed that the concept of solution used by Zaremba in [111] is in the sense of the *contingent derivative*. More precisely, if $u:[0,T]\to\mathbb{R}^n$ is continuous and $t\in[0,T)$, the set

$$Du(t)=\left\{\lim_{m\to\infty}\frac{u(t+t_m)-u(t)}{t_m};\ t_m\downarrow 0\right\}$$

is called the *contingent derivative* of u at t. We say that $u:[0,T]\to\mathbb{D}$ is a *contingent solution* of (6.2.1) if

$$\emptyset\neq Du(t)\subseteq F\big(u(t)\big) \tag{8.1.1}$$

for each $t\in[0,T)$. In 1961, Ważewski [107] proved that, if F is u.s.c. with nonempty, compact and convex values, u is a contingent solution to (6.2.1) if and only if u is a Carathéodory solution to (6.2.1). So, Zaremba's existence result in [111] is nothing than the multivalued counterpart of Peano's local existence theorem 1.1.1. In the same spirit, the viability result of Bebernes and Schuur is the multivalued version of Nagumo's viability theorem 1.1.3. Proposition 8.1.1 below, due to Ważewski [107], is in fact equivalent to the necessity part of Theorem 6.2.1.

PROPOSITION 8.1.1. *Let $\mathbb{K}\subseteq\mathbb{R}^n$ be nonempty and locally closed and let $F:\mathbb{K}\rightsquigarrow\mathbb{R}^n$ be u.s.c. with nonempty, convex, compact values. Then, for every $\xi\in\mathbb{K}$, and every solution $u:[0,T]\to\mathbb{K}$ to* (6.2.1), *with $u(0)=\xi$, there exist $\eta\in F(\xi)$ and a sequence $(t_m)_m$ in $(0,T)$ convergent to* 0 *such that the sequence $(\frac{1}{t_m}(u(t_m)-\xi))_m$ converges to η.*

It is interesting to notice that, by using the viability theory developed in Section 6.1 for the locally closed set $\mathbb{K}=\{(t,u(t));\ t\in[0,T]\}$ and the multifunction $\{1\}\times F$, we can prove (see [22]) that condition (8.1.1) is also equivalent to each one of the following:

(i) $Du(t)\cap F(u(t))\neq\emptyset$ for each $t\in[0,T)$, or

(ii) $\overline{\mathrm{co}}\,Du(t)\cap F(u(t))\neq\emptyset$ for each $t\in[0,T)$.

In 1981, Haddad [52] obtained the first result on viability of preorders. In fact, Haddad adapted the proof of viability of sets in order to obtain viability of preorders. Cârjă and Ursescu [25] showed that the viability, as well as the invariance of preorders can be completely described in terms of viability, or invariance of sets. See Section 3.8.

The structure of the set of viable solutions of a differential inclusion on a subset $\mathbb{K}$, for which $\Pi_{\mathbb{K}}$ has continuous selections, was studied by Plaskacz [82]. See also the references therein.

As far as the infinite dimensional case is concerned, i.e. the case in which instead of $\mathbb{R}^n$ we are considering an infinite-dimensional Banach space X, we mention the pioneering contribution of Gauthier [48]. First, he used Zorn's Lemma in order to get approximate solutions defined on an a priori given interval. Second, he used a sufficient weak tangency condition of the form: *for each* $\xi \in \mathbb{K}$, *there exist* $\eta \in F(\xi)$, *a sequence* $(h_m)_m$ *decreasing to* 0, *a sequence* $(q_m)_m$ *weakly convergent to* 0 *satisfying* $\|q_m + \eta\| \leqslant 2\|\eta\|$ *and* $\xi + h_m(\eta + q_m) \in \mathbb{K}$ *for each* $m \in \mathbb{N}$. This tangency is far from being necessary for the viability of $\mathbb{K}$. A necessary and sufficient condition for the viability of $\mathbb{K}$ in a more general setting has been obtained by Cârjă and Vrabie [27] by means of the so-called "bounded weak tangency condition".

8.2. *The case of Carathéodory mappings*

The first viability result in the case of a single-valued Carathéodory right-hand side is due to Ursescu [98]. Theorem 5.2.1 is an extension of Ursescu's result in [98] which contains only the equivalence between (ii) and (vi) in the above mentioned theorem.

Although not presented here, the Carathéodory case for differential inclusions is well developed and there exists a rather large literature on the subject. Among the first notable results in this direction we mention those of Tallos [92], Ledyaev [64], Frankowska et al. [45]. In all these papers, theorems of Scorza Dragoni type are the main tools. Results of the same kind as in Theorem 5.2.1, but for differential inclusions, can be found in Cârjă and Monteiro Marques [22] in the finite-dimensional case, and in Cârjă and Monteiro Marques [23] in the infinite-dimensional setting. There, a technique of approximation (as in Theorem 5.2.1) of the multifunction through the Aumann integral mean is used.

8.3. *The lower semicontinuous case*

The next example shows that the convexity condition on the values of F is essential in obtaining the viability of a locally closed set $\mathbb{K}$ with respect to a u.s.c. multifunction $F : \mathbb{K} \rightsquigarrow \mathbb{R}^n$ by means of the tangency condition $\mathcal{T}_{\mathbb{K}}(\xi) \cap F(\xi) \neq \emptyset$ for all $\xi \in \mathbb{K}$.

EXAMPLE 8.3.1 (Aubin and Cellina [3, p. 202]). Let $n = 2$, $\mathbb{K} = B(0, 1)$ and $F : \mathbb{K} \to \mathbb{R}^2$, defined by $F(\xi) = \{(-1, 0), (1, 0)\}$ for each $\xi \in \mathbb{K}$. Then, one may easily see that $\mathbb{K}$ is locally closed (in fact closed and convex), F is u.s.c., satisfies the tangency condition, but nevertheless $\mathbb{K}$ is not viable with respect to F.

The lack of convexity of the values of F can be however counterbalanced by an l.s.c. extra-assumption combined with the stronger tangency condition

$$F(\xi) \subseteq \mathcal{T}_{\mathbb{K}}(\xi) \tag{8.3.1}$$

for each $\xi \in \mathbb{K}$. More precisely, we have:

THEOREM 8.3.1 (Aubin and Cellina [3, p. 198]). *Let $\mathbb{K}$ be locally closed and let $F:\mathbb{K}\rightsquigarrow\mathbb{R}^n$ be both u.s.c. and l.s.c. with nonempty and closed values. If* (8.3.1) *is satisfied, then $\mathbb{K}$ is viable with respect to F. Moreover, x' is a regulated function.*[5]

We emphasize that even for the convex-valued case, l.s.c. and (8.3.1) do not ensure invariance (see the case of functions), but implies viability via the Michael's selection theorem [67,68].

8.4. *The semilinear single-valued case*

Another direction was to consider a larger framework in order to handle semilinear partial differential equations as well. To this aim, let us consider an infinite dimensional Banach space X with norm $\|\cdot\|$ and let $\mathcal{L}(X)$ be the space of all linear bounded operators from X to X, endowed with the usual *operator norm* $\|\cdot\|_{\mathcal{L}(X)}$. We recall that $\{S(t);\ t\geqslant 0\}\subseteq\mathcal{L}(X)$ is a C_0*-semigroup* if

(i) $S(0)=I$;
(ii) $S(t+s)=S(t)S(s)$ for all $t,s\geqslant 0$;
(iii) $\lim_{h\downarrow 0}S(h)\xi=\xi$ for each $\xi\in X$.

The *infinitesimal generator* of $\{S(t);\ t\geqslant 0\}$ is the possibly unbounded linear operator $A:D(A)\subseteq X\to X$, defined by

$$\begin{cases} D(A)=\left\{\xi\in X;\ \exists\lim\limits_{h\downarrow 0}\dfrac{1}{h}\big[S(h)\xi-\xi\big]\right\},\\ A\xi=\lim\limits_{h\downarrow 0}\dfrac{1}{h}\big[S(h)\xi-\xi\big]\quad\text{for }\xi\in D(A).\end{cases}$$

Further, if $\{S(t);\ t\geqslant 0\}$ is a C_0-semigroup, $A:D(A)\subseteq X\to X$ is its infinitesimal generator and $\xi\in D(A)$, then the mapping $t\mapsto S(t-\tau)\xi$ is the unique classical, i.e., C^1, solution to the ordinary homogeneous differential equation

$$\begin{cases} u'=Au,\\ u(\tau)=\xi,\end{cases}\tag{8.4.1}$$

defined on $[\tau,+\infty)$. Moreover, since $D(A)$ is dense in X, it follows that, for each $\xi\in X$, the mapping $t\mapsto S(t-\tau)\xi$, which may fail to be C^1, can be approximated uniformly on compact subsets in $[\tau,+\infty)$ by classical solutions to problem (8.4.1). Thus, for each $\xi\in X$, the mapping above can be considered as a generalized solution of (8.4.1), called the *semigroup solution*. Furthermore, inspired from the variation of constants formula, we may define the so-called *mild, or C^0 solution*, $u:[\tau,a)\to X$, of the nonhomogeneous problem

$$\begin{cases} u'=Au+f(t),\\ u(\tau)=\xi,\end{cases}$$

---

[5] A function is *regulated* if it is uniform limit of step functions.

by

$$u(t) = S(t-\tau)\xi + \int_\tau^t S(t-s)f(s)\,\mathrm{d}s$$

for each $t \in [\tau, a)$. Here $f : [\tau, a) \to X$ is a given continuous, or even locally integrable, function and $a > \tau$ is finite or not. So, if $\mathbb{D} \subseteq X$ is a nonempty and open subset in X, $\mathbb{I}$ is a given nonempty open interval, $f : \mathbb{I} \times \mathbb{D} \to X$, $\tau \in \mathbb{I}$ and $\xi \in \mathbb{D}$, we may consider the *semilinear differential equation*

$$\begin{cases} u' = Au + f(t,u), \\ u(\tau) = \xi. \end{cases} \tag{8.4.2}$$

As expected, a *mild, or C^0 solution* of (8.4.2) is a continuous function $u : [\tau, T) \to \mathbb{D}$, with $T \in \mathbb{I}$, and satisfying

$$u(t) = S(t-\tau)\xi + \int_\tau^t S(t-s)f\big(s, u(s)\big)\,\mathrm{d}s \tag{8.4.3}$$

for each $t \in [\tau, T)$. For details on C_0-semigroups see Vrabie [105]. Further, if $\mathbb{K} \subseteq \mathbb{D}$ is locally closed, we redefine the concepts of right viability and right invariance of $\mathbb{K}$ with respect to $A + f$ by using mild or C^0 solutions. The main problem here is that, in general, $\mathbb{K} \cap D(A)$ is very narrow, even empty (we emphasize that $D(A)$ is only dense and could have empty interior), and therefore a necessary, or/and sufficient condition for viability of the type (1.1.3), i.e.

$$\liminf_{h \downarrow 0} \frac{1}{h} \operatorname{dist}\big(\xi + h\big[A\xi + f(t,\xi)\big]; \mathbb{K}\big) = 0, \tag{8.4.4}$$

for each $(t, \xi) \in \mathbb{I} \times [\mathbb{K} \cap D(A)]$, turns out to be unrealistic for nonsmooth semigroups. An example which justifies the remark above is that one when $\mathbb{K}$ is the trajectory of a nowhere differentiable right mild solution to (8.4.2), case in which $\mathbb{K} \cap D(A) = \emptyset$. The case of smooth semigroups (even nonlinear) was considered recently by Barbu and Pavel [5] by means of the tangency condition (8.4.4) with "lim" instead of "lim inf". It is the merit of Pavel [75] to observe that, in general, the really useful sufficient, and, very often, even necessary, condition for right viability which works in this case is

$$\lim_{h \downarrow 0} \frac{1}{h} \operatorname{dist}\big(S(h)\xi + hf(t,\xi); \mathbb{K}\big) = 0, \tag{8.4.5}$$

for each $(t, \xi) \in \mathbb{I} \times \mathbb{K}$. We notice that, whenever $\xi \in \mathbb{K} \cap D(A)$, (8.4.5) is equivalent to (8.4.4) which is a stronger form, i.e. with "lim inf" replaced by "lim", of the classical Nagumo's tangency condition (1.1.3) with $A + f$ instead of f. We re-emphasize that whenever $\mathbb{K}$ is not included in $D(A)$, or even $\mathbb{K} \cap D(A)$ is empty, the only tangency condition which could be of some use is (8.4.5). Namely, the main result in Pavel [75] is:

THEOREM 8.4.1 (Pavel [75]). *If $S(t)$ is compact for each $t > 0$ and $f : \mathbb{I} \times \mathbb{K} \to X$ is continuous, then a necessary and sufficient condition in order that $\mathbb{K}$ be right viable with respect to $(t, u) \mapsto Au + f(t, u)$ is* (8.4.5).

We notice that the compactness of $S(t)$ for each $t > 0$ is a parabolicity condition. The simplest nontrivial example of a C_0 semigroup satisfying this condition is that one occurring in the study of the heat flow. Namely, let $S(t) : L^2(0, \pi) \to L^2(0, \pi)$ be defined by

$$\big[S(t)\xi\big](x) = \sqrt{\frac{2}{\pi}} \sum_{k=1}^{\infty} a_k(\xi) \mathrm{e}^{-k^2 t} \sin kx$$

for $x \in (0, \pi)$, where $a_k(\xi)$ are the Fourier coefficients of ξ with respect to the orthogonal system $\big\{\sqrt{\frac{2}{\pi}} \sin x, \sqrt{\frac{2}{\pi}} \sin 2x, \ldots, \sqrt{\frac{2}{\pi}} \sin kx, \ldots\big\}$, i.e.

$$a_k(\xi) = \sqrt{\frac{2}{\pi}} \int_0^{\pi} \xi(y) \sin ky \, \mathrm{d}y.$$

It is well known that this semigroup defines the $L^2(0, \pi)$-solutions to the one-dimensional heat equation

$$\begin{cases} u_t = u_{xx} & \text{for } (t, x) \in [0, +\infty) \times (0, \pi), \\ u(t, 0) = u(t, \pi) & \text{for } t \in [0, +\infty), \\ u(0, x) = \xi & \text{for } x \in (0, \pi). \end{cases}$$

Moreover, for each $t > 0$, $S(t)$ is a compact, in fact even a Hilbert–Schmidt operator, because it is the limit in the operator norm of a sequence of finite-dimensional range operators. Indeed, a simple example of such sequence $(S_m(t))_m$ is

$$\big[S_m(t)\xi\big](x) = \sqrt{\frac{2}{\pi}} \sum_{k=1}^{m} a_k(\xi) \mathrm{e}^{-k^2 t} \sin kx$$

for $m = 1, 2, \ldots$ and each $\xi \in X$, where $a_k(\xi)$ are as above.

The extension of Pavel's Theorem 8.4.1 to Carathéodory perturbations has been considered by Cârjă and Monteiro Marques [21] for constant in time domains, and by Necula [71] for time dependent domains.

8.5. *The semilinear multivalued case*

As far as we know, the true semilinear and multivalued case, i.e. A linear unbounded and F multivalued, has been analyzed first by Pavel and Vrabie [78,79] by the end of seventies. A good source of references in this respect is Pavel [76]. Shi [91] considers the semilinear case (8.4.2) in which f is replaced by a multifunction, redefines the concept of viability by using strong solutions, i.e. continuous functions $u : [0, T) \to X$, $T \in (0, +\infty]$, which are absolutely continuous on $(0, T)$ and satisfy: $u'(t) = Au(t) + f(t)$ a.e. for $t \in [0, T)$, where $f \in L^\infty(0, T; X)$, $f(t) \in F(t, u(t))$ a.e. for $t \in (0, T)$, and proves a characterization of global viability, i.e. of right viability involving only strong solutions defined on $[0, +\infty)$.

THEOREM 8.5.1 (Shi [91]). *Let X be reflexive, $\mathbb{K}$ a compact subset of X, $F : X \rightsquigarrow X$ a nonempty, convex and compact valued upper semicontinuous mapping, and let $A : D(A) \subseteq X \to X$ be the infinitesimal generator of a differentiable C_0-semigroup $\{S(t);\ t \geqslant 0\}$ with $S(t)$ compact for all $t > 0$. Then a necessary and sufficient condition in order that $\mathbb{K}$ be right global viable with respect to F is the following tangency condition: for each $\xi \in \mathbb{K}$ there exist $\eta \in F(\xi)$, a sequence $(h_m)_m$ decreasing to 0 and a sequence $(p_m)_m$ strongly convergent to 0 such that*

$$S(h_m)\xi + h_m(\eta + p_m) \in \mathbb{K}$$

holds for each $m \in \mathbb{N}$.

As concerns sufficient conditions for global viability, we mention:

THEOREM 8.5.2 (Shi [91]). *Let X be reflexive, $\mathbb{K}$ a compact subset of X, $F : X \rightsquigarrow X$ a nonempty, bounded, closed and convex valued upper semicontinuous mapping, and let $A : D(A) \subseteq X \to X$ be the infinitesimal generator of a differentiable C_0-semigroup $\{S(t);\ t \geqslant 0\}$ with $S(t)$ compact for all $t > 0$. Then, a sufficient condition in order that $\mathbb{K}$ be right global viable with respect to F is the following tangency condition: for each $\xi \in \mathbb{K}$ and each $t > 0$ there exist $\eta \in F(\xi)$, a sequence $(h_m)_m$ decreasing to 0 and a sequence $(p_m)_m$ strongly convergent to 0 such that*

$$S(h_m)\xi + h_m\big(S(t)\eta + p_m\big) \in S(t)\mathbb{K}$$

holds for each $m \in \mathbb{N}$.

We note that the tangency condition in Theorem 8.5.2 is equivalent to:

$$AS(t)\xi + S(t)F(\xi) \subset \mathcal{T}_{S(t)\mathbb{K}}\big(S(t)\xi\big)$$

holds for each $\xi \in \mathbb{K}$ and each $t > 0$.

Clearly, in this case the general assumptions on $\mathbb{K}$ and F are significantly stronger than those in Pavel and Vrabie [78,79]. We note that for instance, in the infinite-dimensional setting, the compactness of $F(\xi)$ for each $\xi \in D$ is not satisfied if F is a superposition operator which is not single-valued. On the other hand, this “weakness” of the general setting of Shi [91] is well counterbalanced by the tangency condition which is quite close to its finite-dimensional counterpart.

The existence of monotone solutions has been considered in this context by Chiş-Şter [31].

8.6. *The nonlinear perturbed single-valued case*

Theorem 8.4.1 was partially extended by Vrabie [103] for the fully nonlinear case, i.e. when A m-dissipative. Namely, Vrabie [103] shows that, if A generates a nonlinear semigroup of nonexpansive compact, for $t > 0$, operators and $f : \mathbb{I} \times \mathbb{K} \to X$ is continuous on the locally closed subset $\mathbb{K}$, a sufficient condition for viability is

$$\lim_{h \downarrow 0} \frac{1}{h} \operatorname{dist}\bigl(u\bigl(t+h, t, \xi, f(t, \xi)\bigr); \mathbb{K}\bigr) = 0 \tag{8.6.1}$$

uniformly for $(t, \xi) \in \mathbb{I} \times \mathbb{K}$, where $u(\cdot, t, \xi, \eta)$ denotes the unique mild solution to the problem $u'(s) \in Au(s) + \eta$ satisfying $u(t, t, \xi, \eta) = \xi$. Subsequent contributions in this context are due to Bothe [14] who allowed $\mathbb{K}$ to depend on t as well. In particular, in case $\mathbb{K}$ independent of t, Bothe [14] showed that the, possibly non uniform, tangency condition (8.6.1) satisfied for each $(t, \xi) \in \mathbb{I} \times \mathbb{K}$ is necessary and sufficient for viability. The case of Carathéodory perturbations defined on time dependent domains has been considered by Necula [72].

8.7. *The multivalued perturbed nonlinear case*

The more delicate case in which one allows A to be nonlinear, as well as f to be multivalued, has been considered by Bressan and Staicu [17] who used the tangency condition proposed in Vrabie [103] after reducing the multivalued case to the single-valued one by means of a continuous selection argument. The case in which the multifunction F is strongly-weakly u.s.c. has been considered by Cârjă and Vrabie [28] by using a weak variant of the tangency condition in Vrabie [103], while the possibly nonconvex valued case has been analyzed by Necula and Vrabie [74] by using a selection theorem due to Fryszkowski [46].

8.8. *Applications*

The problem of finding Liapunov functions for differential inclusions is discussed in Aubin [1, Chapter 9], and Aubin and Cellina [3, Chapter 6]. As we have pointed out in Section 7.6, the characteristics method, when applied in control theory, is related to the dynamic programming method. Theorem 7.6.2 says that there is an equivalence between the Bellman equations (7.6.3) and (7.6.4) and the Bellman optimality conditions (C_1) and (C_2). If we want to study the uniqueness properties of the Bellman equation, we have to add appropriate boundary conditions and to use conditions (C_1) and (C_2) in order to get that a possible solution is necessarily the value function of the given control problem. The notion of solution for the couple of inequalities (7.6.3) and (7.6.4), defined in Section 7.6, is usually called *contingent solution*. It is interesting to notice that this kind of solution is equivalent with that of viscosity solution developed by Crandall and Lions [39]. However, the standard technique to get uniqueness is different here. Among the first contributions in this area we mention Frankowska [44] and Cârjă and Ursescu [26]. See also the more recent article of Aubin [2] and the references therein. For other applications of viability and invariance techniques to Control Theory and Game Theory, we refer the reader to Aubin and Cellina [3, Chapter 5]. As concerns the existence theory for periodic problems via viability and/or invariance see Aubin and Cellina [3, Chapter 5] and Vrabie [104]. For applications to the existence of constrained solutions to nonlinear partial differential inclusions see Bothe [15].

Finally, it should be noticed that the theory of invariant sets pertains a very elegant approach to the study of orbital motions of a mass particle in a given force field. This was pointed out by means of a second order tangency concept by Pavel and Ursescu [77].

Appendix

THEOREM A.1.1. *Let* $p \in [1, +\infty)$ *and let* $[\tau, T] \subseteq \mathbb{I}$. *A subset* $\mathcal{F}$ *in* $L^p_{\mathrm{loc}}(\mathbb{I}; \mathbb{R}^n)$ *is relatively compact in* $L^p(\tau, T; \mathbb{R}^n)$ *if and only if it is bounded and*

$$\lim_{\theta \downarrow 0} \int_\tau^{T-\theta} \left\| f(s+\theta) - f(s) \right\| \mathrm{d}s = 0$$

uniformly for $f \in \mathcal{F}$.

See Vrabie [105, Theorem A.4.1, p. 305]. For the proof of the next result see Dunford and Schwartz [40, Theorem 15, p. 150].

THEOREM A.1.2 (Vitali). *Let* $1 \leqslant p < +\infty$, *let* $\mathbb{S}$ *be a Lebesgue measurable subset in* $\mathbb{R}$, *and let* $(f_m)_m$ *be a sequence in* $L^p(\mathbb{S}; \mathbb{R}^n)$ *converging almost everywhere to a function* f. *Then* $f \in L^p(\mathbb{S}; \mathbb{R}^n)$ *and*

$$\lim_m \| f_m - f \|_{L^p(\mathbb{S};\mathbb{R}^n)} = 0$$

if and only if:

(i) $\{\| f_m \|^p;\ m \in \mathbb{N}\}$ *is uniformly integrable in* $L^1(\mathbb{S}; \mathbb{R}^n)$;

(ii) *for each* $\varepsilon > 0$ *there exists a Lebesgue measurable set* $\mathbb{E}_\varepsilon \subseteq \mathbb{S}$ *with* $\lambda(\mathbb{E}_\varepsilon) < +\infty$ *and such that, for each* $m \in \mathbb{N}$, *we have*

$$\int_{\mathbb{S} \setminus \mathbb{E}_\varepsilon} \left\| f_m(s) \right\|^p \mathrm{d}s < \varepsilon.$$

We recall that a subset $\mathcal{F}$ in $L^1(\mathbb{S}; \mathbb{R}^n)$ is uniformly integrable if

$$\lim_{\lambda(\mathbb{E}) \downarrow 0} \int_{\mathbb{E}} \left\| f(s) \right\| \mathrm{d}s = 0$$

uniformly for $f \in \mathcal{F}$. We notice that whenever $\mathbb{S}$ is a finite length interval, the condition (ii) is automatically satisfied.

References

[1] J.P. Aubin, *Viability Theory*, Birkhäuser, Boston (1991).

[2] J.P. Aubin, *Boundary-value problems for systems of Hamilton–Jacobi–Bellman inclusions with constraints*, SIAM J. Control Optim. **41** (2002), 425–456.

[3] J.P. Aubin and A. Cellina, *Differential Inclusions*, Springer-Verlag, Berlin (1984).

[4] J.P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser (1990).

[5] V. Barbu and N.H. Pavel, *Flow-invariant closed sets with respect to nonlinear semigroup flows*, NoDEA Nonlinear Differential Equations Appl. **10** (2003), 57–72.

[6] W. Bebernes and I.D. Schuur, *The Ważewski topological method for contingent equations*, Ann. Mat. Pura Appl. **87** (1970), 271–278.

[7] R. Bellman, *Bottleneck problems and dynamic programming*, Proc. Natl. Acad. Sci. USA **39** (1953), 947–951.

[8] R. Bellman, *Bottleneck problems, functional equations and dynamic programming*, Econometrica **23** (1955), 73–87.

[9] R. Bellman, *Dynamic Programming*, Princeton Univ. Press, Princeton, NJ (1957).

[10] H. Berliocchi and J.M. Lasry, *Intégrandes normales et mesures paramétrées en calcul des variations*, Bull. Soc. Math. France **101** (1973), 129–184.

[11] V.G. Boltyanskii, *Sufficient optimality conditions* (in Russian), Dokl. Akad. Nauk SSSR **140** (1961), 994–997.

[12] V.G. Boltyanskii, *Sufficient conditions for optimality and the justification of the programming method* (in Russian), Izv. Akad. Nauk SSSR Ser. Mat. **28** (1964), 481–514.

[13] J.M. Bony, *Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés*, Ann. Inst. Fourier (Grenoble) **19** (1969), 277–304.

[14] D. Bothe, *Flow invariance for perturbed nonlinear evolution equations*, Abstr. Appl. Anal. **1** (1996), 417–433.

[15] D. Bothe, *Reaction diffusion systems with discontinuities. A viability approach*, Nonlinear Anal. **30** (1997), 677–686.

[16] H. Bouligand, *Sur les surfaces dépourvues de points hyperlimités*, Ann. Soc. Polon. Math. **9** (1930), 32–41.

[17] A. Bressan and V. Staicu, *On nonconvex perturbations of maximal monotone differential inclusions*, Set-Valued Anal. **2** (1994), 415–437.

[18] H. Brezis, *On a characterization of flow-invariant sets*, Comm. Pure Appl. Math. **23** (1970), 261–263.

[19] H. Brezis and F. Browder, *A general principle on ordered sets in nonlinear functional analysis*, Adv. Math. **21** (1976), 355–364.

[20] O. Cârjă, *Unele metode de analiză funcţională neliniară*, MATRIXROM, Bucureşti (2003).

[21] O. Cârjă and M.D.P. Monteiro Marques, *Viability for nonautonomous semilinear differential equations*, J. Differential Equations **166** (2000), 328–346.

[22] O. Cârjă and M.D.P. Monteiro Marques, *Viability results for nonautonomous differential inclusions*, J. Convex Anal. **7** (2000), 437–443.

[23] O. Cârjă and M.D.P. Monteiro Marques, *Weak tangency, weak invariance and Carathédory mappings*, J. Dynam. Control Systems **8** (2002), 445–461.

[24] O. Cârjă, M. Necula and I.I. Vrabie, *Local invariance via comparison functions*, Electron. J. Differential Equations **2004**(50) (2004), 1–14.

[25] O. Cârjă and C. Ursescu, *The characteristics method for a first order partial differential equation*, An. Ştiin. Univ. Al. I. Cuza Iaşi, Secţ. I a Mat. **39** (1993), 367–396.

[26] O. Cârjă and C. Ursescu, *Viscosity solutions and partial differential inequalities*, Evolution Equations, Control Theory, and Biomathematics, Lecture Notes in Pure and Appl. Math., Vol. 155, Marcel Dekker, New York (1994), 39–44.

[27] O. Cârjă and I.I. Vrabie, *Some new viability results for semilinear differential inclusions*, NoDEA Nonlinear Differential Equations Appl. **4** (1997), 401–424.

[28] O. Cârjă and I.I. Vrabie, *Viability results for nonlinear perturbed differential inclusions*, Panamer. Math. J. **9** (1999), 63–74.

[29] C. Carathéodory, *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, Teubner, Leipzig–Berlin (1935).

[30] L. Cesari, *Optimization—Theory and Applications*, Springer, New York (1983).

[31] I. Chiş-Şter, *Existence of monotone solutions for semilinear differential inclusions*, NoDEA Nonlinear Differential Equations Appl. **6** (1999), 63–78.

[32] F.H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.

[33] F.H. Clarke, Yu.S. Ledyaev and R.J. Stern, *Invariance, monotonicity and applications*, Nonlinear Analysis, Differential Equations and Control, F.H. Clarke and R.J. Stern, eds., Kluwer Academic Publishers, Dordrecht (1999), 207–305.

[34] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern and P.R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer-Verlag, Berlin (1998).

[35] F.H. Clarke and R.B. Vinter, *Local optimality conditions and lipschitzean solutions to the Hamilton–Jacobi equation*, SIAM J. Control Optim. **21** (1983), 856–870.

[36] B. Cornet, *Regularity properties of open tangent cones*, Nonlinear Analysis and Optimization, Louvain-la-Neuve, 1983, Math. Programming Stud. **30** (1987), 17–33.

[37] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik*, Zweiter Band, Julius Springer, Berlin (1937).

[38] M.G. Crandall, *A generalization of Peano's existence theorem and flow-invariance*, Proc. Amer. Math. Soc. **36** (1972), 151–155.

[39] M.G. Crandall and P.L. Lions, *Viscosity solutions of Hamilton–Jacobi equations*, Trans. Amer. Math. Soc. **277** (1983), 1–42.

[40] N. Dunford and J.T. Schwartz, *Linear Operators*, Part I: *General Theory*, Interscience Publishers, New York (1958).

[41] H. Federer, *Curvature measures*, Trans. Amer. Math. Soc. **93** (1959), 418–491.

[42] S. Feferman, *Independence of the axiom of choice from the axiom of independent choices*, J. Symbolic Logic **29** (1967), 226–226.

[43] A.F. Filippov, *Classical solutions of differential equations with multi-valued right-hand side*, SIAM J. Control **5** (1967), 609–621.

[44] H. Frankowska, *Equations d'Hamilton–Jacobi contingentes*, C. R. Acad. Sci. Paris Sér. I Math. **304** (1987), 295–298.

[45] H. Frankowska, S. Plaskacz and T. Rzeżuchowski, *Measurable viability theorems and Hamilton–Jacobi–Bellman equations*, J. Differential Equations **116** (1995), 265–305.

[46] A. Fryszkowski, *Continuous selections for a class of non-convex multivalued maps*, Studia Math. **76** (1983), 163–174.

[47] M. Fukuhara[6], *Sur les systèmes des équations différentielles ordinaires II*, Japan. J. Math. **6** (1929), 269–299.

[48] S. Gautier, *Equations differentielles multivoques sur un fermé*, Publications de l'Université de Pau (1973).

[49] I.V. Girsanov, *Lekcii po Matematičeskoi Teorii Èkstremal'nyh Zadač*, Moscow University, Moscow (1970), (English translation: *Lecture on Mathematical Theory of Extremum Problems*, Lectures Notes in Econom. and Math. Systems, Vol. 67, Springer-Verlag, New York (1972)).

[50] R. Gonzales, *Sur l'existence d'une solution maximale de l'équation de Hamilton–Jacobi*, C. R. Acad. Sci. Paris Sér. I Math. **282** (1976), 1287–1290.

[51] E. Goursat, *Lecons sur l'Integration des Équation Partielles du Premier Ordre*, Hermann, Paris (1891).

[52] G. Haddad, *Monotone trajectories of differential inclusions and functional differential inclusions with memory*, Israel J. Math. **39** (1981), 83–100.

[53] O. Hájek, *On differentiability of the minimal time function*, Funkcial. Ekvac. **20** (1977), 97–114.

[54] P. Hartman, *Ordinary Differential Equations*, Wiley, New York (1964).

[55] P. Hartman, *On invariant sets and on a theorem of Ważewski*, Proc. Amer. Math. Soc. **32** (1972), 511–520.

[56] E.W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, Vol. 1, third edition, Cambridge Univ. Press, Cambridge (1927).

[57] P. Howard and J.E. Rubin, *The Boolean prime ideal theorem plus countable choice do [does] not imply dependent choice*, Math. Logic Quart. **42** (1996), 410–420.

[58] H. Ishii, *A simple, direct proof of uniqueness for solutions of the Hamilton–Jacobi equations of eikonal type*, Proc. Amer. Math. Soc. **100** (1987), 247–251.

---

[6] The old English transliteration of the Japanese name Hukuhara.

[59] E. Kamke, *Differentialgleichungen reeler Funktionen*, Akademische Verlagsgesellschaft, Leipzig (1930).

[60] N. Kenmochi and T. Takahashi, *Nonautonomous differential equations in Banach spaces*, Nonlinear Anal. **4** (1980), 1109–1121.

[61] H. Kneser, Siz.-Ber. d. Preuss Akademie d. Wissensch., Phys.-Math. Klasse (1923), 171–174.

[62] J. Kobayashi, *Difference approximation of Cauchy problems for quasi-dissipative operators and generation of nonlinear semigroups*, J. Math. Soc. Japan **27** (1975), 640–665.

[63] A. Kucia, *Scorza Dragoni type theorems*, Fund. Math. **138** (1991), 197–203.

[64] Yu. Ledyaev, *Criteria for viability of trajectories of nonautonomous inclusions and their applications*, J. Math. Anal. Appl. **182** (1994), 165–188.

[65] V. Lupulescu and M. Necula, *Viability and local invariance for non-convex semilinear differential inclusions*, Nonlinear Funct. Anal. Appl. **9** (2004), 495–512.

[66] R.H. Martin Jr., *Differential equations on closed subsets of a Banach space*, Trans. Amer. Math. Soc. **179** (1973), 399–414.

[67] E. Michael, *Continuous selectors I*, Ann. of Math. **63** (1956), 361–382.

[68] E. Michael, *Continuous selectors II*, Ann. of Math. **64** (1956), 562–580.

[69] B.S. Mordukhovich, *Maximum principle in the problem of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech. **40** (1976), 960–969.

[70] M. Nagumo, *Über die Lage der Integralkurven gewönlicher Differentialgleichungen*, Proc. Phys. Math. Soc. Japan **24** (1942), 551–559.

[71] M. Necula, *Viability of variable domains for nonautonomous semilinear differential equations*, Panamer. Math. J. **11** (2001), 65–80.

[72] M. Necula, *Viability of variable domains for differential equations governed by Carathéodory perturbations of nonlinear m-accretive operators*, An. Ştiinţ. Univ. Al. I. Cuza Iaşi, Secţ. I a Mat. **48** (2002), 41–60.

[73] M. Necula, Personal communication.

[74] M. Necula and I.I. Vrabie, Paper in preparation.

[75] N.H. Pavel, *Invariant sets for a class of semi-linear equations of evolution*, Nonlinear Anal. **1** (1977), 187–196.

[76] N.H. Pavel, *Differential Equations, Flow Invariance and Applications*, Res. Notes Math., Vol. 113, Pitman, Boston (1984).

[77] N.H. Pavel and C. Ursescu, *Flow-invariant sets for second order differential equations and applications in Mechanics*, Nonlinear Anal. **6** (1982), 35–74.

[78] N.H. Pavel and I.I. Vrabie, *Equations d'évolution multivoques dans des espaces de Banach*, C. R. Acad. Sci. Paris Sér. I Math. **287** (1978), 315–317.

[79] N.H. Pavel and I.I. Vrabie, *Semilinear evolution equations with multivalued right-hand side in Banach spaces*, An. Ştiinţ. Univ. Al. I. Cuza Iaşi Secţ. I a Mat. **25** (1979), 137–157.

[80] G. Peano, *Démonstration de l'intégrabilité des équations différentielles ordinaires*, Math. Ann. **37** (1890), 182–228.

[81] O. Perron, *Partieller Differentialgleichungssysteme im reellen Gebiet*, Math. Z. **27** (1928), 549–564.

[82] S. Plaskacz, *On the solution set for differential inclusions*, Boll. Unione Mat. Ital. Sez. A (7) **6** (1992), 387–394.

[83] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, GIF-ML, Moscow (1961) (in Russian).

[84] M. Quincampoix, *Differential inclusions and target problems*, SIAM J. Control Optim. **30** (1992), 324–335.

[85] R. Redheffer, *The theorem of Bony and Brezis on flow-invariant sets*, Amer. Math. Monthly **79** (1972), 740–747.

[86] E. Roxin, *Stability in general control systems*, J. Differential Equations **1** (1965), 115–150.

[87] G. Scorza Dragoni, *Un teorema sulle funzioni continue rispetto ad una e misurabili rispetto ad un'altra variabile*, Rend. Sem. Mat. Univ. Padova **17** (1948), 102–106.

[88] G. Scorza Dragoni, *Una applicazione della quasicontinuità semiregolare delle funzioni misurabili rispetto ad una e continue rispetto ad un'altra variabile*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (8) **12** (1952), 55–61.

[89] F. Severi, *Su alcune questioni di topologia infinitesimale*, Ann. Polon. Soc. Math. **9** (1930), 97–108.

[90] F. Severi, *Sulla differenziabilità totale delle funzioni di più variabili reali*, Ann. Mat. Pura Appl. **13** (1936), 1–35.

[91] S.Z. Shi, *Viability theorems for a class of differential operator inclusions*, J. Differential Equations **79** (1989), 232–257.

[92] P. Tallos, *Viability problems for nonautonomous differential inclusions*, SIAM J. Control Optim. **29** (1991), 253–263.

[93] J.S. Treiman, *Characterization of Clarke's tangent and normal cones in finite and infinite dimensions*, Nonlinear Anal. T.M.A. **7** (1983), 771–783.

[94] C. Ursescu, *Sur une généralisation de la notion de différentiabilité*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (7) **54** (1973), 199–204.

[95] C. Ursescu, *Tangent sets and differentiable functions*, Math. Control Theory Proc. Conf. Zakopane, 1974, Banach Center Publications (1976), 151–155.

[96] C. Ursescu, *Probleme de tangenţă cu aplicaţii în teoria optimizării*, Internal Report, Grant 11996/25.09.1980, Iaşi (1981).

[97] C. Ursescu, *Tangent sets' calculus and necessary conditions for extremality*, SIAM J. Control Optim. **20** (1982), 563–574.

[98] C. Ursescu, *Carathéodory solutions of ordinary differential equations on locally closed sets in finite dimensional spaces*, Math. Japon. **31** (1986), 483–491.

[99] C. Ursescu, C_0*-semigroups, characteristics method, and generic differential equations*, An. Univ. Timişoara Ser. Mat.-Inform. **34** (1996), 283–290.

[100] C. Ursescu, *Evolutors and invariance*, Set-Valued Anal. **11** (2003), 69–89.

[101] C. Ursescu, Personal communication.

[102] I.I. Vrabie, *Compactness Methods for Nonlinear Evolutions*, second edition, Pitman Monographs Surveys Pure Appl. Math., Vol. 75, Longman, Harlow (1995).

[103] I.I. Vrabie, *Compactness methods and flow-invariance for perturbed nonlinear semigroups*, An. Ştiinţ. Univ. Al. I. Cuza Iaşi Secţ. I a Mat. **27** (1981), 117–125.

[104] I.I. Vrabie, *Periodic solutions for nonlinear evolution equations in a Banach space*, Proc. Amer. Math. Soc. **109** (1990), 653–661.

[105] I.I. Vrabie, C_0*-Semigroups and Applications*, North-Holland Math. Stud., Vol. 191 (2003).

[106] I.I. Vrabie, *Differential Equations. An Introduction to Basic Results, Concepts and Applications*, World Scientific, New Jersey (2004).

[107] T. Ważewski, *Sur un condition équivalent à l'équation au contingent*, Bull. Polon. Acad. Sci. Ser. Sci. Math. Astr. Phys. **9** (1961), 865–867.

[108] J.A. Yorke, *Invariance for ordinary differential equations*, Math. Systems Theory **1** (1967), 353–372.

[109] J.A. Yorke, *Invariance for ordinary differential equations: Correction*, Math. Systems Theory **2** (1968), 381.

[110] T. Yoshizawa, *Stability theory by Liapunov's second method*, The Mathematical Society of Japan (1966).

[111] S.C. Zaremba, *Sur les équations au paratingent*, Bull. Sci. Math. (2) **60** (1936), 139–160.

CHAPTER 4

Monotone Dynamical Systems

M.W. Hirsch*

Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA

Hal Smith†

Department of Mathematics, Arizona State University, Tempe, AZ 85287-1804, USA

Contents

0. Introduction 241
1. Strongly order-preserving semiflows 243
    1.1. Definitions and basic results 243
    1.2. Nonordering of omega limit sets 249
    1.3. Local semiflows 251
    1.4. The limit set dichotomy 252
    1.5. Q is plentiful 255
    1.6. Stability in normally ordered spaces 258
    1.7. Stable equilibria in strongly ordered Banach spaces 261
    1.8. The search for stable equilibria 263
2. Generic convergence and stability 265
    2.1. The sequential limit set trichotomy 265
    2.2. Generic quasiconvergence and stability 271
    2.3. Improving the limit set dichotomy for smooth systems 273
    2.4. Generic convergence and stability 279
3. Ordinary differential equations 281
    3.1. The quasimonotone condition 282
    3.2. Strong monotonicity with linear systems 286
    3.3. Autonomous K-competitive and K-cooperative systems 289
    3.4. Dynamics of cooperative and competitive systems 291
    3.5. Smale's construction 294
    3.6. Invariant surfaces and the carrying simplex 295
    3.7. Systems in $\mathbb{R}^2$ 296
    3.8. Systems in $\mathbb{R}^3$ 297

---

[*] Supported by NSF Grant DMS 9700910.

[†] Supported by NSF Grant DMS 0107160.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 2
Edited by A. Cañada, P. Drábek and A. Fonda
© 2005 Elsevier B.V. All rights reserved

4. Delay differential equations . . . 302
  4.1. The semiflow . . . 302
  4.2. The quasimonotone condition . . . 304
  4.3. Eventual strong monotonicity . . . 308
  4.4. K is an orthant . . . 310
  4.5. Generic convergence for delay differential equations . . . 312
5. Monotone maps . . . 313
  5.1. Background and motivating examples . . . 313
  5.2. Definitions and basic results . . . 316
  5.3. The order interval trichotomy . . . 320
  5.4. Sublinearity and the cone limit set trichotomy . . . 323
  5.5. Smooth strongly monotone maps . . . 327
  5.6. Monotone planar maps . . . 329
6. Semilinear parabolic equations . . . 332
  6.1. Solution processes for abstract ODEs . . . 332
  6.2. Semilinear parabolic equations . . . 338
  6.3. Parabolic systems with monotone dynamics . . . 347
References . . . 348

0. Introduction

This chapter surveys a restricted but useful class of dynamical systems, namely, those enjoying a comparison principle with respect to a closed order relation on the state space. Such systems, variously called monotone, order-preserving or increasing, occur in many biological, chemical, physical and economic models.

The following notation will be used. $\mathbb{Z}$ denotes the set of integers; $\mathbb{N} = \{0, 1, \ldots\}$, the set of natural numbers; $\mathbb{N}_+$ is the set of positive integers, and $\mathbb{R}$ is the set of real numbers. For $u, v \in \mathbb{R}^n$ (= Euclidean n-space), we write

$$\begin{aligned} u \leqslant v \quad &\Longleftrightarrow \quad u_i \leqslant v_i, \\ u < v \quad &\Longleftrightarrow \quad u_i \leqslant v_i, \quad u \neq v, \\ u \ll v \quad &\Longleftrightarrow \quad u_i < v_i, \end{aligned}$$

where $i = 1, \ldots, n$. This relation $\leqslant$ is called the *vector order* in $\mathbb{R}^n$.

The prototypical example of monotone dynamics is a Kolmogorov model of cooperating species,

$$\dot{x}_i = x_i G_i(x), \quad x_i \geqslant 0, \quad i = 1, \ldots, n \tag{0.1}$$

in the positive orthant $\mathbb{R}^n_+ = [0, \infty)^n$, where $G : \mathbb{R}^n_+ \to \mathbb{R}^n$ is continuously differentiable. x_i denotes the population and G_i the *per capita* growth rate of species i. Cooperation means that an increase in any population causes an increase of the growth rates of all the other populations, modeled by the assumption that $\partial G_i / \partial x_j \geqslant 0$ for $i \neq j$. The right-hand side $F_i = x_i G_i$ of (0.1) then defines a *cooperative vector field* $F : \mathbb{R}^n \to \mathbb{R}^n$, meaning that $\partial F_i / \partial x_j \geqslant 0$ for $i \neq j$.

Assume for simplicity that solutions to Eq. (0.1) are defined for all $t \geqslant 0$. Let $\Phi = \{\Phi_t : \mathbb{R}^n_+ \to \mathbb{R}^n_+\}_{t \geqslant 0}$ denote the resulting semiflow in $\mathbb{R}^n_+$ that describes the evolution of states in positive time: the solution with initial value u is given by $x(t) = \Phi_t(u)$. The key to the long-term dynamics of cooperative vector fields is an important differential inequality due to Müller [148] and Kamke [91].

$$u \leqslant v \text{ and } t \geqslant 0 \quad \Longrightarrow \quad \Phi_t(u) \leqslant \Phi_t(v).$$

In other words: *The maps Φ_t preserve the vector order*. A semiflow Φ with this property is called *monotone*. Monotone semiflows and their discrete-time counterparts, order-preserving maps, form the subject of Monotone Dynamics.

Returning to the biological setting, we may make the assumption that each species directly or indirectly affect all the others. This is modeled by the condition that the Jacobian matrices $G'(x)$ are irreducible. An extension of the Müller–Kamke theorem shows that in the *open orthant* $\operatorname{Int} \mathbb{R}^n$, the restriction of Φ is *strongly monotone*: If $u, v \in \operatorname{Int} \mathbb{R}^n$, then

$$u < v \text{ and } t > 0 \quad \Longrightarrow \quad \Phi_t(u) \ll \Phi_t(v).$$

A semiflow with this property is *strongly monotone*.

Similar order-preserving properties are found in other dynamical settings, including delay differential equations and quasilinear parabolic partial differential equations. Typically the state space is a subset of a (real) Banach space Y with a distinguished closed cone $Y_+ \subset Y$. An order relation is introduced by $x \geqslant y \Leftrightarrow x - y \in Y_+$. When Y is a space of real valued functions on some domain, Y_+ is usually (but not always) the cone of functions with values in $\mathbb{R}_+ := [0, \infty)$. When $Y = \mathbb{R}^n$, the cooperative systems defined above use the cone $\mathbb{R}^n_+$.

Equations (0.1) model an ecology of competing species if $\partial G_i/\partial x_j \leqslant 0$ for $i \neq j$. The resulting vector field K with components $K_i = x_i G_i$ is not generally cooperative, but its negative $F = -K$ *is* cooperative. Many dynamical properties of the semiflow of K can be deduced from that of F, which is monotone.

We will see that the long-term behavior of monotone systems is severely limited. Typical conclusions, valid under mild restrictions, include the following:

- If all forward trajectories are bounded, the forward trajectory of almost every initial state converges to an equilibrium.
- There are no attracting periodic orbits other than equilibria, because every attractor contains a stable equilibrium.
- In $\mathbb{R}^3$, every compact limit set that contains no equilibrium is a periodic orbit that bounds an invariant disk containing an equilibrium.
- In $\mathbb{R}^2$, each component of any solution is eventually increasing or decreasing.

Other cones in $\mathbb{R}^n$ are also used, especially the orthants defined by restricting the sign of each coordinate. For example, a system of two competing species can be modeled by ODEs

$$\dot{y}_i = y_i H_i(y); \quad y_i \geqslant 0, \quad i = 1, 2$$

with $\partial H_i/\partial y_j < 0$ for $i \neq j$. The coordinate change $x_1 = y_1$, $x_2 = -y_2$ converts this into a cooperative system in the second orthant K defined by $x_1 \geqslant 0 \geqslant x_2$. This system is thus both competitive and cooperative, albeit for different cones. Not surprisingly, the dynamics are very simple.

In view of such powerful properties of cooperative vector fields, it would be useful to know when a given field F in an open set $D \subset \mathbb{R}^n$ can be made cooperative or competitive by changing coordinates. The following sufficient condition appears to be due to DeAngelis et al. [39]; see also Smith [193], Hirsch [74]. Assume the Jacobian matrices $[a_{ij}(x)] = F'(x)$ have the following two properties:

(1) (Sign stability) If $i \neq j$ then a_{ij} does not change sign in D;

(2) (Sign symmetry) $a_{ij}a_{ji} \geqslant 0$ in D.

Let Γ be the combinatorial labeled graph with nodes $1, \ldots, n$ and an edge e_{ij} joining i and j labeled $\sigma_{ij} \in \{+, -\}$ if and only if $i \neq j$ and there exists $p \in D$ such that $\operatorname{sgn} a_{ij}(p) = \sigma_{ij} \neq 0$. Then F is cooperative (respectively, competitive) relative to some orthant if and only if in every closed loop in Γ the number of negative labels is even (respectively, odd).

Order-preserving dynamics also occur in discrete time systems. Consider a nonautonomous Kolmogorov system $\dot{x}_i = x_i H_i(t, x)$, where the map $H := (H_1, \ldots, H_n) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ has period $\tau > 0$ in t. Denote by $T : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ the Poincaré map, which to $x \in \mathbb{R}^n_+$ assigns $y(\tau)$ where $y(t)$ denotes the solution with initial value x. Then T is monotone provided the $\partial H_i/\partial x_j \geqslant 0$ for $i \neq j$, and strongly monotone in the open orthant when these matrices are also irreducible. Most of the results stated above have analogs for T.

Convergence and stability properties of several kinds of order-preserving semiflows are developed in Sections 1 and 2, in the setting of general ordered metric spaces. Section 3 treats ODEs whose flows preserve the order defined by a cone in $\mathbb{R}^n$. Delay differential equations are studied in Section 4. In Section 5 we present results on order-preserving maps. The final section applies the preceding results to second order quasilinear parabolic equations.

1. Strongly order-preserving semiflows

This section introduces the basic definitions and develops the main tools of monotone dynamics. Several results on density of quasiconvergent points are proved, and used to establish existence of stable equilibria.

1.1. *Definitions and basic results*

The setting is a semiflow $\Phi = \{\Phi_t\}_{0 \leqslant t < \infty}$ in a (partially) ordered metric space that preserves the weak order relation: $x \leqslant y$ implies $\Phi_t(x) \leqslant \Phi_t(y)$. Such semiflows, called *monotone*, have severely restricted dynamics; for example, in $\mathbb{R}^n$ with the vector ordering there cannot be stable periodic orbits other than equilibria. But for generic convergence theorems we need semiflows with the stronger property of being "strongly order preserving," together with mild compactness assumptions. In later sections we will see that these conditions are frequently encountered in applications. The centerpiece of this section is the Limit Set Dichotomy, a fundamental tool for the later theory.

1.1.1. *Ordered spaces* Let Z be a metric space and $A, B \subset Z$ subsets. The closure of A is denoted by $\overline{A}$ and its interior by Int A. The distance from A to B is defined as $\operatorname{dist}(A, B) := \inf_{a \in A, b \in B} d(a, b)$. When B is a singleton $\{b\}$ we may write this as $\operatorname{dist}(A, b) = \operatorname{dist}(b, A)$.

X always denotes an *ordered space*. This means X is endowed with a metric d and an order relation $\mathsf{R} \subset X \times X$. As usual we write $x \leqslant y$ to mean $(x, y) \in \mathsf{R}$, and the order relation is:

(i) reflexive: $x \leqslant x$ for all $x \in X$,
(ii) transitive: $x \leqslant y$ and $y \leqslant z$ implies $x \leqslant z$,
(iii) antisymmetric: $x \leqslant y$ and $y \leqslant x$ implies $x = y$.

In addition, the ordering is compatible with the topology in the following sense:

(iv) if $x_n \to x$ and $y_n \to y$ as $n \to \infty$ and $x_n \leqslant y_n$, then $x \leqslant y$.

This is just to say that R is a closed subset of $X \times X$.

We write $x < y$ if $x \leqslant y$ and $x \neq y$. Given two subsets A and B of X, we write $A \leqslant B$ ($A < B$) when $x \leqslant y$ ($x < y$) holds for each choice of $x \in A$ and $y \in B$. The relation $A \leqslant B$ *does not imply* "$A < B$ or $A = B$"!

The notation $x \ll y$ means that there are open neighborhoods U, V of x, y respectively such that $U \leqslant V$. Equivalently, (x, y) belongs to the interior of R. The relation $\ll$, sometimes referred to as the *strong ordering*, is transitive; in many cases it is empty. We write $x \geqslant y$ to mean $y \leqslant x$, and similarly for $>$ and $\gg$.

We call X an *ordered subspace* of an ordered space X' if $X \subset X'$, and the order and topology on X are inherited from X'. When this is this case, the relation $u < v$ for points $u, v \in X$ means the same thing whether u and v are considered as points of X, or points of X'. But there are simple examples for which $u \ll v$ is true in X', yet false in X.

Let X be an ordered space. The *lower boundary* of a set $U \subset X$ is the set of points x in the boundary of U such that every neighborhood of x contains a point $y \in U$ with $y > x$. The *upper boundary* of U is defined dually.

Two points $x, y \in X$ are *order related* if $x < y$ or $y < x$; otherwise they are *unrelated*. A subset of X is *unordered* if it does not contain order related points. The empty set and singletons are unordered.

The (closed) *order interval* determined by $u, v \in X$ is the closed set

$$[u, v] = [u, v]_X := \{x \in X\colon\ u \leqslant x \leqslant v\}$$

which may be empty. The *open order interval* is the open set

$$[[u, v]] = \{x \in X\colon\ u \ll x \ll v\}.$$

A subset of X is *order bounded* if it lies in an order interval, and *order convex* if it contains $[u, v]$ whenever it contains u and v.

A point $x \in X$ is *accessible from below* if there is a sequence $x_n \to x$ with $x_n < x$; such a sequence is said to approximate x from below. We define *accessible from above* dually, that is, by replacing $<$ with $>$. In most applications there is a dense open subset of points that are accessible from both above and below.

The *supremum* $\sup S$ of a subset $S \subset X$, if it exists, is the unique point a such that $a \geqslant S$ and $x \geqslant S \Rightarrow x \geqslant a$. The *infimum* $\inf S$ is defined dually, i.e., substituting $\leqslant$ for $\geqslant$. A *maximal element* of S is a point $a \in S$ such that $x \in S$ and $x \geqslant a$ implies $x = a$. A *minimal element* is defined dually.

The following basic facts are well known:

LEMMA 1.1. *Assume the ordered space X is compact.*

(i) *Every sequence in X that is increasing or decreasing converges.*
(ii) *If X is totally ordered, it contains a supremum and an infimum.*
(iii) *X contains a maximal element and a minimal element.*

PROOF. (i) If p and q denote subsequential limits, then $p \leqslant q$ and $q \leqslant p$, hence $p = q$.

(ii) For each $x \in X$, the set $B_x := \{y \in X\colon\ y \geqslant x\}$ is compact, and every finite family of such sets has nonempty intersection because X is totally ordered. Therefore there exists $a \in \bigcap_x B_x$, and clearly $a = \sup X$. Similarly, $\inf X$ exists.

(iii) Apply (ii) to a maximal totally ordered subset (using Zorn's lemma). □

An *ordered Banach space* is an ordered space whose underlying metric space is a Banach space Y, and such that the set $Y_+ = \{y \in Y\colon\ y \geqslant 0\}$ is a cone, necessarily closed and convex. Thus Y_+ is a closed subset of Y with the properties:

$$\mathbb{R}_+ \cdot Y_+ \subset Y_+, \qquad Y_+ + Y_+ \subset Y_+, \qquad Y_+ \cap (-Y_+) = \{0\}.$$

We always assume $Y_+ \neq \{0\}$.

When $\operatorname{Int} Y_+$ is nonempty we call Y a *strongly ordered* Banach space. In this case $x \ll y \Leftrightarrow y - x \in \operatorname{Int} Y_+$.

The most important examples of ordered Banach spaces are completions of normed vector spaces of real-valued functions on some set Ω, with the positive cone corresponding to nonnegative functions. This cone defines the *functional ordering*. The simplest case is obtained from $\Omega = \{1, 2, \dots, n\}$: here $Y = \mathbb{R}^n$ and $Y_+ = \mathbb{R}^n_+$, the *standard cone* comprising vectors with all components nonnegative. For the corresponding *vector ordering*, $x \leqslant y$ means that $x_i \leqslant y_i$ for all i. Other function spaces are used in Sections 4 and 6.

When Y is an ordered Banach space, the notation $X \subset Y$ tacitly assumes that X is an ordered subspace of Y (but not necessarily a linear subspace).

A subset S of an ordered Banach space is *p-convex* if it contains the line segment spanned by u, v whenever $u, v \in S$ and $u < v$.

1.1.2. *Semiflows* All maps are assumed to be continuous unless the contrary is indicated.

A *semiflow* on X is a map $\Psi : \mathbb{R}_+ \times X \to X$, $(t, x) \mapsto \Psi_t(x)$ such that:

$$\Psi_0(x) = x, \qquad \Psi_t\bigl(\Psi_s(x)\bigr) = \Psi_{t+s}(x) \quad (t, s \geqslant 0,\ x \in X).$$

Thus Ψ can be viewed as a collection of maps $\{\Psi_t\}_{t \in \mathbb{R}_+}$ such that Ψ_0 is the identity map of X and $\Psi_t \circ \Psi_s = \Psi_{t+s}$, and such that $\Psi_t(x)$ is continuous in (t, x).

A *flow* in a space M is a continuous map $\Psi : \mathbb{R} \times M \to M$, written $\Psi(t, x) = \Psi_t(x)$, such that

$$\Psi_0(x) = x, \qquad \Psi_t\bigl(\Psi_s(x)\bigr) = \Psi_{t+s}(x) \quad (t, s \in \mathbb{R}, x \in X).$$

Restricting a flow to $\mathbb{R}_+ \times M$ gives a semiflow. A C^1 vector field F on a compact manifold M, tangent to the boundary, generates a *solution flow*, for which the trajectory of x is the solution $u(t)$ to the initial value problem $\mathrm{d}u/\mathrm{d}t = F(u)$, $u(0) = x$.

The *trajectory* of x is the map $[0, \infty) \to X$, $t \mapsto \Psi_t(x)$; the image of the trajectory is the *orbit* $O(x, \Psi)$, denoted by $O(x)$ when Ψ is understood. When $O(x) = \{x\}$ then x is an *equilibrium*. The set of equilibria is denoted by E.

x and its orbit are called *T-periodic* if $T > 0$ and $\Psi_T(x) = x$; such a T is a *period* of x. In this case $\Psi_{t+T}(x) = \Psi_t(x)$ for all $t \geqslant 0$, so $O(x) = \Psi([0, T] \times \{x\})$. A periodic point is *nontrivial* if it is not an equilibrium.

A set $A \subset X$ is *positively invariant* if $\Psi_t A \subset A$ for all $t \geqslant 0$. It is *invariant* if $\Psi_t A = A$ for all $t \geqslant 0$. Orbits are positively invariant and periodic orbits are invariant.

A set K is said to *attract* a set S if for every neighborhood U of K there exists $t_0 \geqslant 0$ such that $t > t_0 \Rightarrow \Psi_t(S) \subset U$; when $S = \{x\}$ we say K attracts x. An *attractor* is a nonempty invariant set L that attracts a neighborhood of itself. The union of all such neighborhoods is the *basin* of L. If the basin of an attractor L is all of X then L is a *global attractor*.

The *omega limit set* of $x \in X$ is

$$\omega(x) = \omega(x, \Psi) := \bigcap_{t \geqslant 0} \overline{\bigcup_{s \geqslant t} \Psi_s(x)}.$$

This set is closed and positively invariant. When $\overline{O(x)}$ is compact, $\omega(x)$ is nonempty, compact, invariant and connected and it attracts $\overline{O(x)}$ (see, e.g., Saperstone [175]).

A point $x \in X$ is *quasiconvergent* if $\omega(x) \subset E$; the set of quasiconvergent points is denoted by Q. We call x *convergent* when $\omega(x)$ is singleton $\{p\}$; in this case $\Phi_t(x) \to p \in E$. We sometimes signal this by the abuse of notation $\omega(x) \in E$. The set of convergent points is denoted by C.

When all orbit closures are compact and E is totally disconnected (e.g., countable), then $Q = C$; because in this case every omega limit set, being a connected subset of E, is a singleton. For systems of ordinary differential equations generated by smooth vector fields, the Kupka–Smale theorem gives generic conditions implying that E is discrete (see Peixoto [157]); but in concrete cases it is often difficult to verify these conditions.

1.1.3. *Monotone semiflows* A map $f : X_1 \to X_2$ between ordered spaces is *monotone* if

$$x \leqslant y \quad \Longrightarrow \quad f(x) \leqslant f(y),$$

strictly monotone if

$$x < y \quad \Longrightarrow \quad f(x) < f(y),$$

and *strongly monotone* if

$$x < y \quad \Longrightarrow \quad f(x) \ll f(y).$$

Let Φ denote a semiflow in the ordered space X. We call Φ monotone or strictly monotone according as each map Φ_t has the corresponding property.

We call Φ *strongly order-preserving*, SOP for short, if it is monotone and whenever $x < y$ there exist open subsets U, V of x, y respectively, and $t_0 \geqslant 0$, such that

$$\Phi_{t_0}(U) \leqslant \Phi_{t_0}(V).$$

Monotonicity of Φ then implies that $\Phi_t(U) \leqslant \Phi_t(V)$ for all $t \geqslant t_0$.

We call Φ *strongly monotone* if

$$x < y,\ 0 < t \quad \Longrightarrow \quad \Phi_t(x) \ll \Phi_t(y)$$

and *eventually strongly monotone* if it is monotone and whenever $x < y$ there exists $t_0 > 0$ such that

$$t \geqslant t_0 \quad \Longrightarrow \quad \Phi_t(x) \ll \Phi_t(y).$$

This property obviously holds when Φ is strongly monotone. We shall see in Section 6 that many parabolic equations generate SOP semiflows in function spaces that are not strongly ordered and therefore do not support strongly monotone semiflows.

Strong monotonicity was introduced in Hirsch [68,69], while SOP was proposed later by Matano [133,134] and modified slightly by Smith and Thieme [197,199]. We briefly explore the relation between these two concepts.

PROPOSITION 1.2. *If Φ is eventually strongly monotone, it is SOP. If X is an open subset of a Banach space Y ordered by a cone Y_+, Φ is SOP and the maps $\Phi_t : X \to X$ are open, then Φ is eventually strongly monotone. In particular, Φ is eventually strongly monotone provided Y is finite-dimensional, Φ is SOP and the maps Φ_t are injective.*

PROOF. If $x < y$ and Φ is eventually strongly monotone, then there exists $t_0 > 0$ such that $\Phi_{t_0}(x) \ll \Phi_{t_0}(y)$. Take neighborhoods $\widetilde{U}$ of $\Phi_{t_0}(x)$ and $\widetilde{V}$ of $\Phi_{t_0}(y)$ such that $\widetilde{U} < \widetilde{V}$. By continuity of Φ_{t_0}, there are neighborhoods U of x and V of y such that $\Phi_{t_0}(U) \subset \widetilde{U}$ and $\Phi_{t_0}(V) \subset \widetilde{V}$. Therefore, $\Phi_{t_0}(U) < \Phi_{t_0}(V)$ so Φ is SOP.

Suppose that $X \subset Y$ is open and ordered by Y_+ and Φ is SOP. If $x < y$ and U, V are open neighborhoods as in the definition of SOP, the inequality $\Phi_t(U) \leqslant \Phi_t(V)$ together with the fact that $\Phi_t(U)$ and $\Phi_t(V)$ are open in Y imply that $\Phi_t(x) \ll \Phi_t(y)$. ☐

The following very useful result shows that the defining property of SOP semiflows, concerning points $x < y$, extends to a similar property for compact sets $K < L$:

LEMMA 1.3. *Assume Φ is SOP and K, L are compact subsets of X satisfying $K < L$. Then there exists real numbers $t_1 \geqslant 0$, $\epsilon > 0$ and neighborhoods U, V of K, L respectively such that*

$$t \geqslant t_1 \quad \textit{and} \quad 0 \leqslant s \leqslant \epsilon \quad \Longrightarrow \quad \Phi_{t+s}(U) \leqslant \Phi_t(V).$$

PROOF. Let $x \in K$. For each $y \in L$ there exist $t_y \geqslant 0$, a neighborhood U_y of x, and a neighborhood V_y of y such that $\Phi_t(U_y) \leqslant \Phi_t(V_y)$ for $t \geqslant t_y$ since Φ is strongly order preserving. $\{V_y\}_{y\in L}$ is an open cover of L, so we may choose a finite subcover: $L \subset \bigcup_{i=1}^n V_{y_i} := \widetilde{V}$ where $y_i \in L$, $1 \leqslant i \leqslant n$. Let $\widetilde{U}_x = \bigcap_{i=1}^n U_{y_i}$, which is a neighborhood of x, and let $\tilde{t}_x = \max_{1\leqslant i\leqslant n} t_{y_i}$. Then $\Phi_{\tilde{t}}(\widetilde{U}_x) \subset \Phi_{\tilde{t}}(U_{y_i}) \leqslant \Phi_{\tilde{t}}(V_{y_i})$, so $\Phi_t(\widetilde{U}_x) \leqslant \Phi_t(V_{y_i})$ for $t \geqslant \tilde{t}_x$. It follows that

$$t \geqslant \tilde{t}_x \quad \Longrightarrow \quad \Phi_t\big(\widetilde{U}_x\big) \leqslant \Phi_t\big(\widetilde{V}\big).$$

Extract a finite subcover $\{\widetilde{U}_{x_j}\}$ of K from the family $\{\widetilde{U}_x\}$. Setting $U := \bigcup_j \widetilde{U}_{x_j} \supset K$ and $t_1 := \max_{1\leqslant j\leqslant m} \tilde{t}_{x_j}$, we have

$$t \geqslant t_1 \quad\Longrightarrow\quad \Phi_t(U) = \bigcup_j \Phi_t\big(\widetilde{U}_j\big) \leqslant \Phi_t(V).$$

In order to obtain the stronger conclusion of the lemma, note that for each $z \in K$ there exists $\epsilon_z > 0$ and a neighborhood U'_z of z such that $\Phi([0,\epsilon_z) \times W_z) \subset U$. Choose $z_1, \dots, z_m$ in K so that $K \subset \bigcup_j U'_{z_j}$. Define $U' = \bigcup_j U'_{z_j}$ and $\epsilon = \min_j\{\epsilon_{z_j}\}$. If $x \in U'$ and $0 \leqslant s < \epsilon$ then $x \in U'_{z_j}$ for some j so $\Phi_s(x) \in U$. Thus $\Phi([0,\epsilon) \times U') \subset U$ so $\Phi_s(U') \subset U$. It follows that $\Phi_{t+s}(U') \subset \Phi_t(U) \leqslant \Phi_t(V)$ for $t \geqslant t_1$, $0 \leqslant s < \epsilon$. □

Several fundamental results in the theory of monotone dynamical systems are based on the following sufficient conditions for a solution to converge to equilibrium.

THEOREM 1.4 (Convergence Criterion). *Assume Φ is monotone, $x \in X$ has compact orbit closure, and $T > 0$ is such that $\Phi_T(x) \geqslant x$. Then $\omega(x)$ is an orbit of period T. Moreover, x is convergent if the set of such T is open and nonempty or Φ is SOP and $\Phi_T(x) > x$.*

PROOF. Monotonicity implies that $\Phi_{(n+1)T}(x) \geqslant \Phi_{nT}(x)$ for $n = 1, 2, \dots$ and therefore $\Phi_{nT}(x) \to p$ as $n \to \infty$ for some p by the compactness of the orbit closure. By continuity,

$$\begin{aligned}\Phi_{t+T}(p) &= \Phi_{t+T}\Big(\lim_{n\to\infty} \Phi_{nT}(x)\Big)\\ &= \lim_{n\to\infty} \Phi_{(n+1)T+t}(x)\\ &= \lim_{n\to\infty} \Phi_t\big(\Phi_{(n+1)T}(x)\big)\\ &= \Phi_t(p)\end{aligned}$$

for all $t \geqslant 0$. Hence p is T-periodic.

To prove $\omega(x) = O(p)$, suppose $t_j \to \infty$ and $\Phi_{t_j}(x) \to q \in \omega(x)$ as $j \to \infty$, and write $t_j = n_j T + r_j$ where n_j is a natural number and $0 \leqslant r_j < T$. By passing to a subsequence if necessary, we may assume that $r_j \to r \in [0, T]$. Taking limits as $j \to \infty$ and noting that $n_j \to \infty$, we have by continuity:

$$\lim \Phi_{t_j}(x) = \lim \Phi_{r_j}\big(\lim \Phi_{n_j T}(x)\big) = \lim \Phi_{r_j}(p) = \Phi_r(p) = q.$$

Therefore $\omega(x) \subset O(p)$, and the opposite inclusion holds because $p \in \omega(x)$. This proves the first assertion of the theorem.

Suppose $\Phi_t(x) \geqslant x$ for all t in a nonempty open interval $(T - \epsilon, T + \epsilon)$. The first assertion shows that $\omega(x)$ is an orbit $O(p)$ of period τ for every $\tau \in (T - \epsilon, T + \epsilon)$. All elements of $O(p)$ have the same set G of periods; G is closed under addition and contains $(T - \epsilon, T + \epsilon)$. If $0 \leqslant s < \epsilon$ and $t \geqslant 0$ then

$$\Phi_{t+s}(p) = \Phi_t\big(\Phi_s(p)\big) = \Phi_t\big(\Phi_{s+T}(p)\big) = \Phi_t(p).$$

Hence $[0, \epsilon) \subset G$ and therefore $G = \mathbb{R}_+$, which implies $p \in E$. This proves the second assertion.

If $\Phi_T(x) > x$ and Φ is SOP then there exist neighborhoods U of x and V of $\Phi_T(x)$ and $t_0 > 0$ such that $\Phi_{t_0}(U) \leqslant \Phi_{t_0}(V)$. It follows that $\Phi_{t_0}(x) \leqslant \Phi_{t_0+T+\epsilon}(x)$ for all ϵ sufficiently small. The previous assertion implies $\omega(x) = p \in E$. □

1.2. *Nonordering of omega limit sets*

The next result is the first of several describing the order geometry of limit sets.

PROPOSITION 1.5 (Nonordering of Periodic Orbits). *A periodic orbit of a monotone semiflow is unordered.*

PROOF. Let x have minimal period $s > 0$ under a monotone semiflow Φ. Suppose $x \leqslant z \in O(x)$. By compactness of $O(x)$ there is a maximal $y \in O(x)$ such that $y \geqslant z \geqslant x$. By periodicity and monotonicity $y = \Phi_t(x) \leqslant \Phi_t(y)$, $t > 0$, hence $y = \Phi_t(y)$ by maximality. Therefore t is an integer multiple of s, so $x = \Phi_t(x) = y$, implying $x = z$. □

The following result, which implies (1.5), is a broad generalization of the obvious fact that for ODEs in $\mathbb{R}$, nonconstant solutions are everywhere increasing or everywhere decreasing. Let $J \subset \mathbb{R}$ be an interval and $f : J \to X$ a map. A compact subinterval $[a, b] \subset J$ is *rising* for f provided $f(a) < f(b)$, and *falling* if $f(b) < f(a)$.

THEOREM 1.6. *A trajectory of a monotone semiflow cannot have both a rising interval and a falling interval.*

This originated in Hirsch [67], with improvements in Smith [194], Smith and Waltman [203]. An analog for maps is given in Theorem 5.4.

PROOF. Let Φ be a monotone semiflow in X and fix a trajectory $f : [0, \infty) \to X$, $f(t) := \Phi_t(x)$. Call an interval $[d, d']$ *weakly falling* if $f(d) \geqslant f(d')$. Monotonicity shows that when this holds, the *right translates* of $[d, d']$—the intervals $[d + u, d' + u]$ with $u \geqslant 0$—are also weakly falling.

Proceeding by contradiction, we assume f has a falling interval $[a, a + r]$ and a rising interval $[c, c + q]$. To fix ideas we assume $a \leqslant c$, the case $c \leqslant a$ being similar. Define

$$b := \sup\bigl\{t \in [c, c + q]\colon\ f(t) \leqslant f(c),\ s := c + q - b\bigr\}.$$

Then $[b, b + s]$ is a rising interval in $[c, c + q]$, and

$$b < t \leqslant b + s \quad \Longrightarrow \quad f(t) \nleqslant f(b). \tag{1.1}$$

Claim 1: *No interval* $[b - l, b]$ *is weakly falling.* Assume the contrary. Then (i) $l > s$, and (ii) $[b - (l - s), b]$ is weakly falling. To see (i), observe that $f(b + l) \leqslant f(b)$ because

$[b, b+l]$ is a right translate of $[b-l, b]$; hence $l \leqslant s$ would entail $b < b+l \leqslant b+s$, contradicting (1.1) with $t = b+l$. To prove (ii), note that right translation of $[b-l, b]$ shows that $[b-l+s, b+s]$ is weakly falling, implying $f(b-(l-s)) \geqslant f(b+s) > f(b)$; hence $[b-(l-s), b]$ is falling. Repetition of this argument with l replaced by $l-s, l-2s, \ldots$ leads by induction on n to the absurdity that $l - ns > s$ for all $n \in \mathbb{N}$.

Claim 2: $r > s$. For $f(b+r) \leqslant f(b)$ because $[b, b+r]$ is falling, as it is a right translate of $[a, a+r]$. Therefore $r > s$, for otherwise $b < b+r \leqslant b+s$ and (1.1) leads to a contradiction.

As $b+s \geqslant a+r$, we can translate $[a, a+r]$ to the right by $(b+s)-(a+r)$, obtaining the weakly falling interval $[b+s-r, b+s]$. Note that $b+s-r < b$ by Claim 2. From $f(b+s-r) \geqslant f(b+s) > f(b)]$ we conclude that $[b-(r-s), b]$ is falling. But this contradicts Claim 1 with $l = r-s$. □

LEMMA 1.7. *An omega limit set for a monotone semiflow Φ cannot contain distinct points x, y having respective neighborhoods U, V such that $\Phi_r U \leqslant \Phi_r V$ for some $r \geqslant 0$.*

PROOF. We proceed by contradiction. Suppose there exist distinct points $x, y \in \omega(z)$ having respective neighborhoods U, V such that $\Phi_r U \leqslant \Phi_r V$ for some $r \geqslant 0$. Then $\omega(z)$ is not a periodic orbit, for otherwise from $\Phi_r(x) \leqslant \Phi_r(y)$ we infer $x \leqslant y$ and hence $x < y$, violating Nonordering of Periodic Orbits.

There exist real numbers $a < b < c$ be such that $\Phi_a(z) \in U$, $\Phi_b(z) \in V$, $\Phi_c(z) \in U$. Therefore the properties of r, U and V imply

$$\Phi_{a+r}(z) \leqslant \Phi_{b+r}(z), \qquad \Phi_{b+r}(z) \geqslant \Phi_{c+r}(z).$$

As $\omega(z)$ is not periodic, the semiflow is injective on the orbit of z; hence the order relations above are strict. But this contradicts Theorem 1.6. □

It seems to be unknown whether omega limit sets of monotone semiflows must be unordered. This holds for SOP semiflows by the following theorem due to Smith and Thieme [197, Proposition 2.2]; the strongly monotone case goes back to Hirsch [66]. This result is fundamental to the theory of monotone semiflows:

THEOREM 1.8 (Nonordering of Omega Limit Sets). *Let $\omega(z)$ be an omega limit set for a monotone semiflow Φ.*

(i) *No points of $\omega(z)$ are related by $\ll$.*

(ii) *If $\omega(z)$ is a periodic orbit or Φ is SOP, no points of $\omega(z)$ are related by $<$.*

PROOF. Assume $x, y \in \omega(z)$. If $\omega(z)$ is a periodic orbit then x, y are unrelated (Proposition 1.5). If $x \ll y$ or $x < y$ and Φ is SOP, there are respective neighborhoods U, V of x, y such that $\Phi_r(U) \leqslant \Phi_r(V)$ for some $r \geqslant 0$; but this violates Lemma 1.7. □

COROLLARY 1.9. *Assume Φ is SOP.*

(i) *If an omega limit set has a supremum or infimum, it reduces to a single equilibrium.*

(ii) *If the equilibrium set is totally ordered, every quasiconvergent point with compact orbit closure is convergent.*

PROOF. Part (i) follows from Theorem 1.8(ii), since the supremum or infimum, if it exists, belongs to the limit set. Part (ii) is a consequence of (i). □

1.3. *Local semiflows*

For simplicity we have assumed trajectories are defined for all $t \geqslant 0$, but there are occasions when we need the more general concept of a *local semiflow* in X. This means a map $\Psi : \Omega \to X$, with $\Omega \subset [0, \infty) \times X$ an open neighborhood of $\{0\} \times X$, such that the maps

$$\Psi_t : D_t \to X,\ x \mapsto \Psi(t, x) \quad (0 \leqslant t < \infty)$$

satisfy the following conditions: D_t is an open, possibly empty set in X, Ψ_0 is the identity map of X, and $\Psi_{s+t} = \Psi_s \circ \Psi_t$ in the sense that $D_{s+t} = D_t \cap \Psi_t^{-1}(D_s)$ and $\Psi_{s+t}(x) = \Psi_s(\Psi_t(x))$ for $x \in D_{s+t}$.

The trajectory of x is defined as the map

$$I_x \to X,\ t \mapsto \Psi_t(x), \quad \text{where } I_x = \{t \in \mathbb{R}_+ :\ x \in D_t\}.$$

The composition law implies I_x is a half open interval $[0, \tau_x)$; we call $\tau_x \in (0, \infty]$ the *escape time* of x. It is easy to see that every point with compact orbit closure has infinite escape time. Thus a local semiflow with compact orbit closures is a semiflow. In dealing with local semiflows we adopt the convention that the notations $\Psi_t(x)$ and $\Psi_t(U)$ carry the assumptions that $t \in I_x$ and $U \subset D_t$. The image of I_x under the trajectory of x is the orbit $O(x)$. The omega limit set $\omega(x)$ is defined as $\omega(x) = \bigcap_{t \in I_x} \overline{O(\Psi_t(x))}$.

A *local flow* is a map $\Theta : \Lambda \to X$ where $\Lambda \subset \mathbb{R} \times X$ is an open neighborhood of $\{0\} \times X$, and the (possibly empty) maps

$$\Theta_t : D_t \to X,\ x \mapsto \Theta(t, x) \quad (-\infty \leqslant t < \infty)$$

satisfy the following conditions: Θ_0 is the identity map of $D_0 := X$, Θ_t is a homeomorphism of D_t onto D_{-t} with inverse Θ_{-t}, and

$$x \in (\Theta_s)^{-1} D_r \quad \Longrightarrow \quad \Theta_r \circ \Theta_s(x) = \Theta_{r+s}(x).$$

Θ is a *flow* provided $D_t = X$ for all t.

The set $J_x := \{t \in \mathbb{R}:\ x \in D_t\}$ is an open interval around 0. The positive and negative *semiorbits* of x are the respective sets

$$\gamma^+(x) = \gamma^+(x, \Theta) := \bigl\{\Theta_t(x):\ t \in J_x, t \geqslant 0\bigr\},$$

$$\gamma^-(x) = \gamma^-(x, \Theta) := \bigl\{\Theta_t(x):\ t \in J_x, t \leqslant 0\bigr\}.$$

The *time-reversal* of Θ is the local flow $\widetilde{\Theta}$ defined by $\widetilde{\Theta}(t,x)=\Theta(-t,x)$.

The omega limit set $\omega(x)$ (for Θ) is defined to be $\omega(x)=\bigcap_{t\in I_x,t\geqslant 0}\overline{O(\Psi_t(x))}$. The *alpha limit set* $\alpha(x)=\alpha(x,\Theta)$ of x is defined as the omega limit set of x under the time-reversal of Θ^+.

Let F be a locally Lipschitz vector field F on a manifold M tangent along the boundary. Denote by $t\mapsto u(t;x)$ the maximally defined solution to $\dot{u}=F(u)$, $u(0,x)=x$. There is a local flow Θ^F on M such that $\Theta_t(x)=u(t;x)$. The time-reversal of Θ^F is Θ^{-F}. When M is compact, Θ^F is a flow. If we assume that F, rather than being tangent to the boundary, is transverse inward, we obtain a local semiflow.

Our earlier results are readily adapted to monotone local semiflows. In particular, omega limit sets are unordered. Theorems 1.8 and 1.6 have the following extension:

THEOREM 1.10. *Let Φ be a monotone local semiflow.*

(a) *No trajectory has both a rising and a falling interval.*
(b) *No points of an omega limit set are related by $\ll$, or by $<$ if Φ is SOP.*
(c) *The same holds for alpha limit sets provided Φ is a local flow.*

PROOF. The proofs of Theorems 1.6 and 1.8 also prove (a) and (b), and (c) follows by time reversal. □

1.4. *The limit set dichotomy*

Throughout the remainder of Section 1 we adopt the following assumptions:

(H) Φ is a strongly order preserving semiflow in an ordered space X, with every orbit closure compact.

Our goal now is to prove the important Limit Set Dichotomy:

If $x<y$ then either $\omega(x)<\omega(y)$, or $\omega(x)=\omega(y)\subset E$.

LEMMA 1.11 (Colimiting Principle). *Assume $x<y$, $t_k\to\infty$, $\Phi_{t_k}(x)\to p$ and $\Phi_{t_k}(y)\to p$ as $k\to\infty$. Then $p\in E$.*

PROOF. Choose neighborhoods U of x and V of y and $t_0>0$ such that $\Phi_{t_0}(U)\leqslant\Phi_{t_0}(V)$. Let $\delta>0$ be so small that $\{\Phi_s(x)\colon\ 0\leqslant s\leqslant\delta\}\subset U$ and $\{\Phi_s(y)\colon\ 0\leqslant s\leqslant\delta\}\subset V$. Then $\Phi_s(x)\leqslant\Phi_r(y)$ whenever $t_0\leqslant r,\ s\leqslant t_0+\delta$. Therefore,

$$\Phi_{t_k-t_0}\big(\Phi_s(x)\big)\leqslant\Phi_{t_k-t_0}\big(\Phi_{t_0}(y)\big)=\Phi_{t_k}(y)\tag{1.2}$$

for all $s\in[t_0,t_0+\delta]$ and all large k. As

$$\Phi_{t_k-t_0}\big(\Phi_s(x)\big)=\Phi_{s-t_0}\big(\Phi_{t_k}(x)\big)=\Phi_r\big(\Phi_{t_k}(x)\big),$$

where $r = s - t_0 \in [0, \delta]$ if $s \in [t_0, t_0 + \delta]$, we have

$$\Phi_r\big(\Phi_{t_k}(x)\big) \leqslant \Phi_{t_k}(y)$$

for large k and $r \in [0, \delta]$. Passing to the limit as $k \to \infty$ we find that $\Phi_r(p) \leqslant p$ for $0 \leqslant r \leqslant \delta$. If, in (1.2), we replace $\Phi_s(x)$ by $\Phi_{t_0}(x)$ and replace $\Phi_{t_0}(y)$ by $\Phi_s(y)$, and argue as above then we find that $p \leqslant \Phi_r(p)$ for $0 \leqslant r \leqslant \delta$. Evidently, $\Phi_r(p) = p$, $0 \leqslant r \leqslant \delta$ and therefore for all $r \geqslant 0$, so $p \in E$. $\square$

THEOREM 1.12 (Intersection Principle). *If $x < y$ then $\omega(x) \cap \omega(y) \subset E$. If $p \in \omega(x) \cap \omega(y)$ and $t_k \to \infty$, then $\Phi_{t_k}(x) \to p$ if and only if $\Phi_{t_k}(y) \to p$.*

PROOF. If $p \in \omega(x) \cap \omega(y)$ then there exists a sequence $t_k \to \infty$ such that $\Phi_{t_k}(x) \to p$ and $\Phi_{t_k}(y) \to q \in \omega(y)$, and $p \leqslant q$ by monotonicity. If $p < q$ then we contradict the Nonordering of Limit Sets since $p, q \in \omega(y)$. Hence $p = q$. The Colimiting Principle then implies $p \in E$. $\square$

The proof of the next result has been substantially simplified over previous versions.

LEMMA 1.13. *Assume $x < y$, $t_k \to \infty$, $\Phi_{t_k}(x) \to a$, and $\Phi_{t_k}(y) \to b$ as $k \to \infty$. If $a < b$ then $O(a) < b$ and $O(b) > a$.*

PROOF. The set $W := \{t \geqslant 0\colon \Phi_t(a) \leqslant b\}$ contains 0 and is closed. We prove $W = [0, \infty)$ by showing that W is also open. Observe first that if $t \in W$, then $\Phi_t(a) < b$. For equality implies $b \in \omega(x) \cap \omega(y) \subset E$, and then the Intersection Principle entails $\Phi_{t_k}(x) \to b$, giving the contradiction $a = b$.

Suppose $\bar{t} \in W$ is positive. By SOP there are open sets U, V with $\Phi_{\bar{t}}(a) \in U$, $b \in V$ and $t_1 \geqslant 0$ such that $\Phi_t(U) \leqslant \Phi_t(V)$ for $t \geqslant t_1$. There exists $\delta \in (0, \bar{t}/2)$ such that $\Phi_s(a) \in U$ for $|s - \bar{t}| \leqslant \delta$, so we can find an integer $\kappa > 0$ such that $\Phi_s(\Phi_{t_k}(x)) \in U$ for $k \geqslant \kappa$. Choose $k_0 \geqslant \kappa$ such that $\Phi_{t_{k_0}}(y) \in V$. Then we have $\Phi_{t+s+t_{k_0}}(x) \leqslant \Phi_{t+t_{k_0}}(y)$ for $t \geqslant t_1$. Setting $t = t_k - t_{k_0}$ for large k in this last inequality yields $\Phi_{t_k+s}(x) \leqslant \Phi_{t_k}(y)$ for large k. Taking the limit as $k \to \infty$ we get $\Phi_s(a) \leqslant b$ for $|s - \bar{t}| \leqslant \delta$. A similar argument in the case $\bar{t} = 0$ considering only $s \in [0, \delta]$ gives the previous inequality for such s. Therefore, W is both open and closed so $W = [0, \infty)$. This proves $O(a) < b$, and $O(b) > a$ is proved dually. $\square$

LEMMA 1.14 (Absorption Principle). *Let $u, v \in X$. If there exists $x \in \omega(u)$ such that $x < \omega(v)$, then $\omega(u) < \omega(v)$. Similarly, if there exists $x \in \omega(u)$ such that $\omega(v) < x$, then $\omega(v) < \omega(u)$.*

PROOF. Apply Lemma 1.3 to obtain open neighborhoods U of x and V of $\omega(v)$ and $t_0 > 0$ such that

$$r \geqslant t_0 \quad \Longrightarrow \quad \Phi_r(U) \leqslant \Phi_r(V),$$

hence $\Phi_r(U) \leqslant \omega(v)$ since $\omega(v)$ is invariant. As $x \in \omega(u)$, there exists $t_1 > 0$ such that $\Phi_{t_1}(u) \in U$. Hence for $\Phi_{t_0+t_1}(u) \leqslant \omega(v)$, and monotonicity implies that $\Phi_{s+t_0+t_1}(u) \leqslant$

$\omega(v)$ for all $s \geqslant 0$. This implies that $\omega(u) \leqslant \omega(v)$. If $z \in \omega(u) \cap \omega(v)$ then $z = \sup \omega(u) = \inf \omega(v)$, whence $\{z\} = \omega(u) = \omega(v)$ by Corollary 1.9(ii). But this is impossible since $x < \omega(v)$ and $x \in \omega(u)$, so we conclude that $\omega(u) < \omega(v)$. ☐

LEMMA 1.15 (Limit Set Separation Principle). *Assume $x < y$, $a < b$ and there is a sequence $t_k \to \infty$ such that $\Phi_{t_k}(x) \to a$, $\Phi_{t_k}(y) \to b$. Then $\omega(x) < \omega(y)$.*

PROOF. By Lemma 1.13, $O(a) < b$, and therefore $\omega(a) \leqslant b$. If $b \in \omega(a)$ then Corollary 1.9 implies that $\omega(a) = b \in E$. Applying the Absorption Principle with $u = x$, $v = a$, $x = a$, we have $a \in \omega(x)$, $a < \omega(a) = b$ which implies that $\omega(x) < \omega(a)$. This is impossible as $\omega(a) \subset \omega(x)$. Consequently, $\omega(a) < b$. By the Absorption Principle again (with $u = a$, $v = y$), we have $\omega(a) < \omega(y)$. Since $\omega(a) \subset \omega(x)$, the Absorption Principle gives $\omega(x) < \omega(y)$. ☐

We now prove the fundamental tool in the theory of monotone dynamics, stated for strongly monotone semiflows in Hirsch [66,68].

THEOREM 1.16 (Limit Set Dichotomy). *If $x < y$ then either*
(a) $\omega(x) < \omega(y)$, *or*
(b) $\omega(x) = \omega(y) \subset E$.
If case (b) *holds and $t_k \to \infty$ then $\Phi_{t_k}(x) \to p$ if and only if $\Phi_{t_k}(y) \to p$.*

PROOF. If $\omega(x) = \omega(y)$ then $\omega(x) \subset E$ by the Intersection Principle, Theorem 1.12, which also establishes the final assertion. If $\omega(x) \neq \omega(y)$ then we may assume that there exists $q \in \omega(y) \setminus \omega(x)$, the other case being similar. There exists $t_k \to \infty$ such that $\Phi_{t_k}(y) \to q$. By passing to a subsequence if necessary, we can assume that $\Phi_{t_k}(x) \to p \in \omega(x)$. Monotonicity implies $p \leqslant q$ and, in fact, $p < q$ since $q \notin \omega(x)$. By the Limit Set Separation Principle, $\omega(x) < \omega(y)$. ☐

Among the many consequences of the Convergence Criterion is that a monotone semiflow in a strongly ordered Banach space cannot have a periodic orbit γ that is *attracting*, meaning that γ attracts all points in some neighborhood of itself (Hadeler [55], Hirsch [69]). The following consequence of the Limit Set Dichotomy implies the same conclusion for periodic orbits of SOP semiflows:

THEOREM 1.17. *Let γ be a nontrivial periodic orbit, some point of which is accessible from above or below. Then γ is not attracting.*

The accessibility hypothesis is used to ensure that there are points near p that are order-related to p but different from p. Some such hypothesis is required, as otherwise we could simply take $X = \gamma$, and then γ is attracting!

PROOF. Suppose $\gamma \subset W$ attracts an open set W. By hypothesis there exists $p \in \gamma$ and $x \in W$ such that $x > p$ or $x < p$ and $\omega(x) = \gamma$. To fix ideas we assume $x > p$. Then

$p \in \omega(x)$, so the Limit Set Dichotomy implies $p \in E$. Hence the contradiction that γ contains an equilibrium. □

It turns out that the periodic orbits γ considered above are not only not attracting; they enjoy the strong form of instability expressed in the next theorem.

A set $K \subset X$ is *minimal* if it is nonempty, invariant, and every orbit it contains is dense in K.

THEOREM 1.18. *Let K be a compact minimal set that is not an equilibrium, some point of which is accessible from below or above. Then there exists $\delta > 0$ with the following property*: *Every neighborhood of K contains a point x comparable to some point of K, such that* $\operatorname{dist}(\Phi_t(x), K) > \delta$ *for all sufficiently large t.*

PROOF. We may assume there exists a sequence $\tilde{x}_n \to p \in K$ with $\tilde{x}_n > p$. Suppose there is no such δ. Then there exist a subsequence $\{x_n\}$ and points $y_n \in \omega(x_n)$ such that $y_n \to q \in K$. Minimality of K implies $\omega(p) = \omega(q) = K$. Since $x_n > p$, the Limit Set Dichotomy implies $\omega(x_n) \geqslant \omega(p)$; therefore $y_n \geqslant K$, so $q \geqslant K$. It follows that $q = \sup K$, and Corollary 1.9 implies the contradiction that K is a singleton. □

A stronger form of instability for periodic orbits is given in Theorem 2.6.

1.5. *Q is plentiful*

One of our main goals is to find conditions that make quasiconvergent points generic in various senses. The first such results are due to Hirsch [66,73]; the result below is an adaptation of Smith and Thieme [199, Theorem 3.5].

We continue to assume Φ is an SOP semiflow with compact orbit closures.

A *totally ordered arc* is the homeomorphic image of a nontrivial interval $I \subset \mathbb{R}$ under a map $f : I \to X$ satisfying $f(s) < f(t)$ whenever $s, t \in I$ and $s < t$.

THEOREM 1.19. *If $J \subset X$ is a totally ordered arc, $J \setminus Q$ is at most countable.*

Stronger conclusions are obtained in Theorems 2.8 and 2.24.

The following global convergence theorem is adapted from Hirsch [73, Theorem 10.3].

COROLLARY 1.20. *Let Y be an ordered Banach space. Assume $X \subset Y$ is an open set, a closed order interval, or a subcone of Y_+. If $E = \{p\}$, every trajectory converges to p.*

PROOF. If X is open in Y, there exists a totally ordered line segment $J \subset X$ and quasiconvergent points $u, v \in J$ with $u < x < v$, by Theorem 1.19. Therefore $\Phi_t(u) \to p$ and $\Phi_t(v) \to p$, so monotonicity and closedness of the order relation imply $\Phi_t(x) \to p$.

If $X = [a, b]$, the trajectories of a and b converge to p by the Convergence Criterion 1.4, and the previous argument shows all trajectories converge to p. Similarly if X is a subcone of Y_+. □

PROOF OF THEOREM 1.19. Let $W = \overline{\Phi([0,\infty) \times J)}$. Continuity of Φ implies that W is a separable metric space which is positively invariant under Φ. Therefore we may as well assume that X is a separable metric space.

We show that if $x \in J$ and

$$\inf\bigl\{\operatorname{dist}\bigl(\omega(x), \omega(y)\bigr)\colon\ y \in J, y \neq x\bigr\} = 0,$$

then $x \in Q$. Choose a sequence $x_n \in J$, $x_n \neq x$ such that $\operatorname{dist}(\omega(x), \omega(x_n)) \to 0$. We may assume that $x_n < x$ for all n. Taking a subsequence, we conclude from the Limit Set Dichotomy: Either some $\omega(x_n) = \omega(x)$, or every $\omega(x_n) < \omega(x)$.

In the first case, $x \in Q$. In the second case, choose $y_n \in \omega(x_n)$, $z_n \in \omega(x)$ such that $d(y_n, z_n) \to 0$. After passing to subsequences, we assume $y_n, z_n \to z \in \omega(x)$. Because $y_n \leqslant \omega(x)$, we conclude that $z \leqslant \omega(x)$. As $z \in \omega(x)$, Corollary 1.9 implies $\omega(x) = \{z\}$. Hence $x \in Q$ in this case as well.

It follows that for every $x \in J \setminus Q$, there exists an open set U_x containing $\omega(x)$ such that $U_x \cap \omega(y) = \emptyset$ for every $y \in J \setminus \{x\}$. By the axiom of choice we get an injective mapping

$$J \setminus Q \to X, \quad x \mapsto p_x \in \omega(x) \subset U_x.$$

The separable metric space X has a countable base $\mathcal{B}$. A second application of the axiom of choice gives a map

$$J \setminus Q \to \mathcal{B}, \quad x \mapsto V_x \subset U_x, \quad p_x \in V_x.$$

This map is injective. For if x, y are distinct points of $J \setminus Q$, then $V_x \neq V_y$ because V_x, being contained in U_x, does not meet $\omega(y)$; but $p_y \in V_y \cap \omega(y)$. This proves $J \setminus Q$ is countable. □

Let Y be an ordered Banach space and assume $X \subset Y$ is an ordered subspace (not necessarily linear). When Y is finite-dimensional, Theorem 1.19 implies $X \setminus Q$ has Lebesgue measure zero, hence almost every point is quasiconvergent. For infinite-dimensional Y there is an analogous result for Gaussian measures (Hirsch [73, Lemma 7.7]). The next result shows that in this case Q is also plentiful in the sense of category.

A subset of a topological space S is *residual* if it contains the intersection of countably many dense open subsets of S. When S is a complete metric space every residual set is dense by the Baire category theorem.

The assumption on X in the following result holds for many subsets of an ordered Banach space, including all convex sets and all sets with dense interior.

THEOREM 1.21. *Assume X is a subset of an ordered Banach space Y, and a dense open subset $X_0 \subset X$ is covered by totally ordered line segments. Then Q is residual in X.*

PROOF. It suffices to show that the set $Q_1 := Q \cup (Y \setminus X_0)$ is residual in Y. Note that $Y \setminus Q_1 = X_0 \setminus Q$. Let $L \subset Y$ be the 1-dimensional space spanned by some positive vector. Every translate $y + L$ meets $Y \setminus Q_1$ in a finite or countably infinite set by Theorem 1.19, hence $(y + L) \cap Q_1$ is residual in the line $y + L$. By the Hahn–Banach theorem there is a closed linear subspace $M \subset Y$ and a continuous linear isomorphism $F\colon Y \approx M \times L$ such that $F(x + L) = \{x\} \times L$ for each $x \in M$. Therefore $F(Q_1) \cap (\{x\} \times L)$ is residual in $\{x\} \times L$ for all $x \in X_0$, whence $F(Q_1)$ is residual in $M \times L$ by the Kuratowski–Ulam Theorem (Oxtoby [154]). This implies Q_1 is residual in Y. □

Additional hypotheses seem to be necessary in order to prove density of Q in general ordered spaces. The next theorem obtains the stronger conclusion that Q has dense interior. A different approach will be explored in Section 2.

A point x is *doubly accessible from below* (respectively, above) if in every neighborhood of x there exist f, g with $f < g < x$ (respectively, $x < f < g$).

Consider the following condition on a semiflow satisfying (H):

(L) Either every omega limit set has an infimum in X and the set of points that are doubly accessible from below has dense interior, or every omega limit set has a supremum in X and the set of points that are doubly accessible from above has dense interior.

This holds when X is the Banach space of continuous functions on a compact set with the usual ordering, for then every compact set has a supremum and infimum, and every point is doubly accessible from above and below.

THEOREM 1.22. *If* (L) *holds, then* $X \setminus Q \subset \overline{\operatorname{Int} C}$, *and* $\operatorname{Int} Q$ *is dense.*

The proof is based on the following result. For $p \in E$ define $C(p) := \{z \in X\colon \omega(z) = \{p\}\}$. Note that $C = \bigcup_{p \in E} C(p)$.

LEMMA 1.23. *Suppose* $x \in X \setminus Q$ *and* $a = \inf \omega(x)$. *Then* $\omega(a) = \{p\}$ *with* $p < \omega(x)$, *and* $x \in \overline{\operatorname{Int} C(p)}$ *provided* x *is doubly accessible from below.*

PROOF. Fix an arbitrary neighborhood M of x. Note that $a < \omega(x)$ because $\omega(x)$ is unordered (Theorem 1.8). By invariance of $\omega(x)$ we have $\Phi_t a \leqslant \omega(x)$, hence $\Phi_t a \leqslant a$. Therefore the Convergence Criterion Theorem 1.4 implies $\omega(a)$ is an equilibrium $p \leqslant a$. Because $p < \omega(x)$, SOP yields a neighborhood N of $\omega(x)$ and $s \geqslant 0$ such that $p \leqslant \Phi_t N$ for all $t \geqslant s$. Choose $r \geqslant 0$ with $\Phi_t x \in N$ for $t \geqslant r$. Then $p \leqslant \Phi_t x$ if $t \geqslant r + s$. The set $V := (\Phi_{r+s})^{-1}(N) \cap M$ is a neighborhood of x in M with the property that $p \leqslant \Phi_t V$ for all $t \geqslant r + 2s$. Hence:

$$u \in V \quad \Longrightarrow \quad p \leqslant \omega(u). \tag{1.3}$$

Now assume x doubly accessible from below and fix $y_1, y \in V$ with $y_1 < y < x$. By the Limit Set Dichotomy $\omega(y) < \omega(x)$, because $\omega(x) \not\subset E$. By SOP we fix a neighborhood $U \subset V$ of y_1 and $t_0 > 0$ such that $\Phi_{t_0} u \leqslant \Phi_{t_0} y$ for all $u \in U$. The Limit Set Dichotomy implies $\omega(u) = \omega(y)$ or $\omega(u) < \omega(y)$; as $\omega(y) < \omega(x)$, we therefore have:

$$u \in U \quad \Longrightarrow \quad \omega(u) < \omega(x). \tag{1.4}$$

For all $u \in U$, (1.4) implies $\omega(u) \leqslant \omega(a) = \{p\}$, while (1.3) entails $p \leqslant \omega(u)$. Hence $U \subset C(p) \cap M$, and the conclusion follows. □

PROOF OF THEOREM 1.22. To fix ideas we assume the first alternative in (L), the other case being similar. Let X_0 denote a dense open set of points doubly accessible from below. Lemma 1.23 implies $X_0 \subset Q \cup \overline{\operatorname{Int} C} \subset Q \cup \overline{\operatorname{Int} Q}$, hence the open set $X_0 \setminus \overline{\operatorname{Int} Q}$ lies in Q. This prove $X_0 \setminus \overline{\operatorname{Int} Q} \subset \operatorname{Int} Q$, so $X_0 \setminus \overline{\operatorname{Int} Q} = \emptyset$. Therefore $\overline{\operatorname{Int} Q} \supset X_0$, hence $\overline{\operatorname{Int} Q} \supset \overline{X_0} = X$. □

EXAMPLE 1.24. An example in Hirsch [73] shows that generic quasiconvergence and the Limit Set Dichotomy need not hold for a monotone semiflow that does not satisfy SOP. Let X denote the ordered Banach space $\mathbb{R}^3$ whose ordering is defined by the “ice-cream” cone $X_+ = \{x \in \mathbb{R}^3\colon\ x_3 \geqslant \sqrt{x_1^2 + x_2^2}\}$. The linear system $x_1' = -x_2,\ x_2' = x_1,\ x_3' = 0$ generates a flow Φ with global period 2π which merely rotates points about the x_3-axis. Evidently X_+ is invariant, so linearity of Φ implies monotonicity. On the other hand, Φ is not strongly order preserving: If $a = (1, 0, 1)$ (or any other point on ∂Y_+ except the origin 0), SOP would require $\Phi_t(a) \gg 0$ for $t > 0$ because Φ_t is a homeomorphism, but this fails for all $t > 0$. The Limit Set Dichotomy fails to hold: For $a = (1, 0, 1)$ and $b = (2, 0, 2)$ it is easy to see that $a < b$ (for the ordering defined by X_+) and $\omega(a) \cap \omega(b) = \emptyset$, but $\omega(a) \not< \omega(b)$. As $E = C = Q = \{x\colon\ x_1 = x_2 = 0\}$ and most points belonging to periodic orbits of minimal period 2π, quasiconvergence is rare. In fact, the set of nonquasiconvergent points—the complement of the x_3-axis—is open and dense. It is not known whether there is a similar example with a polyhedral cone.

1.6. *Stability in normally ordered spaces*

We continue to assume the semiflow Φ is SOP with compact orbit closures.

The *diameter* of a set Z is $\operatorname{diam} Z := \sup_{x,y\in Z} d(x, y)$.

We now introduce some familiar stability notions. A point $x \in X$ is *stable* (relative to $R \subset X$) if for every $\epsilon > 0$ there exists a neighborhood U of x such that $\operatorname{diam} \Phi_t(U \cap R) < \epsilon$ for all $t \geqslant 0$. The set of stable points is denoted by S.

Suppose x_0 is stable. Then omega limit sets of nearby points are close to $\omega(x_0)$, and if all orbit closures are compact, the map $x \mapsto \omega(x)$ is continuous at x_0 for the Hausdorff metric on the space of compact sets.

x is *stable from above* (respectively, *from below*) if x is stable relative to the set of points $\geqslant x$ (resp., $\leqslant x$). The set of points stable from above (resp., below is denoted by S_+ (resp., S_-).

The *basin of* x in R is the union of all subsets of R of the form $V \cap R$ where $V \subset X$ is an open neighborhood of x such that

$$\lim_{t\to\infty} \operatorname{diam} \Phi_t(V \cap R) = 0.$$

Notice that $\omega(x) = \omega(y)$ for all y in the basin.

If the basin of x in R is nonempty, we say x is *asymptotically stable* relative to R. This implies x is stable relative to R. If x is asymptotically stable relative to X we say x is *asymptotically stable*. The set of asymptotically stable points is an open set denoted by A.

x is *asymptotically stable from above* (respectively, *below)* if it is asymptotically stable relative to the set of points $\geqslant x$ (resp., $\leqslant x$). The basin of x relative to this set is called the *upper* (resp., *lower*) basin of x. The set of such x is denoted by A_+ (resp., A_-).

Note that continuity of Φ shows that asymptotic stability relative to R implies stability relative to R. In particular, $A \subset S$, $A_+ \subset S_+$ and $A_- \subset S_-$.

These stability notions for x depend only on the topology of X, and not on the metric, provided the orbit of x has compact closure.

The metric space X is *normally ordered* if there exists a *normality constant* $\kappa > 0$ such that $d(x, y) \leqslant \kappa d(u, v)$ whenever $u, v \in X$ and $x, y \in [u, v]$. In a normally ordered space order intervals are bounded and the diameter of $[u, v]$ goes to zero with $d(u, v)$. Many common function spaces, including L^p spaces and the Banach space of continuous functions with the uniform norm, are normally ordered by the cone of nonnegative functions. But spaces whose norms involve derivatives are not normally ordered. Normality is required in order to wring the most out of the Sequential Limit Set Trichotomy. The propositions that follow record useful stability properties of SOP dynamics in normally ordered spaces.

PROPOSITION 1.25. *Assume X is normally ordered.*

(a) *$x \in S_+$ (respectively, S_-) provided there exists a sequence $y_n \to x$ such that $y_n > x$ (resp., $y_n < x$) and* $\lim_{n\to\infty} \sup_{t>0} d(\Phi_t(x), \Phi_t(y_n)) = 0$.

(b) *$x \in S$ provided $x \in S_+ \cap S_-$ and x is accessible from above and below.*

(c) *$x \in A$ provided $x \in A_+ \cap A_-$ and x is accessible from above and below.*

(d) *Suppose $a < b$ and $\omega(a) = \omega(b)$. Then $a \in A_+$ and $b \in A_-$. If $a < x < b$ then $x \in A$ and the basin of x includes $[a, b] \setminus \{a, b\}$.*

In particular, (d) shows that an equilibrium e is in A_+ if $x > e$ and $\Phi_t(x) \to e$ (provided X is normally ordered); and dually for A_-.

PROOF. We prove (a) for the case $y_n > x$. Given $\epsilon > 0$, choose m and t_0 so that

$$t > t_0 \implies d\big(\Phi_t(x), \Phi_t(y_m)\big) < \epsilon.$$

By SOP there exists a neighborhood W of x and $t_1 > t_0$ such that

$$t > t_1,\ v \in W \implies \Phi_t(v) < \Phi_t(y_m).$$

Fixing t_1, we shrink W to a neighborhood W_ϵ of x so that

$$0 < t \leqslant t_1,\ v \in W_\epsilon \implies d\big(\Phi_t(x), \Phi_t(v)\big) < \kappa\epsilon,$$

where $\kappa > 0$ is the normality constant. If $x < v \in W_\epsilon$ and $t > t_1$ then $\Phi_t(x) \leqslant \Phi_t(v) \leqslant \Phi_t(y_m)$, and therefore

$$t > t_1,\ x < v \in W_\epsilon \implies d\big(\Phi_t(x), \Phi_t(v)\big) \leqslant \kappa d\big(\Phi_t(x), \Phi_t(y_m)\big) \leqslant \kappa\epsilon.$$

Hence we have proved

$$0 < t < \infty,\ v \in W_\epsilon \quad \Longrightarrow \quad d\big(\Phi_t(x), \Phi_t(v)\big) < \kappa\epsilon.$$

As ϵ is arbitrary, this proves $x \in S_+$.

To prove (b), let $u_n, v_n \to x$ with $u_n < x < v_n$. Because $x \in S_+ \cap S_-$, for any $\epsilon > 0$ there exists $\delta > 0$ such that if $d(y, x) < \delta$ and $y < x$ or $y > x$, then $\sup_{t>0} d(\Phi_t(y), \Phi_t(x)) < \epsilon$. Choose k such that $d(u_k, x) < \delta$ and $d(v_k, x) < \delta$. By SOP there is a neighborhood W_ϵ of x such that $\Phi_t(u_k) \leqslant \Phi_t(W_\epsilon) \leqslant \Phi_t(v_k)$ for sufficiently large t. Normality implies that for such t,

$$\begin{aligned}\kappa^{-1} \operatorname{diam} \Phi_t(W_e) &\leqslant d\big(\Phi_t(u_k), \Phi_t(v_k)\big)\\ &\leqslant d\big(\Phi_t(u_k), \Phi_t(x)\big) + d\big(\Phi_t(x), \Phi_t(v_k)\big) < 2\epsilon.\end{aligned}$$

As κ is constant and ϵ is arbitrary, this proves x is stable.

The proofs of (c) and (d) are similar. □

PROPOSITION 1.26. *Assume X is normally ordered, $p \in E$, and $\{K_n\}$ is a sequence of nonempty compact invariant sets such that $K_n < p$ and* $\operatorname{dist}(K_n, p) \to 0$. *Then*:

(a) *p is stable from below.*

(b) *If z is such that $\omega(z) = p$, then z is stable from below.*

In particular, if p is the limit of a sequence of equilibria $< p$ then p is stable from below.

PROOF. (a) Given $\epsilon > 0$, fix m such that $\operatorname{dist}(K_m, p) < \epsilon$. By Lemma 1.3 there is a neighborhood W of p and $t_0 > 0$ such that $t > t_0 \Longrightarrow \Phi_t(W) \geqslant K_m$, and therefore

$$t > t_0,\ v \in W,\ p > v \quad \Longrightarrow \quad d\big(\Phi_t(p), \Phi_t(v)\big) \leqslant \kappa d\big(\Phi_t(p), \Phi_t(K_m)\big) \leqslant \kappa\epsilon.$$

Pick a neighborhood $W_\epsilon \subset W$ of p so small that

$$0 \leqslant t \leqslant t_0,\ v \in W_\epsilon \quad \Longrightarrow \quad d\big(\Phi_t(p), \Phi_t(v)\big) < \kappa\epsilon.$$

Then

$$0 \leqslant t < \infty,\ v \in W_\epsilon,\ v < p \quad \Longrightarrow \quad d\big(\Phi_t(p), \Phi_t(v)\big) < \kappa\epsilon.$$

This proves $p \in S_-$, because ϵ is arbitrary.

(b) Choose a neighborhood U of z and $t_1 \geqslant 0$ such that $\Phi_{t_1}(U) \subset W$. Assume $y \in U$, $y < z$. If $t \geqslant t_1 + t_0$, then $K_m \leqslant \Phi_t(y) \leqslant \Phi_t(z)$, and therefore by normality, $d(\Phi_t(y), \Phi_t(z)) \leqslant \kappa \operatorname{dist}(K_m, \Phi_t(z))$. As $\Phi_t(z) \to p$, there exists $t_2 \geqslant t_1 + t_0$ such that

$$t \geqslant t_2 \quad \Longrightarrow \quad d\big(\Phi_t(y), \Phi_t(z)\big) \leqslant \kappa \operatorname{dist}(K_m, p) < \kappa\epsilon.$$

Fix this t_2. By continuity of Φ there is a neighborhood $U_1 \subset U$ of z so small that

$$0 \leqslant t \leqslant t_2 \implies d\big(\Phi_t(y), \Phi_t(z)\big) \leqslant \kappa\epsilon.$$

As ϵ is arbitrary, this implies $z \in S_-$. □

1.7. *Stable equilibria in strongly ordered Banach spaces*

In spaces that are not normally ordered we cannot directly use the results of the previous subsection to characterize stable equilibria. For strongly monotone semiflows in strongly ordered Banach spaces we work around this by introducing a weaker norm that makes the order normal, and for which the semiflows are continuous and SOP. This permits use of the earlier results.

Let Y be a strongly ordered Banach space. The *order topology* on Y is the topology generated by open order intervals. An *order norm* on the topological vector space $\widehat{Y}$ is defined by fixing $u \gg 0$ and assigning to x the smallest ϵ such that $x \in [-\epsilon u, \epsilon u]$. It is easy to see that $\widehat{Y}$ is normally ordered by the order norm, with normality constant 1. Every order neighborhood of p in $\widehat{Y}$ contains $[p - \epsilon u, p + \epsilon u]$ for all sufficiently small numbers $\epsilon > 0$. For example, $Y = C^1([0, 1])$ with the usual C^1-norm and with Y_+ the cone of nonnegative functions is strongly ordered but not normally ordered; putting $u := 1$, the order norm becomes the usual supremum.

The induced topology on any subset $Z \subset Y$ is also referred to as the order topology, and the resulting topological space is denoted by $\widehat{Z}$. A neighborhood in $\widehat{Z}$ is an *order neighborhood*.

Every open subset of $\widehat{Z}$ is open in Z, i.e., the identity map of Z is continuous from Z to $\widehat{Z}$. Therefore $\widehat{Z} = Z$ as topological spaces when Z is compact. As shown below, if Ψ is a monotone local semiflow in Z, it is also a local semiflow in $\widehat{Z}$, denoted by $\widehat{\Psi}$. Evidently Ψ and $\widehat{\Psi}$ have the same orbits and the same invariant sets.

LEMMA 1.27. *Let Ψ be a monotone local semiflow in a subset X of a strongly ordered Banach space Y, that extends to a monotone local semiflow in an open subset of Y. Then*:

(a) *$\widehat{\Psi}$ is a monotone local semiflow.*

(b) *If Ψ is a strongly monotone, then $\widehat{\Psi}$ is SOP.*

PROOF. It suffices to prove (a) and (b) when X is open in Y, which condition is henceforth assumed.

$\widehat{\Psi}$ is monotone because Ψ is monotone. To prove continuity of $\widehat{\Psi}$, let $N = [[a, b]]_Y \cap X$ and $(t_0, x_0) \in \widehat{\Psi}^{-1}(N)$. As the latter is open in $\mathbb{R}_+ \times X$, there exists $\epsilon > 0$ and U, an open neighborhood of x_0 in X, such that

$$\big[(t_0 - \epsilon, t_0 + \epsilon) \cap \mathbb{R}_+\big] \times U \subset \widehat{\Psi}^{-1}(N).$$

We may choose $u, v \in U$ such that $x \in [[u, v]]_Y$. If $z \in [[u, v]]_Y \cap X$ and $|t - t_0| < \epsilon$ then by monotonicity and $u, v \in U$ we have $a \ll \widehat{\Psi}_t(u) \leqslant \widehat{\Psi}_t(z) \leqslant \widehat{\Psi}_t(v) \ll b$. Thus,

$$\big[(t_0 - \epsilon, t_0 + \epsilon) \cap \mathbb{R}_+\big] \times \big([[u, v]]_Y \cap X\big) \subset \widehat{\Psi}^{-1}(N),$$

proving the continuity of $\widehat{\Psi}$.

Assume $x, y \in X$, $x < y$ and let $t_0 > 0$ be given. By strong monotonicity of Ψ there are respective open neighborhoods $U, V \subset X$ of x, y such that $\Psi_{t_0}(U) \leqslant \Psi_{t_0}(V)$ (see Proposition 1.2). Choose $w, u, v, z \in X$ such that $u, w \in U$, $v, z \in V$ and

$$w \ll x \ll u, \qquad v \ll y \ll z$$

so that $[[w, u]]_Y \cap X$ and $[[v, z]]_Y \cap X$ are order neighborhoods in X of x, y respectively. Monotonicity of Ψ implies

$$\Psi_{t_0}\big([[w, u]]_Y \cap X\big) \leqslant \Psi_{t_0}\big([[v, z]]_Y \cap X\big). \qquad \square$$

An equilibrium p for $\Psi : \mathbb{R}^+ \times X \to X$ is *order stable* (respectively, *asymptotically order stable* if p is stable (respectively, asymptotically stable) for $\widehat{\Psi}$.

PROPOSITION 1.28. *Let Ψ be a monotone local semiflow in a subset X of a strongly ordered Banach space Y, that extends to a monotone local semiflow in some open subset of Y. Assume p is an equilibrium having a neighborhood W that is attracted to a compact set $K \subset X$. If p is order stable* (*respectively, asymptotically order stable*), *it is stable* (*respectively, asymptotically stable*).

PROOF. Suppose p is order stable and let U be a neighborhood of p. As $\widehat{K} = K$, there is a closed order neighborhood N_0 of p such that $N_0 \cap K \subset U \cap K$. By order stability there exists an order neighborhood N_1 of p such that $O(N_1) \subset N_0$. Compactness of $N_0 \cap K$ implies there is an open set $V \supset K$ there is an open set $V \supset K$ such that $N_0 \cap V \subset U$. Because K attracts W, there is a neighborhood $U_2 \subset W$ of p and $r \geqslant 0$ such that

$$t \geqslant r \quad \Longrightarrow \quad \Psi_t(U_2) \subset V.$$

By continuity of Ψ_r at $p = \Psi_r(p)$ there is a neighborhood $U_3 \subset U_2$ of p such that

$$0 \leqslant t \leqslant r \quad \Longrightarrow \quad \Psi_t(U_3) \subset V.$$

and thus $O(U_3) \subset V$. Therefore $N_1 \cap U_3$ is a neighborhood of p such that

$$O(N_1 \cap U_3) \subset O(N_1) \cap O(U_3) \subset N_0 \cap V \subset U.$$

This shows p is stable.

Assume p is asymptotically order stable and choose an order neighborhood $M \subset X$ of p that is attracted to p by $\widehat{\Psi}$. We show that $M \cap W$ is in the basin of p for Ψ. Consider arbitrary sequences $\{x_k\}$ in $M \cap W$ and $t_k \to \infty$ in $[0, \infty)$. Fix $u \gg 0$. By the choice of M there are positive numbers $\epsilon_k \to 0$ such that

$$p - \epsilon_k u \ll \Psi_{t_k}(x_k) \ll p + \epsilon_k u.$$

This implies $\Psi_{t_k}(x_k) \to p$ in X, because the order relation on X is closed and $\{\Psi_{t_k}(x_k)\}$ is precompact in X by the choice of W and compactness of K. $\square$

1.8. *The search for stable equilibria*

The following results illustrate the usefulness of a dense set of quasiconvergent points. Φ denotes a strongly order preserving semiflow in X; Hypothesis (H) of Section 1.4 is still in force.

PROPOSITION 1.29. *Assume Q is dense. Let $p, q \in E$ be such that $p < q$, p is accessible from above, and q is accessible from below. Then there exists $z \in X$ satisfying one of the following conditions*:

(a) $p < z < q$, *and* $\Phi_t(z) \to p$ *or* $\Phi_t(z) \to q$;

(b) $p < z < q$ *and* $z \in E$;

(c) $z > p$ *and* $p \in O(z)$, *or* $z < q$ *and* $q \in O(z)$.

PROOF. By SOP there are open neighborhoods U, V of p, q respectively and $t_0 \geqslant 0$ such that $\Phi_t U \leqslant \Phi_t V$ for $t \geqslant t_0$. Choose sequences $x_n \to p$ in U and $y_n \to q$ in V with $p < x_n$, $y_n < q$. We assume $p \notin O(x_n)$ and $q \notin O(y_n))$, as otherwise (c) is satisfied. Then

$$t \geqslant t_0 \quad \Longrightarrow \quad p < \Phi_t(x_n) \leqslant \Phi_t(y_n) < q.$$

Choose open neighborhoods U_1, W, V_1 of $p, \Phi_{t_0}(y_1), q$ respectively such that for some $t_1 \geqslant t_0$:

$$t \geqslant t_1 \quad \Longrightarrow \quad \Phi_t(U_1) \leqslant \Phi_t(W) \leqslant \Phi_t(V_1).$$

Choose $w \in Q \cap W$ and a sequence $s_k \to \infty$, $s_k \geqslant t_1$ such that $\Phi_{s_k}(w) \to e \in E$. Fix m so large that $x_m \in U_1$, $y_m \in V_1$. Then for sufficiently large k,

$$p < \Phi_{s_k}(x_m) \leqslant \Phi_{s_k}(w) \leqslant \Phi_{s_k}(y_m) < q.$$

It follows that $p \leqslant e \leqslant q$. If $e = p$ or q then $\omega(\Phi_{s_k}(w)) = p$ or q by the Convergence Criterion 1.4, giving (a) with $z = \Phi_{s_k}(w)$. Therefore if (a) does not hold, (b) holds with $z = e$. □

The assumption in Proposition 1.29 that Q is dense can be considerably weakened, for example, to p (or q) being interior to $\overline{Q}$: Assume $y_1 \in \overline{\operatorname{Int} Q}$ and set $w = \Phi_{s_k}(w_0)$, $w_0 \in (\operatorname{Int} Q) \cap \Phi_{s_k}{}^{-1}(W)$, etc. In fact, density of Q can be replaced with the assumption that p or q lies in the interior of the set $Q_\#$ of points x such that there is a sequence $x_i \to x$ with $\lim_{i\to\infty} \operatorname{dist}(\omega(x_i), E) = 0$. Clearly $Q_\#$ is closed and contains Q, so density of Q implies $Q_\# = X$.

THEOREM 1.30. *Suppose X is normally ordered and the following three conditions hold*:

(a) *Q is dense*;

(b) *if $e \in E$ and e is not accessible from above* (*below*) *then $e = \sup X$* ($e = \inf X$);

(c) *there is a maximal totally ordered subset $R \subset E$ that is nonempty and compact.*

Then R contains a stable equilibrium, an asymptotically stable equilibrium if R is finite.

PROOF. By Lemma 1.1, $\sup R$ ($\inf R$) exists and is a maximal (minimal) element of E. We first prove that every maximal equilibrium q is in A_+. This holds vacuously when $q = \sup X$. Suppose $q \neq \sup X$. If q is in the orbit of some point $> q$ then $q \in A_+$ by Proposition 1.25(d). Hence we can assume:

$$t \geqslant 0,\ y > q \quad \Longrightarrow \quad \Phi_t(y) > q.$$

By hypothesis we can choose $y > q$. By SOP there is an open neighborhood U of q and $s > 0$ such that $\Phi_s(y) \geqslant \Phi_s(U)$. By hypothesis we can choose $z \in U$ such that $\Phi_s(y) \neq \Phi_s(z)$ and $z > q$. Set $x_2 = \Phi_s(y)$, $x_1 = \Phi_s z$. Then $x_2 > x_1 > q$, By SOP and the assumption above there is a neighborhood V_2 of x_2 and $t_0 \geqslant 0$ such that

$$t > t_0 \quad \Longrightarrow \quad q < \Phi_t(x_1) \leqslant \Phi_t(V_2).$$

Choose $v \in V_2 \cap Q$. Then $q < \Phi_t(v)$ for $t \geqslant t_0$, hence $q \leqslant \omega(v) = \omega(\Phi_{t_0}(v)) \subset E$. Therefore $\Phi_t(v) \to q$ by maximality of q, so and Proposition 1.25(d) implies $q \in A_+$, as required. The dual argument shows that every minimal equilibria is in A_-.

Assumption (c) and previous arguments establish that $q = \sup R$ and $p = \inf R$ satisfy $p \leqslant q$ and $q \in A_+$, $p \in A_-$.

Suppose $p = q$; in this case we prove $q \in A$. As q is both maximal and minimal in E, we have $q \in A_+ \cap A_-$. If q is accessible from above and below then $q \in A$ by Proposition 1.25(b). If q is not accessible from above then by hypothesis $q = \sup X$, in which case the fact that $q \in A_-$ implies $q \in A$. Similarly, $q \in A$ if q is not accessible from below.

Henceforth we assume $p < q$. As R is compact and $R \cap S_- \neq \emptyset$ because $p \in R$, it follows that R contains the equilibrium $r := \sup(R \cap S_-)$. Note that $r \in S_-$, because this holds by definition of r if r is isolated in $\{r' \in R\colon r' \leqslant r\}$, and otherwise $r \in S_-$ by Proposition 1.26(a). If $r = q$ a modification of the preceding paragraph proves $q \in S$.

Henceforth we assume $r < q$; therefore r is accessible from above.

If r is not accessible from below then $r = p = \inf X$ so $r \in S$ and we are done; so we may as well assume r is accessible from below as well as from above. If r is the limit of a sequence of equilibria $> r$ then $r \in S_+$ by the dual of Proposition 1.26, hence $r \in S$ by Proposition 1.25(b). Therefore we can assume R contains a smallest equilibrium $r_1 > r$. Note that $r_1 \notin S_-$ by maximality of r. We apply Proposition 1.29 to r, r_1: among its conclusions, the only one possible here is that $z > r$ and $\Phi_t(z) \to r$ (and perhaps $r \in O(z)$). Therefore $r \in S_+$ by Proposition 1.25(a), whence $r \in S$ by 1.25(b). When R is finite, a modification of the preceding arguments proves $\max(R \cap A_-) \subset A$. □

Assumption (b) in the Theorem 1.30 holds for many subsets X of an ordered Banach space Y, including open sets, subcones of Y_+, closed order intervals, and so forth. This result is similar to Theorem 10.2 of Hirsch [73], which establishes equilibria that are merely order stable, but does not require normality.

Assumption (c) holds when E is compact, and also in the following situation: $X \subset Y$ where Y is an L^p space, $1 \leqslant p < \infty$, and E is a nonempty, closed, and order bounded subset of X; then every order bounded increasing or decreasing sequence converges. If (c) holds and some Φ_t is real analytic with spatial derivatives that are compact and strongly positive operators, then R is finite. This follows from the statements and proofs of Lemma 3.3 and Theorem 2 in Jiang and Yu [90].

For related results on stable equilibria see Jiang [86], Mierczyński [138,139], and Hirsch [69].

THEOREM 1.31. *Let Φ be a semiflow in a subset X of a strongly ordered Banach space Y, that extends to a strongly monotone local semiflow in some open subset of Y. Assume hypotheses* (a), (b), (c) *of Theorem* 1.30 *hold, and every equilibrium has a neighborhood attracted to a compact set. Let $R \subset E$ be as in* 1.30(c). *Then R contains a stable equilibrium, and an asymptotically stable equilibrium when R is finite.*

PROOF. Our strategy is to apply Theorem 1.30 to the semiflow $\widehat{\Phi}$ in $\widehat{X}$ (see Section 1.7). Give $\widehat{X}$ the metric coming from an order norm on $\widehat{Y}$; this makes $\widehat{X}$ is normally ordered. Lemma 1.27 shows that $\widehat{\Phi}$ is SOP. Therefore R contains an equilibrium p that is stable for $\widehat{\Phi}$, by Theorem 1.30. This means p is order stable for Φ, whence Proposition 1.28 shows that p is stable for Φ. The final assertion follows similarly. ☐

Stable equilibria are found under various assumptions in Theorems 2.9, 2.10, 2.11, 2.26, 3.14, 4.12.

2. Generic convergence and stability

2.1. *The sequential limit set trichotomy*

Throughout Section 2 we assume Hypothesis (H) of Section 1.4:

> *Φ is a strongly order preserving semiflow in an ordered space X, with all orbit closures compact.*

The main result is that the typical orbit of an SOP semiflow is stable and approaches the set E of equilibria. Existence of stable equilibria is established under additional compactness assumptions.

The index n runs through the positive integers.

A point x is *strongly accessible from below* (*respectively*, above) if there exists a sequence $\{y_n\}$ converging to x such that $y_n < y_{n+1} < x$ (resp., $y_n > y_{n+1} > x$). In this case we say $\{y_n\}$ *strongly approximates x from below* (resp., *from above*).

The sequence $\{x_n\}$ is *omega compact* if $\bigcup_n \omega(x_n)$ is compact.

Define sets $BC, AC \subset X$ as follows:

$x \in BC \iff x$ is strongly accessible from below by an omega compact sequence,

$x \in AC \iff x$ is strongly accessible from above by an omega compact sequence.

In this notation “B” stands for “below,” “A” for above, and “C” for “compact.”

We will also use the following condition on a set $W \subset X$:

(C) Every sequence $\{w_n\}$ in W that strongly approximates a point of W from below or above is omega compact.

This does not assert that any point is strongly accessible from below or above. But if every point of W is accessible from above and W satisfies (C), then $W \subset AC$; and similarly for BC.

The next two propositions imply properties stronger than (C). Recall that a map $f : X \to X$ is *completely continuous* provided $\overline{f(B)}$ is compact for every bounded set $B \subset X$; and f *conditionally completely continuous* provided $\overline{f(B)}$ is compact whenever B and $f(B)$ are bounded subsets of X.

The *orbit* of any set $B \subset X$ is $O(B) = \bigcup_t \Phi_t(B)$.

Proposition 2.1. *Assume the following two conditions*:

(a) *every compact set has a bounded orbit, and*

(b) Φ_s *is conditionally completely continuous for some* $s > 0$.

If $L \subset X$ *is compact, then* $\overline{\bigcup_{x\in L} \omega(x)}$ *is compact and this implies* X *has property* (C).

Proof. $O(L)$ is a bounded set by (a), and positively invariant, so (b) implies compactness of $\overline{\Phi_s(O(L))}$. As the latter set contains $\omega(x)$ for all $x \in L$, the first assertion is proved. The second assertion follows from precompactness of $\{x_n\}$. □

Proposition 2.2. *Assume* $W \subset X$ *has the following property*: *For every* $x \in W$ *there is a neighborhood* $U_x \subset X$ *and a compact set* M_x *that attracts every point in* U_x. *Then* $\overline{O(x)}$ *is compact for every* $x \in W$, *and* $\overline{\bigcup_{y\in U_x} \omega(y)}$ *is compact*. *If* $z_n \to x \in W$ *then* $\overline{\bigcup_n \omega(z_n)}$ *is compact, therefore* W *has property* (C).

Proof. It is easy to see that $\overline{O(x)}$ is compact and $\overline{\bigcup_{y\in U_x} \omega(y)}$ is compact because it lies in M_x. Fix $k \geqslant 0$ such that $z_n \in U_x$ for all $n \geqslant k$. Then

$$\overline{\bigcup_n \omega(z_n)} = \bigcup_{1\leqslant n\leqslant k} \omega(z_n) \cup M_x,$$

which is the union of finitely many compact sets, hence compact. Condition (C) follows trivially. □

The key to stronger results on generic quasiconvergence and stability is the following result of Smith and Thieme [197]:

Theorem 2.3 (Sequential Limit Set Trichotomy). *Let* $\{\tilde{x}_n\}$ *be an omega compact sequence strongly approximating* $z \in BC$ *from below. Then there is a subsequence* $\{x_n\}$ *such that exactly one of the following three conditions holds*:

(a) *There exists* $u_0 \in E$ *such that*

$$\omega(x_n) < \omega(x_{n+1}) < \omega(z) = \{u_0\}$$

and

$$\lim_{n\to\infty} \operatorname{dist}\bigl(\omega(x_n), u_0\bigr) = 0.$$

In this case $z \in C$.

(b) *There exists* $u_1 = \sup\{u \in E\colon\ u < \omega(z)\}$ *and*

$$\omega(x_n) = \{u_1\} < \omega(z).$$

In this case $z \in \overline{\operatorname{Int} C}$. *Moreover* z *has a neighborhood* W *such that if* $w \in W$, $w < z$ *then* $\Phi_t(w) \to u_1$ *and* $\Phi_t(w) > u_1$ *for sufficiently large* t.

(c) $\omega(x_n) = \omega(z) \subset E$.
In this case $z \in \overline{\operatorname{Int} Q}$. *Moreover* $\omega(w) = \omega(z) \subset E$ *for every* $w < z$ *sufficiently near* z.

Note that z is convergent in (a), and strongly accessible from below by convergent points in (b). In (c), z is quasiconvergent and strongly accessible from below by quasiconvergent points.

If $z \in AC$ there is an analogous dual result, obtained by reversing the order relation in X. Although we do not state it formally, we will use it below. If $z \in AC \cap BC$ then both results apply. See Proposition 3.6 in Smith and Thieme [197].

PROOF OF THEOREM 2.3. By the Limit Set Dichotomy 1.16, either there exists a positive integer j such that $\omega(\tilde{x}_n) = \omega(\tilde{x}_m)$ for all $m, n \geqslant j$, or else there exists a subsequence $\{\tilde{x}_{n_i}\}$ such that $\omega(\tilde{x}_{n_i}) < \omega(\tilde{x}_{n_{i+1}})$ for all i. Therefore there is a subsequence $\{x_n\}$ such that $\omega(x_n) < \omega(x_{n+1})$ for all n, or $\omega(x_n) = \omega(x_{n+1})$ for all n.

Case I: $\omega(x_n) < \omega(x_{n+1})$. We will see that (a) holds. The Limit Set Dichotomy 1.16 implies $\omega(x_n) \leqslant \omega(z)$. In fact, that $\omega(x_n) < \omega(z)$. Otherwise $\omega(x_k) \cap \omega(z) \neq \emptyset$ for some k, and the Limit Set Dichotomy implies the contradiction $\omega(x_k) = \omega(z) \geqslant \omega(x_{k+1}) > \omega(x_k)$.

Define $K = \overline{\bigcup \omega(x_n)}$, a nonempty compact invariant set. Consider the set

$$\Lambda = \Bigl\{y\colon\ y = \lim_{n\to\infty} y_n,\ \ y_n \in \omega(x_n)\Bigr\} \subset K.$$

Clearly Λ is invariant and closed, and compactness of K implies Λ is compact and nonempty. We show that Λ is a single equilibrium. Suppose $y, v \in \Lambda$, so that $y_n \to y$, $v_n \to v$ with $y_n, v_n \in \omega(x_n)$. Since $y_n < v_{n+1}$ and $v_n < y_{n+1}$, we have $y \leqslant v$ and $v \leqslant y$, so $v = y$. Thus we can set $\Lambda = \{u_0\}$, and invariance implies $u_0 \in E$.

The definition of Λ and compactness of K imply $\lim_{n\to\infty} \operatorname{dist}(\omega(x_n), u_0) = 0$. From $\omega(x_n) < \omega(x_{n+1}) < \omega(z)$ we infer

$$\omega(x_n) < u_0 \leqslant \omega(z).$$

If $u_0 \in \omega(z)$ then $\omega(z) = \{u_0\}$ by Corollary 1.9, yielding (a).

We show that $u_0 < \omega(z)$ gives a contradiction. Choose a neighborhood W of $\omega(z)$ and $t_0 \geqslant 0$ such that $u_0 \leqslant \Phi_t(W)$ for all $t \geqslant t_0$ (by Lemma 1.3). There exists $t_1 > 0$ such that $\Phi_{t_1}(z) \in W$, and by continuity of Φ_{t_1} there exists m such that $\Phi_{t_1}(x_m) \in W$. It follows that $u_0 \leqslant \Phi_t(x_m)$ for $t \geqslant t_0 + t_1$. As $u_0 \in E$, we have $u_0 \leqslant \omega(x_m)$. But this contradicts $\omega(x_m) < u_0$. Thus (a) holds in Case I.

Case II: $\omega(x_n) = \omega(x_{n+1}) \subset E$. Since $x_n < z$, the Limit Set Dichotomy implies that either $\omega(x_n) = \omega(z)$, which gives (c), or else $\omega(x_n) < \omega(z)$, which we now assume. Choose an equilibrium $u_1 \in \omega(x_1)$. By Lemma 1.3 there exists an open set W containing $\omega(z)$ and $t_0 \geqslant 0$ such that $u_1 \leqslant \Phi_t(W)$ for all $t \geqslant t_0$. Arguing as in Case I, we obtain $u_1 \leqslant \Phi_t(x_m)$ for some m and all large t. Since $u_1 \in \omega(x_m)$, it follows that $\omega(x_m) = u_1$ by Corollary 1.9, and therefore $\omega(x_n) = \{u_1\}$ as asserted in case (b). Finally, if $u \in E$ and $u < \omega(z)$, we argue as above that $\omega(x_m) \geqslant u$ for some m, which implies $u_1 \geqslant u$.

To prove $z \in \overline{\operatorname{Int} Q}$, use SOP to obtain a neighborhood U_n of x_n such that $\Phi_t(x_{n-1}) \leqslant \Phi_t(U_n) \leqslant \Phi_t(x_{n+1})$ for all large t, implying $U_n \subset Q$. A similar argument proves the analogous assertion in (b). ☐

The following addendum to the Sequential Limit Set Trichotomy provides important stability information. In essence, it associates various kinds of stable points to arbitrary elements $z \in BC$:

PROPOSITION 2.4. *Assume X is normally ordered. In cases* (a), (b) *and* (c) *of the Sequential Limit Set Trichotomy, the following statements are valid respectively*:

(a) *z and u_0 are stable from below*;

(b) *z is not stable from below, $\omega(z)$ is unstable from below, and u_1 is asymptotically stable from above*;

(c) *z is asymptotically stable from below, and $z \in \overline{A}$.*

PROOF. (a) follows from Proposition 1.26(a) and (b).

(b) The first two assertions are trivial. To prove $u_1 \in A_+$, take $w = x_n$ for some large n in the last assertion of (b) and apply 1.25(d) with $a = u_1$.

(c) follows from 1.25(d), taking $b = z$. ☐

We expect in real world systems that observable motions are stable trajectories. Our next result implies stable trajectories approach equilibria.

PROPOSITION 2.5. $S \cap (BC \cup AC) \subset Q$.

PROOF. When $z \in S \cap (BC \cup AC)$, only (a) and (c) of the Sequential Limit Set Trichotomy are possible, owing to continuity at z of the function $x \mapsto \omega(x)$. In both cases $z \in Q$. ☐

The inclusion $S \subset Q$ suggests trajectories issuing from nonquasiconvergent points are unlikely to be observed; the next result implies that their limit sets are, not surprisingly, unstable. There are as many concepts of instability as there are of stability, but for our purposes the following very strong property suffices: A set $M \subset X$ is *unstable from above* provided there is an equilibrium $u > M$ such that $\omega(x) = \{u\}$ if $u > x > y$, $y \in M$. Such an equilibrium u is unique, and SOP implies it attracts all points $< u$ in some neighborhood of u. *Unstable from below* is defined dually.

THEOREM 2.6. *Assume $z \in BC \setminus Q$ (respectively, $z \in AC \setminus Q$). Then $\omega(z)$ is unstable from below (resp., above).*

PROOF. To fix ideas we assume $z \in BC \setminus Q$. Then there exists a sequence $x_n \to z$ and an equilibrium u_1 as in conclusion (b) of the Sequential Limit Set Trichotomy. Suppose $u_1 < x < y$, $y \in \omega(z)$. SOP implies there exist open sets W_x and W_y containing x and y, respectively, and $t_0 \geqslant 0$, such that $\Phi_t(W_x) \leqslant \Phi_t(W_y)$ for all $t \geqslant t_0$. As $\Phi_s(z) \in W_y$ for some large s, by continuity $\Phi_s(x_n) \in W_y$ for some large n. Thus $u_1 \leqslant \Phi_t(x) \leqslant \Phi_{t+s}(x_n)$ for all $t \geqslant t_0$. Letting $t \to \infty$ and using the fact that $\omega(x_n) = \{u_1\}$, we find that $\omega(x) = \{u_1\}$. □

A set is *minimal* if it is nonempty, closed and invariant, and no proper subset has these three properties. Every positively invariant nonempty compact set contains a minimal set (by Zorn's Lemma). A minimal set containing more than one point is called *nontrivial*.

COROLLARY 2.7. *A compact, nontrivial minimal set M that meets BC (respectively, AC) is unstable from below (resp., above).*

PROOF. Suppose $z \in M \cap BC$. The assumptions on M imply $M = \omega(z)$ and $M \cap E = \emptyset$. Therefore $z \in BC \setminus Q$, and instability follows from Theorem 2.6. □

When X is a convex subset of a vector space, an alternative formulation of Theorem 2.6 is that $\omega(z)$ belongs to the upper boundary of the basin of attraction of the equilibrium u_1. Corollary 2.7 implies that periodic orbits are unstable. Theorem 2.6 is motivated by Theorem 1.6 in Hirsch [79].

The following sharpening of Theorem 1.19, due to Smith and Thieme [199], is an immediate corollary of the Sequential Limit Set Trichotomy.

THEOREM 2.8. *If $J \subset X$ is a totally ordered arc having property* (C), *then $J \setminus Q$ is a discrete, relatively closed subset of J; hence it is countable, and finite when J is compact.*

PROOF. Every limit point z of $J \setminus Q$ is strongly accessible from above or below by a sequence $\{\tilde{x}_n\}$ in $J \setminus Q$. As Property (C) implies $J \subset BC \cup AC$, there is a sequence $\{x_n\}$ satisfying (a), (b) or (c) of Theorem 2.3 (or its dual result), all of which imply $x_n \in Q$. Thus $J \setminus Q$ contains none of its limit points, which implies the conclusion. □

The following result sharpens Theorems 1.30 and 2.8:

PROPOSITION 2.9. *Assume X is normally ordered and every point is accessible from above and below. Let $J \subset X$ be a totally ordered compact arc having property* (C), *with* *endpoints $a < b$ such that $\omega(a)$ is an equilibrium stable from below and $\omega(b)$ is an equilibrium stable from above. Then J contains a point whose trajectory converges to a stable equilibrium.*

PROOF. Denote by C_s (respectively: C_+, C_-) the set of convergent points whose omega limits belong to S (resp.: to S_+, S_-). Then $C_+ \cap C_- = C_s$ by Proposition 1.25(b).

Set $\sup(J \cap C_-) = z \in J$.

Case 1: $z \notin C_-$. Then $z > a$. Choose a sequence $x_1 < x_2 < \cdots < z$ in $J \cap C_-$ such that $x_n \to z$. By the Sequential Limit Set Trichotomy 2.3 it suffices to consider the following three cases:

(a) There exists $u_0 \in E$ such that

$$\omega(x_n) < \omega(x_{n+1}) < \omega(z) = \{u_0\}.$$

This is not possible, because $u_0 \in S_-$ by Proposition 2.4(a), yielding the contradiction $z \in C_-$.

(b) There exists $u_1 = \sup\{u \in E\colon\ u < \omega(z)\}$, and for all n we have

$$\omega(x_n) = \{u_1\} < \omega(z).$$

Now 2.4(b) has $u_1 \in S_+$, hence $x_n \in C_+$. Therefore $x_n \in C_+ \cap C_- = C_s$, as required.

(c) $\omega(x_n) = \omega(z)$. This is not possible because $x_n \in C_-$ and $\omega(z) = \omega(x_n)$ implies the contradiction $z \in C_-$.

Thus (b) holds, validating the conclusion when $z \notin C_-$.

Case 2: $z \in C_-$. If $z = b$ then $z \in C_+ \cap C_- = C_s$ and there is nothing more to prove. Henceforth we assume $z < b$.

The closed subinterval $K \subset J$ with endpoints z, b satisfies the hypotheses of the theorem. Set $\inf(K \cap C_+) = w \in K$. The dual of the reasoning above shows that if $w \notin C_+$ then the conclusion of the theorem is true.

From now on we assume $w \in C_+$. If $w = z$ there is nothing more to prove, so we also assume $w > z$. Let $L \subset K$ be the closed subinterval with endpoints w and z. Let $\{\bar{x}_n\}$ be a sequence in L converging to w from below.

One of the conclusions (a), (b) or (c) of 2.3 holds. Referring to the corresponding parts of 2.4, we see in case (a) that $\omega(w)$ is an equilibrium $\bar{u}_0$ that is stable from below; but $w > z$, so this contradicts the definition of z. If (b) holds, $\omega(\bar{x}_n)$ is an equilibrium $\bar{u}_1$ stable from above. But $\bar{x}_n < w$, so this contradicts the definition of w. In case (c) we have for all n that $\omega(\bar{x}_n) = \omega(w)$, which is an equilibrium stable from above. But $\bar{x}_n < w$ for $n > 1$, again contradicting the definition of w. □

In the following result the assumption on equilibria holds when Φ has a global compact attractor.

PROPOSITION 2.10. *Assume X is an open subset of a strongly ordered Banach space, Φ is strongly monotone, and every equilibrium has a neighborhood attracted to a compact set. Let $J \subset X$ be a totally ordered compact arc, with endpoints $a < b$ such that $\omega(a)$ is* *an equilibrium stable from below and* $\omega(b)$ *is an equilibrium stable from above. Then* J *contains a point whose trajectory converges to a stable equilibrium.*

PROOF. Apply Proposition 2.9 to the to the SOP semiflow $\widehat{\Phi}$ in the normally ordered space $\widehat{X}$ (see Section 1.7), to obtain an equilibrium p that is stable for $\widehat{\Phi}$. This means p is order stable for Φ, hence stable for Φ by Proposition 1.28. □

COROLLARY 2.11. *Let* X *be a p-convex open set in an ordered Banach space* Y. *Assume* Φ *has a compact global attractor. Suppose that either* Y *is normally ordered, or* Y *is strongly ordered and* Φ *is strongly monotone. Then*:

(i) *There is a stable equilibrium.*

(ii) *Let* $u, v \in X$ *be such that* $u < v$ *and there exist real numbers* $r, s > 0$ *such that* $u < \Phi_r(u)$, $\Phi_s(v) < v$. *Then there is a stable equilibrium in* $[u, v]$.

In case (ii) *with* Y *normally ordered, the hypothesis of a global attractor can be replaced the assumption that the line segment joining* u *to* v *from satisfies condition* (C).

PROOF. We first prove (ii). Monotonicity shows that $\omega(x) \subset [u, v]$ for all $x \in [u, v]$. The Convergence Criterion implies

$$\Phi_t(u) \to a \in E \cap [u, v], \qquad \Phi_t(v) \to b \in E \cap [u, v].$$

We claim that $a \in S_-$ and $b \in S_+$, and $a \in S$ is stable if $a = b$. When Y is normal this follows from Propositions 1.25(b) and (d), and it is easy to prove directly when Φ is strongly monotone. Suppose $a < b$. By p-convexity and Theorems 2.9 and 2.10, the line segment from a to b lies in $[u, v] \cap X$ and contains a point whose x such that $\omega(x)$ is a stable equilibrium z. As noted above, $z \in [u, v]$.

We prove (i) by finding u and v as in (ii). By Theorem 2.8 and compactness of the global attractor, there is a minimal equilibrium p and a maximal equilibrium $q > p$. As X is open, it contains a totally ordered line segment $J < p$. By Theorem 1.19 J contains a quasiconvergent point $u < p$. As $\omega(u) \leqslant p$, minimality of p implies $\Phi_t(u) \to p$. Similarly there exist $v > q$ with $\Phi_t(v) \to q$. It follows from SOP that $u < \Phi_r(u)$, $\Phi_s(v) < v$ for some $r, s > 0$. □

For strongly monotone semiflows, the existence of order stable equilibria in attractors was treated in Hirsch [68,69,73].

2.2. *Generic quasiconvergence and stability*

The following result adapted from Smith and Thieme [197] refines Theorems 1.22 and 1.21:

THEOREM 2.12. (i) $AC \cup BC \subset \overline{\operatorname{Int} Q} \cup C$. *Therefore if* $AC \cup BC$ *is dense, so is* Q.

(ii) $(\operatorname{Int} AC) \cup (\operatorname{Int} BC) \subset \overline{\operatorname{Int} Q}$. *Therefore if* $(\operatorname{Int} AC) \cup (\operatorname{Int} BC)$ *is dense, so is* $\operatorname{Int} Q$.

PROOF. Every $z \in BC$ is the limit of an omega compact sequence $x_1 < x_2 < \cdots$ such that (a), (b) or (c) of the Sequential Limit Set Trichotomy Theorem 2.3 holds, and $z \in \overline{\operatorname{Int} Q} \cup C$ in each case; the proof for AC is similar.

To prove (ii), assume $z \in \operatorname{Int} BC$. If (a) holds for every point of a neighborhood W of z, then $W \subset C$, whence $z \in \operatorname{Int} Q$. If there is no such W, every neighborhood of z contains a point for which (b) or (c) holds, hence $z \in \overline{\operatorname{Int} Q}$. Similarly for $z \in \operatorname{Int} AC$. □

The next result extends Theorems 8.10 and 9.6 of Hirsch [73] and Theorem 3.9 of Smith and Thieme [197]:

THEOREM 2.13. *Assume X is normally ordered and* $\operatorname{Int}(BC \cup AC)$ *is dense. Then* $A \cup \operatorname{Int} C$ *is dense.*

PROOF. We argue by contradiction. If $A \cup \operatorname{Int} C$ is not dense, there exists an open set U such that

$$U \cap \overline{A} = \emptyset = U \cap \overline{\operatorname{Int} C}.$$

Suppose $z \in U \cap BC$, and let $\{x_n\}$ be a sequence in U strongly approximating z from below. Conclusion (b) of the Sequential Limit Set Trichotomy 2.3 is not possible because $z \notin \overline{\operatorname{Int} C}$, and conclusion (c) is ruled out because $z \notin \overline{A}$ (see Proposition 2.4(c)). Therefore conclusion (a) holds, which makes z convergent; likewise when $z \in U \cap AC$. Thus we have $C \supset U \cap (BC \cup AC)$, so $\operatorname{Int} C \supset U \cap \operatorname{Int}(BC \cup AC)$. But the latter set is nonempty by the density hypothesis, yielding the contradiction $U \cap \operatorname{Int} C \neq \emptyset$. □

The following theorem concludes that generic trajectories are not only quasiconvergent, but also stable. Its full force will come into play in the next subsection, under assumptions entailing a dense open set of convergent points.

THEOREM 2.14. *If X is normally ordered and* $\operatorname{Int}(BC \cap AC)$ *is dense, then* $\operatorname{Int}(Q \cap S)$ *is dense.*

PROOF. $\operatorname{Int} Q$ is dense by Theorem 2.12. To prove density of $\operatorname{Int} S$, it suffices to prove that if $z \in \operatorname{Int}(BC \cap AC)$, then every open neighborhood U of z meets $\operatorname{Int} S$. We can assume $z \notin \overline{A}$ because $A \subset \operatorname{Int} S$. Let $\{x_n\}$ be an omega compact sequence strongly approximating z from below. Suppose (b) or (c) of the Sequential Limit Set Trichotomy 2.3 holds. Then $x_m \in U$ for $m \geqslant m_0$. Fix $m \geqslant m_0$. It follows from Proposition 1.25(d) (with $a = x_m$, $x = x_{m+1}$, $b = x_{m+2}$) that $x_{m+1} \in A$, hence $z \in \overline{A}$; this is proved similarly when $\{x_n\}$ strongly approximates z from above.

Henceforth we can assume z belongs to the open set $W = \operatorname{Int}(BC \cap AC) \setminus \overline{A}$, and consequently that there are omega compact sequences $\{x_n\}$, $\{y_n\}$ strongly approximating z from below and above respectively, for which Theorem 2.3(a) and its dual hold respectively. Then Proposition 1.26 implies $z \in S_+ \cap S_-$, whence $z \in S$ by Proposition 1.25(b). Thus the open set W is contained in $\operatorname{Int} S$, and we have proved $\operatorname{Int} S$ is dense. It follows that $\operatorname{Int} S \cap \operatorname{Int} Q$ is dense. □

2.3. *Improving the limit set dichotomy for smooth systems*

The aim now is to strengthen the Limit Set Dichotomy with additional hypotheses, especially smoothness, in order to obtain the following property:

(ILSD) A semiflow satisfies the *Improved Limit Set Dichotomy* if $x_1 < x_2$ implies that either
- (a) $\omega(x_1) < \omega(x_2)$, or
- (b) $\omega(x_1) = \omega(x_2) = e \in E$.

We begin with some definitions.

Let X be a subset of the Banach space Y. A map $f : X \to Y$ is said to be *locally C^1 at* $p \in X$ if there exists a neighborhood U of p in X and a continuous *quasiderivative* map $f' : U \to L(Y)$, where $L(Y)$ is the Banach space of bounded operators on Y, such that

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + \phi(x, x_0)|x - x_0|, \quad x, x_0 \in U$$

with $\phi(x, x_0) \to 0$ as $x \to x_0$. The following result gives a setting where the quasiderivative is uniquely determined by f. We denote the open ball in Y of center p and radius r by $B_Y(p, r) := \{y \in Y\colon |y - p| < r\}$.

LEMMA 2.15. *Let $p \in X \subset Y$ where Y is a strongly ordered Banach space. Assume $f : X \to Y$ is locally C^1 at p, and suppose that either $B_Y(p, r) \cap Y_+ \subset X$ or $B_Y(p, r) \cap (-Y_+) \subset X$ for some $r > 0$. Then $f'(p)$ is uniquely defined.*

PROOF. Suppose $B_Y(p, r) \cap Y_+ \subset X$, the other case being similar. Fix $w \gg 0$ and let $y \in Y$. As $w + y/n := k_n \geqslant 0$ for large n, $y = n(k_n - w)$ so $Y = Y_+ - Y_+$.

Assume

$$f(x) - f(p) = A(x - p) + \phi(x, p)|x - p| = B(x - p) + \psi(x, p)|x - p|,$$

where $A, B \in L(Y)$ and $\phi, \psi \to 0$ as $x \to p$ in X. It suffices to show that $Av = Bv$ for all $v \geqslant 0$. The segment $x = p + sv \in X$ for all small $s \geqslant 0$. Inserting it in the formula above, dividing by s, and letting $s \to 0$ yields the desired result. □

Let Φ be a monotone semiflow on the subset X of the strongly ordered Banach space Y. Concerning X and the set of equilibria E, we assume the following condition on the pair (Y, X):

(OC) Either X is an order convex subset of Y or $E \subset \operatorname{Int} X$. For each $e \in E$ there exists $r > 0$ such that either $B_Y(e, r) \cap Y_+ \subset X$ or $B_Y(e, r) \cap (-Y_+) \subset X$.

This relatively minor restriction is automatically satisfied if X is an open set, an order interval, or the cone Y_+. The second assertion of (OC) trivially holds if $E \subset \operatorname{Int} X$.

We will also assume the following two conditions hold for some $\tau > 0$. A compact, strongly positive linear operator is called a *Krein–Rutman* operator.

(M) $x_1 < x_2 \Longrightarrow \Phi_\tau(x_1) \ll \Phi_\tau(x_2)$
(D*) Φ_τ is locally C^1 at each $e \in E$, with $\Phi'_\tau(e)$ a Krein–Rutman operator.

As motivation for (D*), consider the case that X is an open set in Y and Φ_τ is C^1. If $x \in X$, $y \in Y_+$, $h > 0$, and $x + hy \in X$, then $(\Phi_\tau(x+hy) - \Phi_\tau(x))/h \geqslant 0$ by monotonicity; on taking the limit as $h \to 0$, we get $\Phi'_\tau(x)y \geqslant 0$. Consequently, $\Phi'_\tau(x)Y_+ \subset Y_+$, and hence the assumption that $\Phi'_\tau(x)$ is strongly positive is not such a severe one. Typically, one usually must verify it anyway to prove that Φ_τ is strongly monotone.

Observe that (M) implies that Φ is strongly order preserving on X.

THEOREM 2.16 (Improved Limit Set Dichotomy). *Let Φ be a monotone semiflow on a subset X of the strongly ordered Banach space Y for which* (OC), (M), *and* (D*) *are satisfied. Then* (ILSD) *holds.*

In particular, (ILSD) holds if X is open, the semiflow Φ continuously differentiable and strongly monotone, and the derivative $\Phi'_t(e)$ is a Krein–Rutman operator at each $e \in E$.

Before giving the proof, we explore the spectral and dynamical implications of (D*).

2.3.1. *The Krein–Rutman theorem* The spectrum of a linear operator $A : Y \to Y$ is denoted by $\mathrm{Spec}(A)$. When A is compact (i.e., completely continuous), $\mathrm{Spec}(A)$ consists of a countable set of eigenvalues and perhaps 0, and the eigenvalues have no accumulation point except possibly 0.

Let $\rho(A)$ be the *spectral radius* of A, that is, $\rho(A) = \max\{|\lambda|: \lambda \in \mathrm{Spec}(A)\}$. Denote the null space of A by $N(A)$ and the range by $\mathrm{Im}(A)$.

The set $\mathrm{KR}(Y)$ of Krein–Rutman operators on Y is given the metric induced by the uniform norm.

THEOREM 2.17 (Krein–Rutman). *Let $A \in \mathrm{KR}(Y)$ and set $r = \rho(A)$. Then Y decomposes into a direct sum of two closed invariant subspaces Y_1 and Y_2 such that $Y_1 = N(A - rI)$ is spanned by $z \gg 0$ and $Y_2 \cap Y_+ = \{0\}$. Moreover, the spectrum of $A|Y_2$ is contained in the closed ball of radius $\nu < r$ in the complex plane.*

See Krein and Rutman [104], Takáč [214] or Zeidler [244] for proofs.

It follows that each $A \in \mathrm{KR}(Y)$ has a unique unit eigenvector $z(A) \in Y_+$, and $z(A) \in \mathrm{Int}\, Y_+$, $Az(A) = \rho(A)z(A)$.

LEMMA 2.18. *$\rho(A)$ and $z(A)$ are continuous functions of $A \in \mathrm{KR}(Y)$.*

PROOF. The upper semicontinuity of the spectral radius follows from the upper semicontinuity of the spectrum as a function of the operator (Kato [92]). The lower semicontinuity follows from the lower semicontinuity of isolated parts of the spectrum (Kato [92, Chapter IV, Theorem 3.1, Remark 3.3, Theorem 3.16]). Let P_A be the projection onto $N(A) - \rho(A)I)$ along $\operatorname{Im}(A - \rho(A)I)$. Continuity of $A \mapsto P_A$ is proved in [92, Chapter IV, Theorem 3.16]. Let $A_n \to A$ in $\mathrm{KR}(Y)$ and set $z_n = z(A_n)$, $z = z(A)$. Then $(I - P_A))z_n = (P_{A_n} - P_A)z_n \to 0$ as $n \to \infty$, and $\{P_A z_n\}$ is precompact, so $\{z_n\}$ is precompact. If $z_{n_i} \to u$ for some subsequence, then

$$P_A u = \lim A_{n_i} z_{n_i} = \lim \rho(A_{n_i}) z_{n_i} = \rho(A)u.$$

Uniqueness of the positive eigenvector for A (Theorem 2.17) implies $u = z$ and $z_n \to z$. □

For technical reasons it is useful to employ a norm that is more compatible with the partial order. If $w \in \operatorname{Int} Y_+$ is fixed, then the set $U = \{y \in Y\colon -w \ll y \ll w\}$ is an open neighborhood of the origin. Consequently, if $y \in Y$, then there exists $t_0 > 0$ such that $t_0^{-1} y \in U$, hence, $-t_0 w \ll y \ll t_0 w$. Define the w-norm by

$$\|y\|_w = \inf\{t > 0\colon -tw \leqslant y \leqslant tw\}.$$

Since $w \in \operatorname{Int} Y_+$, there exists $\delta > 0$ such that for all $y \in Y \setminus \{0\}$ we have $w \pm \delta \frac{y}{|y|} \in Y_+$. Thus

$$\|y\|_w \leqslant \delta^{-1} |y|$$

holds for all $y \in Y$, implying that the w-norm is weaker than the original norm. In fact, the two norms are equivalent if Y_+ is normal, but we will have no need for this result. See Amann [6] and Hirsch [73] for more results in this direction. It will be useful to renormalize the positive eigenvector $z(A)$ for $A \in \mathrm{KR}(Y)$. The next result says this can be done continuously. Continuity always refers to the original norm topology on Y unless the contrary is explicitly stated.

LEMMA 2.19. *Let* $Z(A) = z(A)/\|z(A)\|_w$ *and* $\beta(A) = \sup\{\beta > 0\colon Z(A) \geqslant \beta w\}$. *Then* $\beta(A) > 0$, $Z(A) \geqslant \beta(A)w$, *and the maps* $A \to Z(A)$ *and* $A \to \beta(A)$ *are continuous on* $\mathrm{KR}(Y)$.

PROOF. Since the w-norm is weaker than the original norm, the map $A \mapsto \|z(A)\|_w$ is continuous. This implies that $Z(A)$ is continuous in A. It is easy to see that $\beta(A) > 0$. Let $\epsilon > 0$ satisfy $2\epsilon < \beta(A)$ and let $A_n \to A$ in $\mathrm{KR}(Y)$. Then $-\epsilon w \leqslant Z(A) - Z(A_n) \leqslant \epsilon w$ for all large n by continuity of Z and because the w-norm is weaker than the original norm. Therefore, $Z(A_n) = Z(A_n) - Z(A) + Z(A) \geqslant (\beta(A) - \epsilon)w$, so $\beta(A_n) \geqslant \beta(A) - \epsilon$ for all large n. Similarly, $Z(A) = Z(A) - Z(A_n) + Z(A_n) \geqslant (\beta(A_n) - \epsilon)w$ for all large n, so $\beta(A) \geqslant \beta(A_n) - \epsilon$ for all large n. Thus, $\beta(A) - \epsilon \leqslant \beta(A_n) \leqslant \beta(A) + \epsilon$ holds for all large n, completing the proof. □

The key to improving the Limit Set Dichotomy is to show that the omega limit set of a point x that is quasiconvergent but not convergent, is uniformly unstable in the linear approximation. The direction of greatest instability at $e \in \omega(x)$ is the positive direction $z(e) := z(\Phi'_\tau(e))$. The number $\rho(e) := \rho(\Phi'_\tau(e))$ gives a measure of the instability.

Nonordering of Limit Sets means that positive directions are, in some rough sense, "transverse" to the limit set. Thus our next result means that the limit set is uniformly unstable in a transverse direction.

LEMMA 2.20. *Assume* (D*). *Let x be quasiconvergent but not convergent. Then $\rho(e) > 1$ for all $e \in \omega(x)$.*

PROOF. Fix $e \in \omega(x)$. Since $\omega(x)$ is connected, e is the limit of a sequence $\{e_n\}$ in $\omega(x) \cap U \setminus \{e_0\}$, where U is the neighborhood of e in the definition of Φ_τ is locally C^1 at e. Then

$$e_0 - e_n = \Phi_\tau(e) - \Phi_\tau(e_n) = \Phi'_\tau(e)(e - e_n) + \mathrm{o}\big(|e - e_n|\big),$$

where $\mathrm{o}(|e - e_n|)/|e - e_n| \to 0$ as $n \to \infty$. Put $v_n = (e - e_n)/|e - e_n|$. Then

$$v_n = \Phi'_\tau(e)v_n + r_n, \quad r_n \to 0, \quad n \to \infty.$$

The compactness of $\Phi'_\tau(e)$ implies that v_n has a convergent subsequence v_{n_i}; passing to the limit along this subsequence leads to $v = \Phi'_\tau(e)v$ for some unit vector v. Thus $\rho(e) \geqslant 1$. If $\rho(e) = 1$, then the Krein–Rutman Theorem implies $v = rz(e)$ where $r = \pm 1$. Consequently,

$$(e - e_{n_i})/|e - e_{n_i}| \to rz(e)$$

as $i \to \infty$. It follows that $e \ll e_{n_i}$ or $e \gg e_{n_i}$ for all large i, contradicting the Nonordering of Limit Sets. □

PROOF OF THEOREM 2.16. By the Limit Set Dichotomy (Theorem 1.16), it suffices to prove: If $x_1 < x_2$ and $\omega(x_1) = \omega(x_2) = K \subset E$, then K is a singleton. K is compact and connected, unordered by the Nonordering of Limit Sets, and consists of fixed points of Φ_τ. Arguing by contradiction, we assume K is not a singleton.

Set $v_n = \Phi_{n\tau}(x_1)$, $u_n = \Phi_{n\tau}(x_2)$. Then $\mathrm{dist}(K, u_n) \to 0$ and $\mathrm{dist}(K, v_n) \to 0$ as $n \to \infty$. Moreover (M) and the final assertion of the Limit Set Dichotomy imply

$$u_n - v_n \gg 0, \qquad u_n - v_n \to 0.$$

Fix $w \gg 0$ and define real numbers

$$\alpha_n = \sup\{\alpha \in \mathbb{R}\colon \alpha \geqslant 0,\ \alpha w \leqslant u_n - v_n\}.$$

Then $\alpha_n > 0$ and $\alpha_n \to 0$.

To simplify notation, define $S : X \to X$ by $S(x) := \Phi_\tau(x)$. Choose $e_n \in K$ such that $v_n - e_n \to 0$ as $n \to \infty$. By Lemma 2.20, local smoothness of Φ_τ and compactness of K, there exists $r > 1$ such that $\rho(e) > r$ for all $e \in K$. Let $z_n = Z(e_n)$ be the normalized positive eigenvector for $S'(e_n) = \Phi'_\tau(e_n)$ so $\|z_n\|_w = 1$ and $z_n \leqslant w$. By Lemma 2.19, there exists $\epsilon > 0$ such that $\beta(e_n) \geqslant \epsilon$ for all n. In particular, $w \geqslant z_n \geqslant \epsilon w$ for all n.

Fix a positive integer l such that $r^l\epsilon > 1$.

For each $e \in K$, by (D*) we can choose an open neighborhood W_e of e in X and a continuous map $S' : W_e \to L(Y)$ such that for $x, x_0 \in W_e$ we have

$$Sx - Sx_0 = S'(x_0)(x - x_0) + \phi(x, x_0)|x - x_0|, \qquad \lim_{x \to x_0} \phi(x, x_0) = 0.$$

Putting $x_0 = e$ and estimating norms, one easily sees that there exists a convex open neighborhood $U_e \subset W_e$ of e such that $S^i(U_e) \subset W_e$ for $1 \leqslant i \leqslant l$. Furthermore, a simple induction argument implies that S^l is locally C^1 at e with quasiderivative

$$\bigl(S^l\bigr)' : U_e \to L(Y), \qquad \bigl(S^l\bigr)'(x) = S'\bigl(S^{l-1}x\bigr) \circ S'\bigl(S^{l-2}x\bigr) \circ \cdots \circ S'(x).$$

By compactness of K there is a finite subset $\{e_1, \dots, e_\nu\} \subset K$ such that the sets U_{e_j} cover K. Set $U_j = U_{e_j}$, $W_j = W_{e_j}$. Then

$$K \subset \bigcup_{j=1}^{\nu} U_j, \qquad S^i(U_j) \subset W_j \quad (1 \leqslant i \leqslant l),$$

and for $z, z_0 \in U_j$

$$S^l(z) - S^l(z_0) = \bigl(S^l\bigr)'(z_0)(z - z_0) + \phi_{l,j}(z, z_0)|z - z_0|, \qquad \lim_{z \to z_0} \phi_{l,j}(z, z_0) = 0,$$

and the usual chain rule expresses $(S^l)'$ in terms of S'.

By (OC), either X is order convex in Y or $E \subset \operatorname{Int} X$. In the order convex case, from $v_n \ll v_n + \alpha_n w \leqslant u_n$ we infer that $v_n + s\alpha_n w \in X$ for all $s \in [0, 1]$. Since $v_n - e_n \to 0$, $v_n - u_n \to 0$, and $\alpha_n \to 0$, for sufficiently large n there exists $j(n) \in \{1, \dots, \nu\}$ such that $U_{j(n)}$ contains the points v_n, u_n, e_n, and $v_n + s\alpha_n w$ for all $s \in [0, 1]$. When $E \subset \operatorname{Int} X$ the same conclusion holds, and we can take U_j, V_j to be open in Y.

Lemma 2.15 justifies the application of the fundamental theorem of calculus to the map $[0, 1] \to X$, $s \mapsto S^l(v_n + s\alpha_n w)$, leading to

$$S^l(v_n + \alpha_n w) - S^l(v_n) = \bigl(S^l\bigr)'(e_n)(\alpha_n w) + \alpha_n \delta_n$$

and

$$\delta_n = \int_0^1 \bigl[\bigl(S^l\bigr)'(v_n + \eta\alpha_n w) - \bigl(S^l\bigr)'(e_n)\bigr] w \, \mathrm{d}\eta.$$

Using that $v_n + \alpha_n w - e_n \to 0$, K is compact, and $(S^l)'$ is continuous, it is easy to show that

$$\lim_{n \to \infty} \max_{0 \leqslant \eta \leqslant 1} \bigl|\bigl[\bigl(S^l\bigr)'(v_n + \eta\alpha_n w) - \bigl(S^l\bigr)'(e_n)\bigr] w\bigr| = 0.$$

It follows that $d_n := \|\delta_n\|_w \to 0$ as $n \to \infty$. Because $w \geqslant z_n \geqslant \epsilon w \gg 0$ and $\delta_n \geqslant -d_n w$, for sufficiently large n we have:

$$\begin{aligned} S^l(v_n + \alpha_n w) - S^l(v_n) &\geqslant \big[(S^l)'(e_n)\big]\alpha_n w - \alpha_n d_n w \\ &\geqslant \big[S'(e_n)\big]^l \alpha_n w - \alpha_n d_n w \\ &\geqslant r^l \alpha_n z_n - \alpha_n d_n w \\ &\geqslant (r^l\epsilon - d_n)\alpha_n w \\ &\geqslant \alpha_n w, \end{aligned}$$

and therefore

$$u_{n+l} = S^l(u_n) \geqslant S^l(v_n + \alpha_n w) \geqslant S^l v_n + \alpha_n w = v_{n+l} + \alpha_n w.$$

Thus $\alpha_n w \leqslant u_{n+l} - v_{n+l}$, so the definition of α_{n+l} implies $\alpha_{n+l} \geqslant \alpha_n > 0$ for all sufficiently large n. Therefore the sequence $\{\alpha_i\}_{i\in\mathbb{N}_+}$, which converges to 0, contains a nondecreasing positive subsequence $\{\alpha_{n+kl}\}_{k\in\mathbb{N}_+}$. This contradiction implies K is a singleton. □

A drawback of the Improved Limit Set Dichotomy, Theorem 2.16, is that the topology on X comes from a strongly ordered Banach space $Y \supset X$, severely limiting its application to infinite-dimensional systems. The following extension permits use of (ILSD) in more general spaces:

PROPOSITION 2.21. *Let X^1, X^0 be ordered spaces such that $X^1 \subset X^0$ and the inclusion map $j : X^1 \hookrightarrow X^0$ is continuous and order preserving. For $k = 0, 1$ let Φ^k be a monotone semiflow on X^k with compact orbit closures. Assume for all $t > 0$ that Φ^0_t maps X^0 continuously into X^1, and $\Phi^0_t|X^1 = \Phi^1_t$. If* (ILSD) *holds for Φ^1, it also holds for Φ^0.*

PROOF. Denote the closure in X^k of any $S \subset X^k$ by $\mathsf{C}_k S$. For $k \in \{0, 1\}$ and $x \in X^k$, let $O_k(x)$ and $\omega_k(x)$ respectively denote the orbit and omega limit set of x.

The hypotheses imply that the compact set $\mathsf{C}_0 O_0(x)$, which is positively invariant for Φ^0, is mapped homeomorphically by Φ^0_1 onto $\mathsf{C}_1 O_1(y) \subset X^1$, which is positively invariant for Φ^1. As Φ^0 and Φ^1 coincide in X^1, we see that $\omega_0(x) = \omega_1(y)$ as compact sets. Hence Φ^0 and Φ^1 have the same collection of omega limit sets, which implies the conclusion. □

THEOREM 2.22 (Improved Sequential Limit Set Trichotomy). *Assume* (ILSD). *Let $\{\tilde{x}_n\}$ be a sequence approximating $z \in BC$ from below, with $\overline{\bigcup_n \omega(\tilde{x}_n)}$ compact. Then there is a subsequence $\{x_n\}$ such that exactly one of the following three conditions holds for all n:*

(a) *There exists $u_0 \in E$ such that*

$$\omega(x_n) < \omega(x_{n+1}) < \omega(z) = \{u_0\}$$

and

$$\lim_{n\to\infty} \operatorname{dist}\big(\omega(x_n), u_0\big) = 0.$$

(b) *There exists* $u_1 = \sup\{u \in E\colon\ u < \omega(z)\}$, *and*

$$\omega(x_n) = \{u_1\} < \omega(z).$$

In this case $z \in \overline{\operatorname{Int} C}$. *Moreover* z *has a neighborhood* W *such that if* $w \in W$, $w < z$ *then* $\Phi_t(w) \to u_1$ *and* $\Phi_t(w) > u_1$ *for sufficiently large* t.

(c′) *There exists* $u_2 \in E$ *such that* $\omega(x_n) = \omega(x_0) = u_2$.

Note that z is convergent in (a), strongly accessible from below by convergent points in (b), and convergent in (c′).

PROOF. Conclusions (a) and (b) are the same as in the Sequential Limit Set Trichotomy, Theorem 2.3. If 2.3(c) holds, then (c′) follows from (ILSD). □

PROPOSITION 2.23. *Assume* (ILSD). *If* $x \in BC \setminus C$ *then* $\omega(x)$ *is unstable from below. If* $x \in AC \setminus C$ *then* $\omega(x)$ *is unstable from above.*

PROOF. This is just Theorem 2.6 if $x \notin Q$. If $x \in BC \cap (Q \setminus C)$, we must have conclusion (b) of Theorem 2.22. This provides $u_1 \in E$ such that $\omega(x_n) = \{u_1\}$ for all n, and the remainder of the proof mimics that of Theorem 2.6. □

A consequence of Proposition 2.23 is that if $x \in BC \cap AC$ is nonconvergent, then $\omega(x)$ lies in both the upper boundary of the basin of attraction of an equilibrium u_0 and the lower boundary of the basin of attraction of an equilibrium v_0, where $u_0 < L < v_0$. Thus $\omega(x)$ forms part of a separatrix separating the basins of attraction of u_0 and v_0.

2.4. *Generic convergence and stability*

The following result concludes that the set C of convergent points is dense and open in totally ordered arcs:

THEOREM 2.24. *Assume* (ILSD) *and let* $J \subset X$ *be a totally ordered arc having property* (C). *Then* $J \setminus C$ *is a discrete, relatively closed subset of* J; *hence it is countable, and finite when* J *is compact.*

PROOF. The proof is like that of Theorem 2.8, using the Improved Limit Set Trichotomy 2.22 instead of the Sequential Limit Set Trichotomy 2.3. □

We can now prove the following generic convergence and stability results:

THEOREM 2.25. *Assume* (ILSD).

(a) $AC \cup BC \subset \overline{\operatorname{Int} C} \cup C$. *In particular, if* $AC \cup BC$ *is dense, so is* $\operatorname{Int} C$ *is dense.*

(b) *If* $\operatorname{Int}(BC \cap AC)$ *is dense and* X *is normally ordered, then* $\operatorname{Int}(C \cap S)$ *is dense.*

PROOF. The proof of (a) is similar to that of Theorem 2.12: take $p \in X \setminus \operatorname{Int} C$ and use the Improved Limit Set Trichotomy (Theorem 2.22), instead of the Limit Set Trichotomy, to show that $p \in \overline{\operatorname{Int} C} \cup C$. Conclusion (b) follows from (a) and Theorem 2.14. □

THEOREM 2.26. *Assume X is a subset of a strongly ordered Banach space Y, and a dense open subset of X is covered by totally ordered line segments. Let* (M) *and* (D*) *hold. Then*:

(a) *The set of convergent points has dense interior.*

(b) *Suppose Y is normally ordered. Then the set of stable points has dense interior.*

(c) *Assume Y is normally ordered*; *X is open or order convex or a subcone of Y_+*; *and every closed totally ordered subset of E is compact. Then there is a stable equilibrium, and an asymptotically stable equilibrium when E is finite.*

PROOF. The assumption in (a) implies $BC \cap AC$ has dense interior and condition (OC) holds. Therefore the Improved Limit Set Dichotomy (ILSD) holds by Theorem 2.16, so (a) and (b) follow from Theorem 2.25(a). Conclusion (c) is a consequence of (a) and Theorem 1.30. □

As most orbits with compact closure converge to an equilibrium, it is natural to investigate the nature of the convergence. It might be expected that most trajectories converging to a stable equilibrium are eventually increasing or decreasing. We quote a theorem of Mierczyński that demonstrates this under quite general conditions for smooth strongly monotone dynamical systems, including cases when the equilibrium is not asymptotically stable in the linear approximation. Mierczyński assumes the following hypothesis:

(M_1) X is an open set in a strongly ordered Banach space Y. Φ is C^1 on $(0,\infty) \times X$ and strongly monotone, $\Phi_t'(x)$ is strongly positive for all $t > 0, x \in X$, and $\Phi_1'(x)$ is compact.

The following local trichotomy due to Mierczyński [138] builds on earlier work of Poláčik [161]:

THEOREM 2.27. *Assume* (M_1). *Then each equilibrium e satisfying $\rho(\Phi_1'(e)) \leqslant 1$ belongs to a locally invariant submanifold Σ_e of codimension one that is smooth and unordered and has the following property. If $\lim_{t\to\infty} \Phi_t(x) = e$, there exists $t_0 \geqslant 0$ such that one of the following holds as $t \to \infty$, $t \geqslant t_0$*:

(i) *$\Phi_t(x)$ decreases monotonically to e*;

(ii) *$\Phi_t(x)$ increases monotonically to e*;

(iii) *$\Phi_t(x) \in \Sigma_e$.*

Mierczyński also provides further important information: The trajectories in cases (i) and (ii) lie in curves tangent at e to the one-dimensional principle eigenspace Y_1 of $\Phi_1'(e)$ described in the Krein–Rutman Theorem 2.17. The hypersurface Σ_e is locally unique in a neighborhood of e. Its tangent space is the closed complementary subspace Y_2, hence Σ_e is transverse to $z = z(\Phi_1'(e)) \gg 0$ at e. Strong monotonicity implies that when (i) or (ii) holds, e is asymptotically stable for the induced local flow in Σ_e, even when e is not stable.

2.4.1. *Background and related results* Smith and Thieme [197,199] introduced the compactness hypothesis (C) and obtained the Sequential Limit Set Trichotomy. This tool streamlines many of the arguments and leads to stronger conclusions so the presentation here follows [197,199]. Takáč [210] extends the compactness hypothesis, which leads to additional stability concepts.

The results of Smith and Thieme [199] on generic convergence for SOP semiflows were motivated by earlier work of Poláčik [160], who obtained such results for abstract semilinear parabolic evolution systems assuming less compactness but more smoothness than Smith and Thieme.

The set A of asymptotically stable points can be shown to be dense under suitable hypotheses. See, e.g., Hirsch [73, Theorem 9.6]; Smith and Thieme [197, Theorems 3.13 and 4.1].

Hirsch [69] shows that if K is a nonempty compact, invariant set that attracts all points in some neighborhood of itself, then K contains an order-stable equilibrium.

It is not necessary to assume, as we have done here, that the semiflow is globally defined, that is, that trajectories are defined for all $t \geqslant 0$; many of the results adapt to local semiflows. See Hirsch [73], Smith and Thieme [199].

3. Ordinary differential equations

Throughout this section $\mathbb{R}^n$ is ordered by a cone K with nonempty interior. Our first objective is to explore conditions on a vector field that make the corresponding local semiflow monotone with respect to the order defined by K. It is convenient to work with time-dependent vector fields. We then investigate the long-term dynamics of autonomous vector fields f that are K-cooperative, meaning that K is invariant under the forward flow of the linearized system. These results are applied to competitive vector fields by the trick of time-reversal. In fairly general circumstances, limit sets of cooperative or competitive systems in $\mathbb{R}^n$ are invariant sets for systems in $\mathbb{R}^{n-1}$. This leads to particularly sharp theorems for $n = 2$ and 3.

A cone is *polyhedral* if it is the intersection of a finite family of closed half spaces. For example, the standard cone $\mathbb{R}^n_+$ is polyhedral, while the ice-cream cone is not.

The *dual cone* to K is the closed cone K^* in the dual space $(\mathbb{R}^n)^*$ of linear functions on $\mathbb{R}^n$, defined by

$$K^* = \left\{\lambda \in \left(\mathbb{R}^n\right)^*\colon\ \lambda(K) \geqslant 0\right\}.$$

To $\lambda \in K^*$ we associate the vector $a \in \mathbb{R}^n$ such that $\lambda(x) = \langle a, x\rangle$ where $\langle a, x\rangle$ denotes the standard inner product on $\mathbb{R}^n$. Under this association K^* is canonically identified with a cone in $\mathbb{R}^n$, namely, the set of vectors a such that a is normal to a supporting hyperplane H of K, and a and K lie in a common halfspace bounded by H.

We use the following simple consequence of general results on the separation of two closed convex sets:

$$x \in K \quad \Longleftrightarrow \quad \lambda(x) \geqslant 0 \quad (\lambda \in K^*).$$

See, e.g., Theorem 1.2.8 of Berman et al. [18].

PROPOSITION 3.1. *If $x \in K$, then $x \in \operatorname{Int} K$ if and only if $\lambda(x) > 0$ for all $\lambda \in K^* \setminus \{0\}$.*

PROOF. Suppose $x \in \operatorname{Int} K$, $\lambda \in K^* \setminus \{0\}$, and $v \in X$ satisfies $\lambda(v) \neq 0$. Then $x \pm \epsilon v \in K$ for sufficiently small $\epsilon > 0$, so

$$\lambda(x \pm \epsilon v) = \lambda(x) \pm \epsilon \lambda(v) \geqslant 0,$$

implying that $\lambda(x) > 0$.

To prove the converse, assume $\mu(x) > 0$ for all functionals μ in the compact set $\Gamma = \{\lambda \in K^*\colon \|\lambda\| = 1\}$. As $\inf\{\mu(x)\colon \mu \in \Gamma\} > 0$, continuity of the map $(x, \lambda) \mapsto \lambda(x)$ implies $\mu(y) > 0$ for all y in some neighborhood U of x and all $\mu \in \Gamma$. If $\lambda \in K^*$ then $\|\lambda\|^{-1}\lambda \in \Gamma$ and therefore $\lambda(y) > 0$ for all $y \in U$. This proves $U \subset K$. □

An immediate consequence of Proposition 3.1 is that if $x \in \partial K$, then there exists a nontrivial $\lambda \in K^*$ such that $\lambda(x) = 0$.

3.1. *The quasimonotone condition*

Let $J \subset \mathbb{R}$ be a nontrivial open interval, $D \subset \mathbb{R}^n$ an open set and $f : J \times D \to \mathbb{R}^n$ a locally Lipschitz function. We consider the ordinary differential equation

$$x' = f(t, x). \tag{3.1}$$

For every $(t_0, x_0) \in J \times D$, the initial value problem $x(t_0) = x_0$ has a unique noncontinuable solution defined on an open interval $J(t_0, x_0) \subset \mathbb{R}$. We denote this solution by $t \mapsto x(t, t_0, x_0)$. The notation $x(t, t_0, x_0)$ will carry the tacit assumption that $(t_0, x_0) \in J \times D$ and $t \in J(t_0, x_0)$. For fixed s_0, t_0 the map $x_0 \mapsto x(s_0, t_0, x_0)$ is a homeomorphism between open subsets of $\mathbb{R}^n$, the inverse being $x_0 \mapsto x(t_0, s_0, x_0)$.

System (3.1) is called *monotone* if $x_0 \leqslant x_1 \Longrightarrow x(t, t_0, x_0) \leqslant x(t, t_0, x_1)$.

The time-dependent vector field $f : J \times D \to \mathbb{R}^n$ satisfies the *quasimonotone condition* in D if for all $(t, x), (t, y) \in J \times D$ and $\phi \in K^*$ we have:

(QM) $x \leqslant y$ and $\phi(x) = \phi(y)$ implies $\phi(f(t, x)) \leqslant \phi(f(t, y))$.

The quasimonotone condition was introduced by Schneider and Vidyasagar [177] for finite-dimensional, autonomous linear systems and used later by Volkmann [224] for nonlinear infinite-dimensional systems. The following result is inspired by a result of Volkmann [224] and work of W. Walter [227]. See also Uhl [221], Walcher [226].

THEOREM 3.2. *Assume f satisfies* (QM) *in D, $t_0 \in J$, and $x_0, x_1 \in D$. Let $\prec$ denote any one of the relations $\leqslant, <, \ll$. If $x_0 \prec x_1$ then $x(t, t_0, x_0) \prec x(t, t_0, x_1)$, hence* (3.1) *is monotone. Conversely, if* (3.1) *is monotone then f satisfies* (QM).

PROOF. Assume that $x(t, t_0, x_i)$, $i = 0, 1$ are defined for $t \in [t_0, t_1]$ and $x_0 \leqslant x_1$. Let $v \gg 0$ be fixed and define $x_\epsilon := x_1 + \epsilon v$ and $f_\epsilon(t, x) := f(t, x) + \epsilon v$ for $\epsilon > 0$. Denote by $x(t) := x(t, t_0, x_0)$ and let $y_\epsilon(t) := x(t, t_0, x_\epsilon, \epsilon)$ denote the solution of the initial value problem $x'(t) = f_\epsilon(t, x)$, $x(t_0) = x_\epsilon$. It is well known that $y_\epsilon(t)$ is defined on $[t_0, t_1]$ for all sufficiently small ϵ. We show that $x(t) \ll y_\epsilon(t)$ for $t_0 \leqslant t \leqslant t_1$ and all sufficiently small $\epsilon > 0$. If not, then as $x(t_0) \ll y_\epsilon(t_0)$, there would exist $\epsilon > 0$ and $s \in (t_0, t_1]$ such that $x(t) \ll y_\epsilon(t)$ for $t_0 \leqslant t < s$ and $y_\epsilon(s) - x(s) \in \partial K$. By Proposition 3.1, there exists a nontrivial $\phi \in K^*$ such that $\phi(y_\epsilon(s) - x(s)) = 0$ but $\phi(y_\epsilon(t) - x(t)) > 0$ for $t_0 \leqslant t < s$. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigl[\phi\bigl(y_\epsilon(t)\bigr) - \phi\bigl(x(t)\bigr)\bigr]\bigg|_{t=s} \leqslant 0,$$

hence

$$\phi\bigl(f\bigl(s, y_\epsilon(s)\bigr)\bigr) < \phi\bigl(f\bigl(s, y_\epsilon(s)\bigr)\bigr) + \epsilon\phi(v) = \phi\bigl(f_\epsilon\bigl(s, y_\epsilon(s)\bigr)\bigr) \leqslant \phi\bigl(f\bigl(s, x(s)\bigr)\bigr),$$

where the last inequality follows from the one above. On the other hand, by (QM) we have

$$\phi\bigl(f\bigl(s, y_\epsilon(s)\bigr)\bigr) \geqslant \phi\bigl(f\bigl(s, x(s)\bigr)\bigr).$$

This contradiction proves that $x(t) \ll y_\epsilon(t)$ for $t_0 \leqslant t \leqslant t_1$ and all small $\epsilon > 0$. Since $y_\epsilon(t) = x(t, t_0, x_\epsilon, \epsilon) \to x(t, t_0, x_1)$ as $\epsilon \to 0$, by taking the limit we conclude that $x(t, t_0, x_0) \leqslant x(t, t_0, x_1)$ for $t_0 \leqslant t \leqslant t_1$.

Fix t_0 and $t \in J(t_0, x_0)$. As the map $h : x_0 \mapsto x(t, t_0, x_0)$ is injective, from $x_0 < x_1$ we infer $x(t, t_0, x_0) < x(t, t_0, x_1)$. Note that $h(D \cap [x_0, x_1]) \subset [x(t, t_0, x_0), x(t, t_0, x_1)]$. Therefore the relation $x_0 \ll x_1$ implies $\operatorname{Int} D \cap [x_0, x_1] \neq \emptyset$. Injectivity of h and invariance of domain implies $\operatorname{Int}[x(t, t_0, x_0), x(t, t_0, x_1)] \neq \emptyset$, which holds if and only if $x(t, t_0, x_0) \ll x(t, t_0, x_1)$.

Conversely, suppose that (3.1) is monotone, $t_0 \in J$, $x_0, x_1 \in D$ with $x_0 \leqslant x_1$ and $\phi(x_0) = \phi(x_1)$ for some $\phi \in K^*$. Since $x(t, t_0, x_0) \leqslant x(t, t_0, x_1)$ for $t \geqslant t_0$ we conclude that $\frac{\mathrm{d}}{\mathrm{d}t}\phi[x(t, t_0, x_1) - x(t, t_0, x_0)]|_{t=t_0} \geqslant 0$, or $\phi(f(t_0, x_1)) \geqslant \phi(f(t_0, x_0))$. Thus (QM) holds. □

Theorem 3.2 has been stated so as to minimize technical details concerning the domain $J \times D$ by assuming that J and D are open. In many applications, D is a closed set, for example, $D = K$ or $D = [a, b]$ where $a \ll b$. The proof can be modified to handle these (and other) cases. If $D = K$ and K is positively invariant for (3.1), the proof is unchanged because whenever $x \in D$ then $x + \epsilon v \in D$ for small positive ϵ, and because K is also positively invariant for the modified equation. If $D = [a, b]$, then the result follows by applying Theorem 3.2 to $f|J \times [[a, b]]$ and using continuity.

A set S is called *positively invariant* under (3.1) if $S \subset D$ and solutions starting in S stay in S, or more precisely:

$$(t_0, x_0) \in J \times S \quad \text{and} \quad t \in J(t_0, x_0),\ t \geqslant t_0 \quad \Longrightarrow \quad x(t, t_0, x_0) \in S.$$

It will be useful to have the following necessary and sufficient condition for invariance of K:

PROPOSITION 3.3. *The cone K is positively invariant under* (3.1) *if and only if $K \subset D$ and for each $t \in J$*

(P) $\lambda \in K^*, x \in \partial K, \lambda(x) = 0 \Longrightarrow \lambda(f(t,x)) \geqslant 0$.

PROOF. The proof that (P) implies positive invariance of K is similar to that of Theorem 3.2. Given $x_1 \in K$, we pass immediately to $x_\epsilon \gg x_1$ and the solution $y_\epsilon(t)$ of the perturbed equation defined in the proof of Theorem 3.2 and show that $y_\epsilon(t) \gg 0$ for $t_0 \leqslant t \leqslant t_1$ by an argument similar to the one used in the aforementioned proof. The result $x(t, t_0, x_1) \geqslant 0$ for $t \geqslant t_0$ is obtained by passage to the limit as $\epsilon \to 0$. The converse is also an easy modification of the converse argument given in the proof of Theorem 3.2. □

Since we will have occasion to apply (P) to systems other than (3.1), it will be convenient to refer to (P) by saying that (P) holds for $f : J \times D \to \mathbb{R}^n$ where $K \subset D$. Hypothesis (P) says that the time-dependent vector field $f(t,x)$ points into K at points $x \in \partial K$.

Let $A(t)$ be a continuous $n \times n$ matrix-valued function defined on the interval J containing t_0 and consider the linear initial value problem for the matrix solution X:

$$X' = A(t)X, \quad X(t_0) = I. \tag{3.2}$$

Observe that (P) and (QM) are equivalent for linear systems; therefore we have:

COROLLARY 3.4. *The matrix solution $X(t)$ satisfies $X(t)K \subset K$ for $t \geqslant t_0$ if and only if for all $t \in J$,* (P) *holds for the function $x \to A(t)x$. In fact,* (P) *implies that $X(t)$ maps $K \setminus \{0\}$ and* Int K *into themselves for all $t > t_0$.*

A matrix A is *K-positive* if $A(K) \subset K$. Corollary 3.4 implies that $X(t)$ is K-positive for $t \geqslant t_0$ if (P) holds.

If for every $t \in J$, there exists $\alpha \in \mathbb{R}$ such that $A + \alpha I$ is K-positive, then (P) holds for A. Indeed, if $\lambda \in K^*$ satisfies $\lambda(x) = 0$ then application of λ to $(A + \alpha I)x \geqslant 0$ yields that $\lambda(A(t)x) \geqslant 0$. The converse is false for general cones but true for polyhedral cones by Theorem 8 of Schneider and Vidyasagar [177]. See also Theorem 4.3.40 of Berman and Neumann [18]. Lemmert and Volkmann [118] give the following example of a matrix

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

which satisfies (P) for the ice-cream cone above but $A + \alpha I$ is not K-positive for any α.

Recall that the domain D is *p-convex* if for every $x, y \in D$ satisfying $x \leqslant y$ the line segment joining them also belongs to D. Let $\frac{\partial f}{\partial x}(t,x)$ be continuous on $J \times D$. We say that f (or system (3.1)) is *K-cooperative* if for all $t \in J$, $y \in D$, (P) holds for the function $x \to \frac{\partial f}{\partial x}(t, y)x$. By Corollary 3.4 applied to the variational equation

$$X'(t) = \frac{\partial f}{\partial x}\bigl(t, x(t, t_0, x_0)\bigr)X, \quad X(t_0) = I$$

we conclude that if f is K-cooperative then $X(t) = \frac{\partial x}{\partial x_0}(t, t_0, x_0)$ is K-positive.

THEOREM 3.5. *Let $\frac{\partial f}{\partial x}(t, x)$ be continuous on $J \times D$. Then* (QM) *implies that f is K-cooperative. Conversely, if D is p-convex and f is K-cooperative, then* (QM) *holds.*

PROOF. Suppose that (QM) holds, $x \in D$, $h \in \partial K$, and $\phi \in K^*$ satisfies $\phi(h) = 0$. Since $x \leqslant x + \epsilon h$ and $\phi(x) = \phi(x + \epsilon h)$ for small $\epsilon > 0$, (QM) implies that $\phi(f(t, x)) \leqslant \phi(f(t, x + \epsilon h))$. Hence,

$$0 \leqslant \phi\left(\frac{f(t, x + \epsilon h) - f(t, x)}{\epsilon}\right)$$

and the desired result holds on taking the limit $\epsilon \to 0$.

Conversely, suppose that f is K-cooperative and D is p-convex. If $x, y \in D$ satisfy $x \leqslant y$ and $\phi(x) = \phi(y)$ for some $\phi \in K^*$, then either $\phi = 0$ or $y - x \in \partial K$. Consequently

$$\phi\bigl(f(t, y) - f(t, x)\bigr) = \int_0^1 \phi\left(\frac{\partial f}{\partial x}\bigl(t, sy + (1 - s)x\bigr)(y - x)\right)\mathrm{d}s \geqslant 0$$

because the integrand is nonnegative. □

If for each $(t, x) \in J \times D$ there exists α such that $(\frac{\partial f}{\partial x}(t, x) + \alpha I)$ is K-positive, then f is K-cooperative. This is implied by the remark following Corollary 3.4.

In the special case that $K = \mathbb{R}^n_+$, the cone of nonnegative vectors, it is easy to see by using the standard inner product that we may identify K^* with K. The quasimonotone hypothesis reduces to the Kamke–Müller condition [91,148]: $x \leqslant y$ and $x_i = y_i$ for some i implies $f_i(t, x) \leqslant f_i(t, y)$. This holds by taking $\phi(x) = \langle e_i, x\rangle$ (e_i is the unit vector in the x_i-direction) and noting that every $\phi \in K^*$ can be represented as a positive linear combination of these functionals. If f is differentiable, the Kamke–Müller condition implies

$$\frac{\partial f_i}{\partial x_j}(t, x) \geqslant 0, \quad i \neq j. \tag{3.3}$$

Conversely, if $\frac{\partial f}{\partial x}(t, x)$ is continuous on $J \times D$, (3.3) holds and D is p-convex, then the Kamke–Müller condition holds by an argument similar to the one used in the proof of the converse in Theorem 3.5.

Stern and Wolkowicz [206] give necessary and sufficient conditions for (P) to hold for matrix A relative to the ice-cream cone $K = \{x \in \mathbb{R}^n\colon x_1^2 + x_2^2 + \cdots + x_{n-1}^2 \leqslant x_n^2,\ x_n \geqslant 0\}$. Let Q denote the $n \times n$ diagonal matrix with first $n - 1$ entries 1 and last entry -1. Then

(QM) holds for A if and only if $QA + A^T Q + \alpha Q$ is negative semidefinite for some $\alpha \in \mathbb{R}$. Their characterization extends to other ellipsoidal cones.

3.2. *Strong monotonicity with linear systems*

In this section, all matrices are assumed to be square. Recall that the matrix A is strongly positive if $A(K \setminus \{0\}) \subset \operatorname{Int} K$. We introduce the following milder hypothesis on the matrix A, following Schneider and Vidyasagar [177]:

(ST) For all $x \in \partial K \setminus \{0\}$ there exists $\nu \in K^*$ such that $\nu(x) = 0$ and $\nu(Ax) > 0$.

The following result for the case of constant matrices was proved by Elsner [44], answering a question in [177]. Our proof follows that of Theorem 4.3.26 of Berman et al. [18].

THEOREM 3.6. *Let the linear system* (3.2) *satisfy* (P). *Then the fundamental matrix $X(t_1)$ is strongly positive for $t_1 > t_0$ if there exists s satisfying $t_0 \leqslant s \leqslant t_1$ such that* (ST) *holds for $A(s)$.*

PROOF. Observe that the set of all s such that (ST) holds for $A(s)$ is open. If the result is false, there exists $x > 0$ such that the solution of (3.2) given by $y(t) = X(t)x$ satisfies $y(t_1) \in \partial K \setminus \{0\}$. By Corollary 3.4, $y(t) > 0$ for $t \geqslant t_0$ and $y(t) \in \partial K$ for $t_0 \leqslant t \leqslant t_1$. Let $s \in (t_0, t_1]$ be such that (ST) holds for $A(s)$. Then there exists $\nu \in K^*$ such that $\nu(y(s)) = 0$ and $\nu(A(s)y(s)) > 0$. As $\nu \in K^*$ and $y(t) \in K$, $h(t) := \nu(y(t)) \geqslant 0$ for $t_0 \leqslant t \leqslant t_1$. But $h(s) = 0$ and $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=s} h(t) = \nu(A(s)y(s)) > 0$ which, taken together, imply that $h(s - \delta) < 0$ for small positive δ, giving the desired contradiction. ☐

If (3.2) satisfies (P) and if $x \in \partial K$ then for all $\phi \in K^*$ such that $\phi(x) = 0$ we have $\phi(A(t)x) \geqslant 0$. Hypothesis (ST) asserts that if $x \neq 0$ then $\phi(A(t)x) > 0$ for at least one such ϕ. Berman et al. [18] refer to (ST) (they include (P) in their definition) by saying that A is strongly K-subtangential; while we do not use this terminology, our notation is motivated by it.

An example in [18] shows that (P) and (ST) are not necessary for strong positivity. Let K be the ice-cream cone $K = \{x \in \mathbb{R}^3\colon\ x_1^2 + x_2^2 \leqslant x_3^2, x_3 \geqslant 0\}$ and consider the constant coefficient system (3.2) with matrix A given by

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

An easy calculation shows that $(x_1^2 + x_2^2)' = -2x_2^2$ so it follows easily that K is positively invariant, hence (P) holds by Corollary 3.4. The solution satisfying $x(0) = (\cos(\theta), \sin(\theta), 1)^T \in \partial K$ satisfies $x(t) \in \operatorname{Int} K$ for $t > 0$ since the calculation above and the fact that $x_2(t)$ can have only simple zeros implies that $x_1^2 + x_2^2$ is strictly decreasing while x_3 remains unchanged. The linear functional ν, defined $\nu(x) :=$

$(-\cos(\theta), -\sin(\theta), 1)x$ belongs to K^* by an easy calculation and satisfies $\nu(x(0)) = 0$. It is unique, up to positive scalar multiple, with these properties because K is smooth so its positive normal at a point is essentially unique. But $\nu(Ax(0)) = \sin^2(\theta)$ vanishes if $\theta = 0, \pi$. Therefore (ST) fails although $X(t)$ is strongly positive for $t > 0$.

Theorem 3.6 leads to the following result on strong monotonicity for the nonlinear system (3.1).

LEMMA 3.7. *Assume D is p-convex, $\frac{\partial f}{\partial x}(t,x)$ is continuous on $J \times D$ and f is K-cooperative. Let $x_0, x_1 \in D$ satisfy $x_0 < x_1$ and $t > t_0$ with $t \in J(t_0,x_0) \cap J(t_0,x_1)$. If there exists y_0 on the line segment joining x_0 to x_1 and $r \in [t_0,t]$ such that* (ST) *holds for $\frac{\partial f}{\partial x}(r, x(r,t_0,y_0))$ then*

$$x(t,t_0,x_0) \ll x(t,t_0,x_1).$$

PROOF. First, observe that for y_0 on the segment it follows that $t \in J(t_0,y_0)$. We apply the formula

$$x(t,t_0,x_1) - x(t,t_0,x_0) = \int_0^1 \frac{\partial x}{\partial x_0}\big(t,t_0,sx_1 + (1-s)x_0\big)(x_1 - x_0)\,\mathrm{d}s,$$

where $X(t) = \frac{\partial x}{\partial x_0}(t,t_0,y_0)$ is the fundamental matrix for (3.2) corresponding to the matrix $A(t) = \frac{\partial f}{\partial x}(t, x(t,t_0,y_0))$. The left-hand side belongs to $K \setminus \{0\}$ if $x_0 < x_1$ by Theorems 3.5 and 3.2 but we must show it belongs to $\operatorname{Int} K$. For this to be true, it suffices that for each $t > t_0$ there exists $s \in [0,1]$ such that the matrix derivative in the integrand is strongly positive. In fact, this derivative is K-positive by Corollary 3.4 for all values of the arguments with $t \geqslant t_0$, so application of any nontrivial $\phi \in K^*$ to the integral gives a nonnegative numerical result. If there exists s as above, then the application of ϕ to the integrand gives a positive numerical result for all s' near s by continuity and Proposition 3.1 and hence the integral belongs to $\operatorname{Int} K$ by Proposition 3.1. By Theorem 3.6, $\frac{\partial x}{\partial x_0}(t,t_0,y_0)$ is strongly positive for $t > t_0$ if (ST) holds for $A(r) = \frac{\partial f}{\partial x}(r, x(r,t_0,y_0))$ for some $r \in [t_0,t]$. But this is guaranteed by our hypothesis. □

THEOREM 3.8. *D is p-convex, $\frac{\partial f}{\partial x}(t,x)$ is continuous on $J \times D$, and f is K-cooperative. Suppose for every $x_0, x_1 \in D$ with $x_0 < x_1$ and $t_0 \in J$, there exists y_0 on the line segment joining the x_i such that* (ST) *holds for $\frac{\partial f}{\partial x}(t_0,y_0)$. If $x_0, x_1 \in D$, $x_0 < x_1$, and $t > t_0$ then*

$$t \in J(t_0,x_0) \cap J(t_0,x_1) \quad \Longrightarrow \quad x(t,t_0,x_0) \ll x(t,t_0,x_1).$$

PROOF. This is an immediate corollary of Lemma 3.7. □

As the main hypothesis of Theorem 3.8 will be difficult to verify in applications, the somewhat stronger condition of irreducibility may be more useful because there is a large body of theory related to it [18,19]. We now introduce the necessary background. A closed subset F of K that is itself a cone is called *a face* of K if $x \in F$ and $0 \leqslant y \leqslant x$ (inequalities induced by K) implies that $y \in F$. For example, the faces of $K = \mathbb{R}^n_+$ are of the form $\{x \in \mathbb{R}^n_+ \colon x_i = 0,\ i \in I\}$ where $I \subset \{1, 2, \dots, n\}$. For the ice-cream cone $K = \{x \in \mathbb{R}^n \colon x_1^2 + x_2^2 + \dots + x_{n-1}^2 \leqslant x_n^2, x_n \geqslant 0\}$, the faces are the rays issuing from the origin and passing through its boundary vectors. A K-positive matrix A is *K-irreducible* if the only faces F of K for which $A(F) \subset F$ are $\{0\}$ and K. The following is a special case of Theorem 2.3.9 in Berman and Neumann [18]; see Berman and Plemmons [19] for proofs. These references contain additional related results.

THEOREM 3.9. *Let A be an $n \times n$ K-positive matrix. Then the following are equivalent*:

(i) *A is K-irreducible*;

(ii) *No eigenvector of A belongs to ∂K*;

(iii) *A has exactly one unit eigenvector in K and it belongs to* $\operatorname{Int} K$;

(iv) $(I + A)^{n-1}(K \setminus \{0\}) \subset \operatorname{Int} K$.

The famous Perron–Frobenius Theory is developed for K-positive and K-irreducible matrices in the references above. In particular, the spectral radius of A is a simple eigenvalue of A with corresponding eigenvector described in (iii) above.

Below we require the simple observation that if A is K-positive, then the adjoint A^* is K^*-positive. Indeed, if $\nu \in K^*$ then $(A^*\nu)(x) = \nu(Ax) \geqslant 0$ for all $x \in K$ so $A^*\nu \in K^*$. The next result is adapted from Theorem 4.3.17 of Berman et al. [18].

PROPOSITION 3.10. *Let A be an $n \times n$ matrix and suppose that there exists $\alpha \in \mathbb{R}$ such that $B := A + \alpha I$ is K-positive. Then B is K-irreducible if and only if* (ST) *holds for A.*

PROOF. Suppose that $B = A + \alpha I$ is K-positive and (ST) holds. If $Ax = \lambda x$ for some $\lambda \in \mathbb{R}$ and nonzero vector $x \in \partial K$ then there exists $\nu \in K^*$ such that $\nu(x) = 0$ and $\nu(Ax) > 0$. But $\nu(Ax) = \lambda \nu(x) = 0$. Consequently, no eigenvector of B belongs to ∂K so by Theorem 3.9, B is K-irreducible.

Conversely, suppose that B is K-positive and K-irreducible. Let $x \in \partial K$, $x \neq 0$ and let $\nu \in K^*$ satisfy $\nu \neq 0$ and $\nu(x) = 0$. By Theorem 3.9, $C := B + I$ has the property that C^{n-1} is strongly positive so $\nu(C^{n-1}x) > 0$. As C is K-positive, $\nu(C^r x) \geqslant 0$ for $r = 1, 2, \dots, n-1$. Because $\nu(x) = 0$, we may choose $p \in \{1, 2, \dots, n-1\}$ such that $\nu(C^p x) > 0$ but $\nu(C^{p-1}x) = 0$. Let $\tilde{\nu} = (C^*)^{p-1}\nu$. Then $\tilde{\nu} \in K^*$, $\tilde{\nu}(x) = 0$ and $\tilde{\nu}(Cx) > 0$. But then A satisfies (ST) because $\tilde{\nu}(Ax) = \tilde{\nu}(Cx) > 0$. □

Motivated by Proposition 3.10, we introduce the following hypothesis for matrix A.

(CI) There exists $\alpha \in \mathbb{R}$ such that $A + \alpha I$ is K-positive and K-irreducible.

In the special case that $K = \mathbb{R}^n_+$, $n \geqslant 2$, matrix A satisfies (CI) if and only if $a_{ij} \geqslant 0$ for $i \neq j$ and there is no permutation matrix P such that

$$P^T A P = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix},$$

where B and D are square. This is equivalent to the assertion that the incidence graph of A is strongly connected. See Berman and Plemmons [19].

The following is an immediate consequence of Theorem 3.8.

COROLLARY 3.11. *D is p-convex, $\frac{\partial f}{\partial x}(t,x)$ is continuous on $J \times D$ and f is K-cooperative. Suppose that for every $x_0, x_1 \in D$ with $x_0 < x_1$ and $t_0 \in J$, there exists y_0 on the line segment joining the x_i such that* (CI) *holds for $\frac{\partial f}{\partial x}(t_0, y_0)$. If $x_0, x_1 \in D$, $x_0 < x_1$, and $t > t_0$ then*

$$t \in J(t_0, x_0) \cap J(t_0, x_1) \quad \Longrightarrow \quad x(t, t_0, x_0) \ll x(t, t_0, x_1).$$

PROOF. If (CI) holds then, by Proposition 3.10, (ST) holds for $\frac{\partial f}{\partial x}(t,x)$, so the conclusion follows from Theorem 3.8. □

Corollary 3.11 is an improvement of the restriction of Theorem 10 of Kunze and Siegel [111] to the case that K has nonempty interior; their results also treat the case that K has empty interior in $\mathbb{R}^n$ but nonempty interior in some subspace of $\mathbb{R}^n$. Walter [228] gives a sufficient condition for strong monotonicity relative to $K = \mathbb{R}^n_+$ which does not require f to be differentiable.

For polyhedral cones it can be shown that matrix A satisfies (P) and (ST) if and only if there exists $\alpha \in \mathbb{R}$ such that $A + \alpha I$ is K-positive and K-irreducible. See Theorem 4.3.40 of Berman et al. [18]. For the case of polyhedral cones, therefore, Corollary 3.11 and Theorem 3.8 are equivalent.

3.3. *Autonomous K-competitive and K-cooperative systems*

Our focus now is on the autonomous system of ordinary differential equations

$$x' = f(x), \tag{3.4}$$

where f is a vector field on an open subset $D \subset \mathbb{R}^n$; all vector fields are assumed to be continuously differentiable. We change our notation slightly to conform to more dynamical notation, denoting $x(t, 0, x_0)$ by $\Phi_t(x)$, where Φ denotes the dynamical system (= local flow) in D generated by f discussed in Section 1. The notation $\Phi_t(x)$ carries the tacit assumption that $t \in I_x$, the open interval in $\mathbb{R}$ containing the origin on which the trajectory of x under Φ is defined. The *positive semiorbit* (respectively, (*negative semiorbit*) of x is $\gamma^+(x) := \{\Phi_t(x)\colon t \in t \geqslant 0\}$ (respectively, $\gamma^-(x) := \{\Phi_t(x)\colon t \leqslant 0\}$). The limit sets of x can be defined as

$$\omega(x) = \bigcap_{t \geqslant 0} \overline{\bigcup_{\tau \geqslant t} \Phi_\tau(x)}, \qquad \alpha(x) = \bigcap_{t \leqslant 0} \overline{\bigcup_{\tau \leqslant t} \Phi_\tau(x)}.$$

We call f and Eq. (3.4) *K-competitive* in D if the time-reversed system

$$x' = -f(x)$$

is K-cooperative. When K is the standard cone $\mathbb{R}^n_+$, f is competitive if and only if $\partial f_i/\partial x_j \leqslant 0$ for $i \neq j$. Therefore if f is K-competitive with local flow Φ, then $-f$ is K-cooperative with local flow $\widetilde{\Phi}$, where $\widetilde{\Phi}_t(x) = \Phi_{-t}(x)$; and conversely. Thus time-reversal changes K-competitive systems into K-cooperative ones, and vice-versa. This fact will be exploited repeatedly below.

In the remainder of Section 3 we assume $\mathbb{R}^n$ is ordered by a cone $K \subset \mathbb{R}^n$ with nonempty interior.

A map is *locally monotone* if every point in its domain has a neighborhood on which the map is monotone. A local flow or local semiflow Φ is locally monotone if Φ_t is a locally monotone map for all $t > 0$. *Locally strongly monotone* is defined similarly.

THEOREM 3.12. *Let f be a K-cooperative vector field in an open set $D \subset \mathbb{R}^n$, generating the local flow Φ. Then Φ is locally monotone, and monotone when D is p-convex.*

PROOF. If D is p-convex, monotonicity follows from Theorem 3.2 (with $f(t,x) := f(x)$). Suppose D is not p-convex. Denote the domain of Φ_t by D_t.

We first claim: For every $p \in D$ there exists $\tau > 0$ and a neighborhood $N \subset D_\tau$ such that $\Phi_t|N$ is monotone if $t \in [0,\tau]$. But this is obvious since by restricting f to a p-convex neighborhood of p, we can use Theorem 3.2.

Now fix $p \in D$ and let $J(0,p) \cap [0,\infty) = [0,r)$, $0 < r \leqslant \infty$. Let I_p be the set of all nonnegative $s \in [0,r)$ such that there is a neighborhood U_s of p, contained in D_s, such that $\Phi_t|U_s$ is monotone for each $t \in [0,s]$. The previous claim implies that $[0,\tau] \subset I_p$ and, by its definition, I_p is an interval. Furthermore, straightforward applications of the previous claim establish that I_p is both an open and a closed subset of $[0,r)$. It follows that $I_p = [0,r)$. □

The next theorem gives a sufficient condition for strong monotonicity. Define $G(f)$ to be the set of $x \in D$ such that (ST) holds for $A = f'(x)$. Note that $x \in G(f)$ provided (CI) holds for $A = f'(x)$, by Proposition 3.10. If $K = \mathbb{R}^n_+$, a sufficient condition for $x \in G(f)$ is that $f'(x)$ is an irreducible matrix with nonnegative off-diagonal entries.

THEOREM 3.13. *Let f be a K-cooperative vector field in an open set $D \subset \mathbb{R}^n$, generating the local flow Φ. Assume $D \setminus G(f)$ does not contain any totally ordered line segment* (*which holds when $D \setminus G(f)$ is zero dimensional*). *Then Φ is locally strongly monotone, and strongly monotone when D is p-convex.*

PROOF. Suppose D is p-convex, in which case Φ is monotone by from Theorem 3.2. By Theorem 3.8, Φ is strongly monotone.

When D is not p-convex, Φ is locally monotone by Theorem 3.12, and the previous paragraph implies Φ is locally strongly monotone. □

The proof of Theorem 3.13 can be adapted to cover certain nonopen domains D, such as an order interval, a closed halfspace, and the cone K; see the discussion following the proof of Theorem 3.2.

Theorem 3.8 implies that Φ is strongly monotone provided D is p-convex and f satisfies the autonomous version of condition (ST) of Section 3.2, namely:

(ST*) For all $u \in D, x \in \partial K \setminus \{0\}$ there exists $\nu \in K^*$ such that $\nu(x) = 0$ and $\nu(f'(u)x) > 0$.

Without p-convexity of D, condition (ST*) yields local strong monotonicity.

3.4. *Dynamics of cooperative and competitive systems*

We continue to assume $\mathbb{R}^n$ is ordered by a cone K having nonempty interior; all notions involving order refer to that defined by K. For this section, the terms "competitive" and "cooperative" are tacitly understood to mean "K-competitive" and "K-cooperative," and monotonicity refers to the ordering defined by K.

We first apply results from Section 2 to obtain a generic stable convergence theorem for cooperative vector fields.

Let Φ denote the local flow generated by a vector field f on $D \subset \mathbb{R}^n$. We assume D is p-convex throughout this section without further mention. When $\Phi_t(x)$ is defined for all $(t, x) \in [0, \infty) \times D$, as when all positive semiorbits have compact closure in D, the corresponding positive local semiflow Φ^+ is a semiflow. To Φ we associate C, S and E, denoting respectively the sets of convergent, stable and equilibrium points for Φ^+.

THEOREM 3.14. *Let f be a cooperative vector field on an open set $D \subset \mathbb{R}^n$, generating a local flow Φ such that*:

(a) *Every positive semiorbit of Φ has compact closure in D*;

(b) *Condition* (ST*) *above is satisfied, and $D = AC \cup BC$.*

Then Φ has the following properties:

(i) *$C \cap S$ contains a dense open subset of D, consisting of points whose trajectories converge to equilibria*;

(ii) *If E is compact there is a stable equilibrium, and an asymptotically stable equilibrium when E is finite.*

PROOF. Assumption (ST*) makes Φ strongly monotone. The hypothesis of Theorem 2.26, with $X = D$, is fulfilled: D is normally ordered and $D = BC \cup AC$. Therefore Theorem 2.26 implies the conclusion. □

Theorem 3.14, like Theorem 3.13, holds for some more general domains D, including relatively open subsets of V where V denotes a closed halfspace, a closed order interval, or the cone K.

One of the main results of this subsection is that n-dimensional competitive and cooperative systems behave like general systems of one less dimension. Theorems 3.21 and 3.22 illustrate this principle for $n = 2$ in a very strong form. In higher dimensions the principle holds for compact limit sets. The key tool in proving this is the following result due to Hirsch [67]:

THEOREM 3.15. *A limit set of a competitive or cooperative system cannot contain two points related by* $\ll$.

PROOF. By time reversal, if necessary, we assume the system is cooperative, hence the local flow is monotone. Now apply Proposition 1.10. □

A periodic orbit of a competitive or cooperative system is a limit set and consequently it cannot contain two points related by $\ll$. The following sharper result will be useful later:

PROPOSITION 3.16. *Nontrivial periodic orbit of a competitive or cooperative system cannot contain two points related by* $<$.

PROOF. By time-reversal we assume the system is cooperative, and in this case the conclusion follows from Proposition 1.10. □

Let Φ, Ψ be flows in respective spaces A, B. We say Φ and Ψ are *topologically equivalent* if there is a homeomorphism $Q: A \to B$ that is a conjugacy between them, i.e., $Q \circ \Phi_t = \Psi_t \circ Q$ for all $t \in \mathbb{R}$. The relationship of topological equivalence is an equivalence relation on the class of flow; it formalizes the notion of "having the same qualitative dynamics."

A system of differential equations $y' = F(y)$, defined on $\mathbb{R}^k$, is called *Lipschitz* if F is Lipschitz. That is, there exists $K > 0$ such that $|F(y_1) - F(y_2)| \leqslant K|y_1 - y_2|$ for all $y_1, y_2 \in \mathbb{R}^k$. With these definitions, we can state a result of Hirsch [67] that follows directly from Theorem 3.15.

THEOREM 3.17. *The flow on a compact limit set of a competitive or cooperative system in* $\mathbb{R}^n$ *is topologically equivalent to a flow on a compact invariant set of a Lipschitz system of differential equations in* $\mathbb{R}^{n-1}$.

PROOF. Let L be the limit set, $v \gg 0$ be a unit vector and let H_v be the hyperplane orthogonal to v, i.e, $H_v := \{x\colon \langle x, v\rangle = 0\}$. The orthogonal projection Q onto H_v is given by $Qx = x - \langle x, v\rangle v$. By Theorem 3.15, Q is one-to-one on L (this could fail only if L contains two points that are related by $\ll$). Therefore, Q_L, the restriction of Q to L, is a Lipschitz homeomorphism of L onto a compact subset of H_v. We argue by contradiction to establish the existence of $m > 0$ such that $|Q_L x_1 - Q_L x_2| \geqslant m|x_1 - x_2|$ whenever $x_1 \neq x_2$ are points of L. If this were false, then there exists sequences $x_n, y_n \in L$, $x_n \neq y_n$ such that

$$\frac{|Q(x_n) - Q(y_n)|}{|x_n - y_n|} = \frac{|(x_n - y_n) - v\langle v, x_n - y_n\rangle|}{|x_n - y_n|} \to 0$$

as $n \to \infty$. Equivalently, $|w_n - v\langle v, w_n\rangle| \to 0$ as $n \to \infty$ where $w_n = x_n - y_n/|x_n - y_n|$. We can assume that $w_n \to w$ as $n \to \infty$ where $|w| = 1$. Then, $w = v\langle v, w\rangle$ and therefore, $\langle v, w\rangle^2 = 1$ so $w = \pm v$. But then $x_n - y_n/|x_n - y_n| \to \pm v$ as $n \to \infty$ and this implies that $x_n \ll y_n$ or $y_n \ll x_n$ for all large n, contradicting Theorem 3.15. Therefore, Q_L^{-1} is Lipschitz on $Q(L)$. Since L is a limit set, it is an invariant set for (3.4). It follows that the dynamical system restricted to L can be modeled on a dynamical system in H_v. In fact, if $y \in Q(L)$ then $y = Q_L(x)$ for a unique $x \in L$ and $\Psi_t(y) \equiv Q_L(\Phi_t(x))$ is a dynamical system on $Q(L)$ generated by the vector field

$$F(y) = Q_L\bigl(f\bigl(Q_L^{-1}(y)\bigr)\bigr)$$

on $Q(L)$. According to McShane [137], a Lipschitz vector field on an arbitrary subset of H_v can be extended to a Lipschitz vector field on all of H_v, preserving the Lipschitz constant. It follows that F can be extended to all of H_v as a Lipschitz vector field. It is easy to see that $Q(L)$ is an invariant set for the latter vector field. We have established the topological equivalence of the flow Φ on L with the flow Ψ on $Q(L)$. $Q(L)$ is a compact invariant set for the $(n-1)$-dimensional dynamical system on H_v generated by the extended vector field. □

A consequence of Theorem 3.17 is that the flow on a compact limit set, L, of a competitive or cooperative system shares common dynamical properties with the flow of a system of differential equations in one less dimension, restricted to the compact, connected invariant set $Q(L)$. Notice, however, that L may be the limit set of a trajectory not in L, and therefore $Q(L)$ need not be a limit set.

On the other hand, the flow Ψ in a compact limit sets enjoys the topological property of *chain recurrence*, due to Conley [31,30], which will be important in the next subsection. The definition is as follows. Let A be a compact invariant set for the flow Φ. Given two points z and y in A and positive numbers ϵ and t, an (ϵ, t)-chain from z to y in A is an ordered set

$$\{z = x_1, x_2, \dots, x_{m+1} = y; t_1, t_2, \dots, t_m\}$$

of points $x_i \in A$ and times $t_i \geqslant t$ such that

$$\bigl|\Phi_{t_i}(x_i) - x_{i+1}\bigr| < \epsilon, \quad i = 1, 2, \dots, m. \tag{3.5}$$

A is *chain recurrent* for Φ if for every $z \in A$ and for every $\epsilon > 0$ and $t > 0$, there is an (ϵ, t)-chain from z to z in A.

Conley proved that when A is compact and connected, a flow Φ in A is chain recurrent if and only if there are no attractors. This useful condition can be stated as follows: For every proper nonempty compact set $S \subset A$ and all $t > 0$, there exists $s > t$ such that $\Phi_s(S) \not\subset \operatorname{Int} S$.

Compactness of A implies that chain recurrence of the flow in A is independent of the metric, and thus holds for any topologically equivalent flow.

It is intuitively clear that, as Conley proved, flows in compact alpha and omega limit sets are chain recurrent. Indeed, orbit segments of arbitrarily long lengths through point x repeatedly pass near any point of $\omega(x) \cup \alpha(x)$. Of course these segments do not necessarily belong to $\omega(x)$; but by taking suitable limits of points in these segments, one can find enough (ϵ, t)-chains in $\omega(x)$ and $\alpha(x)$ to prove the flows in these sets chain recurrent. For a rigorous proof, see Smith [194].

3.5. *Smale's construction*

Smale [182] showed that it is possible to embed essentially arbitrary dynamics in a competitive or cooperative irreducible system. His aim was to warn population modelers that systems designed to model competition could have complicated dynamics. His result is also very useful for providing counterexamples to conjectures in the theory of monotone dynamics, since by time reversal his systems are cooperative. In this section, competitive and cooperative are with respect to the usual cone.

Smale constructed special systems of Kolmogorov type

$$x_i' = x_i M_i(x), \quad 1 \leqslant i \leqslant n, \tag{3.6}$$

in $\mathbb{R}^n_+$ where the M_i are smooth functions satisfying

$$\frac{\partial M_i}{\partial x_j} < 0 \tag{3.7}$$

for all i, j; all sums are understood to be from 1 to n. We refer to such systems as *totally competitive*. They are simple models of competition between n species, where M_i is interpreted as the *per capita* growth rate of species i.

Smale's object was to choose the M_i so that the standard $(n-1)$-simplex $\Sigma_n = \{x \in \mathbb{R}^n_+\colon \sum x_i = 1\}$ is an attractor in which arbitrary dynamics may be specified.

In order to generate a dynamical system on Σ_n, let H denote the tangent space to Σ_n, that is, $H = \{x \in \mathbb{R}^n\colon \sum x_i = 0\}$, and let $h : \Sigma_n \to H$ be a smooth vector field on Σ_n, meaning that all partial derivatives of h exist and are continuous on Σ_n. We also assume that $h = (h_1, h_2, \ldots, h_n)$ has the form $h_i = x_i g_i(x)$ where the g_i are smooth functions on Σ_n. Then the differential equation

$$x_i' = h_i(x), \quad 1 \leqslant i \leqslant n \tag{3.8}$$

generates a flow in $\mathbb{R}^n_+$ that leaves Σ_n invariant. The form of the h_i ensures that if $x_i(0) = 0$, then $x_i(t) \equiv 0$ so each lower dimensional simplex forming part of the boundary of Σ_n is invariant.

The goal is to construct a competitive system of the form (3.6) satisfying (3.7) such that its restriction to Σ_n is equivalent to (3.8). Let $p : [0, \infty) \to \mathbb{R}_+$ have continuous derivatives of all orders, be identically 1 in a neighborhood of $s = 1$, and vanish outside the interval $[1/2, 3/2]$. As g is a smooth vector field on Σ_n, it has a smooth extension to $\mathbb{R}^n_+$ which we denote by g in order to conserve notation. An example of such an extension is the map $x \mapsto P(\sum x_j) g(x / \sum x_j)/P(1)$, where $P(u) = \int_0^u p(s)\, \mathrm{d}s$.

For $\eta > 0$, define

$$M_i(x) = 1 - S(x) + \eta p\Bigl(\sum x_j\Bigr) g_i(x), \quad 1 \leqslant i \leqslant n.$$

Then (3.7) holds for sufficiently small η since $p(\sum x_j)$ vanishes identically outside a compact subset of $\mathbb{R}^n_+$. Consider the system (3.6) with M as above. $\mathbb{R}^n_+$ is positively invariant; and the function $S(x) = \sum_i x_i$, evaluated along a solution $x(t)$ of (3.6), satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} S\bigl(x(t)\bigr) = S\bigl(x(t)\bigr)\bigl[1 - S\bigl(x(t)\bigr)\bigr]$$

since $\sum x_i g_i(x) = \sum h_i(x) = 0$. Consequently $\varSigma_n$, which is $S^{-1}(1) \cap \mathbb{R}^n_+$, is positively invariant. Moreover if $x(0) \in \mathbb{R}^n_+$ then $S(x(0)) \geqslant 0$. This implies $S(x(t)) \to 1$ as $t \to \infty$, unless $x(t) \equiv 0$, and $\varSigma_n$ attracts all nontrivial solutions of (3.6) in $\mathbb{R}^n_+$. Restricted to $\varSigma_n$, (3.6) becomes

$$x_i' = \eta h_i(x), \quad 1 \leqslant i \leqslant n.$$

Therefore the dynamics of (3.6) restricted to $\varSigma_n$ is equivalent, up to a change in time scale, to the dynamics generated by (3.8).

As noted above, Smale's construction has implications for cooperative and irreducible systems since the time-reversed system corresponding to (3.6) is cooperative and irreducible in $\operatorname{Int}\mathbb{R}^n_+$. Time-reversal makes the simplex a repellor for a cooperative system $\varPhi$ in $\mathbb{R}^n_+$. Therefore every invariant set in the simplex is unstable for $\varPhi$. Each trajectory of $\varPhi$ that is not in the simplex is attracted to the equilibrium at the origin or to the virtual equilibrium at ∞. The simplex is the common boundary between the basins of attraction of these two equilibria.

3.6. *Invariant surfaces and the carrying simplex*

It turns out that the essential features of Smale's seemingly very special construction are found in a large class of totally competitive Kolmogorov systems

$$x_i' = x_i M_i(x), \quad x \in \mathbb{R}^n_+. \tag{3.9}$$

Here and below i and j run from 1 to n. Let $\varPhi$ denote the corresponding local flow. The *unit* $(n-1)$ *simplex* is $\varDelta^{n-1} := \{x \in \mathbb{R}^n_+ \colon \sum x_i = 1\}$.

THEOREM 3.18. *Assume* (3.9) *satisfies the following conditions*:
(a) $\frac{\partial M_i}{\partial x_j} < 0$;
(b) $M_i(0) > 0$;
(c) $M_i(x) < 0$ *for* $|x|$ *sufficiently large*.
Then there exists an invariant compact hypersurface $\varSigma \subset \mathbb{R}^n_+$ *such that*
(i) $\varSigma$ *attracts every point in* $\mathbb{R}^n_+ \setminus \{0\}$;
(ii) $\varSigma \cap \operatorname{Int}\mathbb{R}^n_+$ *is a locally Lipschitz submanifold*;
(iii) $\varSigma \cap \operatorname{Int}\mathbb{R}^n_+$ *is transverse to every line that is parallel to a nonnegative vector and meets* $\varSigma \cap \operatorname{Int}\mathbb{R}^n_+$;
(iv) $\varSigma$ *is unordered*;

(v) *Radial projection defines a homeomorphism* $h : \Sigma \to \Delta^{n-1}$ *whose inverse is locally Lipschitz on the open* $(n-1)$*-cell* $\Delta^{n-1} \cap \operatorname{Int} \mathbb{R}^n_+$. *There is a flow* Ψ *on* Δ^{n-1} *such that* $\Phi_t|\Sigma = h \circ \Phi_t \circ h^{-1}$.

COROLLARY 3.19. *If* $n = 3$, *every periodic orbit in* $\mathbb{R}^3_+$ *bounds an unordered invariant disk.*

Assumption (a) is the condition of total competition; (b) and (c) have plausible biological interpretations. The attracting hypersurface Σ, named the *carrying simplex* by M. Zeeman, is analogous to the carrying capacity K in the one-dimensional logistic equation $\mathrm{d}x/\mathrm{d}t = rx(K-x)$. One can define Σ either as the boundary of the set of points whose alpha limit set is the origin, or as the boundary of the compact global attractor. These sets coincide if and only if Σ is unique, in which case it uniformly attracts every compact set in $\mathbb{R}^n_+ \setminus \{0\}$. Uniqueness holds under mild additional assumptions on the maps M_i (Wang and Jiang [230]). The geometry, smoothness and dynamics of carrying simplices have been investigated by Benaïm [14], Brunovsky [21], Miercyński [140,143,141], Tineo [220], van den Driessche and M. Zeeman [223], Wang and Jiang [230], E. Zeeman [239], E. Zeeman and M. Zeeman [240–242], M. Zeeman [243].

Theorem 3.18 is proved in Hirsch [72] using a general existence theorem for invariant hypersurfaces, of which the following is a generalization:

THEOREM 3.20. *Let* Φ *be a strongly monotone local flow in a p-convex open set* $D \subset \mathbb{R}^n$. *If* $L \subset D$ *is a nonempty compact unordered invariant set,* L *lies in an unordered invariant hypersurface* M *that is a locally Lipschitz submanifold.*

IDEA OF PROOF. Define U to be the set of $x \in D$ such that $\Phi_t(x) \gg y$ for some $t > 0$, and some $y \in L$. Continuity implies U is open, and it is nonempty since it contains $z \in D$ where $z > y \in L$. It can be shown that the lower boundary of U in D (Section 1.1) is a hypersurface with the required properties, by arguments analogous to the proof of Theorem 3.17. □

3.7. *Systems in* $\mathbb{R}^2$

Cooperative and competitive systems in $\mathbb{R}^2$ have particularly simple dynamics. Versions of the following result were proved in Hirsch [67], Theorem 2.7 and Smith [194], Theorem 3.2.2. It is noteworthy that in the next two theorems Φ does not need to be monotone, only locally monotone; hence p-convexity of D is not needed.

THEOREM 3.21. *Let* $D \subset \mathbb{R}^2$ *be an open set and* $g : D \to \mathbb{R}^2$ *a vector field that is cooperative or competitive for the standard cone. Let* $y(t)$ *a nonconstant trajectory defined on an open interval* $I \subset \mathbb{R}$ *containing* 0. *Then there exists* $t_* \in I$ *such that each coordinate* $y_i(t)$ *is nonincreasing or nondecreasing on each connected component of* $I \setminus \{t_*\}$.

PROOF. It suffices to prove that $y_i'(t)$ can change sign at most once. We assume g is cooperative, otherwise reversing time. Let Φ be local flow of g and set $X(t,x) = \frac{\partial \Phi}{\partial x}(t,x)$. The matrix-valued function $X(t,x)$ satisfies the variational equation

$$\frac{\partial}{\partial t} X(t,x) = \frac{\partial g}{\partial x}\bigl(\Phi(t,x)\bigr) \cdot X(t,x), \qquad X(0,x) = I.$$

Cooperativity and Corollary 3.4 show that $X(t,x)$ has nonnegative entries for $t \geqslant 0$, i.e., matrix multiplication by $X(t,x)$ preserves the standard cone. The tangent vector $y'(t)$ to the curve $y(t)$, being a solution of the variational equation, satisfies $y'(t) = X(t, y(0))y'(0)$. Nonnegativity of $X(t,x)$ implies that if $y'(t_0)$ lies in the first or third quadrants, then $y'(t)$ stays in the same quadrant, and hence its coordinates have constant sign, for $t > t_0$. On the other hand if $y'(t)$ for $t \geqslant t_0$ is never in the first or third quadrants, its coordinates again have constant sign. (Note that $y'(t)$ cannot transit directly between quadrants 1 and 3, or 2 and 4, since it cannot pass through the origin.) We have shown that there is at most one $t_0 \in I$ at which $y'(t)$ changes quadrants. If such a t_0 exists, set $t_* = t_0$; otherwise let $t_* \in I$ be arbitrary. □

Variants of the next result have been proved many times for Kolmogorov type population models (Albrecht et al. [1], Grossberg [53], Hirsch and Smale [80], Kolmogorov [97], Rescigno and Richardson [168], Selgrade [178]).

THEOREM 3.22. *Let g be a K-cooperative or K-competitive vector field in a domain $D \subset \mathbb{R}^2$. If $\gamma^+(x)$ (respectively, $\gamma^-(x)$) has compact closure in D, then $\omega(x)$ (respectively, $\alpha(x)$) is a single equilibrium.*

PROOF. For the standard cone, denoted here by P, this follows from Theorem 3.21. The general case follows by making a linear coordinate change $y = Tx$ mapping K onto the standard cone. Here T is any linear transformation that takes a basis for $\mathbb{R}^2$ contained in ∂P into the standard basis, which lies in ∂K. Then we have $u \leqslant_K v$ if and only if $Tu \leqslant_P Tv$; in other words, T is an order isomorphism. It follows that the system $x' = g(x)$ is K-cooperative (respectively, K-competitive) if and only if the system $y' = h(y) := Tg(T^{-1}y)$ is P-cooperative (respectively, P-competitive). Therefore T is a conjugacy between the local flows Φ, Ψ of the two dynamical systems, that is, $T \circ \Phi_t = \Psi_t \circ T$. Consequently the conclusion for P, proved above, implies the conclusion for K. □

3.8. *Systems in $\mathbb{R}^3$*

The following Poincaré–Bendixson theorem for three-dimensional cooperative and competitive systems is the most notable consequence of Theorem 3.17. It was proved by Hirsch [76] who improved earlier partial results [67,187]. The following result from Smith [194] holds for arbitrary cones $K \subset \mathbb{R}^3$ with nonempty interior:

THEOREM 3.23. *Let g be a K-cooperative or K-competitive vector field in a p-convex domain $D \subset \mathbb{R}^3$. Then a compact limit set of g that contains no equilibrium points is a periodic orbit.*

PROOF. Let Φ denote the flow of the system, and L the limit set. By Theorem 3.17, the restriction of Φ to L is topologically equivalent to a flow Ψ, generated by a Lipschitz planar vector field, restricted to the compact, connected, chain recurrent invariant set $Q(L)$. Since L contains no equilibria neither does $Q(L)$. The Poincaré–Bendixson theorem implies that $Q(L)$ consists of periodic orbits and, possibly, entire orbits whose omega and alpha limit sets are periodic orbits contained in $Q(L)$. The chain recurrence of Ψ on $Q(L)$ will be exploited to show that $Q(L)$ consists entirely of periodic orbits.

Let $z \in Q(L)$ and suppose that z does not belong to a periodic orbit. Then $\omega(z)$ and $\alpha(z)$ are distinct periodic orbits in $Q(L)$. Let $\omega(z) = \gamma$ and suppose for definiteness that z belongs to the interior component, V, of $\mathbb{R}^2 \setminus \gamma$ so that $\Psi_t(z)$ spirals toward γ in V. The other case is treated similarly. Then γ is asymptotically stable relative to V. Standard arguments using transversals imply the existence of compact, positively invariant neighborhoods U_1 and U_2 of γ in V such that $U_2 \subset \mathrm{Int}_V U_1$, $z \notin U_1$ and there exists $t_0 > 0$ for which $\Psi_t(U_1) \subset U_2$ for $t \geqslant t_0$. Let $\epsilon > 0$ be such that the 2ϵ-neighborhood of U_2 in D is contained in U_1. Choose t_0 larger if necessary such that $\Psi_t(z) \in U_2$ for $t \geqslant t_0$. This can be done since $\omega(z) = \gamma$. Then any (ϵ, t_0)-chain in $Q(L)$ beginning at $x_1 = z$ satisfies $\Psi_{t_1}(x_1) \in U_2$ and, by (3.5) and the fact that the 2ϵ-neighborhood of U_2 is contained in U_1, it follows that $x_2 \in U_1$. As $t_2 > t_0$, it then follows that $\Psi_{t_2}(x_2) \in U_2$ and (3.5) again implies that $x_3 \in U_1$. Continuing this argument, it is evident that the (ϵ, t_0)-chain cannot return to z. There can be no (ϵ, t_0)-chain in $Q(L)$ from z to z and therefore we have contradicted that $Q(L)$ is chain recurrent. Consequently, every orbit of $Q(L)$ is periodic. Since $Q(L)$ is connected, it is either a single periodic orbit or an annulus consisting of periodic orbits. It follows that L is either a single periodic orbit or a cylinder of periodic orbits.

To complete the proof we must rule out the possibility that $Q(L)$ consists of an annulus of periodic orbits. We can assume that the system is cooperative. The argument will be separated into two cases: $L = \omega(x)$ or $L = \alpha(x)$.

If $L = \omega(x)$ consists of more than one periodic orbit then $Q(L)$ is an annulus of periodic orbits in the plane containing an open subset O. Then there exists $t_0 > 0$ such that $Q(\Phi_{t_0}(x)) \in O$. Let y be the unique point of L such that $Q(y) = Q(\Phi_{t_0}(x))$. $y = \Phi_{t_0}(x)$ cannot hold since this would imply that L is a single periodic orbit so it follows that either $y \ll \Phi_{t_0}(x)$ or $\Phi_{t_0}(x) \ll y$. Suppose that the latter holds, the argument is similar in the other case. Then there exists $t_1 > t_0$ such that $\Phi_{t_1}(x)$ is so near y that $\Phi_{t_0}(x) \ll \Phi_{t_1}(x)$. But then the Convergence Criterion from Chapter 1 implies that $\Phi_t(x)$ converges to equilibrium, a contradiction to our assumption that L contains no equilibria. This proves the theorem in this case.

If $L = \alpha(x)$ and $Q(L)$ consists of an annulus of periodic orbits, let $C \subset L$ be a periodic orbit such that $Q(L)$ contains C in its interior. $Q(C)$ separates $Q(L)$ into two components. Fix a and b in $L \setminus C$ such that $Q(a)$ and $Q(b)$ belong to different components of $Q(L) \setminus Q(C)$. Since $\Phi_t(x)$ repeatedly visits every neighborhood of a and b as $t \to -\infty$, $Q(\Phi_t(x))$ must cross $Q(C)$ at a sequence of times $t_k \to -\infty$. Therefore, there exist $z_k \in C$ such that $Q(z_k) = Q(\Phi_{t_k}(x))$ and consequently, as in the previous case, either $z_k \ll \Phi_{t_k}(x)$

or $\Phi_{t_k}(x) \ll z_k$ holds for each k. Passing to a subsequence, we can assume that either $z_k \ll \Phi_{t_k}(x)$ holds for all k or $\Phi_{t_k}(x) \ll z_k$ holds for all k. Assume the latter, the argument is essentially the same in the other case. We claim that for every $s < 0$ there is a point $w \in C$ such that $\Phi_s(x) > w$. For if $t_k < s$ then

$$\Phi_s(x) = \Phi_{s-t_k} \circ \Phi_{t_k}(x) < \Phi_{s-t_k}(z_k) \in C.$$

If $y \in L$ then $\Phi_{s_n}(x) \to y$ for some sequence $s_n \to -\infty$. By the claim, there exists $w_n \in C$ such that $\Phi_{s_n}(x) > w_n$. Passing to a subsequence if necessary, we can assume that $w_n \to w \in C$ and $y \geqslant w$. Therefore, every point of L is related by $\leqslant$ to some point of C.

The same reasoning applies to every periodic orbit $C' \subset L$ for which $Q(C')$ belongs to the interior of $Q(L)$: either every point of L is $\leqslant$ some point of C' or every point of L is $\geqslant$ some point of C'. Since there are three different periodic orbits in L whose projections are contained in the interior of $Q(L)$, there will be two of them for which the same inequality holds between points of L and points of the orbit. Consider the case that there are two periodic orbits C_1 and C_2 such that every point of L is $\leqslant$ some point of C_1 and $\leqslant$ some point of C_2. The case that the opposite relations hold is treated similarly. If $u \in C_1$ then it belongs to L so we can find $w \in C_2$ such that $u < w$ (equality can't hold since the points belong to different periodic orbits). But $w \in L$ so we can find $z \in C_1$ such that $w < z$. Consequently, $u, z \in C_1$ satisfy $u < z$, a contradiction to Proposition 3.16. This completes the proof. □

A remarkable fact about three-dimensional competitive or cooperative systems on suitable domains is that the existence of a periodic orbit implies the existence of an equilibrium point inside a certain semi-invariant closed ball having the periodic orbit on its boundary. Its primary use is to locate equilibria, or conversely, to exclude periodic orbits. The construction below is adapted from Smith [187,194] where the case $K = \mathbb{R}^3_+$ was treated; here we treat the general case that K has nonempty interior. The terms "competitive" and "cooperative" will be used to mean K-competitive and K-cooperative for brevity. A related result appears in Hirsch [75]. Throughout the remainder of this section, the system is assumed to be defined on a p-convex subset D of $\mathbb{R}^3_+$.

We can assume the system is competitive. Let γ denote the periodic orbit and assume that there exist p, q with $p \ll q$ such that

$$\gamma \subset [p, q] \subset D. \tag{3.10}$$

Define

$$B = \{x \in \mathbb{R}^3\colon x \text{ is not related to any point } y \in \gamma\} = (\gamma + K)^{\mathrm{c}} \cap (\gamma - K)^{\mathrm{c}}.$$

Here we use the notation A^{c} for the complement of the subset A in $\mathbb{R}^3$. Observe that in defining B we ignored the domain D of (3.4), viewing γ as a subset of $\mathbb{R}^3$. Another way to define B is to express its complement as $B^{\mathrm{c}} = (\gamma + K) \cup (\gamma - K)$.

A 3-*cell* is a subset of $\mathbb{R}^3$ that is homeomorphic to the open unit ball.

THEOREM 3.24. *Let γ be a nontrivial periodic orbit of a competitive system in $D \subset \mathbb{R}^3$ and suppose that* (3.10) *holds. Then B is an open subset of $\mathbb{R}^3$ consisting of two connected components, one bounded and one unbounded. The bounded component, $B(\gamma)$, is a* 3*-cell contained in $[p,q]$. Furthermore, $B(\gamma)$ is positively invariant and its closure contains an equilibrium.*

Combining this result with Theorem 3.23 leads to the following dichotomy from Hirsch [75].

COROLLARY 3.25. *Assume the domain $D \subset \mathbb{R}^3$ of a cooperative or competitive system contains $[p,q]$ with $p \ll q$. Then one of the following holds*:

(i) *$[p,q]$ contains an equilibrium*;

(ii) *the forward and backward semi-orbits of every point of $[p,q]$ meet $D \setminus [p,q]$.*

PROOF. We take the system to be competitive, otherwise reversing time. Assume (ii) is false. Then $[a,b]$ contains a compact limit set L. If L is not a cycle, it contains an equilibrium by Theorem 3.23. If L is a cycle, (i) follows from Theorem 3.24. □

PROOF SKETCH OF THEOREM 3.24. That B is open is a consequence of the fact that $\gamma + K$ and $\gamma - K$ are closed. We show that $B \cap D$ is positively invariant. If $x \in B \cap D$, $y \in \gamma$ and $t > 0$ then $\Phi_{-t}(y) \in \gamma$ so x is not related to it. Since the forward flow of a competitive system preserves the property of being unrelated, $\Phi_t(x)$ is unrelated to y. Therefore, $\Phi_t(x) \in B \cap D$.

As in the proof of Theorem 3.17, for $v > 0$, H_v denotes the hyperplane orthogonal to v and Q the orthogonal projection onto H_v along v. Q is one-to-one on γ so $Q(\gamma)$ is a Jordan curve in H_v. Let H_{i} and H_{e} denote the interior and exterior components of $H_v \setminus Q(\gamma)$. If $x \in Q^{-1}(Q(\gamma))$ then $Q(x) = Q(y)$ for some $y \in \gamma$ and therefore either $x = y$, $x \ll y$ or $y \ll x$. In any case, $x \notin B$. Hence,

$$B = \big(B \cap Q^{-1}(H_{\mathrm{i}})\big) \cup \big(B \cap Q^{-1}(H_{\mathrm{e}})\big).$$

Set $B(\gamma) = B \cap Q^{-1}(H_{\mathrm{i}})$.

Given $z \in H_{\mathrm{i}}$, let $A_z^+ := \{s \in \mathbb{R}\colon z + sv \in \gamma + K\}$ and $A_z^- := \{s \in \mathbb{R}\colon z + sv \in \gamma - K\}$. A_z^+ clearly contains all large s by compactness of γ and it is closed because $\gamma + K$ is closed. If $s \in A_z^+$, there exists $y \in \gamma$ and $k \in K$ such that $z + sv = y + k$ so $z + (s+r)v = y + k + rv$, implying that $s + r \in A_z^+$ for all $r \geqslant 0$. It follows that $A_z^+ = [s_+(z), \infty)$, and similarly, $A_z^- = (-\infty, s_-(z)]$. If $s_-(z) \geqslant z_+(z)$ so $A_z^+ \cap A_z^-$ is nonempty, then there exists $s \in \mathbb{R}$, $k_i \in K$, and $y_i \in \gamma$ such that $z + sv = y_1 + k_1 = y_2 - k_2$. We must have $k_1 = k_2 = 0$ or else $y_2 > y_1$, a contradiction to Proposition 3.16, but then $z + sv = y_1$ so $z = Qy_1$ contradicting that $z \in H_{\mathrm{i}}$. We conclude that $s_-(z) < z_+(z)$ and that $z + sv \in B(\gamma)$ if and only if $s_-(z) < s < s_+(z)$. It follows that

$$B(\gamma) = \big\{z + sv\colon\ z \in H_{\mathrm{i}},\ s \in \big(s_-(z), s_+(z)\big)\big\}.$$

It is easy to show that the maps $z \mapsto s_{\pm}(z)$ are continuous and satisfy $s_+(z) - s_-(z) \to 0$ as $z \to y \in \gamma$ and this implies that $B(\gamma)$ is a 3-cell. See the argument given in Smith [187, 194].

To prove $B(\gamma) \subset [p, q]$, we identify K^* as the set of x such that $\langle x, k\rangle \geqslant 0$ for all $k \in K$ (where $\langle x, k\rangle$ denotes inner product). Schneider and Vidyasagar [177] proved the elegant result that every vector x has a unique representation

$$x = k - w, \quad k \in K,\ w \in K^*,\ \langle w, k\rangle = 0.$$

Choose any $z \in B \cap (\mathbb{R}^3 \setminus [p, q])$ and write

$$z - p = k - w, \quad k \in K, w \in K^*,\ \langle w, k\rangle = 0,$$

$$q - z = k' - w', \quad k' \in K, w' \in K^*,\ \langle w', k'\rangle = 0.$$

Observe that $w > 0$, $w' > 0$ because $z \in B$.

Either $k > 0$ or $k' > 0$. For if $k = k' = 0$ then $q - p = -(w + w')$, so

$$0 \leqslant \langle w + w', q - p\rangle = -\|w + w'\|^2 \leqslant 0.$$

This entails $w + w' = 0$ and thus $p = q$, a contradiction.

We assume $k > 0$, as the case $k' > 0$ is similar, and even follows formally by replacing K with $-K$. Then $w > 0$. Consider the ray $R = \{z + tk\colon\ t \geqslant 0\}$. If $y \in \gamma$, then

$$\langle w, z + tk - y\rangle = \langle w, z - p\rangle + \langle w, p - y\rangle \leqslant \langle w, z - p\rangle = -\|w\|^2 < 0.$$

Because z and u are unrelated, there exists $u \in K^*$ such that $\langle u, z - y\rangle > 0$. So

$$\langle u, z + tk - y\rangle = \langle u, z - y\rangle + t(z, k) \geqslant \langle u, z - y\rangle > 0.$$

This shows that no point of R is related to any point of γ. Therefore R and hence z are in the unbounded component of B.

As $B(\gamma)$ is a connected component of the positively invariant set B, it is positively invariant. Consequently its closure is a positively invariant set homeomorphic to the closed unit ball in $\mathbb{R}^3$. It therefore contains an equilibrium by a standard argument using the Brouwer Fixed Point Theorem (see, e.g., Hale [57, Theorem I.8.2]). □

If $B(\gamma)$ contains only nondegenerate equilibria $x_1, x_2, \ldots, x_m$, then standard topological degree arguments imply that m is odd and that $1 = \sum_{i=1}^{m} (-1)^{s_i}$ where $s_i \in \{0, 1, 2, 3\}$ is the number of positive eigenvalues of $Df(x_i)$. See Smith [187] for the proof and further information on equilibria in $B(\gamma)$.

There are many papers devoted to competitive Lotka–Volterra systems in $\mathbb{R}^3$, largely stimulated by the work of M. Zeeman. See for example [82,223,237,239,243,240,242] and references therein. The paper of Li and Muldowney [115] contains an especially nice application to epidemiology. Additional results for three-dimensional competitive and cooperative systems can be found in references [66,75–77,41,194,196,247,248].

The recent paper of Ortega and Sánchez [153] is noteworthy for employing a cone related to the ice-cream cone and observing that results for competitive systems are valid for general cones with nonempty interior. They show that if P is a symmetric matrix of dimension n having one positive eigenvalue λ_+ with corresponding unit eigenvector e_+, and $n-1$ negative eigenvalues, then (3.4) is monotone with respect to the order generated by the cone $K := \{x \in \mathbb{R}^n\colon\ \langle Px, x\rangle \geqslant 0,\ \langle x, e_+\rangle \geqslant 0\}$ if and only if there exists a function $\mu : \mathbb{R}^n \to \mathbb{R}$ such that the matrix $P \cdot Df_x + (Df_x)^T \cdot P + \mu(x)P$ is positive semidefinite for all x. They use this result to show that one of the results of R.A. Smith [204] on the existence of an orbitally stable periodic orbit, in the special case $n = 3$, follows from the results for competitive systems. It is not hard to see that if (3.4) satisfies the conditions above then after a change of variables in (3.4), the resulting system is monotone with respect to the standard ice-cream cone.

For applications of competitive and cooperative systems, see for example Benaïm [15], Benaïm and Hirsch [16,17], Hirsch [69,74] Hofbauer and Sandholm [81], Hsu and Waltman [84], Smith [194,196], Smith and Waltman [202].

4. Delay differential equations

4.1. *The semiflow*

The aim of the present section is to apply the theory developed in Sections 1 and 2 to differential equations containing delayed arguments. Such equations are often referred to as delay differential equations or functional differential equations. Since delay differential equations contain ordinary differential equations as a special case, when all delays are zero, the treatment is quite similar to the previous section. The main difference is that a delay differential equation generally can't be solved backward in time and therefore there is not a well-developed theory of competitive systems with delays.

Delay differential equations generate infinite-dimensional dynamical systems and there are several choices of state space. We restrict attention here to equations with bounded delays and follow the most well-developed theory (see Hale and Verduyn Lunel [61]). If r denotes the maximum delay appearing in the equation, then the space $\mathcal{C} := C([-r, 0], \mathbb{R}^n)$ is a natural choice of state space. Given a cone K in $\mathbb{R}^n$, $\mathcal{C}_K$ contains the cone of functions which map $[-r, 0]$ into K. The section begins by identifying sufficient conditions on the right hand side of the delay differential equation for the semiflow to be monotone with respect to the ordering induced by this cone. This quasimonotone condition reduces to the quasimonotone condition for ordinary differential equations when no delays are present. Our main goal is to identify sufficient conditions for a delay differential equation to generate an eventually strongly monotone semiflow so that results from Sections 1 and 2 may be applied.

In order to motivate fundamental well-posedness issues for delay equations, it is useful to start with a consideration of a classical example that has motivated much research in the field (see, e.g., Krisztin et al. [105] and Hale and Verduyn Lunel [61]), namely the equation

$$x'(t) = -x(t) + h\big(x(t-r)\big), \quad t \geqslant 0, \tag{4.1}$$

where h is continuous and $r > 0$ is the delay. It is clear that $x(t)$ must be prescribed on the interval $[-r, 0]$ in order that it be determined for $t \geqslant 0$. A natural space of initial conditions is the space of continuous functions on $[-r, 0]$, which we denote by $\mathcal{C}$, where $n = 1$ in this case. $\mathcal{C}$ is a Banach space with the usual uniform norm $|\phi| = \sup\{|\phi(\theta)|\colon -r \leqslant \theta \leqslant 0\}$. If $\phi \in \mathcal{C}$ is given, then it is easy to see that the equation has a unique solution $x(t)$ for $t \geqslant 0$ satisfying

$$x(\theta) = \phi(\theta), \quad -r \leqslant \theta \leqslant 0.$$

If the state space is $\mathcal{C}$, then we need to construct from the solution $x(t)$, an element of the space $\mathcal{C}$ to call the state of the system at time t. It should have the property that it uniquely determines $x(s)$ for $s \geqslant t$. The natural choice is $x_t \in \mathcal{C}$, defined by

$$x_t(\theta) = x(t+\theta), \quad -r \leqslant \theta \leqslant 0.$$

Then, $x_0 = \phi$ and $x_t(0) = x(t)$.

The general autonomous functional differential equation is given by

$$x'(t) = f(x_t), \tag{4.2}$$

where $f : D \to \mathbb{R}^n$, D is an open subset of $\mathcal{C}$ and f is continuous. In the example above, f is given by $f(\phi) = -\phi(0) + h(\phi(-r))$ for $\phi \in \mathcal{C}$. Observe that (4.2) includes the system of ordinary differential equations

$$x' = g(x),$$

where $g : \mathbb{R}^n \to \mathbb{R}^n$, as a special case. Simply let $f(\phi) = g(\phi(0))$ so that $f(x_t) = g(x_t(0)) = g(x(t))$.

It will always be assumed that (4.2), together with the initial condition $x_0 = \phi \in D$ has a unique, maximally defined solution, denoted by $x(t,\phi)$, on an interval $[0,\sigma)$. The state of the system is denoted by $x_t(\phi)$ to emphasize the dependence on the initial data. Uniqueness of solutions holds if, for example, f is Lipschitz on compact subsets of D (see Hale and Verduyn Lunel [61]). This holds, for example, if $f \in C^1(D)$ has locally bounded derivative. If uniqueness of solutions of initial value problems hold, then the map $(t,\phi) \to x_t(\phi)$ is continuous. Therefore, a (local) semiflow on D can be defined by

$$\Phi_t(\phi) = x_t(\phi). \tag{4.3}$$

In contrast to the case of ordinary differential equations, $x(t,\phi)$ cannot usually be defined for $t \leqslant 0$ as a solution of (4.2) and consequently, Φ_t need not be one-to-one.

It will be convenient to have notation for the natural embedding of $\mathbb{R}^n$ into $\mathcal{C}$. If $x \in \mathbb{R}^n$, let $\hat{x} \in \mathcal{C}$ be the constant function equal to x for all values of its argument. The set of equilibria for (4.2) is given by

$$E = \{\hat{x} \in D\colon\ x \in \mathbb{R}^n \text{ and } f(\hat{x}) = 0\}.$$

4.2. *The quasimonotone condition*

Given that $\mathcal{C}$ is a natural state space for (4.2), we now consider what sort of cones in $\mathcal{C}$ will yield useful order relations. The most natural such cones are those induced by cones in $\mathbb{R}^n$. Let K be a cone in $\mathbb{R}^n$ with nonempty interior and K^* denote the dual cone. All inequalities hereafter are assumed to be those induced on $\mathbb{R}^n$ by K. The cone K induces a cone $\mathcal{C}_K$ in the Banach space $\mathcal{C}$ defined by

$$\mathcal{C}_K = \{\phi \in \mathcal{C}\colon\ \phi(\theta) \geqslant 0,\ -r \leqslant \theta \leqslant 0\}.$$

It has nonempty interior in $\mathcal{C}$ given by $\operatorname{Int} \mathcal{C}_K = \{\phi \in \mathcal{C}_K\colon\ \phi(\theta) \gg 0, \theta \in [-r, 0]\}$. The usual notation $\leqslant, <, \ll$ will be used for the various order relations on $\mathcal{C}$ generated by $\mathcal{C}_K$. In particular, $\phi \leqslant \psi$ holds in $\mathcal{C}$ if and only if $\phi(s) \leqslant \psi(s)$ holds in $\mathbb{R}^n$ for every $s \in [-r, 0]$. The same notation will also be used for the various order relations on $\mathbb{R}^n$ but hopefully the context will alert the reader to the appropriate meaning. Cones in $\mathcal{C}$ that are not induced by a cone in $\mathbb{R}^n$ have also proved useful. See Smith and Thieme [198,200,194].

An immediate aim is to identify sufficient conditions on f for the semiflow Φ to be a monotone semiflow. The following condition should seem natural since it generalizes the condition (QM) for ordinary differential equations in the previous section. We refer to it here as the *quasimonotone condition*, (QMD) for short. “D” in the notation, standing for delay, is used so as not to confuse the reader with (QM) of the previous section. We follow this pattern in several definitions in this section.

(QMD) $\phi, \psi \in D, \phi \leqslant \psi$ and $\eta(\phi(0)) = \eta(\psi(0))$ for some $\eta \in K^*$, implies $\eta(f(\phi)) \leqslant \eta(f(\psi))$.

For the special case $K = \mathbb{R}^n_+$, (QMD) becomes:

$$\phi, \psi \in D,\ \phi \leqslant \psi \text{ and } \phi_i(0) = \psi_i(0) \quad \text{implies} \quad f_i(\phi) \leqslant f_i(\psi).$$

As in Section 3, it is convenient to consider the nonautonomous equation

$$x'(t) = f(t, x_t), \tag{4.4}$$

where $f : \Omega \to \mathbb{R}^n$ is continuous on Ω, an open subset of $\mathbb{R} \times \mathcal{C}$. Given $(t_0, \phi) \in \Omega$, we write $x(t, t_0, \phi, f)$ and $x_t(t_0, \phi, f)$ for the maximally defined solution and state of the system at time t satisfying $x_{t_0} = \phi$. We assume this solution is unique, which will be the case if f is Lipschitz in its second argument on each compact subset of Ω. We drop the last argument f from $x(t, t_0, \phi, f)$ when no confusion over which f is being considered will result.

$f : \Omega \to \mathbb{R}^n$ is said to satisfy (QMD) if $f(t, \cdot)$ satisfies (QMD) on $\Omega_t \equiv \{\phi \in \mathcal{C}\colon (t, \phi) \in \Omega\}$ for each t.

The next theorem not only establishes the desired monotonicity of the semiflow Φ but also allows comparisons of solutions between related functional differential equations. It generalizes Theorem 3.2 of Chapter 3 to functional differential equations and is a generalization of Proposition 1.1 of [190] and Theorem 5.1.1 of [194] where $K = \mathbb{R}^n_+$ is considered. The quasimonotone condition for delay differential equations seems first to have appeared in the work of Kunisch and Schappacher [109], Martin [128], and Ohta [152].

THEOREM 4.1. *Let $f, g : \Omega \to \mathbb{R}^n$ be continuous, Lipschitz on each compact subset of Ω, and assume that either f or g satisfies* (QMD). *Assume also that $f(t, \phi) \leqslant g(t, \phi)$ for all $(t, \phi) \in \Omega$. Then*

$$\phi, \psi \in \Omega_{t_0},\ \phi \leqslant \psi,\ t \geqslant t_0, \quad \Longrightarrow \quad x(t, t_0, \phi, f) \leqslant x(t, t_0, \psi, g)$$

for all t for which both are defined.

PROOF. Assume that f satisfies (QMD), a similar argument holds if g satisfies (QMD). Let $e \in \mathbb{R}^n$ satisfy $e \gg 0$, $g_\epsilon(t, \phi) := g(t, \phi) + \epsilon e$ and $\psi_\epsilon := \psi + \epsilon \hat{e}$, for $\epsilon \geqslant 0$. If $x(t, t_0, \psi, g)$ is defined on $[t_0 - r, t_1]$ for some $t_1 > t_0$, then $x(t, t_0, \psi_\epsilon, g_\epsilon)$ is also defined on this same interval for all sufficiently small positive ϵ and

$$x(t, t_0, \psi_\epsilon, g_\epsilon) \to x(t, t_0, \psi, g), \quad \epsilon \to 0,$$

for $t \in [t_0, t_1]$ by Hale and Verduyn Lunel [61, Theorem 2.2.2]. We will show that $x(t, t_0, \phi, f) \ll x(t, t_0, \psi_\epsilon, g_\epsilon)$ on $[t_0 - r, t_1]$ for small positive ϵ. The result will then follow by letting $\epsilon \to 0$. If the assertion above were false for some ϵ, then applying the remark below Proposition 3.1, there exists $s \in (t_0, t_1]$ such that $x(t, t_0, \phi, f) \ll x(t, t_0, \psi_\epsilon, g_\epsilon)$ for $t_0 \leqslant t < s$ and $\eta(x(s, t_0, \phi, f)) = \eta(x(s, t_0, \psi_\epsilon, g_\epsilon))$ for some nontrivial $\eta \in K^*$. As $\eta(x(t, t_0, \phi, f)) < \eta(x(t, t_0, \psi_\epsilon, g_\epsilon))$ for $t_0 \leqslant t < s$, by Proposition 3.1, we conclude that $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=s}\eta(x(s, t_0, \phi, f)) \geqslant \frac{\mathrm{d}}{\mathrm{d}t}|_{t=s}\eta(x(s, t_0, \psi_\epsilon, g_\epsilon))$. But

$$\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=s} \eta\big(x(s, t_0, \psi_\epsilon, g_\epsilon)\big) &= \eta\big(g\big(s, x_s(t_0, \psi_\epsilon, g_\epsilon)\big)\big) + \epsilon\eta(e) \\
&> \eta\big(f\big(s, x_s(t_0, \psi_\epsilon, g_\epsilon)\big)\big) \\
&\geqslant \eta\big(f\big(s, x_s(t_0, \phi, f)\big)\big) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=s} \eta\big(x(s, t_0, \phi, f)\big),
\end{aligned}$$

where the last inequality follows from (QMD). This contradiction implies that no such s can exist, proving the assertion. □

In the case of the autonomous system (4.2), taking $f = g$ in Theorem 1.1 implies that $x_t(\phi) \leqslant x_t(\psi)$ for $t \geqslant 0$ such that both solutions are defined. In other words, the semiflow Φ defined by (4.3) is monotone. In contrast to Theorem 3.2 of the previous section, if $\phi < \psi$ we cannot conclude that $x(t,\phi) < x(t,\psi)$ or $x_t(\phi) < x_t(\psi)$ since Φ_t is not generally one-to-one. A simple example is provided by the scalar equation (4.2) with $r = 1$ and $f(\phi) := \max \phi$, which satisfies (QMD). Let $\phi < \psi$ be strictly increasing on $[-1,-1/2]$, $\phi(-1) = \psi(-1) = 0$, $\phi(-1/2) = \psi(-1/2) = 1$, and $\phi(\theta) = \psi(\theta) = -2\theta$ for $-1/2 < \theta \leqslant 0$. It is easy to see that $x(t,\phi) = x(t,\psi)$ for $t \geqslant 0$.

It is useful to have sufficient conditions for the positive invariance of K. By this we mean that $t_0 \in J$ and $\phi \geqslant 0$ implies $x(t,t_0,\phi) \geqslant 0$ for all $t \geqslant t_0$ for which it is defined. The following result provides the expected necessary and sufficient condition. The proof is similar to that of Theorem 4.1; the result is the delay analog of Proposition 3.3.

THEOREM 4.2. *Assume that $J \times K \subset \Omega$ where J is an open interval. Then K is positively invariant for* (4.4) *if and only if for all $t \in J$*

(PD) $\phi \geqslant 0$, $\lambda \in K^*$ *and* $\lambda(\phi(0)) = 0$ *implies* $\lambda(f(t,\phi)) \geqslant 0$

holds.

Let $L: J \to L(\mathcal{C},\mathbb{R}^n)$ be continuous, where $L(\mathcal{C},\mathbb{R}^n)$ denotes the space of bounded linear operators from $\mathcal{C}$ to $\mathbb{R}^n$, and consider the initial value problem for the linear nonautonomous functional differential equation

$$x' = L(t)x_t, \quad x_{t_0} = \phi. \tag{4.5}$$

Observing that (PD) and (QMD) are equivalent for linear systems, we have the following corollary.

COROLLARY 4.3. *Let $x(t,t_0,\phi)$ be the solution of* (4.5). *Then $x(t,t_0,\phi) \geqslant 0$ for all $t \geqslant t_0$ and all $\phi \geqslant 0$ if and only if for each $t \in J$,* (PD) *holds for $L(t)$.*

As in the case of ordinary differential equations, a stronger condition than (PD) for linear systems is that for every $t \in J$, there exists $\alpha \in \mathbb{R}$ such that $L(t)\phi + \alpha\phi(0) \geqslant 0$ whenever $\phi \geqslant 0$.

It is useful to invoke the Riesz Representation Theorem [171] in order to identify $L(t)$ with a matrix of signed Borel measures $\eta(t) = (\eta(t)_{ij})$:

$$L(t)\phi = \int_{-r}^{0} \mathrm{d}\eta(t)\phi. \tag{4.6}$$

The Radon–Nikodym decomposition of η_{ij} with respect to the Dirac measure δ with unit mass at 0 gives $\eta_{ij}(t) = a_{ij}(t)\delta + \tilde{\eta}_{ij}(t)$ where a_{ij} is a scalar and $\tilde{\eta}_{ij}(t)$ is mutually singular with respect to δ. In particular, the latter assigns zero mass to $\{0\}$. Therefore,

$$L(t)\phi = A(t)\phi(0) + \tilde{L}(t)\phi, \quad \tilde{L}(t)\phi := \int_{-r}^{0} \mathrm{d}\tilde{\eta}(t)\phi. \tag{4.7}$$

Continuity of the map $t \to A(t)$ follows from continuity of $t \to L(t)$. The decomposition (4.7) leads to sharp conditions for (PD) to hold for $L(t)$.

PROPOSITION 4.4. (PD) *holds for* $L(t)$ *if and only if*

(a) $A(t)$ *satisfies* (P) *of Proposition* 3.3, *and*

(b) $\tilde{L}(t)\phi \geqslant 0$ *whenever* $\phi \geqslant 0$.

PROOF. If (a) and (b) hold, $\phi \geqslant 0$, $\lambda \in K^*$ and $\lambda(\phi(0)) = 0$ then $\lambda(L(t)\phi) = \lambda(A(t)\phi(0)) + \lambda(\tilde{L}(t)\phi) \geqslant 0$ because each summand on the right is nonnegative.

Conversely, if (PD) holds for $L(t)$, $v \in \partial K$, $\lambda \in K^*$, and $\lambda(v) = 0$, define $\phi_n(\theta) = \mathrm{e}^{n\theta} v$ on $[-r, 0]$. Then $\phi_n \geqslant 0$ and ϕ_n converges point-wise to zero, almost everywhere with respect to $\tilde{\eta}(t)$. By (PD),

$$\lambda(L(t)\phi_n) = \lambda(A(t)v + \tilde{L}(t)\phi_n) \geqslant 0.$$

Letting $n \to \infty$, we get $\lambda(A(t)v) \geqslant 0$ implying that (P) holds for $A(t)$. Let $\phi \geqslant 0$ be given and define $\phi_n(\theta) = [1 - \mathrm{e}^{n\theta}]\phi(\theta)$ on $[-r, 0]$, $n \geqslant 1$. ϕ_n converges point-wise to $\phi\chi$, where χ is the indicator function of the set $[-r, 0)$, and $\phi\chi = \phi$ almost everywhere with respect to $\tilde{\eta}(t)$. If $\lambda \in K^*$, then $\lambda(\phi_n(0)) = 0$ so applying (PD) we get $0 \leqslant \lambda(L(t)\phi_n) = \tilde{L}(t)\phi_n$. Letting $n \to \infty$ we get (b). □

For the remainder of this section, we suppose that $\Omega = J \times D$ where J is a nonempty open interval and $D \subset \mathcal{C}$ is open. Suppose that $\frac{\partial f}{\partial \phi}(t, \psi)$ exists and is continuous on $J \times D$ to $L(\mathcal{C}, \mathbb{R}^n)$. In that case, $x(t, t_0, \phi)$ is continuously differentiable in its last argument and $y(t, t_0, \chi) = \frac{\partial x}{\partial \phi}(t, t_0, \phi)\chi$ satisfies the variational equation

$$y'(t) = \frac{\partial f}{\partial \phi}\bigl(t, x_t(t, \phi)\bigr)y_t, \quad y_{t_0} = \chi. \tag{4.8}$$

See Theorem 2.4.1 of Hale and Verduyn Lunel [61]. We say that f (or (4.4)) is *K-co-operative* if for all $(t, \chi) \in J \times D$ the function $\psi \to \frac{\partial f}{\partial \phi}(t, \chi)\psi$ satisfies (PD). By Corollary 4.3 applied to the variational equation we have the following analog of Theorem 3.5 for functional differential equations. The proof is essentially the same.

THEOREM 4.5. *Let* $\frac{\partial f}{\partial \phi}(t, \psi)$ *exist and be continuous on* $J \times D$. *If* (QMD) *holds for* (4.4), *then* f *is* K*-cooperative. Conversely, if* D *is p-convex and* f *is* K*-cooperative, then* (QMD) *holds for* f.

Consider the nonlinear system

$$x'(t) = g\bigl(x(t), x(t - r_1), x(t - r_2), \ldots, x(t - r_m)\bigr), \tag{4.9}$$

where $g(x, y^1, y^2, \ldots, y^m)$ is continuously differentiable on $\mathbb{R}^{(m+1)n}$ and $r_{j+1} > r_j > 0$. Then

$$\frac{\partial f}{\partial \phi}(\psi) = \frac{\partial g}{\partial x}(x, Y)\delta + \sum_k \frac{\partial g}{\partial y^k}(x, Y)\delta_{-r_k}, \tag{4.10}$$

where δ_{-r_k} is the Dirac measure with unit mass at $\{-r_k\}$ and $x = \psi(0)$, $y^k = \psi(-r_k)$ and $(x, Y) := (x, y^1, y^2, \ldots, y^m)$. By Theorem 4.5, Corollary 4.3, and Proposition 4.4, (QMD) holds if and only if for each (x, Y), $\frac{\partial g}{\partial x}(x, Y)$ satisfies condition (P) and $\frac{\partial g}{\partial y^k}(x, Y)$ is K-positive. If $K = \mathbb{R}^n_+$, the condition becomes $\frac{\partial g_i}{\partial x_j}(x, Y) \geqslant 0$, for $i \neq j$ and $\frac{\partial g_i}{\partial y^k_j}(x, Y) \geqslant 0$ for all i, j, k; if, in addition, $n = 1$ then $\frac{\partial g}{\partial y^k}(x, Y) \geqslant 0$ for all k suffices.

4.3. *Eventual strong monotonicity*

We begin by considering the linear system (4.5). The following hypothesis for the continuous map $L : J \to L(\mathcal{C}, \mathbb{R}^n)$ reduces to (ST) of the previous section when $r = 0$:

(STD) for all $t \in J$ and $\phi \geqslant 0$ with $\phi(0) \in \partial K$ satisfying one of the conditions
(a) $\phi(-r) > 0$ and $\phi(0) = 0$, or
(b) $\phi(s) > 0$ for $-r \leqslant s \leqslant 0$,
there exists $\nu \in K^*$ such that $\nu(\phi(0)) = 0$ and $\nu(L(t)\phi) > 0$.

The following result is the analog of Theorem 3.6 of the previous section for delay differential equations.

THEOREM 4.6. *Let linear system* (4.5) *satisfy* (PD) *and* (STD) *and let $t_0 \in J$. Then*

$$\phi > 0,\ t \geqslant t_0 + 2r \quad \Longrightarrow \quad x(t, t_0, \phi) \gg 0.$$

In particular, $x_t(t_0, \phi) \gg 0$ for $t \geqslant t_0 + 3r$.

PROOF. By Corollary 4.3, we have that $x(t) := x(t, t_0, \phi) \geqslant 0$ for all $t \geqslant t_0$ that belong to J. There exists $t_1 \in (t_0, t_0 + r)$ such that $x(t_1 - r) = \phi(t_1 - r) = x_{t_1}(-r) > 0$ since $\phi > 0$. If $x(t_1) = 0$, then (STD)(a) implies the existence of $\nu \in K^*$ such that $\nu(L(t_1)x_{t_1}) > 0$. As $\nu(x(t)) \geqslant 0$ for $t \geqslant t_0$ and $\nu(x(t_1)) = 0$ we conclude that $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=t_1}\nu(x(t)) \leqslant 0$. But $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=t_1}\nu(x(t)) = \nu(L(t_1)x_{t_1}) > 0$, a contradiction. Therefore, $x(t_1) > 0$.

Now, by (4.7)

$$x' = A(t)x + \tilde{L}(t)x_t$$

from which we conclude

$$x(t) = X(t, t_1)x(t_1) + \int_{t_1}^{t} X(t, r)\tilde{L}(r)x_r \,\mathrm{d}r,$$

where $X(t, t_0)$ is the fundamental matrix for $y' = A(t)y$ satisfying $X(t_0, t_0) = I$. From (a) of Proposition 4.4 and Corollary 3.4, it follows that $X(t, t_0)$ is K positive for $t \geqslant t_0$. This, the fact that $x_r \geqslant 0$, and (b) of Proposition 4.4 imply that the integral belongs to K so we conclude that

$$x(t) \geqslant X(t, t_1)x(t_1) > 0, \quad t \geqslant t_1.$$

We claim that $x(t) \gg 0$ for $t \geqslant t_1 + r$. If not, there is a $t_2 \geqslant t_1 + r$ such that $x(t_2) = x_{t_2}(0) \in \partial K$ but $x_{t_2}(s) > 0$ for $-r \leqslant s \leqslant 0$. Then (STD) implies the existence of $\nu \in K^*$ such that $\nu(x(t_2)) = 0$ and $\nu(L(t_2)x_{t_2}) > 0$. Since $\nu(x(t)) \geqslant 0$ for $t \geqslant t_0$ we must have $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=t_2}\nu(x(t)) \leqslant 0$. But $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=t_2}\nu(x(t)) = \nu(L(t_2)x_{t_2}) > 0$, a contradiction. We conclude that $x(t) \gg 0$ for $t \geqslant t_1 + r$. □

In a sense, (STD)(a) says that r has been correctly chosen; (STD)(b) is more fundamental. The next result gives sufficient conditions for it to hold.

PROPOSITION 4.7. *If $L(t)$ satisfies* (PD) *and either*
(a) *$A(t)$ satisfies* (ST), *or*
(b) $\phi > 0 \Longrightarrow \tilde{L}(t)\phi \gg 0$
then (STD)(b) *holds.*

PROOF. This is immediate from the definitions, the decomposition (4.7), Proposition 4.4, and the expression $\nu(L(t)\phi) = \nu(A(t)\phi(0)) + \nu(\tilde{L}(t)\phi)$. □

Theorem 4.6 leads immediately to a result on eventual strong monotonicity for the nonlinear system (4.4) where we assume that $\Omega = J \times D$ as above.

THEOREM 4.8. *Let D be p-convex, $\frac{\partial f}{\partial \phi}(t, \psi)$ exist and be continuous on $J \times D$ to $L(\mathcal{C}, \mathbb{R}^n)$, and f be K-cooperative. Suppose that* (STD) *holds for $\frac{\partial f}{\partial \phi}(t, \psi)$, for each $(t, \psi) \in J \times D$. Then*

$$\phi_0, \phi_1 \in D,\ \phi_0 < \phi_1 \quad \Longrightarrow \quad x(t, t_0, \phi_0) \ll x(t, t_0, \phi_1)$$

for all $t \geqslant t_0 + 2r$ for which both solutions are defined.

PROOF. By Theorem 4.5, we have $x(t, t_0, \phi_0) \leqslant x(t, t_0, \phi_1)$ for $t \geqslant t_0$ for which both solutions are defined. We apply the formula

$$x(t, t_0, \phi_1) - x(t, t_0, \phi_0) = \int_0^1 \frac{\partial x}{\partial \phi}\big(t, t_0, s\phi_1 + (1 - s)\phi_0\big)(\phi_1 - \phi_0)\,\mathrm{d}s.$$

Here, for $\psi \in D$ and $\beta \in \mathcal{C}$, $y(t, t_0, \beta) := \frac{\partial x}{\partial \phi}(t, t_0, \psi)\beta$ satisfies the variational equation (4.5) where $\phi = \beta$ and $L(t) = \frac{\partial f}{\partial \phi}(t, x_t(t_0, \psi))$. See Theorem 2.4.1 of Hale and Verduyn Lunel [61]. The desired conclusion will follow if we show that $y(t, t_0, \beta) \gg 0$ for

$t \geqslant t_0 + 2r$ for $\psi = s\phi_1 + (1-s)\phi_0$ and $\beta = \phi_1 - \phi_0 > 0$. By Theorem 4.6, it suffices to show that $L(t)$ satisfies (PD) and (STD). But this follows from our hypotheses. □

In the next result, Theorem 4.8 is applied to system (4.9). We make use of notation introduced below Theorem 4.5.

COROLLARY 4.9. *Let* $g : \mathbb{R}^{(m+1)n} \to \mathbb{R}^n$ *be continuously differentiable and satisfy*

(a) $\frac{\partial g}{\partial x}(x, Y)$ *satisfies* (P) *for each* $(x, Y) \in Z$;

(b) *for each* k, $\frac{\partial g}{\partial y^k}(x, Y)$ *is* K *positive*;

(c) *either* $\frac{\partial g}{\partial x}(x, Y)$ *satisfies* (ST) *or some* $\frac{\partial g}{\partial y^k}(x, Y)$ *is strongly positive on* K.

Then the hypotheses of Theorem 4.8 *hold for* (4.9).

PROOF. Recalling (4.10), it is evident that (a) and (b) imply that (4.9) is K-cooperative. Hypothesis (c) and Proposition 4.7 imply that (STD) holds. □

In the special case that (4.9) is a scalar equation, $m = 1$ and $K = \mathbb{R}_+$, then $\frac{\partial g}{\partial y}(x, y) > 0$ suffices to ensure an eventually strongly monotone semiflow.

4.4. *K is an orthant*

Our results can be improved in the case that K is a product cone such as $\mathbb{R}^n_+ = \prod_{i=1}^n \mathbb{R}_+$, i.e., an orthant. The following example illustrates the difficulty with our present set up.

$$x_1'(t) = -x_1(t) + x_2(t - 1/2),$$
$$x_2'(t) = x_1(t-1) - x_2(t).$$

Observe that (PD) holds for the standard cone. For initial data, take $\phi = (\phi_1, \phi_2) \in \mathcal{C}$ $(r = 1)$ where $\phi_1 = 0$ and $\phi_2(\theta) > 0$ for $\theta \in (-1, -2/3)$ and $\phi_2(\theta) = 0$ elsewhere in $[-1, 0]$. The initial value problem can be readily integrated by the method of steps of length $1/2$ and one sees that $x(t) = 0$ for all $t \geqslant -2/3$. In the language of semiflows, $\phi > 0$ yet $\Phi_t(\phi) = \Phi_t(0) = 0$ for all $t \geqslant 0$. The problem is that $C([-1, 0], \mathbb{R}^2)$ is not the optimal state space; a better one is the product space $X = C([-1, 0], \mathbb{R}) \times C([-1/2, 0], \mathbb{R})$. Obviously, an arbitrary cone in $\mathbb{R}^2$ will not induce a cone in the product space X.

For the remainder of this section we focus on the standard cone but the reader should observe that an analogous construction works for any orthant $K = \{x\colon\ (-1)^{m_i} x_i \geqslant 0\}$. Motivated by the example in the previous paragraph, let $r = (r_1, r_2, \ldots, r_n) \in \mathbb{R}^n_+$ be a vector of delays, $R = \max r_i$ and define

$$\mathcal{C}_r = \prod_{i=1}^n C\big([-r_i, 0], \mathbb{R}\big).$$

Note that we allow some delays to be zero. We write $\phi = (\phi_1, \phi_2, \ldots, \phi_n)$ for a generic point of $\mathcal{C}_r$. $\mathcal{C}_r$ is a Banach space with the norm $|\phi| = \sum |\phi_i|$. Let

$$\mathcal{C}_r^+ = \prod_{i=1}^n C\big([-r_i, 0], \mathbb{R}_+\big)$$

denote the cone of functions in $\mathcal{C}_r$ with nonnegative components. It has nonempty interior given by those functions with strictly positive components. As usual, we use the notation $\leqslant, <, \ll$ for the corresponding order relations on $\mathcal{C}_r$ induced by $\mathcal{C}_r^+$. If $x_i(t)$ is defined on $[-r_i, \sigma)$, $1 \leqslant i \leqslant n$, $\sigma > 0$ then we may redefine $x_t \in \mathcal{C}_r$ as $x_t = (x_t^1, x_t^2, \ldots, x_t^n)$ where $x_t^i(\theta) = x_i(t + \theta)$ for $\theta \in [-r_i, 0]$. Notice that now, the subscript signifying a particular component will be raised to a superscript when using the subscript "t" to denote a function.

If $D \subset \mathcal{C}_r$ is open, J is an open interval and $f : J \times D \to \mathbb{R}^n$ is given, then the standard existence and uniqueness theory for the initial value problem associated with (4.4) is unchanged. Furthermore, Theorems 4.1 and 4.2, and Corollary 4.3 remain valid in our current setting where, of course, we need only make use of the coordinate maps $\eta(x) = x_i$, $1 \leqslant i \leqslant n$ in (QMD) and (PD). Our goal now is to modify (STD) so that we may obtain a result like Theorem 4.6 that applies to systems such as the example above. We begin by considering the linear system (4.5) where $L : J \to L(\mathcal{C}_r, \mathbb{R}^n)$ is continuous and let $L_i(t)\phi := \langle e_i, L(t)\phi\rangle$, $1 \leqslant i \leqslant n$.

In our setting, $L(t)$ satisfies (PD) if and only if:

$$\phi \geqslant 0 \text{ and } \phi_i(0) = 0 \quad \text{implies} \quad L_i(t)\phi \geqslant 0.$$

THEOREM 4.10. *Let linear system* (4.5) *satisfy* (PD) *and*

(i) $t \in J, r_j > 0, \phi \geqslant 0, \phi_j(-r_j) > 0 \Longrightarrow L_i(t)\phi > 0$ *for some* i;

(ii) *for every proper subset* I *of* $N := \{1, 2, \ldots, n\}$, *there exists* $j \in N \setminus I$ *such that* $L_j(t)\phi > 0$ *whenever* $\phi \geqslant 0$, $\phi_i(s) > 0$, $-r_i \leqslant s \leqslant 0$, $i \in I$.

Then $x(t, \phi, t_0) \gg 0$ *if* $\phi > 0$ *for all* $t \geqslant t_0 + nR$.

PROOF. By (PD) and Corollary 4.3 we have $x(t) \geqslant 0$ for $t \geqslant t_0$. An application of the Riesz Representation Theorem and Radon–Nikodym Theorem implies that for $i = 1, 2, \ldots, n$, we have

$$L_i(t)\phi = a_i(t)\phi_i(0) + \sum_{j=1}^n \int_{-r_j}^0 \phi_j(\theta)\, \mathrm{d}_\theta \eta_{ij}(t, \theta) = a_i(t)\phi_i(0) + \bar{L}_i(t)\phi,$$

where $\eta_{ij}(t)$ is a positive Borel measure on $[-r_j, 0]$, $a_i(t) \in \mathbb{R}$ and $\bar{L}_i(t)\phi \geqslant 0$ whenever $\phi \geqslant 0$. Moreover, $t \to \eta_{ij}(t)$ and $t \to a_i(t)$ are continuous. See Smith [190,194] for details. The representation of L_i in terms of signed measures, $\bar{\eta}_{ij}$, is standard; (PD) implies that $\eta_{ij} := \bar{\eta}_{ij}$ must be positive for $i \neq j$ and that $\bar{\eta}_{ii}$ has the Lebesgue decomposition $\bar{\eta}_{ii} = a_i\delta + \eta_{ii}$ with respect to δ, the Dirac measure of unit mass at zero, and η_{ii} is a positive measure which is mutually singular with respect to δ.

If $x_i(t_1) > 0$ for some i and $t_1 > t_0$ then from $x_i'(t) = a_i(t)x_i(t) + \bar{L}_i(t)x_t \geqslant a_i(t)x_i(t)$ we conclude from standard differential inequality arguments that $x_i(t) > 0$ for $t \geqslant t_1$.

As $\phi > 0$, there exists j such that $\phi_j > 0$. If $r_j = 0$ then $x_j(t_0) > 0$; if $r_j > 0$ then $x_j(t_1 - r_j) > 0$ for some $t_1 \in (t_0, t_0 + r_j)$. In this case, it follows from (i) that $x_i'(t_1) = L_i(t_1)\phi > 0$ for some i and hence $x_i(t_1) > 0$. Hence, $x_i(t) > 0$ for $t \geqslant t_1$ by the previous paragraph. Applying (ii) with $I = \{i\}$ and $t = t_2 = t_1 + r_i$ we may find $k \neq i$ such that $x_k'(t_2) = L_k(t_2)x_{t_2} > 0$ because $x_{t_2}^i(s) > 0$, $-r_i \leqslant s \leqslant 0$. Therefore, we must have $x_k(t_2) > 0$ and hence $x_k(t) > 0$ for $t \geqslant t_2$. Obviously, we may continue in this manner until we have all components positive for $t \geqslant t_0 + nR$ as asserted. □

Theorem 4.10 leads directly to a strong monotonicity result for the nonlinear nonautonomous delay differential equation (4.4) in the usual way. We extend the definition of K-cooperativity of f to our present setup with state space $\mathcal{C}_r$ exactly as before.

THEOREM 4.11. *Let $D \subset \mathcal{C}_r$ be p-convex, $\frac{\partial f}{\partial \phi}(t, \psi)$ exist and be continuous on $J \times D$ to $L(\mathcal{C}_r, \mathbb{R}^n)$, and f be K-cooperative. Suppose that for all $(t, \psi) \in J \times D$, $L(t) := \frac{\partial f}{\partial \phi}(t, \psi)$ satisfies the conditions of Theorem* 4.10. *Then*

$$\phi_0, \phi_1 \in D,\ \phi_0 < \phi_1,\ t \geqslant t_0 + nR \quad \Longrightarrow \quad x(t, t_0, \phi_0) \ll x(t, t_0, \phi_1).$$

The biochemical control circuit with delays, modeled by the system

$$\begin{aligned} x_1'(t) &= g\big(x_n(t - r_n)\big) - \alpha_1 x_1(t), \\ x_j'(t) &= x_{j-1}(t - r_{j-1}) - \alpha_j x_j(t), \quad 2 \leqslant j \leqslant n \end{aligned} \tag{4.11}$$

with decay rates $\alpha_j > 0$ and delays $r_i \geqslant 0$ with $R > 0$ provides a good application of Theorem 4.11 which cannot be obtained by Theorem 4.8 if the delays are distinct. We assume the $g : \mathbb{R}_+ \to \mathbb{R}_+$ is continuously differentiable and $g' > 0$. Equation (4.11) is an autonomous system for which $\mathcal{C}_r^+$ is positively invariant by Theorem 4.2. See Smith [191, 194] for more on this application.

4.5. *Generic convergence for delay differential equations*

The aim of this section is to apply Theorem 4.8 and Theorem 4.11 to the autonomous delay differential equation (4.2) to conclude that the generic solution converges to equilibrium. To Φ, defined by (4.3), we associate C, S and E, denoting respectively the sets of convergent, stable and equilibrium points. The main result of this section is the following.

THEOREM 4.12. *Let $f \in C^1(D)$,* (4.2) *be cooperative on the p-convex open subset D of $\mathcal{C}$ or $\mathcal{C}_r$ and satisfy*:

(a) *The hypotheses of Theorem* 4.8 *or of Theorem* 4.11 *hold*;

(b) *Every positive semiorbit of Φ has compact closure in D and $D = AC \cup BC$.*

Then

(i) *$C \cap S$ contains a dense open subset of D, consisting of points whose trajectories converge to equilibria*;

(ii) *If E is compact there is a stable equilibrium, and an asymptotically stable equilibrium when E is finite.*

PROOF. For definiteness, suppose that (4.2) is cooperative on the p-convex open subset D of $\mathcal{C}$ and that the hypotheses of Theorem 4.8 hold. The other case is proved similarly. Assumption (a) ensures that Φ is eventually strongly monotone. Moreover, the derivative of $\Phi_t(\phi)$ with respect to ϕ exists and $\Phi'_\tau(\phi)\chi = y_\tau(t_0, \chi)$, where $y(t, t_0, \chi)$ is the solution of the variational equation (4.8). As our hypotheses ensure that $L(t) = \frac{\partial f}{\partial \phi}(x_t(\phi))$ satisfies (STD), we conclude from Theorem 4.6 that $\Phi'_\tau(\phi)$ is strongly positive for $\tau \geqslant 3r$. Compactness of $\Phi'_\tau(\phi) : \mathcal{C} \to \mathcal{C}$ for $\tau \geqslant r$ follows from the fact that a bound for $y_\tau(t_0, \chi)$, uniform for χ belonging to a bounded set $B \subset \mathcal{C}$, can be readily obtained so, using (4.8), we may also find a uniform bound for $y'(t, t_0, \chi)$, $\tau - r \leqslant t \leqslant \tau$. See, e.g., Hale [58, Theorem 4.1.1] for more detail.

The hypotheses of Theorem 2.26, with $X = D$, are fulfilled: D is normally ordered and $D = BC \cup AC$; while (M) and (D^*) hold as noted above. Therefore Theorem 2.26 implies the conclusion. □

In the special case that (4.2) is scalar ($n = 1$) we note that the set E of equilibria is totally ordered in $\mathcal{C}_r$ or $\mathcal{C}$ so the set of quasiconvergent points coincides with the set of convergent points: $Q = C$. The classical scalar delay differential equation (4.1) has been thoroughly investigated in the case of monotone delayed feedback ($f(0) = 0$ and $f' > 0$) by Krisztin et al. [105]. They characterize the closure of the unstable manifold of the trivial solution in case it is three-dimensional and determine in remarkable detail the dynamics on this invariant set.

Smith and Thieme [198,200,194] introduce an exponential ordering, not induced by a cone in $\mathbb{R}^n$, that extends the scope of application of the theory described here. One of the salient results from this work is that a scalar delay equation for which the product of the delay r and the Lipschitz constant of f is smaller than e^{-1} generates an eventually strongly monotone semiflow with respect to the exponential ordering and therefore the generic orbit converges to equilibrium: the dynamics mimics that of the associated ordinary differential equation obtained by ignoring the delay. See also work of Pituk [159].

We have considered only bounded delays. Systems of delay differential equations with unbounded and even infinite delay are also of interest. See Wu [234] for extensions to such systems. Wu and Freedman [235] and Krisztin and Wu [106–108] extend the theory to delay differential equations of neutral type.

5. Monotone maps

5.1. *Background and motivating examples*

One of the chief motivations for the study of monotone maps is their importance in the study of periodic solutions to periodic quasimonotone systems of ordinary differential equations. See for example the monograph of Krasnosel'skii [99], the much cited paper of de Mottoni and Schiaffino [42], Hale and Somolinos [60], Smith [188,189], Liang and Jiang [121], and Wang and Jiang [229–231]. To fix ideas, let $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz function and consider the ordinary differential equation

$$x' = f(t, x). \tag{5.1}$$

As usual, denote by $x(t, t_0, x_0)$ the noncontinuable solution of the initial value problem $x(t_0) = x_0$, which for simplicity we assumed is defined for all t. If f is periodic in t of period one: $f(t + 1, x) = f(t, x)$ for all (t, x), then it is natural to consider the *period map* $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$T(x_0) = x(1, 0, x_0). \tag{5.2}$$

Its fixed points (periodic points) are in one-to-one correspondence with the periodic (subharmonic) solutions of (5.1). If K is a cone in $\mathbb{R}^n$ for which f satisfies the quasimonotone condition (QM), then it follows from Theorem 3.2 that T is a monotone map: $x \leqslant y$ implies $Tx \leqslant Ty$. Moreover, T has the important property, not shared with general monotone maps, that it is an orientation-preserving homeomorphism.

In a similar way, periodic solutions for second order parabolic partial differential equations with time-periodic data can be analyzed by considering period maps in appropriate function spaces. Here monotonicity comes from classical maximum principles. Hess [63] remains an up-to-date survey. See also Alikakos et al. [3] and Zhao [245]. Remarkable results are known for equations on a compact interval with standard boundary conditions. Chen and Matano [23] show that every forward (backward) bounded solution is asymptotic to a periodic solution; Brunovsky et al. [22] extend the result to more general equations. Chen et al. [24] give conditions for the period map to generate Morse–Smale dynamics and thus be structurally stable. Although monotonicity of the period map is an important consideration in these results, it is not the key tool. The fact that the number of zeros on the spatial interval of a solution of the linearized equation is non-increasing in time is far more important. See Hale [59] for a nice survey.

A different theme in order-preserving dynamics originates in the venerable subject of nonlinear elliptic and parabolic boundary value problems. The 1931 edition of Courant and Hilbert's famous book [34] refers to a paper of Bieberbach in *Göttingen Nachrichten*, 1912 dealing with the elliptic boundary value problem $\Delta u = \mathrm{e}^u$ in Ω, $u|\partial\Omega = f$, in a planar region Ω. A solution is found by iterating a monotone map in a function space. Courant and Hilbert extended this method to a broad class of such problems. Out of this technique grew the method of "upper and lower solutions" (or "supersolutions and subsolutions") for solving, both theoretically and numerically, second order elliptic PDEs (see Amann [4], Keller and Cohen [95], Keller [93,94], Sattinger [176]). Krasnosel'skii and Zabreiko [101] trace the use of positivity in functional analysis—closely related to monotone dynamics—to a 1924 paper by Uryson [222] on concave operators. The systematic use of positivity in PDEs was pioneered Krasnosel'skii and Ladyzhenskaya [100] and Krasnosel'skii [98].

Amann [5] showed how a sequence $\{u_n\}$ of approximate solutions to an elliptic problem can be viewed as the trajectory $\{T^n u_0\}$ of u_0 under a certain monotone map T in a suitable function space incorporating the boundary conditions, with fixed points of T being solutions of the elliptic equation. The dynamics of T can therefore be used to investigate the equation. Thus when T is globally asymptotically stable, there is a unique solution; while if T has two asymptotically stable fixed points, in many cases degree theory yields a third fixed point. As Amann [6] emphasized, a few key properties of T—continuity, monotonicity and some form of compactness—allow the theory to be efficiently formulated in terms of monotone maps in ordered Banach spaces.

Many questions in differential equations are framed in terms of eigenvectors of linear and nonlinear operators on Banach spaces. The usefulness of operators that are positive in some sense stems from the theorem of Perron [158] and Frobenius [51], now almost a century old, asserting that for a linear operator on $\mathbb{R}^n$ represented by a matrix with positive entries, the spectral radius is a simple eigenvalue having a positive eigenvector, and all other eigenvalues have smaller absolute value and only nonpositive eigenvectors. In 1912 Jentzsch [85] proved the existence of a positive eigenfunction with a positive eigenvalue for a homogeneous Fredholm integral equation with a continuous positive kernel.

In 1935 the topologists Alexandroff and Hopf [2] reproved the Perron–Frobenius theorem by applying Brouwer's fixed-point theorem to the action of a positive $n \times n$ matrix on the space of lines through the origin in $\mathbb{R}^n_+$. This was perhaps the first explicit use of the dynamics of operators on a cone to solve an eigenvalue problem. In 1940 Rutman [173] continued in this vein by reproving Jentzsch's theorem by means of Schauder's fixed-point theorem, also obtaining an infinite-dimensional analog of Perron–Frobenius, known today as the Krein–Rutman theorem [104,214]. In 1957 G. Birkhoff [20] initiated the dynamical use of Hilbert's projective metric for such questions.

The dynamics of cone-preserving operators continues to play an important role in functional analysis; for a survey, see Nussbaum [149,150]. One outgrowth of this work has been a focus on purely dynamical questions about such operators; some of these results are presented below. Polyhedral cones in Euclidean spaces have lead to interesting quantitative results, including *a priori* bounds on the number of periodic orbits. For recent work see Lemmens et al. [117], Nussbaum [151], Krause and Nussbaum [102], and references therein.

Monotone maps frequently arise as mathematical models. For example, the discrete Lotka–Volterra competition model (see May and Oster [136]):

$$\begin{aligned}(u_{n+1}, v_{n+1}) &= T(u_n, v_n)\\ &:= \bigl(u_n \exp\bigl[r(1 - u_n - bv_n)\bigr], v_n \exp\bigl[s(1 - cu_n - v_n)\bigr]\bigr)\end{aligned}$$

generates a monotone dynamical system relative to the fourth-quadrant cone only when the intrinsic rate of increase of each population is not too large ($r, s \leqslant 1$) and then only on the order interval $[0, r^{-1}] \times [0, s^{-1}]$ (Smith [192]). Fortunately in this case, every point in the first quadrant enters and remains in this order interval after one iteration. As is typical in ecological models, the Lotka–Volterra map is neither injective nor orientation-preserving or orientation-reversing. For monotone maps as models for the spread of a gene or an epidemic through a population, see Thieme [218], Selgrade and Ziehe [181], Weinberger [232], Liu [123] and the references therein.

5.2. *Definitions and basic results*

A continuous map $T : X \to X$ on the ordered metric space X is *monotone* if $x \leqslant y \Rightarrow Tx \leqslant Ty$, *strictly monotone* if $x < y \Rightarrow Tx < Ty$, *strongly monotone* if $x < y \Rightarrow Tx \ll Ty$, and *eventually strongly monotone* if whenever $x < y$, there exists $n_0 \geqslant 1$ such that $T^n x \ll T^n y$. We call T *strongly order-preserving* (*SOP*) if T is monotone, and whenever $x < y$ there exist respective neighborhoods U, V of x, y and $n_0 \geqslant 1$ such that $n \geqslant n_0 \Rightarrow T^n U \leqslant T^n V$.[1] As with semiflows, eventual strong monotonicity implies the strong order preserving property.

The orbit of x is $O(x) := \{T^n x\}_{n \geqslant 0}$, and the omega limit set of x is $\omega(x) := \bigcap_{k \geqslant 0} \overline{O(T^k x)}$. If $O(x)$ has compact closure, $\omega(x)$ is nonempty, compact, invariant (that is, $T\omega(x) = \omega(x)$) and *invariantly connected*. The latter means that $\omega(x)$ is not the disjoint union of two closed invariant sets [116].

If $T(x) = x$ then x is a *fixed point* or *equilibrium*. E denotes the set of fixed points. More generally, if $T^k x = x$ for some $k \geqslant 1$ we call x *periodic*, or k-periodic. The minimal such k is called the period of x (and $O(x)$).

Let Y denotes an ordered Banach space with order cone Y_+. A linear operator $A \in L(Y)$ is called *positive* if $A(Y_+) \subset Y_+$ (equivalently, A is a monotone map) and *strongly positive* if $A(Y_+ \setminus \{0\}) \subset \operatorname{Int} Y_+$) (equivalently, A is a strongly monotone map).

The following result is useful for proving smooth maps monotone or strongly monotone:

LEMMA 5.1. *Let $X \subset Y$ be a p-convex set and $f : X \to Y$ a locally C^1 map with quasiderivative $h : U \to L(Y)$ defined on an open set $U \subset Y$. If the linear maps $h(x) \in L(Y)$ are positive* (*respectively, strongly positive*) *for all $x \in U$, then f is monotone* (*respectively, strongly monotone*).

PROOF. By p-convexity it suffices to prove that every $p \in X$ has a neighborhood N such that $f|N \cap X$ is monotone (respectively, strongly monotone). We take N to be an open ball in U centered at p. Suppose $p + z \in X \cap N$, $z > 0$. By p-convexity, $X \cap N$ contains the line segment from p to $p + z$. The definition (above Lemma 2.15) of locally C^1 implies that the map $g : [0, 1] \to Y$, $t \mapsto f(p + tz)$ is C^1 with $g'(t) = h(tz)z$. Therefore

$$f(p+z) - f(z) = g(0) - g(1) = \int_0^1 g'(t)\,\mathrm{d}t = \int_0^1 h(tz)z\,\mathrm{d}t.$$

Because $h(tz) \in L(Y)$ is positive and $z > 0$, we have $h(tz)z \in Y_+$, therefore $f(p + z) - f(p) \geqslant 0$. If the operators $h(tz)$ are strongly positive, $f(p+z) - f(p) \gg 0$. $\square$

PROPOSITION 5.2 (Nonordering of Periodic Orbits). *A periodic orbit of a monotone map is unordered.*

---

[1] Our use of “strongly order-preserving” conflicts with Dancer and Hess [38], who use these words to mean what we have defined as “strongly monotone”. Our usage is consistent with that of several authors. Takáč [208,209] uses “strongly increasing” for our SOP.

PROOF. If not, there exists x in the orbit such that $T^k(x) > x$ for some $k > 0$. Induction on n shows that $T^{nk}(x) > x$ for all $n > 0$. But if x has period $m > 0$, induction on k proves that $T^{mk}(x) = x$. ☐

LEMMA 5.3 (Monotone Convergence Criterion). *Assume T is monotone and $O(z)$ has compact closure. If $m \geqslant 1$ is such that $T^m z < z$ or $T^m z > z$ then $\omega(z)$ is an m-periodic orbit.*

PROOF. Consider first the case $m = 1$. Compactness of $\overline{O(z)}$ implies the decreasing sequence $\{T^k z\}$ converges to a point $p = \omega(x)$. Now suppose $m > 1$. Applying the case just proved to the map T^m, we conclude that $\{T^{km} z\}$ converges to a point $p = T^m(p)$. It follows that $\omega(z) = \{p, Tp, T^2 p, \ldots, T^{m-1} p\}$. ☐

Lemma 5.3 yields information on one-sided stability of compact limit sets when T is SOP; see Hirsch [70].

In order to state the following lemma succinctly, we call a set $J \subset \mathbb{N}$ an *interval* if it is nonempty and contains all integers between any two of its members. For $a, b \in \mathbb{N}$ we set $[a, b] = \{j \in \mathbb{N}\colon a \leqslant j \leqslant b\}$ (there will be no confusion with real intervals). Two intervals *overlap* if they have more than one point in common.

Let $J \subset \mathbb{N}$ be an interval and $f : J \to X$ be a map. A subinterval $[a, b] \subset J$, $a < b$ is *rising* if $f(a) < f(b)$, and *falling* if $f(b) < f(a)$.

THEOREM 5.4. *A trajectory of a monotone map cannot have both a rising interval and a falling interval.*

PROOF. Follows from Theorem 1.6. ☐

LEMMA 5.5. *If T is monotone, $\omega(z)$ cannot contain distinct points having respective neighborhoods U, V such that $T^r(U) \leqslant T^r(V)$ for some $r \geqslant 0$.*

PROOF. Follows from Theorem 5.4 (see proof of Lemma 1.7). ☐

The next result is fundamental to the theory of monotone maps:

THEOREM 5.6 (Nonordering Principle). *Let $\omega(z)$ be an omega limit set for a monotone map T.*

(i) *No points of $\omega(z)$ are related by $\ll$.*

(ii) *If $\omega(z)$ is a periodic orbit or T is SOP, no points of $\omega(z)$ are related by $<$.*

PROOF. Follows from Proposition 5.2 and Lemma 5.5 (see the proof of Theorem 1.8). ☐

Call x *convergent* if $\omega(x)$ is a fixed point, and *quasiconvergent* if $\omega(x) \subset E$. Just as for semiflows, Proposition 5.6 leads immediately to a convergence criterion:

COROLLARY 5.7. *Assume Φ is SOP.*

(i) *If an omega limit set has a supremum or infimum, it reduces to a single fixed point.*

(ii) *If the fixed point set is totally ordered, every quasiconvergent point with compact orbit closure is convergent.*

PROOF. Part (i) follows from Theorem 5.6(ii), since the supremum or infimum, if it exists, belongs to the limit set. Part (ii) is a consequence (i). ☐

5.2.1. *Failure of the limit set dichotomy* We now point out a significant difference between strongly monotone maps and semiflows:

The Limit Set Dichotomy fails for strongly monotone maps.

Recall that for an SOP semiflow with compact orbit closures, the dichotomy (Theorem 1.16) states:

If $a < b$, either $\omega(a) < \omega(b)$ or $\omega(a) = \omega(b) \subset E$.

Takáč [211, Theorem 3.10], gives conditions on strongly monotone maps under which $a < b$ implies that either $\omega(a) \cap \omega(b) = \emptyset$ or $\omega(a) = \omega(b)$. He also gives a counterexample showing that $\omega(a) \cap \omega(b) = \emptyset$ does not imply $\omega(a) < \omega(b)$, nor does $\omega(a) = \omega(b)$ imply that these limit sets consist of fixed points (they are period-two orbits in his example). However, the mapping in his example is defined on a disconnected space.

For any map T in a Banach space, having an asymptotically stable periodic point p of period > 1, the Limit Set Dichotomy as formulated above must fail: take a point $q > p$ so near to p that $O(p) = \omega(p) = \omega(q)$. Clearly $\omega(p)$, being a nontrivial periodic orbit, contains no fixed points. Thus the second assertion of the Limit Set Dichotomy fails in this case.

Dancer and Hess [38] gave a simple example in $\mathbb{R}^k$ for prime k of a strongly monotone map with an asymptotically stable periodic point of period k which we describe below. Therefore the second alternative of the Limit Set Dichotomy can be no stronger than that $\omega(a) = \omega(b)$ is a periodic orbit.

The Limit Set Dichotomy fails even for strictly monotone maps in $\mathbb{R}^2$. Let $f(x) = 2\arctan(x)$, let $a > 0$ be its unique positive fixed point, and note that $0 < f'(a) < 1$. Define $T_0 : \mathbb{R}^2 \to \mathbb{R}^2$ by $T_0(x, y) := (f(y), f(x))$. Then $E = \{(-a, -a), (0, 0), (a, a)\}$ since f has no points of period 2. The fixed points of T_0^2 are the nine points obtained by taking all pairings of $-a, 0, a$. An easy calculation shows that $\{(-a, a), (a, -a)\}$ is an asymptotically stable period-two orbit of T_0 because the Jacobian matrix of T_0^2 is $f'(a)^2$ times the identity matrix. T^0 is strictly monotone but not strongly monotone. Now consider the perturbations $T_\epsilon(x, y) := T_0(x, y) + (\epsilon x, \epsilon y)$. It is easy to see that T_ϵ is strongly monotone for $\epsilon > 0$; and by the implicit function theorem, for small $\epsilon > 0$, T_ϵ has an asymptotically stable period-two orbit $O(p_\epsilon)$ with p_ϵ near $(-a, a)$. As noted in [38], this example can be generalized to $\mathbb{R}^k$ for prime k.

Takáč [212] shows that linearly stable periodic points can arise for the period map associated with monotone systems of ordinary and partial differential equations. Other counterexamples for low-dimensional monotone maps can be found in Smith [192,195].

As we have shown, asymptotically stable periodic orbits that are not singletons can exist for monotone, even strongly monotone maps. Later we will show that the generic orbit of a smooth, dissipative, strongly monotone map converges to a periodic orbit. Here, we show that every attractor contains a stable periodic orbit.

Recall that a point p is *wandering* if there exists a neighborhood U of p and a positive integer n_0 such that $T^n(U) \cap U = \emptyset$ for $n > n_0$. The *nonwandering set* Ω, consisting of all points q that are not wandering, contains all limit sets. In the following, we assume that X is an open subset of the strongly ordered Banach space Y and $T : X \to X$ is monotone with compact orbit closures. The following result is adapted from Hirsch [71, Theorems 4.1, 6.3].

THEOREM 5.8. *If T is strongly monotone and K is a compact attractor, then K contains a stable periodic orbit.*

The proof relies on the following result that does not use strong monotonicity nor that K attracts uniformly:

THEOREM 5.9. *Let $p \in K$ be a maximal* (*resp., minimal*) *nonwandering point. Then p is periodic, and every neighborhood of p contains an open set $W \gg p$* (*resp., $W \ll p$*) *such that $\omega(x) = O(p)$ for all $x \in W$.*

PROOF. Suppose K attracts the open neighborhood U of K and fix $y \gg p,\ y \in U$. Since p is nonwandering there exists a convergent sequence $x_i \to p$ and a sequence $n_i \to \infty$ such that $T^{n_i} x_i \to p$. For all large i, $x_i \leqslant y$. Passing to a subsequence, we assume that $T^{n_i} y \to q$. By monotonicity and $x_i \leqslant y$ for large i, we have $q \geqslant p$. But $q \in K \cap \Omega$ and the maximality of p requires $q = p$. Since $p \ll y$ and $T^{n_i} y \to p$ it follows that $T^m y \ll y$ for some m. Lemma 5.3 implies that $\omega(y)$ is an m-periodic orbit containing p. As this holds for every $y \gg p$, the result follows. □

LEMMA 5.10. *Let $p, q \in K$ be fixed points such that $p \ll q$, p is order stable from below, and q is order stable from above. Then $K \cap [p, q]$ contains a stable equilibrium.*

PROOF. Let R be a maximal totally ordered set of fixed points in $K \cap [p, q]$. An argument similar to the one in the proof of Theorem 1.30 shows that the fixed point

$$e := \inf\{z \in E \cap R\colon\ z \text{ is order stable from above}\}$$

is order stable. That e is stable follows from the analog of Proposition 1.28. □

PROOF OF THEOREM 5.8. Theorem 5.9 shows that some iterate T^n, $n \geqslant 0$ has fixed points p, q as in Lemma 5.10, which result therefore implies Theorem 5.8. □

Jiang and Yu [90, Theorem 2] implies that if T is analytic, order compact with strongly positive derivative, then K must contain an asymptotically stable periodic orbit.

5.3. *The order interval trichotomy*

In this section we assume that X is a subset of an ordered Banach space Y with positive cone Y_+, with the induced order and topology. Much of the early work on monotone maps on ordered Banach spaces focused on the existence of fixed points for self maps of order intervals $[a,b]$ such that $a,b \in E$; see especially Amann [6]. The following result of Dancer and Hess [38], quoted without proof, is crucial for analyzing such maps.

Let u,v be fixed points of T. A doubly-infinite sequence $\{x_n\}_{n\in\mathbb{Z}}$ ($\mathbb{Z}$ is the set of all integers) in Y is called an *entire orbit from* u *to* v if

$$x_{n+1} = T(x_n), \qquad \lim_{n\to-\infty} x_n = u, \quad \lim_{n\to\infty} x_n = v.$$

If $x_n \leqslant x_{n+1}$ (respectively, $x_n < x_{n+1}$), the entire orbit is *increasing* (respectively, *strictly increasing*). If $x_n \geqslant x_{n+1}$ (respectively, $x_n > x_{n+1}$), the entire orbit is *decreasing* (respectively, *strictly decreasing*). If the entire orbit $\{x_n\}$ is increasing but not strictly increasing, then $x_n = v$ for all sufficiently large n; and similarly for decreasing.

Consider the following hypothesis:

(G) $X = [a,b]$ where $a,b \in Y$, $a < b$ and $Ta = a$, $Tb = b$. The map $T : X \to X$ is monotone and $T(X)$ has compact closure in X.

THEOREM 5.11 (The Order Interval Trichotomy). *Under hypothesis* (G), *at least one of the following holds*:

(a) *there is a fixed point* c *such that* $a < c < b$;

(b) *there exists an entire orbit from* a *to* b *that is increasing, and strictly increasing if* T *is strictly monotone*;

(c) *there exists an entire orbit from* b *to* a *that is decreasing, and strictly decreasing if* T *is strictly monotone.*

An extension of Theorem 5.11 to allow additional fixed points on the boundary of $[a,b]$ is carried out in Hsu et al. [83]. Wu et al. [236] weaken the compactness condition. See Hsu et al. [83], Smith [192], and Smith and Thieme [201] for applications to generalized two-species competition dynamics. For related results see Hess [63], Matano [133], Poláčik [162], Smith [184,194].

A fixed point q of T is *stable* if every neighborhood of q contains a positively invariant neighborhood of q. An immediate corollary of the Order Interval Trichotomy is:

COROLLARY 5.12. *Assume hypothesis* (G), *and let* a *and* b *be stable fixed points. Then there is a third fixed point in* $[a,b]$.

Corollary 5.14 establishes a third fixed point under different assumptions.

In general, more than one of the alternatives (a), (b), (c) may hold (see [83]). The following complement to the Order Interval Trichotomy gives conditions for exactly one to hold; (iii) is taken from Proposition 2.2 of [83].

Consider the following three conditions:

(a′) there is a fixed point c such that $a < c < b$;
(b′) there exists an entire orbit from a to b;
(c′) there exists an entire orbit from b to a.

PROPOSITION 5.13. *Assume hypothesis* (G).

(i) *If T is strongly order-preserving, exactly one of* (a′), (b′), (c′) *can hold. More precisely*: *Assume $a < y < b$ and y has compact orbit closure. Then $\omega(y) = \{b\}$ if there is an entire orbit from a to b, while $\omega(y) = \{a\}$ if there is an entire orbit from b to a.*
(ii) *If $a \ll b$, at most one of* (b′), (c′) *can hold.*
(iii) *Suppose $a \ll b$, and $E \cap [a, b] \setminus \{a, b\} \neq \emptyset$ implies $E \cap [[a, b]] \setminus \{a, b\} \neq \emptyset$. Then at most one of* (a′), (b′), (c′) *can hold.*

PROOF. For (i), consider an entire orbit $\{x_n\}$ from a to b. There is a neighborhood U of a such that $T^k U \leqslant T^k y$ for sufficiently large k. Choose $x_j \in U$. Then $T^k x_j \leqslant T^k y \leqslant b$ for all large k. As $\lim_{k\to\infty} T^k x_j = b$ and the order relation is closed, b is the limit of every convergent subsequence of $\{T^k y\}$. The case of an entire orbit from b to a is similar.

In (ii), choose neighborhoods U, V of a, b respectively such that $U \ll V$. Fix j so that $x_j \in U$. If $y \in V$ then an argument similar to the proof of (i) shows that $\omega(y) = \{b\}$. Hence there cannot be an entire orbit from b to a, since it would contain a point of V.

Assume the hypothesis of (iii), and note that (ii) makes (b′) and (c′) incompatible. If (a′), there is a fixed point $c \in [[a, b]]$, and arguments similar to the proof of (ii) show that neither (b′) nor (c′) holds. □

COROLLARY 5.14. *In addition to hypothesis* (G), *assume T is strongly order preserving with precompact image. If some trajectory does not converge, there is a third fixed point.*

PROOF. Follows from the Order Interval Trichotomy 5.11 and Proposition 5.13(i). □

A number of authors have considered the question of whether *a priori* knowledge that every fixed point is stable implies the convergence of every trajectory. See Alikakos et al. [3], Dancer and Hess [38], Matano [133] and Takáč [209] for such results. The following theorem is adapted from [38].

A set $A \subset X$ is a *uniform global attractor* for the map $T : X \to X$ if $T(A) = A$ and $\operatorname{dist}(T^n x, A) \to 0$ uniformly in $x \in X$.

THEOREM 5.15. *Let $a, b \in Y$ with $a < b$. Assume $T : [a, b] \to [a, b]$ is strongly order preserving with precompact image, and every fixed point is stable. Then E is a totally ordered arc J that is a uniform global attractor, and every trajectory converges.*

PROOF. We first show that there exists a totally ordered arc of fixed points; this will not use the SOP property. $O(a)$ is an increasing sequence converging to the smallest fixed point in $[a, b]$. Similarly, $O(b)$ is a decreasing sequence converging to the largest fixed point in $[a, b]$. By renaming a and b as these fixed points, we may as well assume that $a, b \in E$. The stability hypothesis and Corollary 5.12 implies there is a fixed point c satisfying $a < c < b$.

The same reasoning applies to $[a,c]$ and $[c,b]$, and can be repeated indefinitely to show that every maximal totally ordered set of fixed points is compact and connected, hence an arc (Wilder [233, Theorem I.11.23]). Thus by Zorn's Lemma there is a totally ordered arc $J \subset E$ joining a to b.

Next we prove: *Every unordered compact invariant set K is a point of J.* This will not use precompactness of $T([a,b])$. Set $q = \inf\{e \in J\colon\ K \leqslant e\}$. It suffices to prove $q \in K$, for then K, being unordered, reduces to $\{q\}$. If $q \notin K$ then $q > k$. By SOP and invariance of K there is a are neighborhoods V of p and $n > 0$ such that $K = T^n(K) \leqslant T(V)$, hence $K \leqslant T^n(V \cap J) = V \cap J$. This gives the contradiction $K \leqslant \inf(V \cap J) < q$.

Every $\omega(x)$ is compact by the precompactness assumption, and unordered by the Nonordering Principle 5.6(ii). Total ordering of J therefore implies $\omega(x)$ is a point of J. This proves every trajectory converges.

To show that J is a global attractor, let N be the open ϵ-neighborhood of J for an arbitrary $\epsilon > 0$. The stability hypothesis implies N contains a positively invariant open neighborhood W of J. It suffices to prove $T^n(X) \subset W$ when n is sufficiently large. Convergence of all trajectories implies that for every $x \in X$ there exists an open neighborhood $U(x)$ of x and $n(x) > 0$ such that $T^n(x) \in W$ for all $n \geqslant n(x)$. Precompactness of $T(X)$ implies $T(X) \subset \bigcup U(x_i)$ for some finite set $\{x_i\}$. Hence $T^n(X) \subset W$ provided $n > \max\{n(x_i)\}$. □

If the map T in Theorem 5.15 is C^1 and strongly monotone, then E is a smooth totally ordered arc by a result of Takáč [211].

5.3.1. *Existence of fixed points*

Dancer [37] obtained remarkable results concerning the dynamics of monotone maps with some compactness properties: Limit sets can always be bracketed between two fixed points, and with additional hypotheses these fixed points can be chosen to be stable. The next two theorems are adapted from [37].

A map $T : Y \to Y$ is *order compact* if it takes each order interval, and hence each order bounded set, into a precompact set.

THEOREM 5.16. *Let X be an order convex subset of Y. Assume that $T : X \to X$ is monotone and order compact, with every orbit having compact closure in X and every omega limit set order bounded. Then for all $z \in Y$ there are fixed points f, g such that $f \leqslant \omega(z) \leqslant g$.*

PROOF. There exists $u \in X$ such that $u \geqslant \omega(z)$ because omega limit sets order bounded. Since $T(\omega(z)) = \omega(z)$, it follows that $\omega(z) \leqslant T^i u$ for all i, hence $\omega(z) \leqslant \omega(u)$. Similarly, there exists $s \in X$ such that $\omega(u) \leqslant \omega(s)$. The set $F := \{x \in Y\colon\ \omega(z) \leqslant x \leqslant \omega(s)\}$ is the intersection of closed order intervals, hence closed and convex, nonempty because it contains $\omega(u)$, and obviously order bounded. Moreover $F \subset X$ because X is order convex. Therefore $T(F)$ is defined and is precompact. Monotonicity of T and invariance of $\omega(z)$ and $\omega(s)$ imply $T(F) \subset F$. It follows from the Schauder fixed point theorem that there is a fixed point $g \in F$, and $g \geqslant \omega(z)$ as required. The existence of f is proved similarly. □

The cone Y_+ is *reproducing* if $Y = Y_+ - Y_+$. This holds for many function spaces whose norms do not involve derivatives. If Y_+ has nonempty interior, it is reproducing: any $x \in Y$

can be expressed as $x = \lambda e - \lambda(e - \lambda^{-1}x) \in Y_+ - Y_+$, where $e \gg 0$ is arbitrary and $\lambda > 0$ is a sufficiently large real number.

THEOREM 5.17. *Let $X \subset Y$ be order convex. Assume $T : X \to X$ is monotone, completely continuous, and order compact. Suppose orbits are bounded and omega limits sets are order bounded.*

(i) *For all $z \in X$ there are fixed points f, g such that $f \leqslant \omega(z) \leqslant g$.*

(ii) *Assume Y_+ is reproducing, $X = Y$ or Y_+, and E is bounded. Then there are fixed points $e_M = \sup E$ and $e_m = \inf E$, and all omega limit sets lie in $[e_m, e_M]$. Moreover, if $x \leqslant e_m$ then $\omega(x) = \{e_m\}$, while if $x \geqslant e_M$ then $\omega(x) = \{e_M\}$.*

(iii) *Assume Y_+ is reproducing, $X = Y$ or Y_+, E is bounded, and T is strongly order preserving. Suppose $z_0 \in Y$ is not convergent. Then there are three fixed points $f < p < g$ such that $f < \omega(z_0) < g$. If T is strongly monotone, f and g can be chosen to be stable.*

PROOF. We prove all assertions except for the stability in (iii). Complete continuity implies that every positively invariant bounded set is precompact. Therefore orbit closures are compact and omega limit sets are compact and nonempty, so (i) follows from Theorem 5.16.

To prove (ii), note that E is compact because it is bounded invariant and closed. Choose a maximal element $e_M \in E$ (Lemma 1.1). We must show that $e_M \geqslant e$ for every $e \in E$. Since the order cone is reproducing, $e_M - e = v - w$ with $v, w \geqslant 0$. Set $u := e + v + w$. Then $u \in X$, $u \geqslant e$, and $u \geqslant e_M$. Monotonicity implies $e_M = T^i e_M \leqslant T^i u$ for all $i \geqslant 0$, hence $e_M \leqslant \omega(u)$. By Theorem 5.16 there exists $g \in E$ such that $\omega(u) \leqslant g$. Hence $e_M \leqslant g$, whence $e_M = g$ by maximality. We now have $e_M \leqslant \omega(u) \leqslant g = e_M$, so $\omega(u) = \{e_M\}$. Monotonicity implies (as above) $e \leqslant \omega(u)$, therefore $e \leqslant e_M$ as required. This proves $e_M = \sup E$, and the dual argument proves $e_m = \inf E$. If $x \leqslant e_m$ then $\omega(x) \leqslant e_m$ by monotonicity; but $\omega(x) \geqslant e_m$ by (i), so $\omega(x) = \{e_m\}$. Similarly for the case $x \geqslant e_M$.

To prove the first assertion of (iii), note that $e_m < \omega(z) < e_M$ by (i) and the Nonordering Principle 5.6(ii). Monotonicity and order compactness of T imply $[e_m, e_M]$ is positively invariant with precompact image. As T is SOP, there is a third fixed point in $[e_m, e_M]$ by Corollary 5.14. □

5.4. *Sublinearity and the cone limit set trichotomy*

Motivated by the problem of establishing the existence of periodic solutions of quasimonotone, periodic differential equations defined on the positive cone in R^n, Krasnosel'skii pioneered the dynamics of sublinear monotone self-mappings of the cone [99]. We will prove Theorem 5.20 below, adapted from the original finite-dimensional version of Krause and Ranft [103].

Let Y denote an ordered Banach space with positive cone Y_+. Denote the interior (possibly empty) of Y_+ by P. A map $T : Y_+ \to Y_+$ is *sublinear* (or "subhomogeneous") if

$$0 < \lambda < 1 \implies \lambda T(x) \leqslant T(\lambda x),$$

and *strongly sublinear* if

$$0 < \lambda < 1,\ x \gg 0 \quad \Longrightarrow \quad \lambda T(x) \ll T(\lambda x).$$

Strong sublinearity is the strong concavity assumption of Krasnosel'skii [99]. It can be verified by using the following result from that monograph:

LEMMA 5.18. *$T : P \to Y$ is strongly sublinear provided T is differentiable and $Tx \gg T'(x)x$ for all $x \gg 0$.*

PROOF. Let $F(s) = s^{-1}T(sx)$ for $s > 0$ and some fixed $x \gg 0$. Then $F'(s) = -s^{-2}T(sx) + s^{-1}T'(sx)x \ll 0$ by our hypothesis. So for $0 < \lambda < 1$, we have

$$\phi\big(Tx - \lambda^{-1}T(\lambda x)\big) = \phi\big(F(1)\big) - \phi\big(F(\lambda)\big) < 0$$

for every nontrivial $\phi \in Y_+^*$, the dual cone in Y^*, because $\frac{\mathrm{d}}{\mathrm{d}s}\phi(F(s)) < 0$. The desired conclusion follows from Proposition 3.1. $\square$

COROLLARY 5.19. *Assume Y is strongly ordered. A continuous map $T : Y_+ \to Y$ is sublinear provided T is differentiable in P and $Tx \geqslant T'(x)x$ for all $x \gg 0$.*

PROOF. By continuity it suffices to prove $T|P$ is sublinear. Fix $e \gg 0$. For each $\delta > 0$ the map $P \to Y$, $x \mapsto Tx + \delta e$ is strongly sublinear by Lemma 5.18. Sending δ to zero implies T is sublinear. $\square$

Krause and Ranft [103] have results establishing sublinearity of some iterate of T, which is an assumption used in Theorem 5.20 below.

The following theorem demonstrates global convergence properties for order compact maps that are monotone and sublinear in a suitably strong sense.

THEOREM 5.20 (Cone Limit Set Trichotomy). *Assume $T : Y_+ \to Y_+$ is continuous and monotone and has the following properties for some $r \geqslant 1$:*

(a) *T^r is strongly sublinear*;
(b) *$T^r x \gg 0$ for all $x > 0$*;
(c) *T^r is order compact.*

Then precisely one of the following holds:

(i) *each nonzero orbit is order unbounded*;
(ii) *each orbit converges to* 0, *the unique fixed point of T*;
(iii) *each nonzero orbit converges to $q \gg 0$, the unique nonzero fixed point of T.*

A key tool in the proofs of such results is Hilbert's projective metric and the related part metric due to Thompson [219]. We define the part metric $\mathsf{p}(x, y)$ here in a very limited way, as a metric on P (which is the "part"). For $x, y \gg 0$, define

$$\mathsf{p}(x, y) := \inf\big\{\rho > 0\colon\ \mathrm{e}^{-\rho}x \ll y \ll \mathrm{e}^{\rho}y\big\}.$$

The family of open order intervals in P forms a base for the topology of the part metric. It is easy to see that the identity map of P is continuous from the original topology on P to that defined by the part metric.

When $Y = \mathbb{R}^n$ with vector ordering, with $P = \mathrm{Int}(\mathbb{R}^n_+)$, the part metric is isometric to the max metric on $\mathbb{R}^n$, defined by $d_{\max}(x, y) = \max_i |x_i - y_i|$, *via* the homeomorphism $\mathrm{Int}(\mathbb{R}^n_+) \approx \mathbb{R}^n$, $x \mapsto (\log x_1, \dots, \log x_n)$. Restricted to compact sets in $\mathrm{Int}(\mathbb{R}^n_+)$, the part metric and the max metric are equivalent in the sense that there exist $\alpha, \beta > 0$ such that $\alpha \mathsf{p}(x, y) \leqslant |x - y|_{\max} \leqslant \beta \mathsf{p}(x, y)$.

The usefulness of the part metric in dynamics stems from the following result. Recall map T between metric spaces is a *contraction* if it has a Lipschitz constant < 1, and it is *nonexpansive* if it has Lipschitz constant 1. We say T is *strictly nonexpansive* if $\mathsf{p}(Tx, Ty) < \mathsf{p}(x, y)$ whenever $x \neq y$.

PROPOSITION 5.21. *Let $T : P \to P$ be a continuous, monotone, sublinear map.*

- (i) *T is nonexpansive for the part metric.*
- (ii) *If T is strongly sublinear, T is strictly nonexpansive for the part metric.*
- (iii) *If T is strongly monotone, $A \subset P$, and no two points of A are linearly dependent, then $T|_A$ is strictly nonexpansive for the part metric.*
- (iv) *Under the assumptions of* (ii) *or* (iii), *if $L \subset A$ is compact* (*in the norm topology*) *and $T(L) \subset L$, then the set $L_\infty = \bigcap_{n>0} T^n(L)$ is a singleton.*

PROOF. Fix distinct points $x, y \in A$ and set $\mathrm{e}^{\mathsf{p}(x,y)} = \lambda > 1$, so that $\lambda^{-1}x \leqslant y \leqslant \lambda x$ and λ is the smallest number with this property. By sublinearity and monotonicity,

$$\lambda^{-1}Tx \leqslant T\big(\lambda^{-1}x\big) \leqslant Ty \leqslant T(\lambda x) \leqslant \lambda Tx \tag{5.3}$$

which implies $\mathsf{p}(Tx, Ty) \leqslant \mathsf{p}(x, y)$.

If T is strongly sublinear, the first and last inequalities in (5.3) can be replaced by $\ll$, which implies $\mathsf{p}(Tx, Ty) < \mathsf{p}(x, y)$.

When x and y are linearly independent, $\lambda^{-1}x < y < \lambda x$. If also T is strongly monotone, (5.3) is strengthened to

$$\lambda^{-1}Tx \leqslant T\big(\lambda^{-1}x\big) \ll Ty \ll T(\lambda x) \leqslant \lambda Tx$$

which also implies $\mathsf{p}(Tx, Ty) < \mathsf{p}(x, y)$.

To prove (iv), observe first that if L is compact in the norm metric, it is also compact in the part metric. In both (ii) and (iii) T reduces the diameter in the part metric of every compact subset of L. Since T maps L_∞ onto itself but reduces its part metric diameter, (iv) follows. □

PROOF OF THE CONE LIMIT SET TRICHOTOMY 5.20. We first work under the assumption that $r = 1$. In this case Proposition 5.21 shows that every compact invariant set in P reduces to a fixed point, and there is at most one fixed point in P. It suffices to consider the orbits of points $x \in P$, by (b).

Suppose there is a fixed point $q \gg 0$. There exist numbers $0 < \lambda < 1 < \mu$ such that $x \in [\lambda q, \mu q] \subset P$. For all n we have

$$0 \ll \lambda q = \lambda T^n q \leqslant T^n(\lambda q) \leqslant T^n x \leqslant T^n(\mu q) \leqslant \mu T^n q = \mu q.$$

Hence $O(x) \subset [\lambda q, \mu q]$, so $O(Tx)$ lies in $T([\lambda p, \mu q])$, which is precompact by (c). Therefore $\omega(x)$ is a compact unordered invariant set in P. Proposition 5.21(iii) implies that $\omega(x) = \{q\}$. This verifies (iii).

Case I: If some orbit $O(y)$ is order unbounded, we prove (i). We may assume $y \gg 0$. There exists $0 < \gamma < 1$ such that $\gamma y \ll x$. Then $\gamma T^n y \leqslant T^n(\gamma y) \leqslant T^n x$, implying $O(x)$ is unbounded.

Case II: If $0 \in \omega(y)$ for some y, we prove (ii). We may assume $y \gg 0$. Fix $\mu > 1$ with $x \ll \mu y$. Then $0 \leqslant T^n x \leqslant T^n(\mu y) \leqslant \mu T^n y \to 0$. Therefore $\overline{O(x)}$ is compact and $T^n x \to 0$.

Case III: If the orbit closure $\overline{O(x)} \subset [a, b] \subset P$, then (iii) holds. For $\overline{O(x)}$ is compact by (c), so $\omega(x)$ is a nonempty compact invariant set. Because $\omega(x) \subset \overline{O(x)} \subset P$, Case I implies (iii).

Cases I, II and III cover all possibilities, so the proof for $r = 1$ is complete. Now assume $r > 1$. One of the statements (i), (ii), (iii) is valid for T^r in place of T. If (i) holds for T^r, it obviously holds for T. Assume (ii) holds for T^r. If $x > 0$ then $\omega(x) = \{0, T(0), \dots, T^{r-1}(0)\}$. As this set is compact and T^r invariant, it reduces to $\{0\}$, verifying (ii) for T. A similar argument shows that if (iii) holds for T^r, it also holds for T. □

The conclusion of the Cone Limit Trichotomy can fail for strongly monotone sublinear maps—simple linear examples in the plane have a line of fixed points. But the following holds:

THEOREM 5.22. *Assume*:

(a) *$T : Y_+ \to Y_+$ is continuous, sublinear, strongly monotone, and order compact*;

(b) *for each $x > 0$ there exists $r \in \mathbb{N}$ such that $T^r x \gg 0$.*

Then:

(i) *either $O(x)$ is not order bounded for all $x > 0$, or $O(x)$ converges to a fixed point for all $x \geqslant 0$;*

(ii) *the set of fixed points > 0 has the form $\{\lambda e\colon\ a \leqslant \lambda \leqslant b\}$ where $e \gg 0$ and $0 \leqslant a \leqslant b \leqslant \infty$.*

PROOF. Let $y > 0$ be arbitrary. If $O(y)$ is not order bounded, or $0 \in \omega(y)$, the proof of (i) follows Cases I and II in the proof of the Cone Limit Set Trichotomy 5.20. If $\overline{O(x)} \subset [a, b] \subset P$, then $\omega(y)$ is a compact invariant set in P, as in Case III of 5.20. As $\omega(y)$ is unordered, every pair of its elements are linearly independent. Therefore Proposition 5.21(iv) implies $\omega(y)$ reduces to a fixed point, proving (i). The same reference shows that all fixed points lie on a ray $R \subset Y_+$ through the origin, which must pass through some

$e \gg 0$ by (b). Suppose p, q are distinct fixed points and $0 \ll p \ll x <\ll q$. There exist unique numbers $0 < \mu < 1 < \nu$ such that $x = \mu p = \mu q$. Then

$$Tx \geqslant \mu Tp = \mu p = x, \qquad Tx \leqslant \nu Tq = \nu q = x$$

proving $Tx = x$. This implies (ii). □

Papers related to sublinear dynamics and the part metric include Dafermos and Slemrod [35], Krause and Ranft [103], Krause and Nussbaum [102], Nussbaum [149,150], Smith [183], and Takáč [208,215]. For interesting applications of sublinear dynamics to higher order elliptic equations, see Fleckinger-Pellé and Takáč [45,46].

5.5. *Smooth strongly monotone maps*

Smoothness together with compactness allows one to settle questions of stability of fixed points and periodic points by examining the spectrum of the linearization of the mapping. Let $T : X \to X$ where X is an open subset of the ordered Banach space Y with cone Y_+ having nonempty interior in Y. Assume that T is a completely continuous, C^1 mapping with a strongly positive derivative at each point. Then T is strongly monotone by Lemma 5.1 and $T'(x)$ is a Krein–Rutman operator so the Krein–Rutman Theorem 2.17 holds for $T'(p)$, $p \in E$. Let ρ be the spectral radius of $T'(p)$, which the reader will recall is a simple eigenvalue which dominates all others in modulus and for which the generalized eigenspace is spanned by an eigenvector $v \gg 0$. Let V_1 be the span of v in Y. There is a complementing closed subspace V_2 such that $Y = V_1 \oplus V_2$ satisfying $T'(p)V_2 \subset V_2$ and $V_2 \cap Y_+ = \{0\}$. Let P denote the projection of Y onto V_2 along v. Finally, let τ denote the spectral radius of $T'(p)|V_2 : V_2 \to V_2$, which obviously satisfies $\tau < \rho$. Mierczyński [139] exploits this structure of the linearized mapping to obtain very detailed behavior of the orbits of points near p. In order to describe his results, define $K := \{x \in X\colon\ T^n x \to p\}$ to be the basin of attraction of p. Let $M_- := \{x \in X\colon\ T^{n+1} \ll T^n x, n \geqslant n_0,\ \text{some } n_0\}$ be the set of eventually decreasing orbits, $M_+ := \{x \in X\colon\ T^n x \ll T^{n+1}x, n \geqslant n_0,\ \text{some } n_0\}$ be the set of eventually increasing orbits, and $M := M_- \cup M_+$ be the set of eventually monotone (in the strong sense) orbits.

The following result is standard but nonetheless important.

THEOREM 5.23 (Principle of Linearized Stability). *If $\rho < 1$, there is a neighborhood U of p such that $T(U) \subset U$ and constants $c > 0$, $\kappa \in (\rho, 1)$ such that for each $x \in U$ and all n*

$$\left\| T^n x - p \right\| \leqslant c\kappa^n \|x - p\|.$$

In the more delicate case that $\rho \leqslant 1$, Mierczyński [139] obtains a smooth hypersurface C, which is an analog for T of the codimension-one linear subspace V_2 invariant under the linearized mapping $T'(p)$:

THEOREM 5.24. *If* $\rho \leqslant 1$ *there exists a codimension-one embedded invariant manifold* $C \subset X$ *of class* C^1 *having the following properties*:

(i) $C = \{x + Pw + R(w)v\colon\ w \in O\}$ *where* $R : O \to \mathbb{R}$ *is a* C^1 *map defined on the relatively open subset* O *of* V_2 *containing* 0, *satisfying* $R(0) = R'(0) = 0$. *In particular,* C *is tangent to* V_2 *at* p.

(ii) C *is unordered.*

(iii) $C = \{p \in X\colon\ \|T^n x - p\|/\kappa^n \to 0\} = \{x \in X\colon\ \|T^n x - p\|/\kappa^n$ *is bounded*$\}$, *for any* κ, $\tau < \kappa < \rho$. *In particular,* $C \subset K$.

(iv) $K \setminus C = \{x \in K\colon\ \|T^n x - p\|/\kappa^n \to \infty\} = \{x \in K\colon\ \|T^n x - p\|/\kappa^n$ *is unbounded*$\}$, *for any* κ, $\tau < \kappa < \rho$.

(v) $K \setminus C = K \cap M$.

Conclusion (v) implies most orbits converging to p do so monotonically, but more can be said. Indeed, $K \cap M_+ = \{x \in K\colon\ (T^n x - p)/\|T^n x - p\| \to -v\}$ and a similar result for $K \cap M_-$ with v replacing $-v$ holds. The manifold C is a local version of the unordered invariant hypersurfaces obtained by Takáč in [209].

Corresponding to the space V_1 spanned by $v \gg 0$ for $T'(p)$, a locally forward invariant, one dimensional complement to the codimension one manifold C is given in the following result.

THEOREM 5.25. *There is* $\epsilon > 0$ *and a one-dimensional locally forward invariant* C^1 *manifold* $W \subset B(p; \epsilon)$, *tangent to* V *at* p. *If* $\rho > 1$, *then* W *is locally unique, and for each* $x \in W$ *there is a sequence* $\{x_{-n}\} \subset W$ *with* $Tx_{-n} = x_{-n+1}$, $x_0 = x$, *and* $\kappa^n \|x_{-n} - p\| \to 0$ *for any* κ, $1 < \kappa < \rho$.

Here $B(p; \epsilon)$ is the open ϵ-ball centered at p. Local forward invariance of W means that $x \in W$ and $Tx \in B(x; \epsilon)$ implies $Tx \in W$. Related results are obtained by Smith [184]. In summary, the above results assert that the dynamical behavior of the nonlinear map T behaves near p like that of its linearization $T'(p)$. Obviously, the above results can be applied at a periodic point p of period k by considering the map T^k which has all the required properties.

Mierczyński [139] uses the results above to classify the convergent orbits of T. Similar results are obtained by Takáč in [210].

It is instructive to consider the sort of stable bifurcations that can occur from a linearly stable fixed point, or a linearly stable periodic point, for a one parameter family of mappings satisfying the hypotheses of the previous results, as the parameter passes through a critical value at which $\rho = 1$. The fact that there is a simple positive dominant eigenvalue of $(T^k)'(p)$ ensures that period-doubling bifurcations from a stable fixed point or from a stable periodic point, as a consequence of a real eigenvalue passing through -1, cannot occur. In a similar way, a Neimark–Sacker [113] bifurcation to an invariant closed curve cannot occur from a stable fixed or periodic point. These sorts of bifurcations can occur from unstable fixed or periodic points but then they will “be born unstable.”

The generic orbit of a smooth strongly order preserving semiflow converges to fixed point but such a result fails to hold for discrete semigroups, i.e., for strongly order preserving mappings. Indeed, such mappings can have attracting periodic orbits of period exceeding one as we have seen. However, Tereščák [217], improving earlier joint work with Poláčik [164,165], and [65], has obtained the strongest result possible for strongly monotone, smooth, dissipative mappings.

THEOREM 5.26 (Tereščák, 1996). *Let $T : Y \to Y$ be a completely continuous, C^1, point dissipative map whose derivative is strongly positive at every point of the ordered Banach space Y having cone Y_+ with nonempty interior. Then there is a positive integer m and an open dense set $U \subset Y$ such that the omega limit set of every point of U is a periodic orbit with period at most m.*

The map P is *point dissipative* (see Hale [58]) provided there is a bounded set B with the property that for every $x \in X$, there is a positive integer $n_0 = n_0(x)$ such that $P^n x \in B$ for all $n \geqslant n_0$. We note that the hypothesis that $T'(x)$ is strongly positive implies that T is strongly monotone by Lemma 5.1.

5.6. *Monotone planar maps*

A remarkable convergence result for planar monotone maps was first obtained by de Mottoni and Schiaffino [42]. They focused on the period-map for the two-species, Lotka–Volterra competition system of ordinary differential equations with periodic coefficients. The full generality of their arguments was recognized and improved upon by Hale and Somolinos [60] and Smith [188,189,192]. We follow the treatment Smith in [192].

In addition to the usual order relations on $\mathbb{R}^2$, $\leqslant, <, \ll$, generated by $\mathbb{R}^2_+$, we have the "southeast ordering" ($\leqslant_K$), generated by the fourth quadrant $K = \{(u, v)\colon\ u \geqslant 0, v \leqslant 0\}$. The map T is *cooperative* if it is monotone relative to $\leqslant$ and *competitive* if it is monotone relative to K.

Throughout this subsection, we assume that $T : A \to A$ is a continuous competitive map on the subset A of the plane. Further hypotheses concerning A will be made below. As noted above, all of the results have obvious analogs in the case of cooperative planar maps (just interchange cones). Competitive planar maps preserve the order relation $\leqslant_K$ by definition, but they also put constraints on the usual ordering, as we show below.

LEMMA 5.27. *Let $T : A \to A$ be a competitive map on $A \subset \mathbb{R}^2$. If $x, y \in A$ satisfy $Tx \ll Ty$, then either $x \ll y$ or $y \ll x$.*

PROOF. If neither $x \ll y$ nor $y \ll x$ hold, then $x \leqslant_K y$ or $y \leqslant_K x$ holds. But $x \leqslant_K y$ implies $Tx \leqslant_K Ty$ which is incompatible with $Tx \ll Ty$. A similar contradiction is obtained from $y \leqslant_K x$. □

Lemma 5.27 suggests placing one of the following additional assumptions on T.

(O_+) If $x, y \in A$ and $Tx \ll Ty$, then $x \leqslant y$.
(O_-) If $x, y \in A$ and $Tx \ll Ty$, then $y \leqslant x$.

As we shall soon see, if T is orientation preserving, then (O_+) holds and if it is orientation reversing, then (O_-) holds. A sequence $\{x_n = (u_n, v_n)\} \subset \mathbb{R}^2$ is *eventually componentwise monotone* if there exists a positive integer N such that either $u_n \leqslant u_{n+1}$ for all $n \geqslant N$ or $u_{n+1} \leqslant u_n$ for all $n \geqslant N$ and similarly for v_n.

In the case of orientation-preserving maps, the following result was first proved by de Mottoni and Schiaffino [42] for the period map of a periodic competitive Lotka–Volterra system of differential equations.

THEOREM 5.28. *If T is a competitive map for which* (O_+) *holds then for all $x \in A$, $\{T^n x\}_{n \geqslant 0}$ is eventually component-wise monotone. If the orbit of x has compact closure in A, then it converges to a fixed point of T. If, instead,* (O_-) *holds then for all $x \in A$, $\{T^{2n} x\}_{n \geqslant 0}$ is eventually component-wise monotone. If the orbit of x has compact closure in A, then its omega limit set is either a period-two orbit or a fixed point.*

PROOF. We first note that if T is competitive and (O_-) holds then T^2 is competitive and (O_+) holds (use Lemma 5.27) so the second conclusion of the theorem follows from the first.

Suppose that (O_+) holds. If $T^n x \leqslant_K T^{n+1} x$ or $T^{n+1} x \leqslant_K T^n x$ holds for some $n \geqslant 1$, then it holds for all larger n so the conclusion is obvious. Therefore, we assume that this is not the case. It follows that for each $n \geqslant 1$ either (a) $T^n x \ll T^{n+1} x$ or (b) $T^{n+1} x \ll T^n x$. We claim that either (a) holds for all n or (b) holds for all n. Assume $x \ll Tx$ (the argument is similar in the other case). If the claim is false, then there is an $n \geqslant 1$ such that $x \ll Tx \ll \cdots \ll T^{n-1} x \ll T^n x$ but $T^{n+1} x \ll T^n x$. But (O_+) implies $T^n x \leqslant T^{n-1} x$ contradicting the displayed inequality. □

Orbits may not converge to a fixed point if (O_-) holds. Consider the map $T : I \to I$ where $I = [-1, 1]^2$ and $T(u, v) = (-v, -u)$ reflects points through the line $v = -u$. It is easy to see using Lemma 5.1 that T is competitive and that (O_-) holds (see below). Fixed points of T lie on the above-mentioned line but all other points in I are period-two points.

The hypotheses (O_+) and (O_-) on T are global in nature and therefore can be difficult to check in specific examples. We now give sufficient conditions for them to hold that may be easier to verify in applications. *A contains order intervals* if $x, y \in A$ and $x \ll y$ implies that $[x, y] \subset A$. Clearly, $A = [a, b]$ contains order intervals. If $A \subset \mathbb{R}^2$ and $T : A \to \mathbb{R}^2$, we say that T is C^1 if for each $a \in A$ there is an open set U in $\mathbb{R}^2$ and a continuously differentiable function $F : U \to \mathbb{R}^2$ that coincides with T on $U \cap A$. We will have occasion to make certain hypotheses concerning $T'(x)$ even though it is not necessarily uniquely defined. What we mean by this is that there exists an F as above such that $T'(x) = DF(x)$ has the desired properties. This abuse of language will lead to no logical difficulties in the arguments below. In the applications, A will typically be $\mathbb{R}^2_+$ or some order interval $[a, b]$ where $a \ll b$ in which case T' is uniquely defined.

Consider the following hypothesis:

(H_+) (a) A contains order intervals and is p-convex with respect to $\leqslant_K$.
(b) $\det T'(x) > 0$ for $x \in A$.

(c) $T'(x)(K) \subset K$ for $x \in A$.
(d) T is injective.

Hypothesis (H_-) is identical except the inequality is reversed in (b).

LEMMA 5.29. *If $T : A \to A$ satisfies* (H_+)*, then T is competitive and* (O_+) *holds. If* (H_-) *holds, then T is competitive and* (O_-) *holds.*

PROOF. T is competitive by hypothesis (c) since A is p-convex with respect to $\leqslant_K$. Assuming that (H_+) holds, $x, y, Tx, Ty \in A$ and $Tx \ll Ty$, we will show that $x \ll y$. According to Lemma 5.27, the only alternative to $x \ll y$ is $y \ll x$ so we assume the latter for contradiction. Let a and b be the northwest and southeast corners of the rectangle $[y, x] \subset A$ so that $a \ll_K b$ and $[y, x] = [a, b]_K$. Since T is competitive on A, $T([y, x]) \subset [Ta, Tb]_K$ and $Tx \ll Ty$ implies that $Ta \ll_K Tb$. Consider the oriented Jordan curve forming the boundary of $[y, x]$ starting at a and going horizontally to x, then going vertically down to b, horizontally back to y and vertically up to a. As T is injective on A, the image of this curve is an oriented Jordan curve. Monotonicity of T implies that the image curve is contained in $[Ta, Tb]_K$, begins at Ta and moves monotonically with respect to $\leqslant_K$ (southwest) through Tx and then monotonically to Tb before moving monotonically (decreasing or northwest) from Tb through Ty and on to Ta. (H_+)(b) implies that T is locally orientation preserving, so upon traversing the first half of the image curve from Ta to Ty to Tb, the curve must make a “right turn” at Tb before continuing on to Tx and to Ta. As the image curve cannot intersect itself, we see that $Tx \ll Ty$ cannot hold, a contradiction. □

In specific examples it is often difficult to check that T is injective. It automatically holds if A is compact and connected and there exists $z \in T(A)$ such that the set $T^{-1}(z)$ is a single point. This is because the cardinality of $T^{-1}(w)$ is finite and constant for $w \in T(A)$ by Chow and Hale [26, Lemma 2.3.4].

The following is an immediate corollary of Theorem 5.28 and Lemma 5.29.

COROLLARY 5.30. *If $T : A \to A$ satisfies* (H_+)*, then $\{T^n x\}$ is eventually component-wise monotone for every $x \in A$. In this case, if an orbit has compact closure in A, then it converges to a fixed point of T. If T satisfies* (H_-)*, then $\{T^{2n} x\}$ is eventually component-wise monotone for every $x \in A$. In this case, if an orbit has compact closure in A, then its omega limit set is either a fixed point or a period-two orbit.*

As an application of Corollary 5.30, we recall the celebrated results of de Mottoni and Schiaffino [42] for the periodic Lotka–Volterra system

$$\begin{aligned} x' &= x\bigl[r(t) - a(t)x - b(t)y\bigr], \\ y' &= y\bigl[s(t) - c(t)x - b(t)y\bigr], \end{aligned} \tag{5.4}$$

where r, s, a, b, c, d are periodic of period one and $a, b, c, d \geqslant 0$. The period map $T : \mathbb{R}^2_+ \to \mathbb{R}^2_+$, defined by (5.2) for (5.4), is strictly monotone relative to the fourth quadrant cone K by virtue of Theorem 3.5. Indeed, (5.4) is a competitive system relative to the cone $\mathbb{R}^2_+$ (the off-diagonal entries of the Jacobian $J = J(t, x, y)$ of the right-hand side are nonpositive), and every such system is monotone relative to K. Observe that $J + \alpha I$, for large enough $\alpha > 0$, has nonnegative diagonal entries so $(J + \alpha I)(u, v)^T \in K$ if $(u, v)^T \in K$ (i.e., $u \geqslant 0, v \leqslant 0$). T is strongly monotone relative to K in $\operatorname{Int} \mathbb{R}^2_+$ if $b, c > 0$ by Corollary 3.11. Because T is injective and orientation preserving by Liouville's theorem, (H_+) holds. Orbits are seen to be bounded by simple differential inequality arguments, e.g., applied to $x' \leqslant x[r(t) - a(t)x]$. Consequently, by Corollary 5.30, all orbits $O(T)$ converge to a fixed point; equivalently, every solution of (5.4) is asymptotic to a period-one solution.

System (5.4) is most interesting when each species can survive in the absence of its competitor, i.e. the time average of r and s are positive. In that case, aside from the trivial fixed point $E_0 := (0, 0)$, there are unique fixed points of type $E_1 := (e, 0)$ and $E_2 := (0, f)$. Of course, $e, f > 0$ give initial data corresponding to the unique nontrivial one-periodic solutions of the scalar equations: $x' = x[r(t) - a(t)x]$ and $y' = y[s(t) - d(t)y]$. The dynamics of the period map for these equations is described by alternative (iii) of Theorem 5.20.

It is shown by de Mottoni and Schiaffino that there is a monotone, relative to K, T-invariant curve joining E_1 to E_2 which is the global attractor for the dynamics of T in $\mathbb{R}^2_+ \setminus \{E_0\}$. This work has inspired a very large amount of work on competitive dynamics. See Hale and Somolinos [60], Smith [188,189], Hess and Lazer [64], Hsu et al. [83], Smith and Thieme [201], Wang and Jiang [230,231,229], Liang and Jiang [121], Zanolin [238].

6. Semilinear parabolic equations

The purpose of this section is to analyze the monotone dynamics in a broad class of second order, semilinear parabolic equations.

For basic theory and further information on many topics we refer the reader to books of Amann [11], Henry [62] Cholewa and Dlotko [25], Hess [63], Lunardi [124] and Martin [125], the papers of Amann [7–10], and the survey article of Poláčik [163].

Solution processes for semilinear parabolic problems have been obtained by many authors; see for example [3,8,38,62,63,124,125,134,147,163,194,174,208,246]. We briefly outline the general procedure, due to Henry, with important improvements by Mora and Lunardi.

To balance the sometimes conflicting goals of order, topology and dynamics, the domain of a solution process must be chosen carefully. We rely on results of Mora [147], refined by Lunardi [124], for solution processes in Banach subspaces $C^k_B(\overline{\Omega}) \subset C^k(\overline{\Omega})$, $k = 0, 1$ determined by the boundary operator B.

6.1. *Solution processes for abstract ODEs*

If Y and X are spaces such that Y is a subset of X and the inclusion map $Y \to X$ is continuous, we write $Y \hookrightarrow X$. When Y and X are ordered Banach space structures, this notation tacitly states that Y is a linear subspace of X and $Y_+ = Y \cap X_+$.

The domain and range of any map h are denoted by $\mathsf{D}(h)$ and $\mathsf{R}(h)$.

6.1.1. *Processes* Let Z be a topological space and set $\widehat{Z} := \{(t, t_0, z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times Z\colon\ t \geqslant t_0\}$. A *process* in Z is a family $\Theta = \{\Theta_{t,t_0}\}_{0 \leqslant t_0 < t}$ of continuous maps

$$\Theta_{t,t_0}\colon D_{t,t_0} \to Z, \quad D_{t,t_0} \text{ open in } Z,$$

where the set $\{(t, t_0, z) \in \widehat{Z}\colon\ z \in D(t, t_0)\}$ is open in $\widehat{Z}$ containing $\{(t, t, z)\colon\ t \geqslant 0, z \in Z\}$, with the properties:

- the map $(t, t_0, z) \mapsto \Theta_{t,t_0}(z)$ is continuous from $\widehat{Z}$ to Z.
- the *cocycle identities* hold:

$$t \geqslant t_1 \geqslant t_0 \quad \Longrightarrow \quad \Theta_{t,t_1} \circ \Theta_{t_1,t_0} = \Theta_{t,t_0}, \quad \Theta_{t,t} = \text{identity map of } Z.$$

Equivalently: there is a local semiflow Λ on $\mathbb{R}_+ \times Z$ such that $\Lambda_t(t_0, u_0) = (t + t_0, \Theta_{t,t_0}(u_0))$. It follows that for each (t_0, z) there is a maximal $\tau := \tau(t_0, z) \in (t_0, \infty]$ such that $z \in D_{t,t_0}$ for all $t \in [t_0, \tau)$. The *trajectory* of (t_0, z) is the parametrized curve $[t_0, \tau) \to Z$, $t \mapsto \Theta_{t,t_0}(z)$, whose image is the *orbit* of (t_0, z). A subset $S \subset Z$ is *positively invariant* if it contains the orbit of every point in $\mathbb{R}_+ \times S$.

A trajectory is *global* if it is defined on $[t_0, \infty)$. The process is called global when all trajectories are global.

Let S be a space such that $S \hookrightarrow Z$. It may be that S is positively invariant under the process Θ, and the maps $\Theta_{t,t_0}\colon S \cap D(t, t_0) \to S$ are continuous respecting the topology on S and furthermore, the map $(t, t_0, s) \to \Theta_{t,t_0}s$ is continuous from $\widehat{S}$ to S. In this case these maps form the *induced process* Θ^S in S.

A process Θ in an ordered space is called (locally) monotone, SOP, Lipschitz, compact, and so forth, provided every map Θ_{t,t_0}, $t > t_0$ has the corresponding property.

6.1.2. *Solution processes* Let X be a Banach space. $\mathcal{A}$ denotes a linear operator (usually unbounded) in X with domain $\mathsf{D}(\mathcal{A}) \subset X$, that is *sectorial* in the following strong sense:

- $\mathcal{A}$ is a densely defined, closed operator generating an analytic semigroup $\{\mathrm{e}^{t\mathcal{A}}\}_{t \geqslant 0}$ in $L(X)$, and the resolvent operators $(\lambda I - \mathcal{A})^{-1} \in L(X)$ are compact for sufficiently large $\lambda \geqslant 0$.

The latter property ensures that $\mathrm{e}^{t\mathcal{A}}$ is compact for $t > 0$ [156, Theorem 2.3.3].

We make $\mathsf{D}(\mathcal{A})$ into a Banach space with the graph norm $\|x\|_{\mathsf{D}(\mathcal{A})} = \|x\| + \|\mathcal{A}x\|$, or any equivalent norm. Then $\mathcal{A}\colon \mathsf{D}(\mathcal{A}) \to X$ is bounded, and $\mathsf{D}(\mathcal{A}) \hookrightarrow X$.

For $0 \leqslant \alpha \leqslant 1$ we define the fractional power domain of $\mathcal{A}^\alpha$ to be $X^\alpha = X^\alpha(\mathcal{A}) := \mathsf{D}(\mathcal{A}^\alpha)$. Thus we have [62]

$$\mathsf{D}(\mathcal{A}) \hookrightarrow X^\alpha \hookrightarrow X, \qquad \overline{\mathsf{D}(\mathcal{A})} = X.$$

Let $F\colon [0, \infty) \times X^\alpha \to X$ be a continuous map that is *locally Lipschitz in the second variable*, i.e.:

- $F|[0, \tau] \times B(r)$ has Lipschitz constant $L(\tau, r)$ in the second variable whenever $[0, \tau] \subset [0, \infty)$ and $B(r)$ is the closed ball of radius r in X^α.

Locally Hölder in the first variable is defined analogously. We say F is C^1 *in the second variable* if $\partial_w F(t, w)$ is continuous.

The data $(X, \mathcal{A}, F)$ determine the abstract initial value problem

$$\begin{cases} u'(t) = \mathcal{A}u(t) + F(t, u(t)) \quad (t > t_0), \\ u(t_0) = u_0 \in X. \end{cases} \tag{6.1}$$

A continuous curve $u : [t_0, \tau) \to X$, $t_0 < \tau \leqslant \infty$ is a (classical) *solution through* (t_0, u_0) if $u(t) \in \mathsf{D}(\mathcal{A})$ for $t_0 < t < \tau$ and (6.1) holds. It is well known (e.g., Lunardi [124, 4.1.2]) that every solution is also a *mild solution*, i.e., it satisfies the integral equation

$$u(t) = \mathrm{e}^{(t-t_0)\mathcal{A}} u_0 + \int_{t_0}^{t} \mathrm{e}^{(t-s)\mathcal{A}} F\big(s, u(s)\big)\, \mathrm{d}s \quad (t_0 \leqslant t < \tau). \tag{6.2}$$

Moreover, every mild solution is a solution provided F is locally Hölder in t (Lunardi [124, Proposition 7.1.3]).

A classical or mild solution is *maximal* if it does not extend to a classical or mild solution on a larger interval in $[t_0, \infty)$; it is then referred to as a *trajectory* at (t_0, u_0), and its image is an *orbit*. When such a trajectory is unique it is denoted by $t \mapsto u(t, t_0, u_0)$. In this case the *escape time* of (t_0, u_0) is $\tau(t_0, u_0) := \tau$. If $\tau = \infty$ the trajectory is called *global*.

The following basic result means that Eq. (6.1) is well-posed in a strong sense, and that solutions enjoy considerable uniformity and compactness.

THEOREM 6.1. *Let* $(t_0, u_0) \in \mathbb{R}_+ \times X^\alpha$. *There is a unique mild trajectory at* (t_0, u_0), *and it is a classical trajectory provided* $F(t, u)$ *is locally Hölder in* t. *If* $t_0 < t_1 < \tau(t_0, u_0)$, *there is a neighborhood* U *of* x_0 *in* X^α *and* $M > 0$ *such that*

$$\|u(t, t_0, u_1) - u(t, t_0, u_2)\|_{X^\alpha} \leqslant M \|u_1 - u_2\|_{X^\alpha}, \quad u_1, u_2 \in U.$$

There exist $C > 0$, $t_0 < t_1 < \tau(t_0, u_0)$, *a bounded neighborhood* N *of* u_0 *in* X *and a continuous map*

$$\Psi : [t_0, t_1] \times N \to X, \qquad (t, v) \mapsto u(t, t_0, v),$$

where $u(t, t_0, v)$ *is a mild solution, such that the following hold. If* $s, t \in (t_0, t_1]$, $0 \leqslant \alpha < 1$ *and* $v, w \in N$:

(i) $\|\Psi(s, v) - \Psi(s, w)\| \leqslant C\|v - w\|$;

(ii) $\|\Psi(s, v) - \Psi(s, w)\|_{X^\alpha} \leqslant (s - t_0)^{-\alpha} C \|v - w\|$;

(iii) $\Psi([s, t_1] \times N)$ *is precompact in* X^α;

(iv) $u(\cdot, t_0, v) : (t_0, t] \to X^\alpha$ *and* $u(\cdot, t_0, v) : [t_0, t] \to X$ *are continuous*;

(v) *trajectories bounded in* X^α *are global.*

PROOF. Lunardi [124, Theorems 7.1.2, 7.1.3 and 7.1.10] proves the first assertion. Items (i), (ii) and (iv) follow from [124, Theorem 7.1.5], and (v) follows from Theorem 7.1.8 (see also Henry [62, 3.3.4]). Fix β with $\alpha < \beta < 1$. As N is bounded in X, $\Psi(s \times N)$ is bounded in X^β by (ii) (with α in (ii) replaced by β). Therefore $\Psi(s \times N)$ is precompact in X^α, and (iii) follows because Ψ defines a local semiflow on $\mathbb{R}_+ \times X^\alpha$. □

Equation (6.1) induces a *solution process* Θ in X, defined by $\Theta_{t,t_0}(u_0) := u(t, t_0, u_0)$. Its restriction to X^α defines an induced solution process on that space. When Eq. (6.1) is autonomous, i.e., $F(t, u) = F(u)$, this solution process boils down to a local semiflow Φ in X^α, defined by $\Phi_t(t_0, u_0) = u(t + t_0, t_0, u_0)$.

When $F(t, u)$ has period $\lambda > 0$ in t, the solution process is λ-*periodic*: $\Theta_{t,t_0} \equiv \Theta_{t+\lambda,\, t_0+\lambda}$. In this case Θ reduces to a local semiflow on $\mathbf{S}^1 \times X^\alpha$, the dynamics of which are largely determined by the Poincaré map $T := \Theta_{\lambda,0}$ which maps an open subset of X^α continuously into X^α.

Let S be a set and Z a Banach space. We use expressions such as "S is bounded in Z" or "$S \subset Z$ is bounded" to mean $S \subset Z$ and $\sup_{u \in S} \|u\|_Z < \infty$. Note that S may also be unbounded in other Banach spaces.

A map defined on a metric space is *compact* if every bounded set in its domain has precompact image. It is *locally compact* if every point of the domain has a neighborhood with precompact image.

A Banach space Y is *adapted* to the data $(X, \mathcal{A}, F)$ if the following two conditions hold:

$$X^\alpha \hookrightarrow Y \hookrightarrow X \tag{6.3}$$

and the map $(t, u_0) \mapsto \Theta_{t,t_0} u_0$ from $[t_0, \tau) \times D(t, t_0) \cap Y$ to Y is continuous. The solution process Θ determines the *induced solution process* Θ^Y in Y. The domain of Θ^Y_{t,t_0} is the open subset $D^Y(t, t_0) := D_{t,t_0} \cap Y$ of Y.

Rather than work with fractional power spaces, one can assume that $F : [0, \infty) \times K \to X$ where K is a suitable subset of X. The subset $K \subset X$ is *locally closed* in the Banach space X if for each $x \in K$ there exists $r > 0$ such that $\{y \in K\colon \|x - y\| \leqslant r\}$ is closed in X. Closed and open subsets K of X are locally closed. Note that the following result gives existence and uniqueness of mild solutions while at the same time giving positive invariance. It is a special case of Theorems VIII.2.1 and VIII.3.1 in Martin [125]. Assumptions on the semigroup $\mathrm{e}^{t\mathcal{A}}$ remain as above.

THEOREM 6.2. *Let K be a nonempty locally closed subset of a Banach space X and let $F : [0, \infty) \times K \to X$ be continuous and satisfy*: *For each $R > 0$ there are $L_R > 0$ and $\gamma \in (0, 1]$ such that for $x, y \in K$, $\|x\|, \|y\| \leqslant R$, $0 \leqslant s, t \leqslant R$*

$$\big\|F(t, x) - F(s, y)\big\| \leqslant L_R\big(|t - s|^\gamma + \|x - y\|\big). \tag{6.4}$$

Suppose also that:

(a) $\mathrm{e}^{t\mathcal{A}}(K) \subset K$ *for all $t \geqslant 0$, and*

(b) $\liminf_{h \searrow 0} \frac{1}{h} \operatorname{dist}(x + hF(t, x), K) = 0$ *for $(t, x) \in [0, \infty) \times K$.*

Then for each $(t_0, u_0) \in [0, \infty) \times K$, there is a unique classical trajectory $u(t, t_0, u_0)$ of (6.2) *defined on a maximal interval $[t_0, \tau)$, and $u(t) \in K$ for $t_0 \leqslant t < \tau$.*

This result is useful for parabolic systems when $X = C^k(\overline{\Omega})$, $k = 0, 1$ but not when $X = L^{\mathrm{p}}(\Omega)$. The substitution operators are well-behaved in the former cases but require very stringent growth conditions for the latter; see Martin [125]. By virtue of the uniqueness assertions of Theorem 6.1 and Theorem 6.2, the solution processes given by the two results agree on K if (6.4) holds.

Hypothesis (a) is obviously required for the positive invariance of K in case $F = 0$. Hypothesis (b), called the *subtangential condition,* is easily seen to be a necessary condition for the positive invariance of K if $\mathcal{A} = 0$. See Martin [125, Theorem VI.2.1]. Both hypotheses are trivially satisfied if $K = X$.

The following result is a special case of [125, Proposition VIII.4.1]:

PROPOSITION 6.3. *Let* $F : [0, \infty) \times X \to X$ *be continuous and satisfy* (6.4) *with* $K = X$ *and let* $u(t) = u(t, t_0, x_0)$ *be the unique classical trajectory defined on a maximal interval* $[t_0, \tau)$ *guaranteed by Theorem* 6.2. *If* $\tau < \infty$ *then* $\lim_{t\to\tau} \|u(t)\| = \infty$.

6.1.3. *Monotone processes* Given our interest in establishing monotonicity properties of solution processes induced by parabolic systems in various functions spaces, there are two approaches one may take. One is to establish the properties on spaces of smooth functions such as fractional power spaces X^α for $\alpha < 1$ near unity and then try to extend the monotonicity to larger spaces, e.g., $C^0(\overline{\Omega})$, by approximation. An alternative is to establish the monotonicity properties on the larger spaces first and then get corresponding properties on the smaller spaces by restriction. We give both approaches here, beginning with the former.

A process Θ is *very strongly order preserving* (= VSOP) if it is monotone and has the following property: Given $t_0 \geqslant 0$, $u > v$, and $\epsilon > 0$, there exist $s \in (t_0, t_0 + \epsilon]$ and neighborhoods U, V of u, v respectively such that

$$t \geqslant s \quad \Longrightarrow \quad \Theta_{t,t_0}(U \cap D_{t,t_0}) > \Theta_{t,t_0}(V \cap D_{t,t_0}).$$

This implies Θ is SOP and strictly monotone.

THEOREM 6.4. *Assume* X *is an ordered Banach space and* $Y \hookrightarrow X$ *an ordered Banach space such that* Y *is dense in* X *and the order cone* $Y_+ := Y \cap X_+$ *is dense in* X_+. *Let* Θ *be a process in* X *that induces a monotone process* Θ^Y *in* Y. *Then*:

(a) Θ *is monotone.*

(b) *Assume* $\mathsf{R}(\Theta_{t,t_0}) \subset Y$ *for all* $t > t_0 \geqslant 0$. *Then* Θ *is strictly monotone if* Θ^Y *is strictly monotone, and* Θ *is VSOP provided* Θ^Y *is strongly monotone and* $\Theta_{t,t_0} : D(t, t_0) \to Y$ *is continuous for* $t > t_0$.

PROOF. (a) Fix u and $v > u$ in X. The closed line segment $\overline{uv}$ spanned by u and v is compact, hence there exists $\rho > t_0$ with $\overline{uv} \subset D_{t_0,\rho}$. By the density assumptions there exist convergent sequences $u_n \to u$, $v_n \to v$ in $D_{t_0,\rho}$ such that $u_n, v_n \in Y$ and $u_n < v_n$. As Θ^Y is induced from Θ, it follows that $u_n, v_n \in D^Y_{t_0,\rho}$. For all $t \in [t_0, \rho)$,

$$\Theta_{t,t_0}(u_n) = \Theta^Y_{t,t_0}(u_n) \leqslant \Theta^Y_{t,t_0}(v_n) = \Theta_{t,t_0}(v_n).$$

Taking limits as $n \to \infty$ proves $\Theta_{t,t_0}(u) \leqslant \Theta_{t,t_0}(v)$. Thus Θ is monotone.

(b) Assume now that Θ^Y is strictly monotone. We show that Θ is strictly monotone. Let $u(t), v(t)$ be local trajectories with $u(t_0) < v(t_0)$. If $r \in (t_0, t_1]$ is sufficiently near t_0, then $u(r), v(r)$ are distinct points of Y, and $u(r) < v(r)$ by (a). Hence $u(t_1) < v(t_1)$ by strict monotonicity of Θ^Y.

To prove Θ is VSOP, let $u(t), v(t)$ be as above with $u(t_0), v(t_0) \in D_{t,t_0}$. If $t_0 < s < r < t$, strict monotonicity implies $u(s) > v(s)$. These points are in Y, Θ^Y is strongly monotone, and Θ agrees with Θ^Y in Y. Therefore there are disjoint neighborhoods $U_1, V_1 \subset Y$ of $u(s), v(s)$ respectively, such that

$$\Theta_{r,s}(U_1 \cap D_{r,s}) \gg \Theta_{r,s}(V_1 \cap D_{r,s})$$

and strict monotonicity implies that

$$t > r \quad \Longrightarrow \quad \Theta_{t,r}(U_1 \cap D_{t,r}) > \Theta_{t,r}(V_1 \cap D_{t,r}). \tag{6.5}$$

As $\Theta_{r,t_0} : D(r, t_0) \to Y$ is continuous, we may define neighborhoods $U, V \subset X$ of $u(t_0), v(t_0)$ respectively by

$$U = \Theta_{r,t_0}^{-1}(U_1), \qquad V = \Theta_{r,t_0}^{-1}(V_1).$$

By (6.5) and the cocycle identities,

$$t > r \quad \Longrightarrow \quad \Theta_{t,t_0}(U \cap D_{t,t_0}) > \Theta_{t,t_0}(V \cap D_{t,t_0}). \qquad \square$$

Let X be an ordered Banach space with positive cone X_+ and K a locally closed subset. The mapping $F : K \to X$ is said to be *quasimonotone* (relative to X_+) if:

(QM) For all $(t, x), (t, y) \in [0, \infty) \times K$ satisfying $x \leqslant y$ we have:

$$\lim_{h \searrow 0} \frac{1}{h} \operatorname{dist}\big(y - x + h\big[F(t, y) - F(t, x)\big], X_+\big) = 0.$$

The next result is due to [125, Proposition VIII.6.1 and Lemma 6.3] (see also [129] in case of abstract delay differential equations).

THEOREM 6.5. *Assume the hypotheses of Theorem* 6.2 *hold*, F *is quasimonotone*, *and* $e^{t\mathcal{A}}$ *is a positive operator for* $t \geqslant 0$. *In addition, suppose one of the following*:

(i) K *is open*.

(ii) $K + X_+ \subset K$.

(iii) X *is a Banach lattice and* $K = [u, v]$ *for some* $u, v \in X \cup \{-\infty, \infty\}$, $u \leqslant v$.

Then

$$x, y \in K, x \leqslant y \quad \Longrightarrow \quad u(t, x) \leqslant u(t, y) \quad \big(0 \leqslant t \leqslant \min\{\tau_x, \tau_y\}\big).$$

By $[-\infty, v]$, $v \in X$, is meant the set $\{x \in X\colon\ x \leqslant v\}$; similarly for other intervals involving $\pm\infty$. Of course, $-\infty \leqslant v \leqslant \infty$ for every $v \in X$. Observe that $K = [u, \infty]$ is covered by both (ii) and (iii).

REMARK 6.6. If F has the property that for each $x, y \in K$ with $x \leqslant y$, there exists $\lambda > 0$ such that $F(t,x) + \lambda x \leqslant F(t,y) + \lambda y$ then F is quasimonotone because

$$\begin{aligned} y - x + h\bigl[F(t,y) - F(t,x)\bigr] &= (1 - \lambda h)(y - x) \\ &\quad + h\bigl[F(t,y) + \lambda y - F(t,x) - \lambda x\bigr] \in X_+ \end{aligned}$$

when $h < \lambda^{-1}$.

REMARK 6.7. It is well-known that $\mathrm{e}^{t\mathcal{A}}$ is a positive operator if and only if $(\lambda I - \mathcal{A})^{-1}$ is a positive operator for all large positive λ. See, e.g., [11, Theorem II 6.4.1] or [125, Proposition 7.5.3]. Indeed, if K is a closed convex subset of X, then $\mathrm{e}^{t\mathcal{A}}K \subset K$ if and only if $(\lambda I - \mathcal{A})^{-1}K \subset K$ for all large positive λ.

A Banach space X is a *Banach lattice* if for each $x, y \in X$, $x \vee y := \sup\{x, y\}$ exists and the norm is monotone in the sense:

$$|x| \leqslant |y| \quad \Longrightarrow \quad \|x\| \leqslant \|y\|,$$

where $|x|$ denotes the absolute value of x: $|x| := (-x) \vee x$ (see Vulikh [225]). Banach lattices are easy to work with due to simple formulas such as

$$\operatorname{dist}(x, X_+) = \|x - x_+\| = \|x_-\|,$$

where $x_+ := x \vee 0$ and $x_- = -(-x)_+$. The requirement that X be a Banach lattice is a rather strong hypothesis which essentially restricts applicability to $X = L^p(\Omega), C^0(\overline{\Omega})$ or $C_0^0(\overline{\Omega})$. However, the latter two will be important for reaction–diffusion systems.

6.2. *Semilinear parabolic equations*

Let $\Omega \subset \mathbb{R}^n$ be the interior of a compact n-dimensional manifold with C^2 boundary $\partial\Omega$. We consider the semilinear system of m coupled equations ($1 \leqslant i \leqslant m$):

$$\begin{aligned} &\frac{\partial u_i}{\partial t} = (A_i u_i)(t,x) + f_i(t,x,u,\nabla u) && (x \in \Omega,\ t > t_0), \\ &(B_i u_i)(t,x) = 0 && (x \in \partial\Omega,\ t > t_0), \\ &u_i(t_0, x) = v_{0,i}(x) && (x \in \overline{\Omega}). \end{aligned} \tag{6.6}$$

Here the unknown function is $u = (u_1, \dots, u_m) : \overline{\Omega} \to \mathbb{R}^m$, and $\nabla u := (\nabla u_1, \dots, \nabla u_m) \in (\mathbb{R}^n)^m$ lists the spatial gradients ∇u_i of the u_i, i.e., $\nabla u_i := (\frac{\partial u_i}{\partial x_1}, \dots, \frac{\partial u_i}{\partial x_n})$. Each $A_i(x)$ is a second order, elliptic differential operator of the form

$$A_i(x) = \sum_{l,j=1}^{n} C_{lj}^{i}(x) \frac{\partial}{\partial x_l} \frac{\partial}{\partial x_j} + \sum_{j=1}^{n} b_j^i(x) \frac{\partial}{\partial x_j} \tag{6.7}$$

with uniformly continuous and bounded coefficients. Each $n \times n$ matrix $C^i(x) := [C^i_{lj}(x)]$ is assumed positive definite:

$$0 < \inf\bigl(\bigl\langle C^i(x)y, y\bigr\rangle\bigr) \quad \bigl(x \in \Omega, y \in \mathbb{R}^n, |y| = 1\bigr),$$

where $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product on $\mathbb{R}^n$.

The function

$$f := (f_1, \dots, f_m) : \mathbb{R}_+ \times \overline{\Omega} \times \mathbb{R}^m \times \bigl(\mathbb{R}^n\bigr)^m \to \mathbb{R}^m$$

is continuous, and $f(t, x, u, \xi)$ is locally Lipschitz in $(u, \xi) \in \mathbb{R}^m \times (\mathbb{R}^n)^m$.

Each boundary operator B_i acts on sufficiently smooth functions $v : [t_0, \tau) \times \overline{\Omega} \to \mathbb{R}$ in one of the following ways, where $x \in \partial\Omega$:

$$\begin{aligned}
&\textit{Dirichlet}: && (B_i v)(t, x) = v(t, x);\\
&\textit{Robin}: && (B_i v)(t, x) = \gamma_i v(t, x) + \frac{\partial v}{\partial \xi_i}(t, x);\\
&\textit{Neumann}: && (B_i v)(t, x) = \frac{\partial v}{\partial \xi_i}(t, x),
\end{aligned}$$

where $\gamma_i : \overline{\Omega} \to [0, \infty)$ is continuously differentiable, and $\xi_i : \overline{\Omega} \to \mathbb{R}^n$ is a continuously differentiable vector field transverse to $\partial\Omega$ and pointing outward from Ω. Note that Neumann is a special case of Robin.

We rewrite (6.6) as an initial-boundary value problem for an unknown vector-valued function $u := (u_1, \dots, u_m) : [t_0, \tau) \times \overline{\Omega} \to \mathbb{R}^m$,

$$\begin{aligned}
&\frac{\partial u}{\partial t} = (Au)(t, x) + f(t, x, u, \nabla u) && (x \in \Omega,\ t > t_0),\\
&(Bu)(t, x) = 0 && (x \in \partial\Omega,\ t > t_0),\\
&u(t_0, x) = u_0(x) && (x \in \overline{\Omega}),
\end{aligned} \tag{6.8}$$

where the operators $A := A_1 \times \cdots \times A_m$ and $B := B_1 \times \cdots \times B_m$ act componentwise on $u = (u_1, \dots, u_m)$. By a *solution process* for Eq. (6.8) we mean a process in some function space on $\overline{\Omega}$, whose trajectories are solutions to (6.8).

Of special interest are *autonomous* systems, for which $f = f(x, u, \nabla u)$; and the *reaction–diffusion* systems, characterized by $f = f(t, x, u)$.

Assume $n < p < \infty$. To Eq. (6.8) we associate an abstract differential Eq. (6.1) in $L^p(\Omega, \mathbb{R}^m)$. The pair of operators (A_i, B_i) has a sectorial realization $\mathcal{A}_i$ in $L^p(\Omega)$ with domain $\mathcal{D}(\mathcal{A}_i) \hookrightarrow L^p(\Omega)$ (Lunardi [124, 3.1.3]). The operator $\mathcal{A} := \mathcal{A}_1 \times \cdots \times \mathcal{A}_m$ is sectorial on $X := L^p(\Omega, \mathbb{R}^m) = [L^p(\Omega)]^m$.

For $\alpha \in [0, 1)$ set $X^\alpha := X^\alpha(\mathcal{A})$. We choose α so that f defines a continuous *substitution operator*

$$F : \mathbb{R}_+ \times X^\alpha \to X, \qquad F(t, u)(x) := f\bigl(t, x, u(x), \nabla u(x)\bigr).$$

It suffices to take α

$$1 > \alpha > \frac{1}{2}\left(1 + \frac{n}{p}\right), \tag{6.9}$$

for then $X^\alpha \hookrightarrow C^1(\overline{\Omega}, \mathbb{R}^m)$ by the Sobolev embedding theorems.

The data $(\mathcal{A}, F)$ thus determine an abstract differential equation $u' = \mathcal{A}u + F(t, u)$ in X, whose trajectories $u(t)$ correspond to solutions $u(t, x) := u(t)(x)$ of (6.6). The assumptions on f make $F(t, u)$ locally Lipschitz in $u \in X^\alpha$.

By Theorem 6.1 and the Sobolev embedding theorem we have:

PROPOSITION 6.8. *Equation* (6.8) *defines a solution process* Θ *on* $X := L^p(\Omega, \mathbb{R}^m)$ *which induces a solution process in* X^β *for every* $\beta \in [0, 1)$ *with* $\beta \geqslant \alpha$.

We quote a useful condition for globality of a solution:

PROPOSITION 6.9. *Assume there are constants* $C > 0$ *and* $0 < \epsilon \leqslant 1$ *such that*

$$\big\| f(t, x, v, \xi)\big\| \leqslant C\big(1 + \|v\| + \|\xi\|^{2-\epsilon}\big) \quad \text{for all } (t, x, v, \xi) \in \mathbb{R}_+ \times \overline{\Omega} \times S \times \mathbb{R}^n. \tag{6.10}$$

If $u : [t_0, \tau) \to L^p(\Omega, \mathbb{R}^m)$ *is a trajectory such that*

$$\limsup_{t \to \tau-} \big\| u(t)\big\|_{L^p(\Omega, \mathbb{R}^m)} < \infty \tag{6.11}$$

then $\tau = \infty$.

PROOF. Follows from Amann [9, Theorem 5.3(i)], taking the constants of that result to be $m = k = p_0 = \gamma_0 = 1$, $\kappa = s_0 = 0$, $\gamma_1 = 2 - \epsilon$. □

Solutions $u : [t_0, \tau) \times \overline{\Omega} \to \mathbb{R}^m$ to (6.8) enjoy considerable smoothness. For example, if the data $\partial\Omega_i, f_i, A_i, B_i$ are smooth of class $C^{2+2\epsilon}$, $0 < 2\epsilon < 1$, then $u \in C^{1+\epsilon, 2+2\epsilon}([t_1, t_2] \times \overline{\Omega}, \mathbb{R}^m)$ for all $t_0 < t_1 < t_2 < \tau$ (Lunardi [124, 7.3.3(iii)]).

While useful for many purposes, solution processes in the spaces X^α suffer from the drawback that X^α and its norm are defined implicitly, leaving unclear the domains of solutions and the meaning of convergence, stability, density and similar topological terms. In addition, the topology of X^α might be unsuitable for a given application. To overcome these difficulties we could appeal to results of Colombo and Vespri [29], Lunardi [124] and Mora [147], establishing induced processes in Banach spaces of continuous, smooth or L^p functions; or we can apply Theorem 6.2. We now define these spaces.

For $r \in \mathbb{N}$ let $C^r(\overline{\Omega})$ denotes the usual Banach space of C^r functions on $\overline{\Omega}$. Set

$$C_0^r\big(\overline{\Omega}\big) := \big\{v \in C^r\big(\overline{\Omega}\big) \colon v|\partial\Omega = 0\big\}.$$

With γ, ξ as in a Robin boundary operator and $r \geqslant 1$, define

$$C^r_{\gamma,\xi}\big(\overline{\Omega}\big) := \left\{ v \in C^r\big(\overline{\Omega}\big) \colon \gamma(x)v(x) + \frac{\partial v}{\partial \xi}(x) = 0 \ (x \in \partial\Omega) \right\}.$$

It is not hard to show that:

- $C^0(\overline{\Omega})$, $C^1(\overline{\Omega})$ and $C^1_{\gamma,\xi}(\overline{\Omega})$ are strongly ordered, with $u \gg 0$ if and only if $u(x) > 0$ for all $x \in \overline{\Omega}$;
- $C^1_0(\overline{\Omega})$ is strongly ordered, with $u \gg 0$ if and only if $u(x) > 0$ for all $x \in \Omega$ and $\partial u/\partial \nu > 0$ where $\nu : \partial\Omega \to \mathbb{R}^n$ is the unit vector field inwardly normal to $\partial\Omega$;
- $C^0_0(\overline{\Omega})$ is not strongly ordered. Both $C^0_0(\overline{\Omega})$ and $C^0(\overline{\Omega})$ are Banach lattices.

In terms of the boundary operators B_i, for $k = 0, 1$ we define Banach spaces

$$C^k_{B_i}\big(\overline{\Omega}\big) := \begin{cases} C^k_0\big(\overline{\Omega}\big) & \text{if } B_i \text{ is Dirichlet,} \\ C^k_{\gamma,\xi}\big(\overline{\Omega}\big) & \text{if } B_i \text{ is Robin and } k = 1, \\ C^0\big(\overline{\Omega}\big) & \text{if } B_i \text{ is Robin and } k = 0. \end{cases}$$

Note that $C^1_{B_i}(\overline{\Omega})$ is strongly ordered, while $C^0_{B_i}(\overline{\Omega})$ is strongly ordered if and only if B_i is Robin; $C^0_{B_i}(\overline{\Omega})$ is a Banach lattice. The ordered Banach space

$$C^k_B\big(\overline{\Omega}, \mathbb{R}^m\big) := \Pi_i C^k_{B_i}\big(\overline{\Omega}\big),$$

with the product order cone, is strongly ordered if $k = 1$, or $k = 0$ and no B_i is Dirichlet. The order cone $L^p(\Omega, \mathbb{R}^m)_+$ is the subset of $L^p(\Omega, \mathbb{R}^m)$ comprising equivalence classes represented by functions $\Omega \to \mathbb{R}^m_+$. Note that $L^p(\Omega, \mathbb{R}^m)$ is normally ordered but not strongly ordered.

It is known that the pair of operators (A_i, B_i) has a sectorial realization $\mathcal{A}_i$ on $C^k(B_i)$ and therefore the product operator $\mathcal{A}$ is sectorial on $C^k_B(\overline{\Omega}, \mathbb{R}^m)$. See Corollary 3.1.24, Theorems 3.1.25, 3.1.26 in [124].

LEMMA 6.10. *For $X = L^p(\Omega, \mathbb{R}^m)$ or $C^k_B(\overline{\Omega}, \mathbb{R}^m)$, the analytic semigroup $\mathrm{e}^{t\mathcal{A}}$ is a positive operator for $t \geqslant 0$ with respect to the cone of componentwise nonnegative functions in X.*

PROOF. As noted in Remark 6.7, it suffices to show that $(\lambda I - \mathcal{A})^{-1}$ is positive for large $\lambda > 0$, or equivalently, that for each i and $f_i \geqslant 0$, the solution $g_i \in D(\mathcal{A}_i)$ of $f_i = \lambda g_i - \mathcal{A}_i g_i$ satisfies $g_i \geqslant 0$. The existence of g_i is not the issue but rather it's positivity. Thus it boils down to $\lambda g_i - A_i g_i \geqslant 0 \Longrightarrow g_i \geqslant 0$. But these follow from standard maximum principle arguments. See Lemma 3.1.4 in [155]. □

With $X = L^p(\Omega, \mathbb{R}^m)$ and $\mathcal{A}$ and α as above, we have a chain of continuous inclusions of ordered Banach spaces

$$\mathsf{D}(\mathcal{A}) \hookrightarrow X^\alpha \hookrightarrow C^1_B\big(\overline{\Omega}, \mathbb{R}^m\big) \hookrightarrow C^0_B\big(\overline{\Omega}, \mathbb{R}^m\big) \hookrightarrow L^p\big(\Omega, \mathbb{R}^m\big),$$

with a solution process in $L^p(\Omega, \mathbb{R}^m)$ and an induced solution process in X^α.

PROPOSITION 6.11. *Let Θ be the solution process in $L^p(\Omega, \mathbb{R}^m)$ for Eq.* (6.1) *with $n < p < \infty$.*

(a) *For all $t > t_0$, Θ_{t,t_0} maps D_{t,t_0} continuously into $C^1_B(\overline{\Omega}, \mathbb{R}^m)$.*
(b) *Θ induces a solution process Θ^1 in $C^1_B(\overline{\Omega}, \mathbb{R}^m)$.*
(c) *Θ induces a solution process Θ^0 in $C^0_B(\overline{\Omega}, \mathbb{R}^m)$ provided $f = f(t, x, u)$.*

PROOF. By uniqueness of solutions it suffices to establish induced solution processes in $C^1_B(\overline{\Omega}, \mathbb{R}^m) \hookrightarrow L^p(\Omega, \mathbb{R}^m)$, and in $C^0_B(\overline{\Omega}, \mathbb{R}^m) \hookrightarrow L^p(\Omega, \mathbb{R}^m)$ when $f = f(t, x, u)$. This is done in Lunardi [124, Proposition 7.3.3] for $m = 1$, and the general case is similar. Part (c) follows from Theorem 6.2. □

Henceforth Θ^k, $k \in \{0, 1\}$, denotes the process Θ^0 or Θ^1 as in Proposition 6.11.

6.2.1. *Dynamics in spaces X_Γ* For any set $\Gamma \subset \mathbb{R}^m$ and $k = 0, 1$ define

$$\begin{aligned} X^k_\Gamma &:= \{u \in C^k_B(\overline{\Omega}, \mathbb{R}^m) \colon u(\overline{\Omega}) \subset \Gamma\}, \\ X_\Gamma &:= \{u \in L^p(\Omega, \mathbb{R}^m) \colon u(\Omega) \subset \Gamma\}. \end{aligned} \tag{6.12}$$

A *rectangle* in $\mathbb{R}^m$ is a set of the form $J = J_1 \times \cdots \times J_m$ where each $J_i \subset \mathbb{R}$ is a nondegenerate closed interval. $\mathbb{R}^m$, $\mathbb{R}^m_+$ and closed order intervals $[a, b]$, $a \leqslant b$ are rectangles.

PROPOSITION 6.12. *Let $J := \Pi_{i=1}^m J_i$ be a rectangle in $\mathbb{R}^m$ such that either $0 \in J_i$ or B_i is Neumann, and the following hold for all $x \in \overline{\Omega}$, $u \in \partial J$:*

$$f_i(t, x, u, 0) \geqslant 0 \quad \text{if } u_i = \inf J_i, \qquad f_i(t, x, u, 0) \leqslant 0 \quad \text{if } u_i = \sup J_i. \tag{6.13}$$

Then:

(i) *In the reaction–diffusion case, X_J is positively invariant for Θ and X^k_J is positively invariant for Θ^k ($k = 0, 1$).*
(ii) *Suppose $k = m = 1$ and $J \subset \mathbb{R}$ is an interval. Then X_J is positively invariant for Θ and X^1_J is positively invariant for Θ^1.*

PROOF. For the reaction–diffusion case we sketch a proof that X^0_J is Θ^0-positively invariant using Theorem 6.2. The proof that X_J is Θ-positively invariant follows from this since $\Theta_{t,t_0}(u)$ is the L^p limit $\lim_k \Theta^0_{t,t_0}(u_k)$ where $u_k \in X^0_J$ approximates $u \in X_J$ in L^p and the facts: $\Theta^0 = \Theta$ on X^0_J, a dense subset of the closed subset X_J. In order to verify the subtangential condition for X^0_J, it suffices to verify the subtangential condition for J:

$$\liminf_{h \searrow 0} \frac{1}{h} \operatorname{dist}\bigl(u + hf(t, x, u), J\bigr) = 0 \tag{6.14}$$

for each $(t, x, u) \in [0, \infty) \times \overline{\Omega} \times J$ by Martin [125, Proposition IX.1.1]. But (6.14) is a necessary condition for J to be positively invariant for the ODE

$$v' = f(t, x, v),$$

where x is a parameter. See, e.g., [125, Theorem VI.2.1]. It is well-known and easy to prove that condition (6.13) implies the positive invariance of J for the ODE (see, e.g., Proposition 3.3, Smith and Waltman [203, Proposition B.7], or Walter [227, Chapter II, Section 12, Theorem II]). It follows that (6.14) holds. Therefore the subtangential condition for X_J^0 holds. Finally, we must verify that $\mathrm{e}^{t\mathcal{A}} X_J^0 \subset X_J^0$ or, equivalently, that $\mathrm{e}^{t\mathcal{A}_i} C_{B_i}^0(\overline{\Omega}, J_i) \subset C_{B_i}^0(\overline{\Omega}, J_i)$. This follows from Remark 6.7 and standard maximum principle arguments. It also follows from standard comparison principles for parabolic equations. See, e.g., Pao [155, Lemma 2.1] or Smith [194, Corollary 2.4].

The case $k = m = 1$ is a special case of [227, Chapter IV, Section 25, Theorem II, Section 31, Corollaries IV and V]. □

Consider the case that (6.6) is autonomous:

$$\begin{aligned} &\frac{\partial u_i}{\partial t} = A_i u_i + f_i(x, u, \nabla u) && (x \in \overline{\Omega},\ t > t_0), \\ &B_i u_i = 0 && (x \in \partial\Omega,\ t > t_0), \end{aligned} \tag{6.15}$$

$i = 1, \ldots, m$. The solution processes Θ, Θ^1, Θ^0 reduce to local semiflows.

We introduce a mild growth condition, trivially satisfied in the reaction–diffusion case:

$$\begin{aligned} &\text{For each } s > 0 \text{ there exists } C(s) > 0 \text{ such that} \\ &|v| \leqslant s \quad \Longrightarrow \quad \bigl|f(x, v, \xi)\bigr| \leqslant C(s)\bigl(1 + |\xi|^{2-\epsilon}\bigr). \end{aligned} \tag{6.16}$$

The following result gives sufficient conditions for solution processes in X_Γ to be global, and to admit compact global attractors:

PROPOSITION 6.13. *Assume system* (6.15) *satisfies* (6.16). *Let $\Gamma \subset \mathbb{R}^m$ be a nonempty compact set such that X_Γ is positively invariant for* (6.15). *Then*:

(a) *There are solution semiflows Φ, Φ^1 in X_Γ, X_Γ^1 respectively. Φ^1 is compact.*

(b) *Assume* (6.15) *is reaction–diffusion. Then there is also a solution semiflow Φ^0 in X_Γ^0. The semiflows Φ, Φ^0, Φ^1 are compact and order compact. There is a compact set $K \subset X_\Gamma^1$ which is the global attractor for all three semiflows.*

PROOF. (a) Let Γ lie in the open ball of radius $R > 0$ about the origin in $\mathbb{R}^m$ and let $h : \mathbb{R}^m \to \mathbb{R}^m$ be any smooth bounded function that agrees with the identity on the open ball of radius R. Define g by $g(x, v, \xi) = f(x, h(v), \xi)$. Every trajectory in X_Γ of (6.15) is also a trajectory of the analogous system in which f is replaced by g (compare Poláčik [163, pp. 842–843]). Nonlinearity g satisfies (6.16) with $C(s)$ constant so (6.10) holds. As

$$\limsup_{t \to \tau -} \bigl\| u(t) \bigr\|_{C^0(\overline{\Omega}, \mathbb{R}^m)} \leqslant R,$$

which implies (6.11), all trajectories are global by Proposition 6.9. Thus, the restrictions of Ψ, Ψ^1 in X and $C^1_B(\overline{\Omega}, \mathbb{R}^m)$ respectively to X_Γ and X^1_Γ define semiflows Φ and Φ^1. As Ψ^1 is compact by Hale [58], Theorem 4.2.2, Φ^1 is compact because X^1_Γ is closed in $C^1_B(\overline{\Omega}, \mathbb{R}^m)$.

(b) In the reaction–diffusion case a similar argument establishes a compact solution semiflow Φ^0 in X^0_Γ; and Φ^0 is order compact because order intervals in X^0_Γ are bounded. To prove Φ^0 order compact, let N be an order interval in X_Γ. For every $t = 2s > 0$, Φ_s maps N continuously into an order interval N' of X^0_Γ. Precompactness in X_Γ of $\Phi_t N$ follows from the precompactness in X^0_Γ of $\Phi^0_s N'$, already established, and the continuous inclusion $\Phi_t N = \Phi^0_s \circ \Phi_s N \subset \Phi^0_s N'$.

To prove order compactness of Φ^1, let $N_1 \subset X^1_\Gamma$ be an order interval. N_1 is contained in an order interval N_0 of X^0_Γ. Let C^k denote closure in X^k_Γ. For all $t > 0$ we have $\mathsf{C}^1(\Phi^1_t N_1) = \mathsf{C}^1(\Phi^0_t N_1) \subset \mathsf{C}^1\mathsf{C}^0(\Phi^0_t N_0)$, and the latter set is compact because Φ^0 is order compact. This proves $\Phi^1_t N_1$ is precompact in X^1_Γ.

X^0_Γ is closed and bounded in X^0, hence $\Phi^0_t X^0_\Gamma$ is precompact in X^0_Γ for all $t > 0$ by (a). Therefore $K := \bigcap_{t>0} \overline{\Phi^0_t X^0_\Gamma}$ is a compact global attractor for Φ^0. Similarly, K (with the same topology) is a compact global attractor for Φ.

We rely on the identity $\Phi^1_t = \Phi^0_t | X^1_\Gamma$ and continuity of $\Phi^0_t : X^0_\Gamma \to X^1_\Gamma$ for all $t > 0$. As K is invariant under Φ^0, it follows that K is a compact subset of X^1_Γ. To prove K a global attractor for Φ^1, it suffices to prove: For arbitrary sequences $\{x(i)\}$ in X^1_Γ, and $t(i) \to \infty$ in $\mathbb{R}_+$ with $t(i) > \epsilon > 0$, there is a sequence $i_k \to \infty$ in $\mathbb{N}$ such that $\{\Phi^1_{t(i_k)} x(i_k)\}$ converges in X^1_Γ to a point of K. Choose $\{i_k\}$ so that $\Phi^0_{t(i_k)-\epsilon}(i_k)$ converges X^0_Γ as $k \to \infty$ to $p \in K$; this is possible because K is a compact global attractor for Φ^0. Then $\Phi^1_{t(i_k)} x(i_k) = \Phi^0_\epsilon \circ \Phi^0_{t(i_k)-\epsilon} x(i_k)$, which converges in X^1_Γ as $k \to \infty$ to $\Phi^0_\epsilon p \in K$. □

EXAMPLE. Let the u_i denote the concentrations or densities of entities such as chemicals or species. Such quantities are inherently positive, so taking the state space to be $L^p(\Omega, \mathbb{R}^m_+)$ or $C^k_B(\overline{\Omega}, \mathbb{R}^m_+)$ is appropriate. We make the plausible assumption that sufficiently high density levels must decrease. Modeling this situation by (a) and (b) below, we get the following result.

PROPOSITION 6.14. *In Eq.* (6.15) *assume* $f = f(x, u)$ *and let the following hold for* $i = 1, \ldots, m$:

(a) $f_i(x, u) \geqslant 0$ *if* $u_i = 0$;

(b) *there exists* $\kappa > 0$ *such that* $f_i(x, u) < 0$ *if* $u_i \geqslant \kappa$.

Then for $k = 0, 1$ *solution processes in the order cones* $L^p(\Omega, \mathbb{R}^m_+)$, $C^k_B(\overline{\Omega}, \mathbb{R}^m_+)$ *are defined by semiflows* Φ, Φ^k *respectively*; *and there is a compact set* $K \subset X^k_{[0,\kappa]^m}$ *that is the global attractor for* Φ, Φ^0 *and* Φ^1.

PROOF. Proposition 6.12 and (a) proves $L^p(\Omega, \mathbb{R}^m_+)$ and $C^k_B(\overline{\Omega}, \mathbb{R}^m_+)$ are positively invariant under the solution process.

Consider the compact rectangles $J(c) := [0, c\kappa]^m \subset \mathbb{R}^m$, $c \geqslant 1$. Assumption (b) and Proposition 6.12 entail positive invariance of $X_{J(c)}$. Proposition 6.13 shows that there are solution semiflows in $X_{J(c)}$ and $X^k_{J(c)}$ having a compact global attractor $K_c \subset X^1_{J(c)}$ in common. As the $J(c)$ are nested and exhaust $\mathbb{R}^m_+$, these semiflows come from solution semiflows Φ, Φ^k as required. Moreover, all the attractors K_c coincide with the compact set $K := K_1 \subset X^1_{J(1)}$. It is easy to see that K is the required global attractor. □

Results on global solutions and positively invariant sets can be found in many places. See for example Amann [9,10], Cholewa and Dlotko [25], Cosner [33], Lunardi [124], Poláčik [163], Smith [194], Smoller [205].

6.2.2. *Monotone solution processes for parabolic equations* We restrict attention here to monotonicity properties with respect to the standard point-wise and component-wise ordering of functions $\Omega \to \mathbb{R}^m$: $f \leqslant g$ if and only if $f_i(x) \leqslant g_i(x)$ for all x and all i. The natural ordering on $L^p(\Omega, \mathbb{R}^m)$ is defined on equivalence classes by the condition on representatives that $f_i(x) \leqslant g_i(x)$ almost everywhere.

Orderings induced by orthants in $\mathbb{R}^n$ other than the positive orthant can be handled easily by change of variables. See Mincheva [144] and [145] for results in the case of polyhedral cones in $\mathbb{R}^n$.

Consider the case $m = 1$ in Eq. (6.8).

THEOREM 6.15. *In Eq.* (6.8), *assume $m = 1$ and f is C^1. Then*:

(i) *Θ is VSOP on $L^p(\Omega, \mathbb{R}^m)$.*

(ii) *Θ^1 is strongly monotone in $C^1_B(\overline{\Omega})$.*

(iii) *If $f = f(t, x, u)$ the induced process Θ^0 on $C^0_B(\overline{\Omega})$ is VSOP, and strongly monotone if all boundary operators are Robin.*

PROOF. Let $u, v : [t_0, t_1] \times \overline{\Omega} \to \mathbb{R}$ be solutions with $v(t_0, x) - u(t_0, x) \geqslant 0$ for all x and > 0 for some x. Then $w := v - u$ is the solution to the problem

$$\begin{aligned}
&\frac{\partial w}{\partial t} = Aw + \sum_{j=1}^{n} b_j \frac{\partial w}{\partial x_j} + cw \quad && (x \in \overline{\Omega},\ t > t_0),\\
&Bw(t, x) = 0 && (x \in \partial\Omega,\ t > t_0),\\
&w(t_0, x) \geqslant 0, \quad w(t_0, x) \not\equiv 0 && (x \in \overline{\Omega})
\end{aligned} \tag{6.17}$$

where $b_j = b_j(t, x)$ and $c_j = c_j(t, x)$ are obtained as follows. Evaluate u, v and their spatial gradients at (t, x), and for $s \in [0, 1]$ set

$$Z(s) = (1 - s)(t, x, u, \nabla u) + s(t, x, v, \nabla v),$$

$$b(t, x) = \big(b_1(t, x), \ldots, b_n(t, x)\big) = \int_0^1 D_4 f\big(Z(s)\big)\,\mathrm{d}s,$$

$$c(t, x) = \int_0^1 D_3 f\big(Z(s)\big)\,\mathrm{d}s$$

where $D_4 f$ and $D_3 f$ denote respectively the derivatives of $f(t,x,y,\xi)$ with respect to $\xi \in \mathbb{R}^n$ and $y \in \mathbb{R}$. By Taylor's theorem

$$f(t,x,v,\nabla v) - f(t,x,u,\nabla u) = b(t,x)(\nabla u - \nabla v) + c(t,x)(u-v),$$

whence (6.17) follows.

The parabolic maximum principle and boundary point lemma ([194, Theorems 7.2.1, 7.2.2]) imply that the function $w(t_1,\cdot)$, considered as an element of $C^1_B(\overline{\Omega})$, is $\gg 0$. This proves (ii), and the first assertion of (iii) follows from Theorem 6.4(b). The proof of strong monotonicity for Robin boundary conditions is similar to the arguments given above. Part (i) follows from strong monotonicity of Θ^1, Theorem 6.4 and continuity of $\Theta_{t,t_0} : L^p(\Omega,\mathbb{R}^m) \to X^\alpha \hookrightarrow C^1_B(\overline{\Omega},\mathbb{R}^m)$. □

For $m \geqslant 2$ we impose further conditions on system (6.6) in order to have a monotone solution process: it must be of reaction–diffusion type, and the vector fields $f(t,x,\cdot)$ on $\mathbb{R}^m$ must be *cooperative*. In other words, $f(t,x,u)$ is C^1 in u and $\partial f_i/\partial u_j \geqslant 0$ for all $i \neq j$. (The latter condition holds vacuously if $m = 1$). When this holds then the system is called cooperative. If in addition, there exists $\bar{x} \in \Omega$ such that the $m \times m$ Jacobian matrix $[\partial f_i/\partial u_j(t,\bar{x},u)]$ is irreducible for all (t,u), we call the system *cooperative and irreducible*

THEOREM 6.16. *If system* (6.15) *is cooperative, then* Θ, Θ^k, $k = 0,1$ *are monotone. If the system is also irreducible, then*:

(i) *Θ is VSOP on $L^p(\Omega,\mathbb{R}^m)$.*
(ii) *Θ^1 is strongly monotone in $C^1_B(\overline{\Omega},\mathbb{R}^m)$.*
(iii) *Θ^0 is VSOP in $C^0_B(\overline{\Omega},\mathbb{R}^m)$ and is strongly monotone when all boundary operators are Robin.*

PROOF. Monotonicity in $C^0_B(\overline{\Omega},\mathbb{R}^m)$ follows directly from Theorem 6.5 and Remark 6.6. Indeed, let $u \leqslant v$ in $C^0_B(\overline{\Omega},\mathbb{R}^m)$ and t be fixed. Then

$$\begin{aligned}&\big[F(t,v) - F(t,u) + \lambda(v-u)\big](x)\\ &\quad = \int_0^1 \left(\frac{\partial f}{\partial u}\big(t,x,su(x)+(1-s)v(x)\big) + \lambda I\right) \mathrm{d}s\,(v-u)(x) \geqslant 0\end{aligned}$$

for some $\lambda > 0$ and all $x \in \overline{\Omega}$ by cooperativity of f and compactness of $\overline{\Omega}$. This implies that (QM) holds. The positivity of $\mathrm{e}^{t\mathcal{A}}$ follows from Lemma 6.10. Monotonicity of Θ in $L^p(\Omega,\mathbb{R}^m)$ follows from monotonicity of Θ^0 and Theorem 6.4.

The proof of VSOP and strong monotonicity for Robin boundary conditions in $C^0_B(\overline{\Omega},\mathbb{R}^m)$ is like that of Theorem 6.15(i), exploiting the maximum principle for weakly coupled parabolic systems (Protter and Weinberger [166, Chapter 3, Theorems 13, 14, 15 and pp. 192, Remark (i)]). See Smith [194, Section 7.4] for a similar proof.

Monotonicity of Θ^1 follows from monotonicity of Θ^0. Strong monotonicity of Θ^1, in the case of Dirichlet boundary conditions, requires exploiting the maximum principle as in the previous case (the same references apply). VSOP of Θ follows from strong monotonicity of Θ^1, Theorem 6.4 and continuity of the composition $\Theta_{t,t_0} : L^p(\Omega, \mathbb{R}^m) \to X^\alpha \to C^1_B(\overline{\Omega}, \mathbb{R}^m)$. □

6.3. *Parabolic systems with monotone dynamics*

We now treat autonomous systems (6.15) having monotone dynamics. Our goal is Theorem 6.17, a sample of the convergence and stability results derivable from the general theory.

In addition to the assumptions for (6.6), we require the following conditions to hold for the solution process Θ in $X := L^p(\Omega, \mathbb{R}^m)$, with p satisfying (6.9) and X^k_Γ defined in (6.12):

(SP) If $m \geqslant 2$ in system (6.15) then $f = f(x,u)$ and the system is cooperative and irreducible. $\Gamma \subset \mathbb{R}^m$ is a nonempty set, either an open set or the closure of an open set. The solution process induces semiflows Φ, Φ^1 in X_Γ, X^1_Γ respectively, and Φ^0 in X^0_Γ for the reaction–diffusion case. These semiflows are assumed to have compact orbit closures.

Simple conditions implying (SP) can be derived from Propositions 6.13.

The following statements follow from (SP), assertions about Φ^0 having the implied hypothesis $f = f(x,u)$:

- X^1_Γ is dense in X^0_Γ and in X_Γ.
- Φ and Φ^0 agree on X^0_Γ, and Φ, Φ^0 and Φ^1 agree on X^1_Γ.
- Φ_t (respectively, Φ^0_t) maps X_Γ (respectively, X^0_Γ) continuously into X^1_Γ for $t > 0$ (Proposition 6.11).
- Φ, Φ^1 and Φ^0 have the same omega limit sets, compact attractors and equilibria.
- If Γ is open or order convex and $f(x,u,\xi)$ is C^1 in (u,ξ), the Improved Limit Set Dichotomy (ILSD) holds for Φ^1 by Theorem 2.16, and for Φ and Φ^0 by Proposition 2.21.
- If Γ is compact then Φ^1 is compact. In the reaction–diffusion case with Γ compact, Φ, Φ^1 and Φ^0 are compact and order compact, and a common compact global attractor (Propositions 6.13).
- Φ^1 is strongly monotone; Φ^0 is VSOP, and strongly monotone if all boundary operators are Robin; Φ is VSOP (Theorem 6.16).

The sets of quasiconvergent, convergent and stable points for any semiflow Ψ are denoted respectively by $Q(\Psi), C(\Psi), S(\Psi)$. References to intrinsic or extrinsic topology of these sets (e.g., closure, density) for Φ, Φ^1 or Φ^0 are to be interpreted in terms of the topology of the corresponding domain X_Γ, X^1_Γ or X^0_Γ.

THEOREM 6.17. *If system* (6.15) *satisfies hypothesis* (SP), *then*:

(i) *The sets $Q(\Phi)$, $Q(\Phi^0)$ and $Q(\Phi^1)$ are residual.*

(ii) *Assume Γ is open or order convex and $f(x,u,\xi)$ is C^1 in (u,ξ). Then the sets $C(\Phi) \cap S(\Phi)$, $C(\Phi^0) \cap S(\Phi^0)$ and $C(\Phi^1)$ have dense interiors.*

(iii) *Assume $f = f(x, u)$ and Γ is compact. Then the semiflows Φ, Φ^0, Φ^1 are compact and order compact, and they have a compact global attractor in common.*

(iv) *Assume Γ is open or order convex and E_Γ is compact. Then some $p \in E_\Gamma$ is stable for Φ. Every such p is also stable for Φ^1, and for Φ^0 in the reaction–diffusion case. When E_Γ is finite, the same holds for asymptotically stable equilibria.*

PROOF. (i) follows from Theorem 1.21.

(ii) for Φ and Φ^0 follows from Theorem 2.25(b). For Φ^1, (ii) follows from Theorem 2.26(a).

(iii) is a special case of Proposition 6.13(b).

In (iv), to find a $p \in E_\Gamma$ having the asserted stability properties for Φ, it suffices to verify the hypotheses of Theorem 1.30: (a) follows from (i), while (b) and (c) holds by the assumptions on Γ and compactness of E. Similarly for Φ^0 in the reaction–diffusion case.

To prove the stability properties for p under Φ^1, it suffices by Theorem 1.31 to show that p has a neighborhood in X^1_Γ that is attracted to a compact set. By (i) and the assumptions on Γ, there are sequences $\{u_k\}$, $\{v_k\}$ in $Q(\Phi^1)$ converging to p in X_Γ, such that

$$u_k \leqslant u_{k+1} \leqslant p \leqslant v_{k+1} \leqslant v_k$$

and

$$p \neq \inf X_\Gamma \quad \Longrightarrow \quad u_k < u_{k+1} < p, \qquad p \neq \sup X_\Gamma \quad \Longrightarrow \quad p < v_{k+1} < v_k.$$

Replacing u_k, v_k by their images under Φ_{ϵ_k} for sufficiently small $\epsilon_k > 0$, we see from strong monotonicity of Φ^1 that we can assume:

$$p \neq \inf X_\Gamma \quad \Longrightarrow \quad u_k \ll u_{k+1} \ll p, \qquad p \neq \sup X_\Gamma \quad \Longrightarrow \quad p \ll v_{k+1} \ll v_k.$$

The sets $N_k := [[u_k, v_k]]_X \cap X_\Gamma$ are positively invariant and form a neighborhood basis at p in X_Γ.

Fix k_0 such that N_k is bounded in X^1_Γ for all $k \geqslant k_0$. By Theorem 6.1(iii), for every $s > 0$ there exists $j \geqslant k_0$ such that $\Phi_s(N_j)$ is precompact in X^α, hence in X^1_Γ. Fix such numbers s and j and let P denote the closure of $\Phi_s(N_j)$ in X^1_Γ. Being compact and positively invariant, P contains the compact global attractor $K := \bigcap_{t>0} \overline{\Phi_t P}$ for the semiflow in $\Phi^1|P$. Then N_j is a neighborhood p in X^1_Γ that is attracted under Φ^1 to K. □

References

[1] F. Albrecht, H. Gatzke, A. Haddad and N. Wax, *The dynamics of two interacting populations*, J. Math. Anal. Appl. **46** (1974), 658–670.

[2] A. Alexandroff and H. Hopf, *Topologie*, Springer, Berlin (1935).

[3] N. Alikakos, P. Hess and H. Matano, *Discrete order-preserving semigroups and stability for periodic parabolic differential equations*, *J. Differential Equations* **82** (1989), 322–341.

[4] H. Amann, *On the existence of positive solutions of nonlinear elliptic boundary value problems*, Indiana Univ. Math. J. **21** (1971), 125–146.

[5] H. Amann, *On the number of solutions of nonlinear equations in ordered Banach spaces*, J. Funct. Anal. **11** (1972), 346–384.

[6] H. Amann, *Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces*, SIAM Rev. **18** (1976), 620–709.

[7] H. Amann, *Invariant sets and existence theorems for semilinear parabolic and elliptic systems*, Math. Anal. Appl. **65** (1978), 432–467.

[8] H. Amann, *Existence and regularity for semilinear parabolic evolution equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **11** (1984), 593–676.

[9] H. Amann, *Global existence for semilinear parabolic systems*, J. Reine Angew. Math. **360** (1985), 47–83.

[10] H. Amann, *Quasilinear evolution equations and parabolic systems*, Trans. Amer. Math. Soc. **293** (1986), 191–227.

[11] H. Amann, *Linear and Quasilinear Parabolic Problems, Vol. 1: Abstract Linear Theory*, Monogr. Math., Vol. 89, Birkhäuser, Boston, MA (1995), xxxvi + 335 pp.

[12] S. Angenent, *Nonlinear analytic semiflows*, Proc. Roy. Soc. Edinburgh Sect. A **115** (1990), 91–107.

[13] C. Batty and C. Robinson, *Positive one-parameter semigroups on ordered spaces*, Acta Appl. Math. **2** (1984), 221–296.

[14] M. Benaïm, *On invariant hypersurfaces of strongly monotone maps*, J. Differential Equations **137** (1997), 302–319.

[15] M. Benaïm, *Convergence with probability one of stochastic approximation algorithms whose average is cooperative*, Nonlinearity **13** (2000), 601–616.

[16] M. Benaïm and M. Hirsch, *Adaptive processes, mixed equilibria and dynamical systems arising from fictitious plays in repeated games*, Games Econom. Behav. **29** (1999), 36–72.

[17] M. Benaïm and M. Hirsch, *Stochastic approximation with constant step size whose average is cooperative*, Ann. Appl. Probab. **9** (1999), 216–241.

[18] A. Berman, M. Neumann and R. Stern, *Nonnegative Matrices in Dynamic Systems*, Wiley, New York (1989).

[19] A. Berman and R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York (1979).

[20] G. Birkhoff, *Extensions of Jentzsch's theorem*, Trans. Amer. Math. Soc. **85** (1957), 219–227.

[21] P. Brunovsky, *Controlling nonuniqueness of local invariant manifolds*, J. Reine Angew. Math. **446** (1994), 115–135.

[22] P. Brunovsky, P. Poláčik and B. Sandstede, *Convergence in general periodic parabolic equations in one space variable*, Nonlinear Anal. **18** (1992), 209–215.

[23] X.-Y. Chen and H. Matano, *Convergence, asymptotic periodicity, and finite-point blow-up in one-dimensional semilinear heat equations*, J. Differential Equations **78** (1989), 160–190.

[24] M. Chen, X. Chen and J.K. Hale, *Structural stability for time-periodic one-dimensional parabolic equations*, J. Differential Equations **96** (1992), 355–418.

[25] J. Cholewa and T. Dlotko, *Global Attractors in Abstract Parabolic Problems*, London Math. Soc. Lecture Note Ser., Vol. 278, Cambridge Univ. Press, Cambridge (2000).

[26] S.N. Chow and J.K. Hale, *Methods of Bifurcation Theory*, Springer, New York (1982).

[27] Ph. Clements, H. Heijmans, S. Angenent, C. van Duijn and B. de Pagter, *One-Parameter Semigroups*, CWI Monographs, Vol. 5, North-Holland, Amsterdam (1987).

[28] E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw–Hill, New York (1955).

[29] F. Colombo and V. Vespri, *Generation of analytic semigroups in* $W^{k,p}$, Differential Integral Equations **9** (1996), 421–436.

[30] C. Conley, *Isolated Invariant Sets and the Morse Index*, CBMS Reg. Conf. Ser. Math., Vol. 38, Amer. Math. Soc., Providence, RI (1978).

[31] C. Conley, *The gradient structure of a flow:* I, IBM Research, RC 3939 (1972) (17806), Yorktown Heights, NY. Also, Ergodic Theory Dynam. Systems **8** (1988), 11–26.

[32] W. Coppel, *Stability and Asymptotic Behavior of Differential Equations*, Heath, Boston (1965).

[33] C. Cosner, *Pointwise a priori bounds for strongly coupled semilinear systems of parabolic partial differential equations*, Indiana Univ. Math. J. **30** (1981), 607–620.

[34] R. Courant and D. Hilbert, *Methoden der mathematischen Physik, Bd. 2*, Springer, Berlin (1937).

[35] C.M. Dafermos and M. Slemrod, *Asymptotic behaviour of nonlinear contractions semigroups*, J. Funct. Anal. **13** (1973), 97–106.

[36] E.N. Dancer, *Upper and lower stability and index theory for positive mappings and applications*, Nonlinear Anal. **17** (1991), 205–217.

[37] E. Dancer, *Some remarks on a boundedness assumption for monotone dynamical systems*, Proc. Amer. Math. Soc. **126** (1998), 801–807.

[38] E. Dancer and P. Hess, *Stability of fixed points for order-preserving discrete-time dynamical systems*, J. Reine Angew. Math. **419** (1991), 125–139.

[39] D.L. DeAngelis, W.M. Post and C.C. Travis, *Positive Feedback in Natural Systems*, Springer, Berlin (1986).

[40] K. Deimling, *Nonlinear Functional Analysis*, Springer, New York (1985).

[41] P. de Leenheer and H.L. Smith, *Virus dynamics: A global analysis*, SIAM J. Appl. Math. (2003).

[42] P. de Mottoni and A. Schiaffino, *Competition systems with periodic coefficients: A geometric approach*, J. Math. Biol. **11** (1982), 319–335.

[43] A. Eden, C. Foias, B. Nicolaenko and R. Temam, *Exponential Attractors for Dissipative Evolution Equations*, Wiley, New York (2000).

[44] L. Elsner, *Quasimonotonie und Ungleichungen in halbgeordneten Raumen*, Linear Algebra Appl. **8** (1974), 249–261.

[45] J. Fleckinger-Pellé and P. Takáč, *Uniqueness of positive solutions for nonlinear cooperative systems with the p-Laplacian*, Indiana Univ. Math. J. **43** (1994), 1227–1253.

[46] J. Fleckinger-Pellé and P. Takáč, *Unicité de la solution d'un système non linéaire strictement coopératif*, C. R. Acad. Sci. Paris Sér. I Math. **319** (1994), 447–450.

[47] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice Hall, Englewood Cliffs, NJ (1964).

[48] A. Friedman, *Remarks on nonlinear parabolic equations*, Proc. Sympos. Applied Math., Vol. XVII, *Application of Nonlinear Partial Differential Equations in Mathematical Physics*, Amer. Math. Soc. (1965).

[49] P. Fife, *Mathematical Aspects of Reacting and Diffusing Systems*, Lecture Notes in Math., Vol. 28, Springer, New York (1979).

[50] P. Fife and M. Tang, *Comparison principles for reaction–diffusion systems: Irregular comparison functions and applications to questions of stability and speed of propagation of disturbances*, J. Differential Equations **40** (1981), 168–185.

[51] G. Frobenius, *Über aus nicht negativen Elementen*, Sitzungsber. Königl. Preuss. Akad. Wiss. (1912), 456–477.

[52] G. Greiner, *Zur Perron–Frobenius Theorie stark stetiger Halbgruppen*, Math. Z. **177** (1981), 401–423.

[53] S. Grossberg, *Competition, decision, and consensus*, J. Math. Anal. Appl. **66** (1978), 470–493.

[54] S. Grossberg, *Biological competition: Decision rules, pattern formation, and oscillations*, Proc. Nat. Acad. Sci. **77** (1980), 2338–2342.

[55] K. Hadeler and D. Glas, *Quasimonotone systems and convergence to equilibrium in a population genetics model*, J. Math. Anal. Appl. **95** (1983), 297–303.

[56] J.K. Hale, *Theory of Functional Differential Equations*, Springer, New York (1977).

[57] J.K. Hale, *Ordinary Differential Equations*, Krieger, Malabar, FL (1980).

[58] J.K. Hale, *Asymptotic Behavior of Dissipative Systems*, Amer. Math. Soc., Providence (1988).

[59] J.K. Hale, *Dynamics of a scalar parabolic equation*, Canad. Appl. Math. J. **5** (1997), 209–306.

[60] J.K. Hale and A.S. Somolinos, *Competition for fluctuating nutrient*, J. Math. Biol. **18** (1983), 255–280.

[61] J.K. Hale and S.M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Springer, New York (1993).

[62] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Math., Vol. 840, Springer, New York (1981).

[63] P. Hess, *Periodic-parabolic Boundary Value Problems and Positivity*, Longman Scientific and Technical, New York (1991).

[64] P. Hess and A.C. Lazer, *On an abstract competition model and applications*, Nonlinear Anal. **16** (1991), 917–940.

[65] P. Hess and P. Poláčik, *Boundedness of prime periods of stable cycles and convergence to fixed points in discrete monotone dynamical systems*, SIAM J. Math. Anal. **24** (1993), 1312–1330.

[66] M.W. Hirsch, *Convergence in ordinary and partial differential equations*, Notes for Colloquium Lectures at University of Toronto, August 23–26, 1982, Amer. Math. Soc., Providence, RI (1982).

[67] M.W. Hirsch, *Systems of differential equations which are competitive or cooperative. I: Limit sets*, SIAM J. Appl. Math. **13** (1982), 167–179.

[68] M.W. Hirsch, *Differential equations and convergence almost everywhere in strongly monotone semiflows*, Contemp. Math. **17** (1983), 267–285.

[69] M.W. Hirsch, *The dynamical systems approach to differential equations*, *Bull. Amer. Math. Soc.* **11** (1984), 1–64.

[70] M.W. Hirsch, *Systems of differential equations which are competitive or cooperative. II: Convergence almost everywhere*, SIAM J. Math. Anal. **16** (1985), 423–439.

[71] M.W. Hirsch, *Attractors for discrete-time monotone dynamical systems in strongly ordered spaces*, Geometry and Topology, J. Alexander and J. Harer, eds., Lecture Notes in Math., Vol. 1167, Springer, New York (1985), 141–153.

[72] M.W. Hirsch, *Systems of differential equations which are competitive or cooperative. III: Competing species*, Nonlinearity **1** (1988), 51–71.

[73] M.W. Hirsch, *Stability and convergence in strongly monotone dynamical systems*, J. Reine Angew. Math. **383** (1988), 1–53.

[74] M.W. Hirsch, *Convergent activation dynamics in continuous time neural networks*, Neural Networks **2** (1989).

[75] M.W. Hirsch, *Systems of differential equations that are competitive or cooperative. V: Convergence in 3-dimensional systems*, J. Differential Equations **80** (1989), 94–106.

[76] M.W. Hirsch, *Systems of differential equations that are competitive or cooperative. IV: Structural stability in three dimensional systems*, SIAM J. Math. Anal. **21** (1990), 1225–1234.

[77] M.W. Hirsch, *Systems of differential equations that are competitive or cooperative. VI: A local C^r closing lemma for 3-dimensional systems*, Ergodic Theory Dynam. Systems **11** (1991), 443–454.

[78] M.W. Hirsch, *Fixed points of monotone maps*, J. Differential Equations **123** (1995), 171–179.

[79] M.W. Hirsch, *Chain transitive sets for smooth strongly monotone dynamical systems*, Differential Equations Dynamical Systems **5** (1999), 529–543.

[80] M.W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, Acad. Press, New York (1974).

[81] J. Hofbauer and W. Sandholm, *On the global convergence of stochastic fictitious play*, Econometrica **70** (2002), 2265–2294.

[82] J. Hofbauer and J.W.-H. So, *Multiple limit cycles for three-dimensional competitive Lotka–Volterra equations*, Appl. Math. Lett. **7** (1994), 65–70.

[83] S.-B. Hsu, H. Smith and P. Waltman, *Competitive exclusion and coexistence for competitive systems on ordered Banach spaces*, Trans. Amer. Math. Soc. **348** (1996), 4083–4094.

[84] S.-B. Hsu and P. Waltman, *Analysis of a model of two competitors in a chemostat with an external inhibitor*, SIAM J. Applied Math. **52** (1992), 528–540.

[85] R. Jentzsch, *Über Integralgleichungen mit positiven Kern*, J. Reine Angew. Math. **141** (1912), 235–244.

[86] J. Jiang, *Attractors for strongly monotone flows*, J. Math. Anal. Appl. **162** (1991), 210–222.

[87] J. Jiang, *On the existence and uniqueness of connecting orbits for cooperative systems*, Acta Math. Sinica (N.S.) **8** (1992), 184–188.

[88] J. Jiang, *Three- and four-dimensional cooperative systems with every equilibrium stable*, J. Math. Anal. Appl. **18** (1994), 92–100.

[89] J. Jiang, *Five-dimensional cooperative systems with every equilibrium stable*, Differential Equations and Control Theory, Wuhan, 1994, Lecture Notes in Pure and Appl. Math., Vol. 176, Dekker, New York (1996), 121–127.

[90] J. Jiang and S. Yu, *Stable cycles for attractors of strongly monotone discrete-time dynamical systems*, J. Math. Anal. Appl. **202** (1996), 349–62.

[91] E. Kamke, *Zur Theorie der Systeme gewöhnlicher Differentialgleichungen II*, Acta Math. **58** (1932), 57–85.

[92] T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin (1976).

[93] H.B. Keller, *Elliptic boundary value problems suggested by nonlinear diffusion processes*, Arch. Ration Mech. Anal **35** (1969), 363–381.

[94] H.B. Keller, *Positive solutions of some nonlinear eigenvalue problems*, J. Math. Mech. **19** (1969), 279–296.

[95] H.B. Keller and D.S. Cohen, *Some positone problems suggested by nonlinear heat generation*, J. Math. Mech. **16** (1967), 327–342.

[96] W. Kerscher and R. Nagel, *Asymptotic behavior of one-parameter semigroups of positive operators*, Acta Appl. Math. **2** (1984), 297–309.

[97] A. Kolmogorov, *Sulla teoria di Volterra della lotta per l'esistenza*, Giorno. Ist. Ital. Attuari **7** (1936), 74–80.

[98] M. Krasnosel'skii, *Positive Solutions of Operator Equations*, Groningen, Noordhoff (1964).

[99] M. Krasnosel'skii, *The Operator of Translation along Trajectories of Differential Equations*, Transl. Math. Monogr., Vol. 19, Amer. Math. Soc., Providence (1968).

[100] M.A. Krasnosel'skii and L. Ladyzhenskaya, *The structure of the spectrum of positive nonhomogeneous operators*, Trudy Moskov. Mat. Obshch. **3** (1954), 321–346 (in Russian).

[101] M.A. Krasnosel'skii and P.P. Zabreiko, *Geometric Methods of Nonlinear Analysis*, Springer, New York (1984).

[102] U. Krause and R.D. Nussbaum, *A limit set trichotomy for self-mappings of normal cones in Banach spaces*, Nonlinear Anal. **20** (1993), 855–870.

[103] U. Krause and P. Ranft, *A limit set trichotomy for monotone nonlinear dynamical systems*, Nonlinear Anal. **19** (1992), 375–392.

[104] M.G. Krein and M.A. Rutman, *Linear operators leaving invariant a cone in a Banach space*, Uspekhi Mat. Nauk **3** (1948), 3–95. AMS Translation No. 26, Amer. Math. Soc., Providence, RI (1950).

[105] T. Krisztin, H.-O. Walther and J. Wu, *Shape, Smoothness and Invariant Stratification of an Attracting Set for Delayed Monotone Positive Feedback*, Fields Inst. Monogr., Vol. 11, Amer. Math. Soc., Providence, RI (1999).

[106] T. Krisztin and J. Wu, *Monotone semiflows generated by neutral equations with different delays in neutral and retarded parts*, Acta Math. Univ. Comenian. (N.S.) **63** (1994), 207–220.

[107] T. Krisztin and J. Wu, *Asymptotic behaviors of solutions of scalar neutral functional differential equations*, Differential Equations Dynam. Systems **4** (1996), 351–366.

[108] T. Krisztin and J. Wu, *Asymptotic periodicity, monotonicity, and oscillations of solutions of scalar neutral functional differential equations*, J. Math. Anal. Appl. **199** (1996), 502–525.

[109] K. Kunisch and W. Schappacher, *Order preserving evolution operators of functional differential equations*, Boll. Un. Mat. Ital. B (6) (1979), 480–500.

[110] H. Kunze and D. Siegel, *Monotonicity with respect to closed convex cones I*, Dynam. Contin. Discrete Impuls. Systems **5** (1999), 433–449.

[111] H. Kunze and D. Siegel, *Monotonicity with respect to closed convex cones II*, Appl. Anal. **77** (2001), 233–248.

[112] H. Kunze and D. Siegel, *A graph theoretic approach to strong monotonicity with respect to polyhedral cones*, Positivity **6** (2002), 95–113.

[113] Y.A. Kuznetsov, *Elements of Applied Bifurcation Theory*, second edition, Springer (1998).

[114] L. Ladyzhenskaya, *Attractors for Semigroups and Evolution Equations*, Cambridge Univ. Press, Cambridge (1991).

[115] M. Li and J. Muldowney, *Global stability for the SEIR model in epidemiology*, Math. Biosci. **125** (1995), 155–164.

[116] J.P. LaSalle, *The Stability of Dynamical Systems*, SIAM, Philadelphia, PA (1976).

[117] B. Lemmens, R.D. Nussbaum and S.M. Verduyn Lunel, *Lower and upper bounds for ω-limit sets of nonexpansive maps*, Indag. Math. (N.S.) **12** (2001), 191–211.

[118] R. Lemmert and P. Volkmann, *On the positivity of semigroups of operators*, Comment. Math. Univ. Carolin. **39** (1998), 483–489.

[119] A.W. Leung, *Systems of Nonlinear Partial Differential Equations: Applications to Biology and Engineering*, Kluwer Academic, Dordrecht (1989).

[120] X. Liang and J. Jiang, *The classification of the dynamical behavior of 3-dimensional type K monotone Lotka–Volterra systems*, Nonlinear Anal. **51** (2002), 749–763.

[121] X. Liang and J. Jiang, *On the finite-dimensional dynamical systems with limited competition*, Trans. Amer. Math. Soc. **354** (2002), 3535–3554.

[122] R. Loewy and H. Schneider, *Positive operators on the n-dimensional ice-cream cone*, J. Math. Anal. Appl. **49** (1975), 375–392.

[123] R. Lui, *A nonlinear integral operator arising from a model in population genetics. IV: Clines*, SIAM J. Math. Anal. **17** (1986), 152–168.

[124] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Progr. Nonlinear Differential Equations Appl., Vol. 16, Birkhäuser, Boston (1995).

[125] R.H. Martin, *Nonlinear Operators and Differential Equations in Banach Spaces*, Wiley, New York (1976).

[126] R.H. Martin, *Asymptotic stability and critical points for nonlinear quasimonotone parabolic systems*, J. Differential Equations **30** (1978), 391–423.

[127] R.H. Martin, *A maximum principle for semilinear parabolic systems*, Proc. Amer. Math. Soc. **14** (1979), 66–70.

[128] R.H. Martin, *Asymptotic behavior of solutions to a class of quasimonotone functional differential equations*, Proc. of Workshop on Functional Differential Equations and Nonlinear Semigroups, Pitman Res. Notes Math. Ser., Vol. 48, Pitman, Boston, MA (1981).

[129] R.H. Martin and H.L. Smith, *Abstract functional differential equations and reaction–diffusion systems*, Trans. Amer. Math. Soc. **321** (1990), 1–44.

[130] R.H. Martin and H.L. Smith, *Reaction–diffusion systems with time-delays: Monotonicity, invariance, comparison and convergence*, J. Reine Angew. Math. **413** (1991), 1–35.

[131] H. Matano, *Convergence of solutions of one-dimensional semilinear parabolic equations*, J. Math. Kyoto Univ. **18** (1978), 221–227.

[132] H. Matano, *Asymptotic behavior and stability of solutions of semilinear diffusion equations*, Publ. Res. Inst. Math. Sci. **19** (1979), 645–673.

[133] H. Matano, *Existence of nontrivial unstable sets for equilibriums of strongly order preserving systems*, J. Fac. Sci. Univ. Tokyo **30** (1984), 645–673.

[134] H. Matano, *Strongly order-preserving local semi-dynamical systems—theory and applications*, Semigroups, Theory and Applications, Vol. 1, H. Brezis, M.G. Crandall and F. Kappel, eds., Res. Notes Math., Vol. 141, Longman Scientific and Technical, London (1986), 178–185.

[135] H. Matano, *Strong comparison principle in nonlinear parabolic equations*, Nonlinear Parabolic Equations: Qualitative Properties of Solutions, L. Boccardo and A. Tesei, eds., Pitman Res. Notes Math. Ser., Longman Scientific and Technical, London (1987), 148–155.

[136] R. May and G. Oster, *Bifurcations and dynamic complexity in simple ecological models*, Amer. Naturalist **110** (1976), 573–599.

[137] E.J. McShane, *Extension of range of functions*, Bull. Amer. Math. Soc. (N.S.) **40** (1934), 837–842.

[138] J. Mierczyński, *On monotone trajectories*, Proc. Amer. Math. Soc. **113** (1991), 537–544.

[139] J. Mierczyński, *P-arcs in strongly monotone discrete-time dynamical systems*, Differential Integral Equations **7** (1994), 1473–1494.

[140] J. Mierczyński, *The C^1 property of carrying simplices for a class of competitive systems of ODEs*, J. Differential Equations **111** (1994), 385–409.

[141] J. Mierczyński, *On peaks in carrying simplices*, Colloq. Math. **81** (1999), 285–292.

[142] J. Mierczyński, *Smoothness of carrying simplices for three-dimensional competitive systems: A counterexample*, Dynam. Contin. Discrete Impuls. Systems **6** (1999), 147–154.

[143] J. Mierczyński, *A remark on M.W. Hirsch's paper "Chain transitive sets for smooth strongly monotone dynamical systems" [Dynam. Contin. Discrete Impuls. Systems 5 (1999), 529–543]*, Dynam. Contin. Discrete Impuls. Systems **7** (2000), 455–461.

[144] M. Mincheva, *Qualitative properties of reaction–diffusion systems in chemical reactions*, Ph.D. Thesis, University of Waterloo (2003).

[145] M. Mincheva and D. Siegel, *Comparison principles for reaction–diffusion systems with respect to proper polyhedral cones*, Dynam. Cont., Discrete and Impulsive Systems **10**, (2003) p. 477.

[146] K. Mischaikow, H.L. Smith and H.R. Thieme, *Asymptotically autonomous semiflows: Chain recurrence and Lyapunov functions*, Trans. Amer. Math. Soc. **347** (1995), 1669–1685.

[147] X. Mora, *Semilinear parabolic problems define semiflows on C^k spaces*, Trans. Amer. Math. Soc. **278** (1983), 21–55.

[148] M. Müller, *Über das Fundamenthaltheorem in der Theorie der gewöhnlichen Differentialgleichungen*, Math. Z. **26** (1926), 619–645.

[149] R.D. Nussbaum, *Hilbert's projective metric and iterated nonlinear maps*, Mem. Amer. Math. Soc. **391** (1988).

[150] R.D. Nussbaum, *Iterated nonlinear maps and Hilbert's projective metric, II*, Memoirs of the Amer. Math. Soc. **401** (1989).

[151] R.D. Nussbaum, *Estimates of the periods of periodic points for nonexpansive operators*, Israel J. Math. **76** (1991), 345–380.

[152] Y. Ohta, *Qualitative analysis of nonlinear quasimonotone dynamical systems described by functional differential equations*, IEEE Trans. Circuits Systems **28** (1981), 138–144.

[153] R. Ortega and L. Sánchez, *Abstract competitive systems and orbital stability in* R^3, Proc. Amer. Math. Soc. **128** (2000), 2911–2919.

[154] J. Oxtoby, *Measure and Category*, Grad. Texts in Math., Vol. 2. Springer, New York (1980).

[155] C.V. Pao, *Nonlinear Parabolic and Elliptic Equations*, Plenum, New York (1992).

[156] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, New York (1983).

[157] M. Peixoto, *On an approximation theorem of Kupka and Smale*, J. Differential Equations **3** (1967), 214–227.

[158] O. Perron, *Zur Theorie der Matrizen*, Math. Ann. **64** (1907), 248–263.

[159] M. Pituk, *Convergence to equilibrium in scalar nonquasimonotone functional differential equations*, J. Differential Equations **193** (2003), 95–130.

[160] P. Poláčik, *Convergence in smooth strongly monotone flows defined by semilinear parabolic equations*, J. Differential Equations **79** (1989), 89–110.

[161] P. Poláčik, *Domains of attraction of equilibria and monotonicity properties of convergent trajectories in parabolic systems admitting strong comparison principle*, J. Reine Angew. Math. **400** (1989), 32–56.

[162] P. Poláčik, *Existence of unstable sets for invariant sets in compact semiflows. Applications in order-preserving semiflows*, Comment. Math. Univ. Carolin. **31** (1990), 263–276.

[163] P. Poláčik, *Parabolic Equations: Asymptotic Behavior and Dynamics on Invariant Manifolds*, Handbook of Dynamical Systems, Vol. 2, B. Fiedler, ed., Elsevier, New York (2002).

[164] P. Poláčik and I. Tereščák, *Convergence to cycles as a typical asymptotic behavior in smooth strongly monotone discrete-time dynamical systems*, Arch. Ration. Mech. Anal. **116** (1992), 339–360.

[165] P. Poláčik and I. Tereščák, *Exponential separation and invariant bundles for maps in ordered Banach spaces with applications to parabolic equations*, J. Dynam. Differential Equations **5** (1993), 279–303.

[166] M.H. Protter and H.F. Weinberger, *Maximum Principles in Differential Equations*, Prentice–Hall, Englewood Cliffs, NJ (1967).

[167] G. Raugel, *Global Attractors in Partial Differential Equations*, Handbook of Dynamical Systems, Vol. 2, B. Fiedler, ed., Elsevier, New York (2002).

[168] A. Rescigno and I. Richardson, *The struggle for life, I: Two species*, Bull. Math. Biophysics **2** (1967), 377–388.

[169] C. Robinson, *Stability theorems and hyperbolicity in dynamical systems*, Rocky Mountain J. Math. **7** (1977), 425–437.

[170] J. Robinson, *Infinite-Dimensional Dynamical Systems*, Cambridge Univ. Press, Cambridge (2001).

[171] W. Rudin, *Real and Complex Analysis*.

[172] D. Ruelle, *Analyticity properties of characteristic exponents of random matrix products*, Adv. Math. **32** (1979), 68–80.

[173] M. Rutman, *Sur les opérateurs totalement continus linéaires laissant invariant un certain cǒne*, Rec. Math. (Mat. Sbornik) N.S. **8** (1940), 77–96.

[174] R. Sacker and G. Sell, *Lifting properties in skew-product flows with applications to differential equations*, Mem. Amer. Math. Soc. **190** (1977).

[175] S.H. Saperstone, *Semidynamical Systems in Infinite-Dimensional Spaces*, Appl. Math. Sci., Vol. 37, Springer, Berlin (1981).

[176] D. Sattinger, *Monotone methods in nonlinear elliptic and parabolic boundary value problems*, Indiana Univ. Math. J. **21** (1972), 979–1000.

[177] H. Schneider and M. Vidyasagar, *Cross-positive matrices*, SIAM J. Numer. Anal. **7** (1970), 508–519.

[178] J. Selgrade, *Asymptotic behavior of solutions to single loop positive feedback systems*, J. Differential Equations **38** (1980), 80–103.

[179] J. Selgrade, *A Hopf bifurcation in single loop positive feedback systems*, Quart. Appl. Math. (1982), 347–351.

[180] J. Selgrade, *On the existence and uniqueness of connecting orbits*, Nonlinear Anal. **7** (1983), 1123–1125.

[181] J. Selgrade and M. Ziehe, *Convergence to equilibrium in a genetic model with differential viability between sexes*, J. Math. Biol. **25** (1987), 477–490.

[182] S. Smale, *On the differential equations of species in competition*, J. Math. Biol. **3** (1976), 5–7.

[183] H.L. Smith, *Cooperative systems of differential equations with concave nonlinearities*, Nonlinear Anal. **10** (1986), 1037–1052.

[184] H.L. Smith, *Invariant curves for mappings*, SIAM J. Math. Anal. **17** (1986), 1053–1067.

[185] H.L. Smith, *On the asymptotic behavior of a class of deterministic models of cooperating species*, SIAM J. Appl. Math. **46** (1986), 368–375.

[186] H.L. Smith, *Competing subcommunities of mutualists and a generalized Kamke theorem*, SIAM J. Appl. Math. **46** (1986), 856–874.

[187] H.L. Smith, *Periodic orbits of competitive and cooperative systems*, J. Differential Equations **65** (1986), 361–373.

[188] H.L. Smith, *Periodic competitive differential equations and the discrete dynamics of competitive maps*, J. Differential Equations **64** (1986), 165–194.

[189] H.L. Smith, *Periodic solutions of periodic competitive and cooperative systems*, SIAM J. Math. Anal. **17** (1986), 1289–1318.

[190] H.L. Smith, *Monotone semiflows generated by functional differential equations*, J. Differential Equations **66** (1987), 420–442.

[191] H.L. Smith, *Oscillation and multiple steady states in a cyclic gene model with repression*, J. Math. Biol. **25** (1987), 169–190.

[192] H.L. Smith, *Planar competitive and cooperative difference equations*, J. Differ. Equations Appl. **3** (1988), 335–357.

[193] H.L. Smith, *Systems of ordinary differential equations which generate an order preserving flow. A survey of results*, SIAM Rev. **30** (1988), 87–113.

[194] H.L. Smith, *Monotone Dynamical Systems, an Introduction to the Theory of Competitive and Cooperative Systems*, Math. Surveys Monogr., Vol. 41, American Mathematical Society, Providence, RI (1995).

[195] H.L. Smith, *Complicated dynamics for low-dimensional strongly monotone maps*, Nonlinear Anal. **30** (1997), 1911–1914.

[196] H.L. Smith, *Dynamics of Competition*, Mathematics Inspired by Biology, Lecture Notes in Math., Vol. 1714, Springer (1999), 191–240.

[197] H.L. Smith and H.R. Thieme, *Quasi convergence for strongly ordered preserving semiflows*, SIAM J. Math. Anal. **21** (1990), 673–692.

[198] H.L. Smith and H.R. Thieme, *Monotone semiflows in scalar nonquasi-monotone functional differential equations*, J. Math. Anal. Appl. **150** (1990), 289–306.

[199] H.L. Smith and H.R. Thieme, *Convergence for strongly ordered preserving semiflows*, SIAM J. Math. Anal. **22** (1991), 1081–1101.

[200] H.L. Smith and H.R. Thieme, *Strongly order preserving semiflows generated by functional differential equations*, J. Differential Equations **93** (1991), 332–363.

[201] H.L. Smith and H.R. Thieme, *Stable coexistence and bi-stability for competitive systems on ordered Banach spaces*, J. Differential Equations **176** (2001), 195–222.

[202] H.L. Smith and P. Waltman, *A classification theorem for three-dimensional competitive systems*, J. Differential Equations **70** (1987), 325–332.

[203] H.L. Smith and P. Waltman, *The Theory of the Chemostat*, Cambridge Stud. Math. Biol., Vol. 13, Cambridge Univ. Press (1995).

[204] R.A. Smith, *Orbital stability for ordinary differential equations*, J. Differential Equations **69** (1987), 265–287.

[205] J. Smoller, *Shock Waves and Reaction Diffusion Equations*, Springer, New York (1983).

[206] R. Stern and H. Wolkowicz, *Exponential nonnegativity on the ice-cream cone*, SIAM J. Matrix Anal. Appl. **12** (1991), 160–165.

[207] H. Stewart, *Generation of analytic semigroups by strongly elliptic operators under general boundary conditions*, Trans. Amer. Math. Soc. **259** (1980), 299–310.

[208] P. Takáč, *Asymptotic behavior of discrete-time semigroups of sublinear, strongly increasing mappings with applications in biology*, Nonlinear Anal. **14** (1990), 35–42.

[209] P. Takáč, *Convergence to equilibrium on invariant d-hypersurfaces for strongly increasing discrete-time semigroups*, J. Math. Anal. Appl. **148** (1990), 223–244.

[210] P. Takáč, *Domains of attraction of generic ω-limit sets for strongly monotone semiflows*, Z. Anal. Anwendungen **10** (1991), 275–317.

[211] P. Takáč, *Domains of attraction of generic ω-limit sets for strongly monotone discrete-time semigroups*, J. Reine Angew. Math. **432** (1992), 101–173.

[212] P. Takáč, *Linearly stable subharmonic orbits in strongly monotone time-periodic dynamical systems*, Proc. Amer. Math. Soc. **115** (1992), 691–698.

[213] P. Takáč, *A construction of stable subharmonic orbits in monotone time-periodic dynamical systems*, Monatsh. Math. **115** (1993), 215–244.

[214] P. Takáč, *A short elementary proof of the Kreĭn–Rutman theorem*, Houston J. Math. **20** (1994), 93–98.

[215] P. Takáč, *Convergence in the part metric for discrete dynamical systems in ordered topological cones*, Nonlinear Anal. **26** (1996), 1753–1777.

[216] P. Takáč, *Discrete monotone dynamics and time-periodic competition between two species*, Differential Integral Equations **10** (1997), 547–576.

[217] I. Tereščák, *Dynamics of C^1 smooth strongly monotone discrete-time dynamical system*, Preprint, Comenius University, Bratislava (1994).

[218] H.R. Thieme, *Density dependent regulation of spatially distributed populations and their asymptotic speeds of spread*, J. Math. Biol. **8** (1979), 173–187.

[219] A.C. Thompson, *On certain contraction mappings in a partially ordered vector space*, Proc. Amer. Math. Soc. **14** (1963), 438–443.

[220] A. Tineo, *On the convexity of the carrying simplex of planar Lotka–Volterra competitive systems*, Appl. Math. Comput. **123** (2001), 93–108.

[221] R. Uhl, *Ordinary differential inequalities and quasimonotonicity in ordered topological vector spaces*, Proc. Amer. Math. Soc. **126** (1998), 1999–2003.

[222] P.S. Uryson, *Mittlehre Breite une Volumen der konvexen Körper im n-dimensionalen Räume*, Mat. Sb. **31** (1924), 477–486.

[223] P. van den Driessche and M. Zeeman, *Three-dimensional competitive Lotka–Volterra systems with no periodic orbits*, SIAM J. Appl. Math. **58** (1998), 227–234.

[224] P. Volkmann, *Gewöhnliche Differentialungleichungen mit quasimonoton wachsenden Funktionen in topologischen Vektorräumen*, Math. Z. **127** (1972), 157–164.

[225] B. Vulikh, *Introduction to the Theory of Partially Ordered Spaces*, Wolters-Noordhoff Scientific Publications, Groningen (1967).

[226] S. Walcher, *On cooperative systems with respect to arbitrary orderings*, J. Math. Anal. Appl. **263** (2001), 543–554.

[227] W. Walter, *Differential and Integral Inequalities*, Springer, Berlin (1970).

[228] W. Walter, *On strongly monotone flows*, Ann. Polon. Math. **LXVI** (1997), 269–274.

[229] Y. Wang and J. Jiang, *The general properties of discrete-time competitive dynamical systems*, J. Differential Equations **176** (2001), 470–493.

[230] Y. Wang and J. Jiang, *Uniqueness and attractivity of the carrying simplex for discrete-time competitive dynamical systems*, J. Differential Equations **186** (2002), 611–632.

[231] Y. Wang and J. Jiang, *The long-run behavior of periodic competitive Kolmogorov systems*, Nonlinear Anal. Real World Appl. **3** (2002), 471–485.

[232] H.F. Weinberger, *Long-time behavior of a class of biological models*, SIAM J. Math. Anal. **13** (1982), 353–396.

[233] R. Wilder, *Topology of Manifolds*, Amer. Math. Soc., Providence, RI (1949).

[234] J. Wu, *Global dynamics of strongly monotone retarded equations with infinite delay*, J. Integral Equations Appl. **4** (1992), 273–307.

[235] J. Wu and H. Freedman, *Monotone semiflows generated by neutral functional differential equations with application to compartmental systems*, Canad. J. Math. **43** (1991), 1098–1120.

[236] J. Wu, H.I. Freedman and R.K. Miller, *Heteroclinic orbits and convergence of order-preserving set-condensing semiflows with applications to integrodifferential equations*, J. Integral Equations Appl. **7** (1995), 115–133.

[237] D. Xiao and W. Li, *Limit cycles for the competitive three-dimensional Lotka–Volterra system*, J. Differential Equations **164** (2000), 1–15.

[238] F. Zanolin, *Permanence and positive periodic solutions for Kolmogorov competing species systems*, Results Math. **21** (1992), 224–250.

[239] E. Zeeman, *Classification of quadratic carrying simplices in two-dimensional competitive Lotka–Volterra systems*, Nonlinearity **15** (2002), 1993–2018.

[240] E.C. Zeeman and M.L. Zeeman, *On the convexity of carrying simplices in competitive Lotka–Volterra systems*, Differential Equations, Dynamical Systems, and Control Science, Lecture Notes in Pure and Appl. Math., Vol. 152, Dekker, New York (1994), 353–364.

[241] E.C. Zeeman and M.L. Zeeman, *An n-dimensional competitive Lotka–Volterra system is generically determined by the edges of its carrying simplex*, Nonlinearity **15** (2002), 2019–2032.

[242] E.C. Zeeman and M.L. Zeeman, *From local to global behavior in competitive Lotka–Volterra systems*, Trans. Amer. Math. Soc. **355** (2003), 713–734 (electronic).

[243] M. Zeeman, *Hopf bifurcations in competitive three-dimensional Lotka–Volterra systems*, Dynam. Stability Systems **8** (1993), 189–217.

[244] E. Zeidler, *Nonlinear Functional Analysis and Its Applications, I: Fixed Point Theorems*, Springer, New York (1986).

[245] X.-Q. Zhao, *Global attractivity and stability in some monotone discrete dynamical systems*, Bull. Austral. Math. Soc. **53** (1996), 305–324.

[246] X.-Q. Zhao and B.D. Sleeman, *Permanence in Kolmogorov competition models with diffusion*, IMA J. Appl. Math. **51** (1993), 1–11.

[247] H.-R. Zhu, *The existence of stable periodic orbits for systems of three-dimensional differential equations that are competitive*, Ph.D. Thesis, Arizona State University (1991).

[248] H.-R. Zhu and H.L. Smith, *Stable periodic orbits for a class of three-dimensional competitive systems*, J. Differential Equations **110** (1994), 143–156.

This page intentionally left blank

CHAPTER 5

Planar Periodic Systems of Population Dynamics

Julián López-Gómez[†]
Departamento de Matemática Aplicada, Universidad Complutense de Madrid, 28040-Madrid, Spain
E-mail: *Lopez_Gomez@mat.ucm.es*

Contents

1. Introduction 361
2. Basic preliminaries. The single logistic equation 364
    2.1. The differential operator $\frac{\mathrm{d}}{\mathrm{d}t} + V : \mathcal{C}^1_T \to \mathcal{C}_T$ 364
    2.2. The single logistic equation 367
    2.3. Point-wise behaviour of the map $\rho \mapsto \theta_{[\rho\alpha,\beta]}$ 369
    2.4. Attractiveness of $\theta_{[\rho\alpha,\beta]}$ 375
3. Linearized stability of the semi-trivial solutions 376
4. The curves of change of stability of the semi-trivial positive solutions 378
    4.1. The species u and v compete ($b > 0$ and $c > 0$) 379
    4.2. The species v preys on u ($b > 0$ and $c < 0$) 386
    4.3. The species u and v cooperate ($b < 0$ and $c < 0$) 387
    4.4. The general case when b and c change of sign 388
5. The existence of coexistence states 392
    5.1. Abstract unilateral bifurcation from semi-trivial states 392
    5.2. The existence of coexistence states in the general case 401
    5.3. Local structure of $\mathfrak{C}^\lambda_\mu$ and $\mathfrak{C}^\mu_\lambda$ at $(\mu_0, \theta_{[\lambda\ell,a]}, 0)$ and $(\lambda_0, 0, \theta_{[\mu m,d]})$ 406
6. The symbiotic model ($b < 0$ and $c < 0$) 411
    6.1. Periodic systems of cooperative type 411
    6.2. Structure of the set of coexistence states 421
    6.3. Stability of the coexistence states along $\mathfrak{C}^\lambda_\mu$ and $\mathfrak{C}^\mu_\lambda$ 426
    6.4. Low and strong symbiosis effects 434
7. The competing species model ($b > 0$ and $c > 0$) 438
    7.1. Periodic systems of quasi-cooperative type 439
    7.2. Structure of the components $\mathfrak{C}^\lambda_\mu$ and $\mathfrak{C}^\mu_\lambda$ 443
    7.3. Stability of the coexistence states along $\mathfrak{C}^\lambda_\mu$ and $\mathfrak{C}^\mu_\lambda$ 448
    7.4. Some bifurcation diagrams 452
    7.5. Dynamics of (1.1) 454
8. The species v preys on u ($b > 0$ and $c < 0$) 455
    8.1. Characterizing the existence of coexistence states 456

[†] Supported by the Ministry of Science and Technology of Spain under grant REN2003-00707.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 2
Edited by A. Cañada, P. Drábek and A. Fonda
© 2005 Elsevier B.V. All rights reserved

8.2. Dynamics of (1.1) 458
References 459

Abstract

This chapter analyzes the dynamics of the positive solutions of a general class of planar periodic systems including those of Lotka–Volterra type and a more general class of models simulating symbiotic interactions within global competitive environments. It seems this is the first occasion where this problematic has been addressed within the context of periodic planar systems. Most of our mathematical analysis is focused towards the study of the existence of coexistence states and the problem of ascertaining the structure, multiplicity and stability of these coexistence states in purely symbiotic, and competitive, environments.

Keywords: Lotka–Volterra systems, symbiosis in competitive environments, coexistence states.

MSC: 34A34, 34C12, 92B05.

1. Introduction

Throughout this chapter, for any $T > 0$ we denote by $\mathcal{C}_T$ the Banach space of real T-periodic continuous functions endowed with the norm

$$\|u\|_{\mathcal{C}_T} := \max_{t\in\mathbb{R}} |u(t)| = \max_{t\in[0,T]} |u(t)|.$$

Similarly, $\mathcal{C}^1_T$ stands for the Banach space of real T-periodic functions of class $\mathcal{C}^1$ endowed with the norm

$$\|u\|_{\mathcal{C}^1_T} := \|u\|_{\mathcal{C}_T} + \|u'\|_{\mathcal{C}_T},$$

where $' = \frac{\mathrm{d}}{\mathrm{d}t}$. The Banach space $\mathcal{C}_T$ is ordered by its cone of non-negative functions

$$P := \{u \in \mathcal{C}_T\colon\ u(t) \geqslant 0,\ t \in \mathbb{R}\}.$$

Note that

$$\operatorname{Int} P = \{u \in \mathcal{C}_T\colon\ u(t) > 0,\ t \in \mathbb{R}\}.$$

Given $u, v \in \mathcal{C}_T$, we shall write $u \geqslant v$ if $u - v \in P$, $u > v$ if $u - v \in P \setminus \{0\}$, and $u \gg v$ if $u - v \in \operatorname{Int} P$.

The main goal of this chapter is analyzing the dynamics of

$$\begin{cases} u'(t) = \lambda\,\ell(t)u(t) - a(t)u^2(t) - b(t)u(t)v(t), \\ v'(t) = \mu\,m(t)v(t) - d(t)v^2(t) - c(t)u(t)v(t), \\ u(0) = x > 0, \quad v(0) = y > 0, \end{cases} \tag{1.1}$$

and, in particular, the existence, attractiveness and multiplicity of the component-wise non-negative T-periodic solutions of

$$\begin{cases} u'(t) = \lambda\ell(t)u(t) - a(t)u^2(t) - b(t)u(t)v(t), \\ v'(t) = \mu m(t)v(t) - d(t)v^2(t) - c(t)u(t)v(t), \end{cases} \tag{1.2}$$

where $\{\ell, m, a, b, c, d\} \subset \mathcal{C}_T$ satisfy

$$\ell > 0, \quad m > 0, \quad a > 0, \quad d > 0, \tag{1.3}$$

and $(\lambda, \mu) \in \mathbb{R}^2$ are regarded as two real parameters. System (1.2) exhibits three different types of non-negative T-periodic solution pairs. Namely, the *trivial state* $(0, 0)$, the *semi-trivial positive states* $(u, 0)$ and $(0, v)$ with $u \gg 0$ and $v \gg 0$, and the *coexistence states*, which are the solution pairs $(u, v) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ with $u \gg 0$ and $v \gg 0$. The interest of analyzing these states coming from the fact that in many circumstances they govern the dynamics of (1.1).

It should be noted that we are not imposing any sign restriction on the coupling coefficient functions $b(t)$ and $c(t)$. Thus, (1.2) is far from falling within the classical family of Lotka–Volterra periodic systems, since b and c might not have a constant sign. Actually, if given $h \in \mathcal{C}_T$ we denote

$$h^+ := \max\{h, 0\}, \quad h^- := h^+ - h, \quad I_+^h := \operatorname{Int}\operatorname{supp} h^+, \quad I_-^h := \operatorname{Int}\operatorname{supp} h^-,$$

and

$$I_0^h := \operatorname{Int} h^{-1}(0) = \mathbb{R} \setminus \big(\bar{I}_+^h \cup \bar{I}_-^h\big),$$

then, the species u and v interact according to the following patterns:

- u and v compete if $t \in I_+^b \cap I_+^c$—periods of time where $b(t) > 0$ and $c(t) > 0$.
- u and v cooperate if $t \in I_-^b \cap I_-^c$—periods of time where $b(t) < 0$ and $c(t) < 0$.
- u preys on v if $t \in I_-^b \cap I_+^c$, and v preys on u if $t \in I_+^b \cap I_-^c$—periods of time where $b(t)c(t) < 0$.
- u is free from the action of v if $t \in I_0^b$, and v is free from the action of u if $t \in I_0^c$—periods of time where $b(t)c(t) = 0$.

Consequently, within our general setting (1.1) allows all different types of interactions between the species u and v as time passes by.

Due to (1.3), we have that $\int_0^T a > 0$ and $\int_0^T d > 0$, and, hence, the change of variable

$$u := \frac{T}{\int_0^T a} U, \quad v := \frac{T}{\int_0^T d} V, \quad \Lambda := \frac{\lambda}{T} \int_0^T \ell, \quad M := \frac{\mu}{T} \int_0^T m,$$

transforms (1.2) into

$$\begin{cases} U'(t) = \Lambda \dfrac{\ell(t)}{\frac{1}{T}\int_0^T \ell} U(t) - \dfrac{a(t)}{\frac{1}{T}\int_0^T a} U^2(t) - \dfrac{b(t)}{\frac{1}{T}\int_0^T d} U(t)V(t), \\ V'(t) = M \dfrac{m(t)}{\frac{1}{T}\int_0^T m} V(t) - \dfrac{d(t)}{\frac{1}{T}\int_0^T d} V^2(t) - \dfrac{c(t)}{\frac{1}{T}\int_0^T a} U(t)V(t). \end{cases}$$

Consequently, without lost of generality, we will throughout assume that the following conditions are satisfied:

$$\frac{1}{T}\int_0^T \ell = \frac{1}{T}\int_0^T m = \frac{1}{T}\int_0^T a = \frac{1}{T}\int_0^T d = 1. \tag{1.4}$$

Therefore,

$$h = 1 \quad \text{if } h \in \{\ell, m, a, d\} \text{ is assumed to be constant.} \tag{1.5}$$

Throughout this chapter, given any function $h \in \mathcal{C}_T$ we shall denote

$$\hat{h} := \frac{1}{T}\int_0^T h(t)\,\mathrm{d}t, \qquad h_M := \max_{\mathbb{R}} h, \qquad h_L := \min_{\mathbb{R}} h. \tag{1.6}$$

Using these notations, conditions (1.3) and (1.4) can be summarized as follows:

$$\{\ell, m, a, d\} \subset P \setminus \{0\}, \qquad \hat{\ell} = \hat{m} = \hat{a} = \hat{d} = 1. \tag{1.7}$$

This work is distributed as follows. In Section 2 we provide some basic preliminaries needed for the subsequent mathematical analysis. Basically, it collects some well known results for the single logistic equation. Section 3 ascertains the linearized stability character of the semi-trivial positive solutions of (1.2).

Section 4 analyzes the minimal complexity of the components of the (λ, μ)-plane determined by the curves of neutral stability of the semi-trivial states. Several different situation cases will be differentiated, according to the nature of the interactions between the species u and v. The mathematical analysis carried out in Section 4 is reminiscent from J. Eilbeck and J. López-Gómez [12], where attention was exclusively focused into the competing species model.

In Section 5 we give an abstract unilateral global bifurcation result for systems. The main theorem of Section 5 is new in its full generality, and it admits a number of applications to the search of coexistence states in wide classes of nonlinear elliptic boundary value problems and planar periodic systems. Here, we will apply it to get all available existence results concerning the existence of coexistence states of (1.2). We point out that the abstract result of Section 5 does not follow as a direct application of the unilateral theory developed by P.H. Rabinowitz [31], but from the updated unilateral theorem of [19].

Section 6 considers the symbiotic prototype model ($b < 0$ and $c < 0$) and uses the theory of monotone periodic systems to show that the set of coexistence states linking the surfaces of the semi-trivial states along their respective curves of neutral stability is a real analytic surface, and that the stability character of the coexistence states changes as one crosses any *turning point* along that surface, while it remains unchanged when an hysteresis point is passed by. The theory developed here goes back to [17], within the context of quasi-cooperative systems, though here we are considerably tidying up the mathematical analysis of [17] and adding all necessary technical details that were omitted there in. Some of our abstract results for linear periodic systems provide with substantial improvements of some of the results of M.A. Krasnosel'skii [16]. Most precisely, the characterization of the attractive character of the system in terms of the existence of a strict positive supersolution seems to be a completely new result. This theorem is reminiscent of the characterization of the strong maximum principle for second order linear elliptic operators found by J. López-Gómez and M. Molina-Meyer [21,20]. Section 7 adapts the mathematical analysis of Section 6 to the competing species model ($b > 0$ and $c > 0$), which possesses a quasi-cooperative structure.

Finally, in Section 8 we briefly collect some of the main results available for predator prey models ($b > 0$ and $c < 0$, or $b < 0$ and $c > 0$). In a forthcoming chapter, we will adapt the theory developed by J. López-Gómez and M. Molina-Meyer [22,23], in order to discuss the effects of strategic symbiosis in competitive environments in order to show that strategic symbiosis allows the species to avoid extinction, while, simultaneously—and quite strikingly—the productivity of the ecosystem grows.

For historical remarks and a precise account of some of the most important works available about these models, the reader is sent to the expository chapter of J. Mawhin [27, 28], which bring an excellent complement to the mathematical analysis carried out in this chapter.

Throughout this chapter, given two real Banach spaces, X and Y, we denote by $\mathcal{L}(X;Y)$ the space of linear continuous operators between X and Y, and given $T \in \mathcal{L}(X;Y)$, $N[T]$, $R[T]$, and $\operatorname{spr} T$ will stand for the null space (kernel), the image (rank), and the spectral radius of T, respectively. The identity in X will be denoted by I_X, and, given $R > 0$ and $x \in X$, $B_R(x)$ will stand for the ball of radius R centered at x.

2. Basic preliminaries. The single logistic equation

In this section we focus our attention into the single logistic equation

$$w'(t) = \rho\alpha(t)w(t) - \beta(t)w^2(t), \tag{2.1}$$

where $\alpha, \beta \in \mathcal{C}_T$ satisfy

$$\hat{\alpha} > 0, \quad \beta > 0, \tag{2.2}$$

and $\rho \in \mathbb{R}$ is regarded as a real parameter. This equation provides us with the semi-trivial states—positive—of (1.2) when $(\alpha, \beta) \in \{(\ell, a), (m, d)\}$.

2.1. *The differential operator* $\frac{\mathrm{d}}{\mathrm{d}t} + V : \mathcal{C}^1_T \to \mathcal{C}_T$

We begin studying some useful properties of the differential operator

$$\frac{\mathrm{d}}{\mathrm{d}t} + V : \mathcal{C}^1_T \longrightarrow \mathcal{C}_T, \quad w \mapsto w' + Vw, \tag{2.3}$$

where $V \in \mathcal{C}_T$. The following result collects its main positivity properties.

PROPOSITION 2.1. *Suppose $V, f \in \mathcal{C}_T$ and consider the periodic problem*

$$w' + Vw = f, \quad w \in \mathcal{C}^1_T. \tag{2.4}$$

Then, the following assertions are true:

(a) (2.4) *has a unique solution for each $f \in \mathcal{C}_T$ if, and only if, $\int_0^T V \neq 0$. In such case, we denote by $\psi_{[V,f]}$ its unique solution.*

(b) *In case $\int_0^T V > 0$, $\psi_{[V,f]} \gg 0$ if $f > 0$, and $\psi_{[V,f]} \ll 0$ if $f < 0$.*

(c) *In case $\int_0^T V < 0$, $\psi_{[V,f]} \ll 0$ if $f > 0$, and $\psi_{[V,f]} \gg 0$ if $f < 0$.*

(d) *In case $\int_0^T V = 0$,* (2.4) *possesses a solution if, and only if,*

$$\int_0^T \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s = 0. \tag{2.5}$$

In particular, (2.4) *cannot admit a solution if either* $f > 0$ *or* $f < 0$. *Actually, under condition* (2.5) *any solution of*

$$w' + Vw = f$$

must be T*-periodic.*

PROOF. For any $w_0 \in \mathbb{R}$, the unique solution of the Cauchy problem

$$\begin{cases} w' + Vw = f, \\ w(0) = w_0, \end{cases}$$

is given by

$$w(t) = \mathrm{e}^{-\int_0^t V}\left[w_0 + \int_0^t \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s\right], \quad t \in \mathbb{R}, \tag{2.6}$$

and it is T-periodic if, and only if, $w(0) = w(T)$. Equivalently,

$$\big(\mathrm{e}^{\int_0^T V} - 1\big)w_0 = \int_0^T \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s. \tag{2.7}$$

Thus, in case $\int_0^T V \neq 0$,

$$w_0 := \frac{\int_0^T \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s}{\mathrm{e}^{\int_0^T V} - 1} \tag{2.8}$$

is the unique initial value to a T-periodic solution, while, if $\int_0^T V = 0$, then identity (2.7) implies (2.5).

As a result from this elementary analysis, substituting (2.8) into (2.6) and rearranging terms, we find that under assumption $\int_0^T V \neq 0$ the function

$$\begin{aligned}\psi_{[V,f]}(t) &:= \frac{\mathrm{e}^{-\int_0^t V}}{\mathrm{e}^{\int_0^T V} - 1}\left[\int_0^T \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s + \big(\mathrm{e}^{\int_0^T V} - 1\big)\int_0^t \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s\right] \\ &= \frac{\mathrm{e}^{-\int_0^t V}}{\mathrm{e}^{\int_0^T V} - 1}\left(\int_t^T \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s + \mathrm{e}^{\int_0^T V}\int_0^t \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s\right), \quad t \in \mathbb{R},\end{aligned}$$

provides us with the unique solution of (2.4). Note that, for any $V \in \mathcal{C}_T$, the auxiliary function

$$\Xi(t) := \int_t^T \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s + \mathrm{e}^{\int_0^T V}\int_0^t \mathrm{e}^{\int_0^s V} f(s)\,\mathrm{d}s, \quad t \in \mathbb{R}, \tag{2.9}$$

satisfies

$$\Xi \gg 0 \quad \text{if } f > 0, \quad \text{and} \quad \Xi \ll 0 \quad \text{if } f < 0.$$

Consequently, in case $\int_0^T V > 0$ we have that, for each $t \in \mathbb{R}$,

$$\psi_{[V,f]}(t) = \frac{\mathrm{e}^{-\int_0^t V}}{\mathrm{e}^{\int_0^T V} - 1} \Xi(t) > 0$$

if $f > 0$, whereas $\psi_{[V,f]}(t) < 0$ if $f < 0$. Similarly, in case $\int_0^T V < 0$, we have that $\psi_{[V,f]} \gg 0$ if $f < 0$, and $\psi_{[V,f]} \ll 0$ if $f > 0$, which concludes the proof of parts (a), (b) and (c).

Finally, suppose $\int_0^T V = 0$. In this case, we already know that (2.5) is necessary for the existence of a solution to (2.4). Suppose (2.5). Then, for each $w_0 \in \mathbb{R}$, we have that (2.6) satisfies

$$w(T) = \mathrm{e}^{-\int_0^T V} \left[w_0 + \int_0^T \mathrm{e}^{\int_0^s V} f(s)\, \mathrm{d}s \right] = w_0 = w(0)$$

and, therefore, it provides us with a solution of (2.4). This concludes the proof. □

As an immediate consequence from Proposition 2.1, in case $\int_0^T V \neq 0$ the linear differential operator (2.3) is a bijection. Clearly, it is continuous. So, it follows from the open mapping theorem that it defines a linear isomorphism between $\mathcal{C}_T^1$ and $\mathcal{C}_T$. Subsequently, we shall denote by J the canonical injection

$$J : \mathcal{C}_T^1 \to \mathcal{C}_T.$$

By Ascoli–Arzela's theorem, J is compact, i.e., it sends bounded sets of $\mathcal{C}_T^1$ into relatively compact subsets of $\mathcal{C}_T$. Therefore, the *resolvent operator*

$$\mathfrak{R}_V := J \left(\frac{\mathrm{d}}{\mathrm{d}t} + V \right)^{-1} : \mathcal{C}_T \to \mathcal{C}_T \tag{2.10}$$

is as well compact. Actually, it is *strongly positive* if $\int_0^T V > 0$, while it is *strongly negative* if $\int_0^T V < 0$, i.e.,

$$\mathfrak{R}_V\big(P \setminus \{0\}\big) \subset \operatorname{Int} P \quad \text{if } \int_0^T V > 0, \tag{2.11}$$

while,

$$\mathfrak{R}_V\big(P \setminus \{0\}\big) \subset -\operatorname{Int} P \quad \text{if } \int_0^T V < 0. \tag{2.12}$$

As a consequence from these positivity properties, one can easily get the main comparison properties associated to Problem (2.4). To state them we need the following elementary concepts.

DEFINITION 2.2. Given $V, f \in \mathcal{C}_T$,

(1) a function $w \in \mathcal{C}_T^1$ is said to be a subsolution (resp. strict subsolution) of (2.4) if $w' + Vw \leqslant f$ (resp. $w' + Vw < f$);

(2) a function $w \in \mathcal{C}_T^1$ is said to be a supersolution (resp. strict supersolution) of (2.4) if $w' + Vw \geqslant f$ (resp. $w' + Vw > f$).

PROPOSITION 2.3. *Suppose* $V, f \in \mathcal{C}_T$, *with* $\int_0^T V \neq 0$. *Then*:

(a) *In case* $\int_0^T V > 0$, *any strict subsolution* w *of* (2.4) *satisfies* $w \ll \psi_{[V,f]}$, *whereas* $w \gg \psi_{[V,f]}$ *for any strict supersolution* w.

(b) *In case* $\int_0^T V < 0$, *any strict subsolution* w *of* (2.4) *satisfies* $w \gg \psi_{[V,f]}$, *whereas* $w \ll \psi_{[V,f]}$ *for any strict supersolution* w.

PROOF. Suppose $\int_0^T V > 0$ and w is a strict subsolution of (2.4). Then, the auxiliary function p defined by

$$p := \left(\frac{\mathrm{d}}{\mathrm{d}t} + V\right)(\psi_{[V,f]} - w)$$

satisfies $p > 0$ and, due to (2.11),

$$\psi_{[V,f]} - w = \mathfrak{R}_V(p) \gg 0.$$

Similarly, when w is a strict supersolution of (2.4),

$$q := \left(\frac{\mathrm{d}}{\mathrm{d}t} + V\right)(w - \psi_{[V,f]}) > 0$$

and (2.11) implies

$$w - \psi_{[V,f]} = \mathfrak{R}_V(q) \gg 0,$$

which concludes the proof of part (a). The proof of part (b) follows the same scheme, though one should use (2.12), instead of (2.11). □

2.2. *The single logistic equation*

The following result characterizes the existence of positive solutions of (2.1).

PROPOSITION 2.4. *Suppose* $\alpha, \beta \in \mathcal{C}_T$ *satisfy* (2.2). *Then,* (2.1) *possesses a positive* T*-periodic solution if, and only if,* $\rho > 0$. *Moreover, it is unique if it exists and if we denote it by* $\theta_{[\rho\alpha,\beta]}$, *then,*

$$\theta_{[\rho\alpha,\beta]} = \frac{1}{\psi_{[\rho\alpha,\beta]}}, \tag{2.13}$$

where $\psi_{[\rho\alpha,\beta]}$ *stands for the unique positive solution of*

$$z' + \rho\alpha z = \beta, \quad z \in \mathcal{C}_T^1. \tag{2.14}$$

PROOF. Note that any positive solution $w \in \mathcal{C}_T^1$ of (2.1) must satisfy $w \gg 0$, since $w' = (\rho\alpha - \beta w)w$. Moreover, the change of variable $w = z^{-1}$ transforms (2.1) into the linear equation (2.14). Thus, the positive T-periodic solutions of (2.1) are in one-to-one correspondence with the positive solutions of problem (2.14).

Suppose $\rho \neq 0$. Then, (2.2) implies $\rho\hat{\alpha} \neq 0$ and, hence, thanks to Proposition 2.1(a), (2.14) possesses a unique solution, $\psi_{[\rho\alpha,\beta]}$. Moreover, thanks to Proposition 2.1(b)–(c), $\psi_{[\rho\alpha,\beta]} \gg 0$ if $\rho > 0$, while $\psi_{[\rho\alpha,\beta]} \ll 0$ if $\rho < 0$.

Finally, note that if $\rho = 0$, then, due to Proposition 2.1(d), (2.14) cannot admit a positive solution, since

$$\int_0^T e^{\rho \int_0^s \alpha} \beta(s)\, ds = \int_0^T \beta(s)\, ds > 0.$$

Combining these features concludes the proof. □

We now analyze some important comparison properties between subsolutions, supersolutions and solutions of (2.1). First, we will introduce some basic concepts.

DEFINITION 2.5. Let $\rho \in \mathbb{R}$ and $\alpha, \beta \in \mathcal{C}_T$. Then,

(1) a function $w \in \mathcal{C}_T^1$ is said to be a subsolution (resp. strict subsolution) of (2.1) if $w' \leqslant \rho\alpha w - \beta w^2$ (resp. $w' < \rho\alpha w - \beta w^2$);

(2) a function $w \in \mathcal{C}_T^1$ is said to be a supersolution (resp. strict supersolution) of (2.1) if $w' \geqslant \rho\alpha w - \beta w^2$ (resp. $w' > \rho\alpha w - \beta w^2$).

PROPOSITION 2.6. *Suppose* $\rho > 0$ *and* $\alpha, \beta \in \mathcal{C}_T$ *satisfy* $\hat{\alpha} > 0$ *and* $\beta > 0$. *Then, any strict subsolution* $\underline{w} \gg 0$ *of* (2.1) *satisfies* $\underline{w} \ll \theta_{[\rho\alpha,\beta]}$, *while any strict supersolution* $\overline{w} \gg 0$ *satisfies* $\overline{w} \gg \theta_{[\rho\alpha,\beta]}$.

PROOF. Suppose $\underline{w} \gg 0$ is a strict subsolution of (2.1). Then, the function $\underline{z} := 1/\underline{w}$ is a strict supersolution of (2.14) and, hence, it follows from Proposition 2.3(a) that $\underline{z} \gg \psi_{[\rho\alpha,\beta]}$. Therefore,

$$\underline{w} \ll \frac{1}{\psi_{[\rho\alpha,\beta]}} = \theta_{[\rho\alpha,\beta]}.$$

Similarly, if $\overline{w} \gg 0$ is a strict supersolution of (2.1), then $\bar{z} := 1/\overline{w}$ is a strict subsolution of (2.14) and, hence, $\bar{z} \ll \psi_{[\rho\alpha,\beta]}$. Consequently, $\overline{w} \gg \theta_{[\rho\alpha,\beta]}$. □

COROLLARY 2.7. *Let, for each* $j \in \{1,2\}$, $\rho_j > 0$ *and* α_j, $\beta_j \in \mathcal{C}_T$ *satisfying* $\alpha_j > 0$ *and* $\beta_j > 0$. *Suppose*

$$\rho_1 \leqslant \rho_2, \quad \alpha_1 \leqslant \alpha_2, \quad \beta_1 \geqslant \beta_2, \quad (\rho_1,\alpha_1,\beta_1) \neq (\rho_2,\alpha_2,\beta_2). \tag{2.15}$$

Then, $\theta_{[\rho_1,\alpha_1,\beta_1]} \ll \theta_{[\rho_2,\alpha_2,\beta_2]}$.

PROOF. Under conditions (2.15), $\theta_{[\rho_1,\alpha_1,\beta_1]} \gg 0$ is a strict subsolution of

$$w' = \rho_2\alpha_2 w - \beta_2 w^2$$

and the conclusion follows from Proposition 2.6. □

2.3. *Point-wise behaviour of the map* $\rho \mapsto \theta_{[\rho\alpha,\beta]}$

In this section we regard to the unique positive solution of (2.1), $\theta_{[\rho\alpha,\beta]}$, as a map depending on the parameter $\rho > 0$ in order to analyze some sharp properties of the global bifurcation diagram of non-negative solutions of (2.1). Our main result reads as follows.

THEOREM 2.8. *Suppose* $\hat{\alpha} > 0$ *and* $\beta > 0$, *and set*

$$\Theta(\rho) := \theta_{[\rho\alpha,\beta]} \in \mathcal{C}^1_T, \quad \rho > 0. \tag{2.16}$$

Then, the map $(0,\infty) \to \mathcal{C}^1_T$, $\rho \mapsto \Theta(\rho)$, *is real analytic and it possesses the following asymptotic expansion at* $\rho = 0$

$$\Theta(\rho) = \frac{\hat{\alpha}}{\hat{\beta}}\left[\rho + \rho^2 I_2 + \mathrm{O}\left(\rho^3\right)\right] \quad \text{as } \rho \downarrow 0, \tag{2.17}$$

where I_2 *stands for the function defined by*

$$I_2(t) := \int_0^t \alpha - \frac{\hat{\alpha}}{\hat{\beta}} \int_0^t \beta - \frac{1}{T}\left(\int_0^T\int_0^t \alpha - \frac{\hat{\alpha}}{\hat{\beta}}\int_0^T\int_0^t \beta\right), \quad t \in \mathbb{R}.$$

In particular,

$$\lim_{\rho\downarrow 0} \left\|\Theta(\rho)\right\|_{\mathcal{C}_T} = 0. \tag{2.18}$$

If, in addition, $\alpha > 0$, *then the map* $(0,\infty) \to \mathcal{C}_T$, $\rho \mapsto \Theta(\rho)$, *is strongly increasing in the sense that* $\Theta(\rho) \gg \Theta(\tilde{\rho})$ *if* $\rho > \tilde{\rho}$. *Further, if* $\alpha \gg 0$ *and* $\beta \gg 0$, *then, for each* $\rho > 0$,

$$\frac{\alpha_L}{\beta_M}\,\rho \leqslant \Theta(\rho) \leqslant \frac{\alpha_M}{\beta_L}\,\rho \tag{2.19}$$

and

$$\lim_{\rho\uparrow\infty} \frac{\Theta(\rho)}{\rho} = \frac{\alpha}{\beta} \quad \text{in } \mathcal{C}_T. \tag{2.20}$$

PROOF. Note that the solutions of (2.1) are the zeroes of the non-linear operator $\mathfrak{F} : \mathbb{R} \times \mathcal{C}_T^1 \to \mathcal{C}_T$ defined by

$$\mathfrak{F}(\rho, w) := w' - \rho\alpha w + \beta w^2, \tag{2.21}$$

which is real analytic. Fix $\rho_0 > 0$. Then,

$$\mathfrak{F}\bigl(\rho_0, \Theta(\rho_0)\bigr) = 0$$

and $D_u\mathfrak{F}(\rho_0, \Theta(\rho_0)) \in \mathcal{L}(\mathcal{C}_T^1; \mathcal{C}_T)$ is given through

$$D_u\mathfrak{F}\bigl(\rho_0, \Theta(\rho_0)\bigr) := \frac{\mathrm{d}}{\mathrm{d}t} + 2\beta\Theta(\rho_0) - \rho_0\alpha. \tag{2.22}$$

Since $\Theta(\rho_0) \gg 0$ satisfies

$$\frac{\mathrm{d}\Theta(\rho_0)}{\mathrm{d}t} = \rho_0\alpha\Theta(\rho_0) - \beta\bigl[\Theta(\rho_0)\bigr]^2,$$

division by $\Theta(\rho_0)$ and integration in $[0,T]$ gives

$$0 = \operatorname{Log}\frac{\Theta(\rho_0)(T)}{\Theta(\rho_0)(0)} = \int_0^T \bigl[\rho_0\alpha - \beta\Theta(\rho_0)\bigr] = 0$$

and, hence,

$$\int_0^T \bigl[2\beta\Theta(\rho_0) - \rho_0\alpha\bigr] = \int_0^T \beta(t)\Theta(\rho_0)(t)\,\mathrm{d}t > 0,$$

since $\beta > 0$. Thus, thanks to Proposition 2.1, $D_u\mathfrak{F}(\rho_0, \Theta(\rho_0))$ is a linear isomorphism between $\mathcal{C}_T^1$ and $\mathcal{C}_T$. Therefore, thanks to the implicit function theorem, there exist $\varepsilon \in (0,\rho_0)$, $\delta > 0$, and a unique real analytic map

$$W : (\rho_0 - \varepsilon, \rho_0 + \varepsilon) \to \bigl\{w \in \mathcal{C}_T^1\colon\ \|w - \Theta(\rho_0)\|_{\mathcal{C}_T^1} < \delta\bigr\}$$

such that $W(\rho_0) = \Theta(\rho_0)$, $\mathfrak{F}(\rho, W(\rho)) = 0$ if $|\rho - \rho_0| < \varepsilon$, and $w = W(\rho)$ if $\mathfrak{F}(\rho, w) = 0$ with $|\rho - \rho_0| < \varepsilon$ and $\|w - \Theta(\rho_0)\|_{\mathcal{C}^1_T} < \delta$. As $W(\rho_0) = \Theta(\rho_0) \gg 0$, necessarily $W(\rho) \gg 0$ for $\rho \sim \rho_0$, and, therefore, $W(\rho) = \Theta(\rho)$, by the uniqueness of the positive solution of (2.1). This shows that $\rho \mapsto \Theta(\rho)$ is real analytic.

Although the fact that the map $(0, \infty) \to \mathcal{C}_T$, $\rho \mapsto \Theta(\rho)$, is strictly increasing if $\alpha > 0$ is a consequence from Corollary 2.7, it should be noted that implicit differentiation in

$$\mathfrak{F}\big(\rho, \Theta(\rho)\big) = 0$$

gives rise to

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + 2\beta\Theta(\rho) - \rho\alpha\right)\frac{\mathrm{d}\Theta}{\mathrm{d}\rho}(\rho) = \alpha\Theta(\rho) > 0$$

and, hence,

$$\frac{\mathrm{d}\Theta}{\mathrm{d}\rho}(\rho) = \left(\frac{\mathrm{d}}{\mathrm{d}t} + 2\beta\Theta(\rho) - \rho\alpha\right)^{-1}\big(\alpha\Theta(\rho)\big) \gg 0,$$

because of (2.11).

To prove (2.17) we will use a celebrated device introduced by M.G. Crandall and P.H. Rabinowitz [6]. Note that

$$D_u\mathfrak{F}(0, 0) = \frac{\mathrm{d}}{\mathrm{d}t} : \mathcal{C}^1_T \to \mathcal{C}_T$$

and, hence,

$$N\big[D_u\mathfrak{F}(0, 0)\big] = \operatorname{span}[1].$$

Moreover,

$$\mathcal{C}^1_T = N\big[D_u\mathfrak{F}(0, 0)\big] \oplus Y, \quad Y := \left\{y \in \mathcal{C}^1_T \colon \int_0^T y = 0\right\}.$$

Indeed, any $w \in \mathcal{C}^1_T$ admits a unique decomposition as

$$w = x + y, \quad x \in \mathbb{R},\ y \in Y.$$

Actually,

$$x = \hat{w}, \quad y = w - \hat{w}.$$

Now, consider the auxiliary operator $\mathfrak{G} : \mathbb{R}^2 \times Y \to \mathcal{C}_T$ defined by

$$\mathfrak{G}(s, \rho, y) := y' - \rho\alpha(1 + y) + \beta s(1 + y)^2, \quad (s, \rho, y) \in \mathbb{R}^2 \times Y. \tag{2.23}$$

It should be noted that $\mathfrak{G}(s,\rho,y)=0$ implies $\mathfrak{F}(\rho,s(1+y))=0$. It is rather clear that $\mathfrak{G}$ is a real analytic operator such that $\mathfrak{G}(0,0,0)=0$. Moreover, the operator $D_{(\rho,y)}\mathfrak{G}(0,0,0):\mathbb{R}\times Y\to\mathcal{C}_T$ is given by

$$D_{(\rho,y)}\mathfrak{G}(0,0,0)(\rho,y)=y'-\rho\alpha,\quad (\rho,y)\in\mathbb{R}\times Y,$$

and, so, it is a linear isomorphism. Indeed, if for some $(\rho,y)\in\mathbb{R}\times Y$ one has that $y'=\rho\alpha$, then integrating in $[0,T]$ gives $\rho\hat{\alpha}=0$ and, hence, $\rho=0$, since $\hat{\alpha}>0$. Thus, $y'=0$ and, therefore, y is constant. Consequently, as it has zero average, we obtain that $y=0$, which concludes the proof of the injectivity. Now, fix $f\in\mathcal{C}_T$. Then, it is straightforward to check that

$$y(t):=-\frac{\hat{f}}{\hat{\alpha}}\int_0^t\alpha+\int_0^t f-\frac{1}{T}\left(-\frac{\hat{f}}{\hat{\alpha}}\int_0^T\int_0^t\alpha+\int_0^T\int_0^t f\right),\quad t\in\mathbb{R},$$

provides us with the unique element of Y solving

$$D_{(\rho,y)}\mathfrak{G}(0,0,0)(\rho,y)=f,\quad \rho:=-\frac{\hat{f}}{\hat{\alpha}}.$$

Therefore, thanks to the implicit function theorem, there exist $s_0>0$, $\delta>0$, and a unique real analytic map

$$(\rho,y):(-s_0,s_0)\to(-\delta,\delta)\times\left\{y\in Y:\ \|y\|_{\mathcal{C}_T^1}<\delta\right\}$$

such that $(\rho(0),y(0))=(0,0)$, $\mathfrak{G}(s,\rho(s),y(s))=0$ for each $s\in(-s_0,s_0)$, and $(\rho,y)=(\rho(s),y(s))$ if $\mathfrak{G}(s,\rho,y)=0$ with $|s|<s_0$, $|\rho|<\delta$, and $\|y\|_{\mathcal{C}_T^1}<\delta$. In particular,

$$\mathfrak{F}\big(\rho(s),s\big(1+y(s)\big)\big)=0,\quad |s|<s_0,$$

and, so, for sufficiently small $s>0$,

$$(\rho,w):=\big(\rho(s),s\big(1+y(s)\big)\big)$$

provides us with a positive T-periodic solution of (2.1). Necessarily, for sufficiently small $s>0$, we have that

$$s\big(1+y(s)\big)=\Theta\big(\rho(s)\big)\tag{2.24}$$

because of the uniqueness of the positive solution of (2.1). Now, in order to find the lowest order terms of the local developments

$$\rho(s)=\sum_{n=1}^{\infty}\rho_n s^n,\quad y(s)=\sum_{n=1}^{\infty}s^n y_n,\quad y_n\in Y,\ n\geqslant 1,\tag{2.25}$$

we substitute (2.25) into the equation of the zeroes of $\mathfrak{G}$. Then, we get

$$\sum_{n=1}^{\infty} s^n y_n' = \left[\alpha \sum_{n=1}^{\infty} \rho_n s^n - s\beta\left(1 + \sum_{n=1}^{\infty} s^n y_n\right)\right]\left(1 + \sum_{n=1}^{\infty} s^n y_n\right)$$

and identifying terms of order one in both members gives

$$y_1' = \rho_1 \alpha - \beta.$$

Thus,

$$\rho_1 = \frac{\hat{\beta}}{\hat{\alpha}}, \quad y_1(t) = \rho_1 \int_0^t \alpha - \int_0^t \beta - \frac{1}{T}\left(\rho_1 \int_0^T \int_0^t \alpha - \int_0^T \int_0^t \beta\right)$$

and, therefore, (2.24) shows that

$$\rho(s) = \rho_1 s + \mathrm{O}(s^2) \quad \text{and} \quad \Theta(\rho(s)) = s\left(1 + s\, y_1(t) + \mathrm{O}(s^2)\right) \tag{2.26}$$

as $s \downarrow 0$. Thanks again to the implicit function theorem, one can eliminate s as a function of ρ at $\rho = 0$ from the identity

$$\rho_1 s + \mathrm{O}(s^2) = \rho.$$

Actually,

$$s = \frac{1}{\rho_1}\rho + \mathrm{O}(\rho^2) = \frac{\hat{\alpha}}{\hat{\beta}}\rho + \mathrm{O}(\rho^2),$$

and, so, substituting this expansion into the second identity of (2.26) shows that

$$\Theta(\rho) = \frac{\hat{\alpha}}{\hat{\beta}}\left[\rho + \frac{\hat{\alpha}}{\hat{\beta}} y_1 \rho^2 + \mathrm{O}(\rho^3)\right]$$

as $\rho \downarrow 0$, which provides us with the asymptotic expansion (2.17).

Subsequently, we suppose that $\alpha \gg 0$ and $\beta \gg 0$. The global estimates (2.19) and (2.20) are consequences from Proposition 2.6. Indeed, setting $\Theta := \Theta(\rho)$, we have that

$$\rho\alpha_L \Theta - \beta_M \Theta^2 \leqslant \Theta' = \rho\alpha\Theta - \beta\Theta^2 \leqslant \rho\alpha_M \Theta - \beta_L \Theta^2,$$

and, hence, for each $\rho > 0$, $\Theta(\rho)$ is a positive subsolution of $w' = \rho\alpha_M w - \beta_L w^2$ and a positive supersolution of $w' = \rho\alpha_L w - \beta_M w^2$. Therefore, (2.19) follows from Proposition 2.6 by taking into account that $\frac{\alpha_M}{\beta_L}\rho$ and $\frac{\alpha_L}{\beta_M}\rho$ provide us with the unique T-periodic positive solutions of each of these autonomous equations.

Now, fix $\delta > 0$ sufficiently small so that $\frac{\alpha}{\beta} \gg \delta$, consider two functions $\underline{\theta}$, $\bar{\theta} \in \mathcal{C}^1_T$ such that

$$\frac{\alpha}{\beta} - \delta < \underline{\theta} < \frac{\alpha}{\beta} - \frac{\delta}{2}, \qquad \frac{\alpha}{\beta} + \frac{\delta}{2} < \bar{\theta} < \frac{\alpha}{\beta} + \delta, \tag{2.27}$$

and set

$$\underline{w} := \rho\,\underline{\theta}, \qquad \overline{w} := \rho\,\bar{\theta}.$$

Then, for each sufficiently large $\rho > 0$, say $\rho \geqslant \rho_0$, we have that

$$\rho\alpha\overline{w} - \beta\overline{w}^2 = \rho^2\beta\bar{\theta}\left(\frac{\alpha}{\beta} - \bar{\theta}\right) < -\frac{\delta}{2}\rho^2\beta\bar{\theta} \leqslant \rho\bar{\theta}' = \overline{w}',$$

and

$$\rho\alpha\underline{w} - \beta\underline{w}^2 = \rho^2\beta\underline{\theta}\left(\frac{\alpha}{\beta} - \underline{\theta}\right) > \frac{\delta}{2}\rho^2\beta\underline{\theta} \geqslant \rho\underline{\theta}' = \underline{w}'.$$

Consequently, for each $\rho \geqslant \rho_0$, $\underline{w}$ and $\overline{w}$ provide with a strict subsolution and a strict supersolution of (2.1), respectively, and, hence, thanks to Proposition 2.6,

$$\underline{w} = \rho\,\underline{\theta} \ll \Theta(\rho) \ll \overline{w} = \rho\,\bar{\theta}.$$

Thus, it follows from (2.27) that, for each $\rho \geqslant \rho_0$,

$$\frac{\alpha}{\beta} - \delta \ll \frac{\Theta(\rho)}{\rho} \ll \frac{\alpha}{\beta} + \delta.$$

Therefore, passing to the limit as $\rho \uparrow \infty$, shows that

$$\frac{\alpha}{\beta} - \delta \leqslant \liminf_{\rho\uparrow\infty} \frac{\Theta(\rho)}{\rho} \leqslant \limsup_{\rho\uparrow\infty} \frac{\Theta(\rho)}{\rho} \leqslant \frac{\alpha}{\beta} + \delta.$$

As these inequalities are valid for any $\delta > 0$, the proof of (2.20) is concluded, as well as the proof of the theorem. □

It should be noted that the strong monotonicity of $\Theta(\rho)$ is lost when α changes of sign. Indeed, suppose that, for some sufficiently small $\delta > 0$, $\beta(t) = 0$ and $\alpha(t) < 0$ for each $t \in [0, \delta]$. Then, for each $t \in [0, \delta]$ we have that

$$\theta_{[\rho\alpha,\beta]}(t) = \mathrm{e}^{\rho\int_0^t \alpha}\theta_{[\rho\alpha,\beta]}(0).$$

Since $\int_0^t \alpha < 0$, the map $\rho \mapsto \theta_{[\rho\alpha,\beta]}(t)$ is decreasing for each $t \in (0, \delta]$.

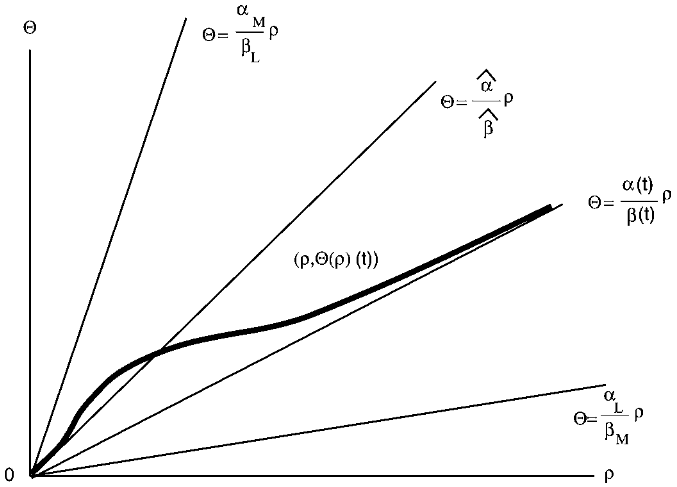


Fig. 1. Graph of $\rho \mapsto \Theta(\rho)(t)$ in case $\alpha \gg 0$ and $\beta \gg 0$.

In Fig. 1 we have represented the bifurcation diagram $(\rho, \Theta(\rho)(t))$ for some $t \in \mathbb{R}$ in the (ρ, Θ)-plane in the special, though important, case when $\alpha \gg 0$ and $\beta \gg 0$. Thanks to (2.19) the curve $\rho \mapsto \Theta(\rho)(t)$ must be confined in between the straight lines $\Theta = \frac{\alpha_L}{\beta_M}\rho$ and $\Theta = \frac{\alpha_M}{\beta_L}\rho$, and, due to (2.17) and (2.20), its tangents at $\rho = 0$ and $\rho = \infty$ are

$$\Theta = \frac{\hat{\alpha}}{\hat{\beta}}\rho \quad \text{and} \quad \Theta = \frac{\alpha(t)}{\beta(t)}\rho,$$

respectively.

2.4. *Attractiveness of* $\theta_{[\rho\alpha,\beta]}$

Note that the linearized equations of (2.1) at $w = 0$ and $w = \Theta(\rho)$ are given by

$$w' = \rho\alpha w \quad \text{and} \quad w' = \big[\rho\alpha - 2\beta\Theta(\rho)\big]w,$$

respectively, whose Floquet multipliers are

$$\mathrm{e}^{\rho\int_0^T \alpha} \quad \text{and} \quad \mathrm{e}^{\int_0^T [\rho\alpha - 2\beta\Theta(\rho)]}.$$

As a result, $w = 0$ is linearly stable if $\rho < 0$, linearly neutrally stable if $\rho = 0$ and linearly unstable if $\rho > 0$. Moreover, in the proof of Theorem 2.8 we have seen that, for each $\rho > 0$,

$$\int_0^T \big[\rho\alpha - \beta\Theta(\rho)\big] = 0, \tag{2.28}$$

and, hence,

$$\mathrm{e}^{\int_0^T [\rho\alpha - 2\beta\Theta(\rho)]} = \mathrm{e}^{-\int_0^T [\beta\Theta(\rho)]} < 1.$$

Therefore, $\Theta(\rho)$ is linearly stable. Actually, it is easy to see that 0 is a global attractor of (2.1) if $\rho \leqslant 0$, whereas $\Theta(\rho)$ attracts to all positive solutions of (2.1) if $\rho > 0$.

3. Linearized stability of the semi-trivial solutions

Thanks to Proposition 2.4, (1.2) possesses a semi-trivial state—throughout the remaining of the chapter, positive—of the form $(u, 0)$ if and only if $\lambda > 0$. Moreover, in such case, $(\theta_{[\lambda\ell,a]}, 0)$ is the unique semi-trivial state of this form. Similarly, $(0, \theta_{[\mu m,d]})$ provides us with the unique semi-trivial state of (2.1) of the form $(0, v)$, and it exists if and only if $\mu > 0$. The following result characterizes the stability of these states and the stability of the trivial state $(0, 0)$.

PROPOSITION 3.1. *The following assertions are true*:

(a) $(0, 0)$ *is linearly stable if* $\lambda < 0$ *and* $\mu < 0$, *and it is linearly unstable if either* $\lambda > 0$ *or* $\mu > 0$.

(b) *Suppose* $\lambda > 0$. *Then,* $(\theta_{[\lambda\ell,a]}, 0)$ *is linearly unstable if and only if*

$$\mu > \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t, \tag{3.1}$$

and linearly stable if and only if

$$\mu < \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t. \tag{3.2}$$

Consequently, the **curve of change of stability** *of* $(\theta_{[\lambda\ell,a]}, 0)$ *in the* (λ, μ)*-plane is given through its curve of neutral stability*

$$\mu = f(\lambda) := \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t. \tag{3.3}$$

(c) *Suppose* $\mu > 0$. *Then,* $(0, \theta_{[\mu m,d]})$ *is linearly unstable if and only if*

$$\lambda > \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t, \tag{3.4}$$

and linearly stable if and only if

$$\lambda < \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t. \tag{3.5}$$

Consequently, the **curve of change of stability** *of* $(0, \theta_{[\mu m,d]})$ *in the* (λ, μ)*-plane is given through its curve of neutral stability*

$$\lambda = g(\mu) := \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t. \tag{3.6}$$

PROOF. The linearization of (1.2) at any solution $(u_0, v_0) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ is given by

$$\begin{cases} u' = (\lambda\,\ell - 2au_0 - bv_0)u - bu_0v, \\ v' = -cv_0u + (\mu\,m - 2dv_0 - cu_0)v. \end{cases} \tag{3.7}$$

When $(u_0, v_0) = (0, 0)$, (3.7) becomes into

$$\begin{cases} u' = \lambda\ell u, \\ v' = \mu m v, \end{cases}$$

whose Poincaré map is given by

$$\mathfrak{P}(x, y) = \begin{pmatrix} \mathrm{e}^{\lambda T} & 0 \\ 0 & \mathrm{e}^{\mu T} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \quad (x, y) \in \mathbb{R}^2,$$

since $\hat{\ell} = \hat{m} = 1$. Thus, its Floquet multipliers, $\mathrm{e}^{\lambda T}$ and $\mathrm{e}^{\mu T}$, are less than one if $\lambda < 0$ and $\mu < 0$, while some of them is greater than one if either $\lambda > 0$ or $\mu > 0$, which concludes the proof of part (a).

When $(u_0, v_0) = (\theta_{[\lambda\ell,a]}, 0)$, (3.7) becomes into

$$\begin{cases} u' = (\lambda\,\ell - 2a\theta_{[\lambda\ell,a]})u - b\theta_{[\lambda\ell,a]}v, \\ v' = (\mu\,m - c\theta_{[\lambda\ell,a]})v, \end{cases} \tag{3.8}$$

and its associated Poincaré map is given by

$$\mathfrak{P}(x, y) = \begin{pmatrix} \mathrm{e}^{\int_0^T (\lambda\ell - 2a\theta_{[\lambda\ell,a]})} & E \\ 0 & \mathrm{e}^{\int_0^T (\mu m - c\theta_{[\lambda\ell,a]})} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \quad (x, y) \in \mathbb{R}^2,$$

for some constant $E \in \mathbb{R}$ whose explicit knowledge is not important here. Thanks to (2.28),

$$\int_0^T (\lambda\ell - a\theta_{[\lambda\ell,a]}) = 0$$

and, hence,

$$0 < \mathrm{e}^{\int_0^T (\lambda\ell - 2a\theta_{[\lambda\ell,a]})} = \mathrm{e}^{-\int_0^T a(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t} < 1.$$

Consequently, the attractive character of $(\theta_{[\lambda\ell,a]},0)$ is given by the remaining Floquet multiplier

$$\nu := \mathrm{e}^{\int_0^T(\mu m - c\theta_{[\lambda\ell,a]})} = \mathrm{e}^{\mu T - \int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t}.$$

As $\nu > 1$ when (3.1) is satisfied, $\nu = 1$ if (3.3) is satisfied, and $0 < \nu < 1$ under condition (3.2), the proof of part (b) is concluded. Part (c) follows by symmetry. This ends the proof. □

4. The curves of change of stability of the semi-trivial positive solutions

In this section we shall study some global properties of the curves of neutral stability (3.3) and (3.6) of the semi-trivial solutions of (1.2). Such analysis is imperative in order to ascertain the shape of each of the following regions within the parameter space (λ,μ):

$$\begin{aligned} \mathrm{R}_{\mathrm{uu}} &:= \bigl\{(\lambda,\mu)\in(0,\infty)^2\colon\ (\theta_{[\lambda\ell,a]},0) \text{ and } (0,\theta_{[\mu m,d]}) \text{ are l.u.}\bigr\},\\ \mathrm{R}_{\mathrm{ss}} &:= \bigl\{(\lambda,\mu)\in(0,\infty)^2\colon\ (\theta_{[\lambda\ell,a]},0) \text{ and } (0,\theta_{[\mu m,d]}) \text{ are l.s.}\bigr\}, \end{aligned} \tag{4.1}$$

$$\begin{aligned} \mathrm{R}_{\mathrm{su}} &:= \bigl\{(\lambda,\mu)\in(0,\infty)^2\colon\ (\theta_{[\lambda\ell,a]},0) \text{ is l.s., } (0,\theta_{[\mu m,d]}) \text{ is l.u.}\bigr\},\\ \mathrm{R}_{\mathrm{us}} &:= \bigl\{(\lambda,\mu)\in(0,\infty)^2\colon\ (\theta_{[\lambda\ell,a]},0) \text{ is l.u., } (0,\theta_{[\mu m,d]}) \text{ is l.s.}\bigr\}, \end{aligned} \tag{4.2}$$

$$\begin{aligned} \mathrm{R}_{\mathrm{u}\star} &:= \bigl\{(\lambda,\mu)\in(0,\infty)\times(-\infty,0]\colon\ (\theta_{[\lambda\ell,a]},0) \text{ is l.u.}\bigr\},\\ \mathrm{R}_{\mathrm{s}\star} &:= \bigl\{(\lambda,\mu)\in(0,\infty)\times(-\infty,0]\colon\ (\theta_{[\lambda\ell,a]},0) \text{ is l.s.}\bigr\}, \end{aligned} \tag{4.3}$$

$$\begin{aligned} \mathrm{R}_{\star\mathrm{u}} &:= \bigl\{(\lambda,\mu)\in(-\infty,0]\times(0,\infty)\colon\ (0,\theta_{[\mu m,d]}) \text{ is l.u.}\bigr\},\\ \mathrm{R}_{\star\mathrm{s}} &:= \bigl\{(\lambda,\mu)\in(-\infty,0]\times(0,\infty)\colon\ (0,\theta_{[\mu m,d]}) \text{ is l.s.}\bigr\}, \end{aligned} \tag{4.4}$$

where *l.u.* stands for *linearly unstable* and *l.s.* for *linearly stable*. As the global behaviour of the curves of neutral stability of the semi-trivial states is strongly based upon the nature of the model under study, i.e., under the sign properties of the interaction coefficient functions $b(t)$ and $c(t)$, in the subsequent discussion we will have to distinguish several different cases. We begin the section by giving a general result which is a straightforward consequence from Theorem 2.8.

THEOREM 4.1. *Suppose $\alpha,\beta,\gamma\in\mathcal{C}_T$ satisfy*

$$\alpha \gg 0,\quad \beta\gg 0,\quad \hat\alpha = \hat\beta = 1, \tag{4.5}$$

and consider the function

$$N(\rho) := \frac{1}{T}\int_0^T \gamma(t)\theta_{[\rho\alpha,\beta]}(t)\,\mathrm{d}t,\quad \rho>0. \tag{4.6}$$

Then,

$$N(\rho) = \hat{\gamma}\,\rho + \frac{1}{T}\int_0^T \gamma(t) I_2(t)\,\mathrm{d}t\,\rho^2 + \mathrm{O}\big(\rho^3\big) \quad \text{as } \rho \downarrow 0, \tag{4.7}$$

where

$$I_2(t) := \int_0^t (\alpha - \beta) - \frac{1}{T}\int_0^T\int_0^t (\alpha - \beta), \quad t \in \mathbb{R}, \tag{4.8}$$

and

$$\lim_{\rho\uparrow\infty} \frac{N(\rho)}{\rho} = \frac{1}{T}\int_0^T \gamma(t)\frac{\alpha(t)}{\beta(t)}\,\mathrm{d}t. \tag{4.9}$$

Moreover, $\rho \mapsto N(\rho)$ *is increasing if* $\gamma > 0$*, decreasing if* $\gamma < 0$*, and, in general, is far from monotone if* γ *changes of sign. In such case,* (4.9) *can be either negative, or positive, or zero, according to the value of the average on its right-hand side.*

PROOF. By (4.5), it follows from Theorem 2.8 that

$$\theta_{[\rho\alpha,\beta]} = \rho + I_2\,\rho^2 + \mathrm{O}\big(\rho^3\big) \quad \text{as } \rho \downarrow 0.$$

Thus,

$$N(\rho) := \frac{1}{T}\int_0^T \gamma(t)\theta_{[\rho\alpha,\beta]}(t)\,\mathrm{d}t = \hat{\gamma}\,\rho + \frac{1}{T}\int_0^T \gamma(t) I_2(t)\,\mathrm{d}t\rho^2 + \mathrm{O}\big(\rho^3\big),$$

which is (4.7). The validity of (4.9) is an immediate consequence from (2.20). The remaining assertions of the theorem are obvious, since $\theta_{[\rho\alpha,\beta]} \gg 0$. □

It should be noted that, in general, nothing can be said about the concavity of $N(\rho)$ at $\rho = 0$. Indeed, since I_2 changes of sign if $I_2 \neq 0$, because $\hat{I}_2 = 0$, the average of the function γI_2 can have any sign we wish by choosing and adequate γ that may be taken either positive, or negative, or changing of sign, in strong contrast with the autonomous case, where $N(\rho) = \gamma\rho$ for every $\rho > 0$.

4.1. *The species u and v compete* ($b > 0$ *and* $c > 0$)

Besides (1.7), throughout this section we assume that

$$\ell \gg 0, \quad m \gg 0, \quad a \gg 0, \quad d \gg 0, \quad b > 0, \quad c > 0, \tag{4.10}$$

though most of our discussion can be easily adapted to cover much weaker requirements. Under assumptions (4.10), the curves of neutral stability of the semi-trivial states

$(\theta_{[\lambda\ell,a]},0)$ and $(0,\theta_{[\mu m,d]})$, (3.3) and (3.6), are increasing in λ and μ, respectively, and satisfy

$$f(\lambda)>0,\quad g(\mu)>0,\quad (\lambda,\mu)\in(0,\infty)^2. \tag{4.11}$$

It should be noted that these are the curves delimiting the regions defined by (4.1) and (4.2). Thanks to Theorem 4.1, the tangents of these curves at $(\lambda,\mu)=(0,0)$ are given by

$$\mu=\hat{c}\lambda \quad\text{and}\quad \lambda=\hat{b}\mu, \tag{4.12}$$

respectively, as for the autonomous counterpart of (1.2)—where these are the curves of neutral stability of $(\lambda,0)$ and $(0,\mu)$—and their tangents at infinity are given by

$$\begin{aligned} \mu&=\alpha\lambda, \qquad \alpha:=\frac{1}{T}\int_0^T c(t)\frac{\ell(t)}{a(t)}\,\mathrm{d}t,\\ \lambda&=\beta\mu, \qquad \beta:=\frac{1}{T}\int_0^T b(t)\frac{m(t)}{d(t)}\,\mathrm{d}t, \end{aligned} \tag{4.13}$$

respectively. In strong contrast with the autonomous case—where (4.12) and (4.13) are the same straight lines—as a consequence from the non-autonomous character of (1.2), the *relative positions* of the straight lines (4.12) and (4.13) can be inter-exchanged at $(\lambda,\mu)=(0,0)$ and at infinity, so entailing the existence of a crossing point between the curves of change of stability of the semi-trivial states; a feature of great significance from the point of view of the applications, as it might allow us designing environments where none of the competitors is driven to extinction by the other. Most precisely, there are admissible choices of the function coefficients a, b, c, d, ℓ and m, for which either

$$\hat{b}\hat{c}<1 \quad\text{and}\quad \alpha\beta=\frac{1}{T}\int_0^T c(t)\frac{\ell(t)}{a(t)}\,\mathrm{d}t\cdot\frac{1}{T}\int_0^T b(t)\frac{m(t)}{d(t)}\,\mathrm{d}t>1, \tag{4.14}$$

or else

$$\hat{b}\hat{c}>1 \quad\text{and}\quad \alpha\beta=\frac{1}{T}\int_0^T c(t)\frac{\ell(t)}{a(t)}\,\mathrm{d}t\cdot\frac{1}{T}\int_0^T b(t)\frac{m(t)}{d(t)}\,\mathrm{d}t<1. \tag{4.15}$$

In such circumstances, as the curves $\mu=f(\lambda)$, $\lambda>0$, and $\lambda=g(\mu)$, $\mu>0$, are real analytic, they must cross at most at a finite number of points—at least one—whose total sum of contact orders must be an odd integer number. To construct examples satisfying any of these requirements one can make the following choices:

$$a=d=1,\quad \ell:=1+\ell_1\sin\left(\frac{2\pi}{T}\,\cdot\right),\quad m:=1+m_1\sin\left(\frac{2\pi}{T}\,\cdot\right), \tag{4.16}$$

for some constants $\ell_1, m_1\in(-1,1)$, and

$$b:=b_0+b_1\sin\left(\frac{2\pi}{T}\,\cdot\right),\quad c:=c_0+c_1\sin\left(\frac{2\pi}{T}\,\cdot\right), \tag{4.17}$$

for some constants

$$b_0 > 0, \quad c_0 > 0, \quad b_1 \in (-b_0, b_0), \quad c_1 \in (-c_0, c_0).$$

For these choices, we have that

$$\hat{b} = b_0, \quad \hat{c} = c_0,$$

and

$$\alpha = \frac{1}{T}\int_0^T c(t)\frac{\ell(t)}{a(t)}\,\mathrm{d}t = c_0 + \frac{1}{2}c_1\ell_1, \quad \beta = \frac{1}{T}\int_0^T b(t)\frac{m(t)}{d(t)}\,\mathrm{d}t = b_0 + \frac{1}{2}b_1 m_1.$$

Thus, (4.14) and (4.15) are equivalent to

$$b_0c_0 < 1 \quad \text{and} \quad \left(c_0 + \frac{1}{2}c_1\ell_1\right)\left(b_0 + \frac{1}{2}b_1m_1\right) > 1, \tag{4.18}$$

and

$$b_0c_0 > 1 \quad \text{and} \quad \left(c_0 + \frac{1}{2}c_1\ell_1\right)\left(b_0 + \frac{1}{2}b_1m_1\right) < 1, \tag{4.19}$$

respectively. We now show that any of these conditions can be reached by an adequate choice of the several constants involved in their settings. Take, for sufficiently small $\varepsilon > 0$,

$$b_0 = c_0 = 1 - \varepsilon, \quad b_1 = c_1 = 1 - 2\varepsilon.$$

Then, the function $p(\varepsilon)$ defined by

$$p(\varepsilon) := \left[1 - \varepsilon + \frac{1}{2}(1 - 2\varepsilon)\ell_1\right]\left[1 - \varepsilon + \frac{1}{2}(1 - 2\varepsilon)m_1\right], \quad \varepsilon > 0,$$

satisfies

$$p(0) = 1 + \frac{1}{2}(\ell_1 + m_1) + \frac{1}{4}\ell_1 m_1$$

and, hence, $p(0) > 1$ if $\ell_1 > 0$ and $m_1 > 0$. Thus, if we choose ℓ_1, $m_1 \in (0, 1)$, then, for sufficiently small $\varepsilon > 0$, we have that $p(\varepsilon) > 1$, by continuity, and, consequently, (4.18) is satisfied. On the contrary, choosing

$$b_0 = c_0 = 1 + \varepsilon, \quad b_1 = c_1 = 1,$$

the auxiliary function $n(\varepsilon)$ defined by

$$n(\varepsilon) := \left[1 + \varepsilon + \frac{1}{2}\ell_1\right]\left[1 + \varepsilon + \frac{1}{2}m_1\right], \quad \varepsilon > 0,$$

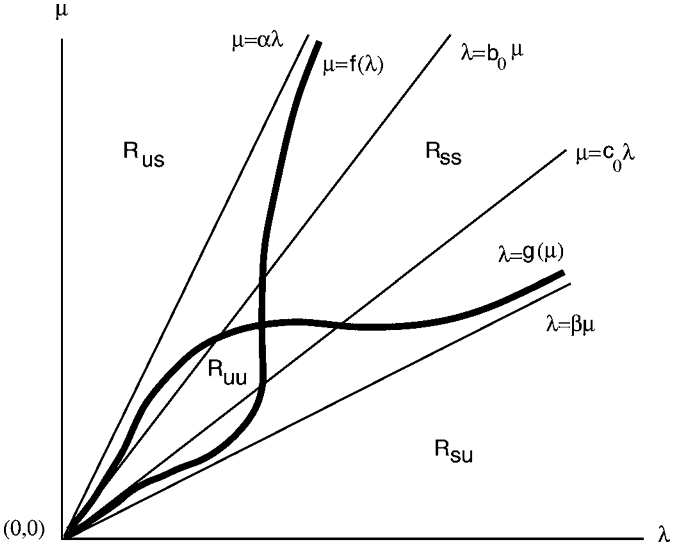


Fig. 2. The curves $\mu = f(\lambda)$ and $\lambda = g(\mu)$ in case (4.18).

satisfies

$$n(0) = 1 + \frac{1}{2}(\ell_1 + m_1) + \frac{1}{4}\ell_1 m_1$$

and, hence, $n(0) < 1$ if $\ell_1 < 0$ and $m_1 < 0$. Thus, if we choose $\ell_1, m_1 \in (-1, 0)$, then, for sufficiently small $\varepsilon > 0$, we have that $n(\varepsilon) < 1$, by continuity, and, consequently, (4.19) holds.

In Fig. 2 we have represented the curves of neutral stability of the semi-trivial states in an admissible situation satisfying (4.18). It should be noted that, for the previous choice, we have that, for each sufficiently small $\varepsilon > 0$,

$$\frac{1}{\beta} = \left(b_0 + \frac{1}{2}b_1 m_1\right)^{-1} = \left(1 - \varepsilon + \frac{1}{2}(1 - 2\varepsilon)m_1\right)^{-1} < c_0 = 1 - \varepsilon < 1$$

and

$$\alpha = c_0 + \frac{1}{2}c_1\ell_1 = 1 - \varepsilon + \frac{1}{2}(1 - 2\varepsilon)\ell_1 > \frac{1}{b_0} = \frac{1}{1 - \varepsilon} > 1$$

since $\ell_1 > 0$ and $m_1 > 0$. As illustrated by Fig. 2, the curves of change of stability of the semi-trivial states must cross at least once, and at most at a finite number of points whose total sum of contact orders must be odd. In the situations described by Fig. 2, one among all admissible ones, these curves divide the first quadrant into four different regions: the two regions enclosed by these curves, and the two exterior regions. In general,

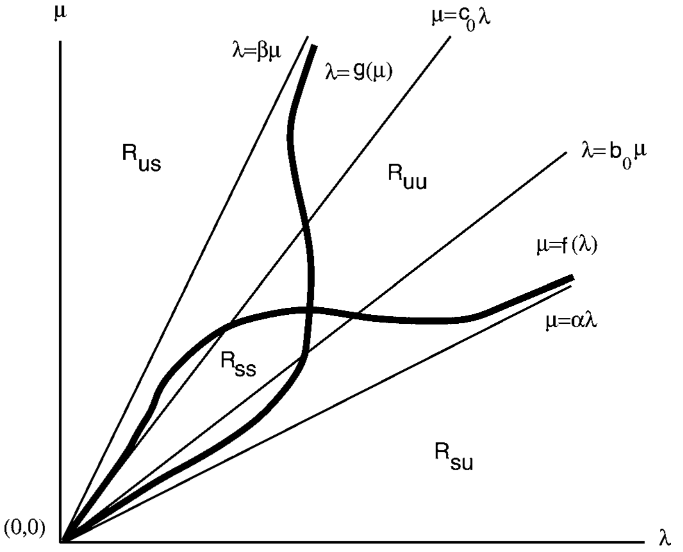


Fig. 3. The curves $\mu = f(\lambda)$ and $\lambda = g(\mu)$ in case (4.19).

the region enclosed between them consists of a finite number of components—one unbounded, and the remaining bounded. Within the bounded components both semi-trivial states are—alternatively—either linearly unstable, or linearly stable. In Fig. 2, both are linearly unstable, because $\mu > f(\lambda)$ and $\lambda > g(\mu)$. Hence, this is the region $\mathrm{R_{uu}}$ defined in (4.1). Independently of the number of bounded components enclosed by the curves of neutral stability of the semi-trivial states, the unbounded component is always a subset of the region $\mathrm{R_{ss}}$ defined in (4.1), where both semi-trivial states are linearly stable. The remaining regions correspond with the regions $\mathrm{R_{su}}$ and $\mathrm{R_{us}}$ defined in (4.2), where one of the semi-trivial states is linearly stable and the other linearly unstable. As in all competing species models, in this example we have that

$$\mathrm{R_{s*}} = (0,\infty)\times(-\infty,0], \quad \mathrm{R_{u*}} = \emptyset,$$
$$\mathrm{R_{*s}} = (-\infty,0]\times(0,\infty), \quad \mathrm{R_{*u}} = \emptyset.$$

In Fig. 3 we have represented an admissible situation adjusted to (4.19) for the choices previously analyzed. Now, for each sufficiently small $\varepsilon > 0$, we have that

$$\alpha := c_0 + \frac{1}{2}c_1\ell_1 = 1+\varepsilon+\frac{1}{2}\ell_1 < \frac{1}{b_0} = \frac{1}{1+\varepsilon} < 1$$

and

$$\frac{1}{\beta} := \left(b_0 + \frac{1}{2}b_1 m_1\right)^{-1} = \left(1+\varepsilon+\frac{1}{2}m_1\right)^{-1} > c_0 = 1+\varepsilon > 1,$$

since $\ell_1 < 0$ and $m_1 < 0$. As the relative positions between the curves of neutral stability of the semi-trivial states suffered a drastic change between the cases sketched in Figs. 2 and 3, in the latest case, within the bounded component enclosed by $\mu = f(\lambda)$ and $\lambda = g(\mu)$ both semi-trivial states must be linearly stable, while in the unbounded component both are linearly unstable. As it is easily realized, this inter-exchange between the relative positions of the curves of neutral stability of the semi-trivial states entails dramatic changes from the point of view of applications.

Although the previous examples exhibit a reminiscent behaviour from the autonomous model for (λ, μ) in a neighborhood of $(0, 0)$, in the sense that, for such range of (λ, μ) *low intensity competition* occurs if $\hat{b}\hat{c} < 1$, while *high intensity competition* occurs if $\hat{b}\hat{c} > 1$, these simple examples illustrate very well how temporal heterogeneities can affect the dynamics of (1.1), as for (λ, μ) sufficiently large the intensity of the competition is measured by α, instead of $\hat{c}$, and β, instead of $\hat{b}$. Actually, for (λ, μ) large, *low intensity competition* occurs if $\alpha\beta < 1$ (cf. Fig. 3), while *high intensity competition* occurs if $\alpha\beta > 1$ (cf. Fig. 2). The observation of this phenomenology goes back to [12].

In the previous examples, a further crossing point, perturbing from $(0, 0)$, might occur if $b_0 c_0 = 1$, as a result of the nature of the second order terms of the asymptotic expansion of (3.3) and (3.6) at $(0, 0)$, but we refrain of performing this sharper analysis here (cf. [13] for the corresponding analysis in the autonomous competing species model with diffusion).

It should be noted that most of the examples treated in the literature did not cover the previous two cases, but, instead, covered some very special situations where the function coefficients were chosen so that the curves of neutral stability of the semi-trivial states cannot meet in the interior of the first quadrant, as in those cases the method of sub- and supersolutions gives the existence of coexistence states in a rather simple way. Most precisely, let $t_L, t_M \in [0, T]$ such that

$$(\theta_{[\lambda\ell,a]})_j = \theta_{[\lambda\ell,a]}(t_j), \quad j \in \{L; M\}.$$

Then, since $\frac{\mathrm{d}\theta_{[\lambda\ell,a]}}{\mathrm{d}t}(t_j) = 0$, we have that

$$\theta_{[\lambda\ell,a]}(t_j) = \frac{\ell(t_j)}{a(t_j)}\lambda, \quad j \in \{L, M\},$$

and, hence,

$$\left(\frac{\ell}{a}\right)_L \lambda \leqslant \theta_{[\lambda\ell,a]} \leqslant \left(\frac{\ell}{a}\right)_M \lambda, \quad \lambda > 0. \tag{4.20}$$

By symmetry,

$$\left(\frac{m}{d}\right)_L \mu \leqslant \theta_{[\mu m,d]} \leqslant \left(\frac{m}{d}\right)_M \mu, \quad \mu > 0. \tag{4.21}$$

Thus, for each $\lambda > 0$,

$$\hat{c}\left(\frac{\ell}{a}\right)_L \lambda \leqslant f(\lambda) = \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t \leqslant \hat{c}\left(\frac{\ell}{a}\right)_M \lambda, \tag{4.22}$$

and, for each $\mu > 0$,

$$\hat{b}\left(\frac{m}{d}\right)_L \mu \leqslant g(\mu) = \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t \leqslant \hat{b}\left(\frac{m}{d}\right)_M \mu. \tag{4.23}$$

Consequently, the curves of neutral stability (3.3) and (3.6) cannot meet at a point $(\lambda, \mu) \neq (0, 0)$ if either

$$\hat{b}\hat{c}\left(\frac{\ell}{a}\right)_M\left(\frac{m}{d}\right)_M < 1, \tag{4.24}$$

or

$$\hat{b}\hat{c}\left(\frac{\ell}{a}\right)_L\left(\frac{m}{d}\right)_L > 1. \tag{4.25}$$

Actually, in any of these cases the curves $\mu = f(\lambda)$ and $\lambda = g(\mu)$ divide the first quadrant into three regions and, necessarily, either $\mathrm{R}_{\mathrm{ss}} = \emptyset$ or $\mathrm{R}_{\mathrm{uu}} = \emptyset$ (cf. Fig. 4). As in the autonomous model, low intensity competition occurs under (4.24), while high intensity competition occurs if (4.25) is satisfied. It should be noted that (4.24) and (4.25) become into $\hat{b}\hat{c} < 1$ and $\hat{b}\hat{c} > 1$, respectively, if $\ell = a$ and $m = d$. These are the unique two possible situations exhibited by the autonomous model.

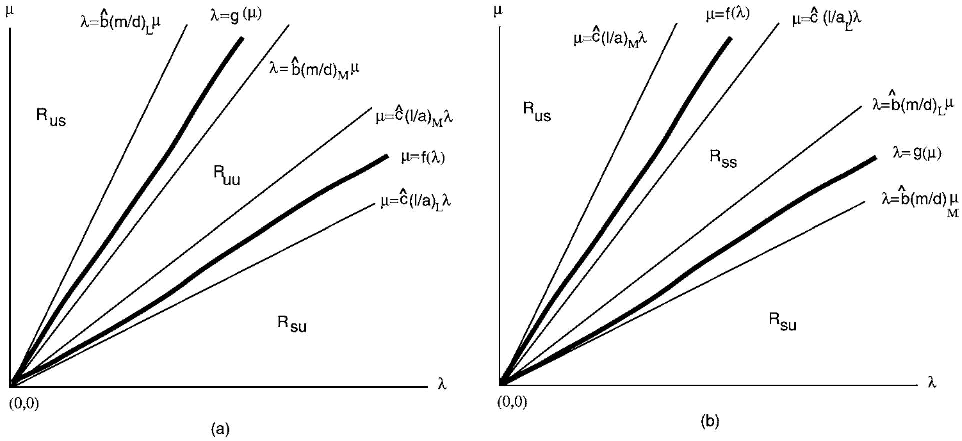

μ
$\lambda=\hat{b}(m/d)_L\mu$
$\lambda=g(\mu)$
$\lambda=\hat{b}(m/d)_M\mu$
$\mu=\hat{c}(l/a)_M\lambda$
$\mu=f(\lambda)$
$\mu=\hat{c}(l/a)_L\lambda$
R_{us}
R_{uu}
R_{su}
λ
(0,0)
(a)
μ
$\mu=\hat{c}(l/a)_M\lambda$
$\mu=f(\lambda)$
$\mu=\hat{c}(l/a)_L\lambda$
$\lambda=\hat{b}(m/d)_L\mu$
$\lambda=g(\mu)$
$\lambda=\hat{b}(m/d)_M\mu$
R_{us}
R_{ss}
R_{su}
λ
(0,0)
(b)

Fig. 4. Curves of neutral stability in cases (4.24) (a) and (4.25) (b).

4.2. *The species v preys on u ($b > 0$ and $c < 0$)*

Then, the curves of neutral stability of the semi-trivial states cannot meet at a point $(\lambda, \mu) \neq (0, 0)$, since

$$f(\lambda) = \frac{1}{T} \int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\, \mathrm{d}t < 0 \quad \text{for each } \lambda > 0,$$

and

$$g(\mu) = \frac{1}{T} \int_0^T b(t)\theta_{[\mu m,d]}(t)\, \mathrm{d}t > 0 \quad \text{for each } \mu > 0.$$

Note that, as a consequence from Theorem 2.8, $\lambda \mapsto f(\lambda)$ is decreasing and $\mu \mapsto g(\mu)$ is increasing, since $\ell > 0$ and $m > 0$ (cf. Fig. 5). If we further assume that $\ell \gg 0$, $m \gg 0$, $a \gg 0$ and $d \gg 0$, then, besides the fact that $f(\lambda)$ and $g(\mu)$ are monotone, the tangents to $\mu = f(\lambda)$ and $\lambda = g(\mu)$ at $(0, 0)$ and at infinity are given by the straight lines (4.12) and (4.13), respectively.

Now, the half-plane $\lambda > 0$ is divided into four supplementary regions according to the existence and character of each of the semi-trivial states. The interior of the first quadrant is divided into the two regions R_{us} and R_{uu} (cf. (4.1) and (4.2)), and the quadrant $\lambda > 0$, $\mu \leqslant 0$, where (1.2) exclusively admits the semi-trivial state $(\theta_{[\lambda\ell,a]}, 0)$, is divided into the region $\mu > f(\lambda)$ $\mathrm{R}_{\mathrm{u}*}$, where $(\theta_{[\lambda\ell,a]}, 0)$ is linearly unstable, and the region $\mu < f(\lambda)$ $\mathrm{R}_{\mathrm{s}*}$, where $(\theta_{[\lambda\ell,a]}, 0)$ is linearly stable. Moreover,

$$\mathrm{R}_{*\mathrm{s}} = (-\infty, 0] \times (0, \infty), \quad \mathrm{R}_{*\mathrm{u}} = \emptyset, \quad \mathrm{R}_{\mathrm{ss}} = \emptyset.$$

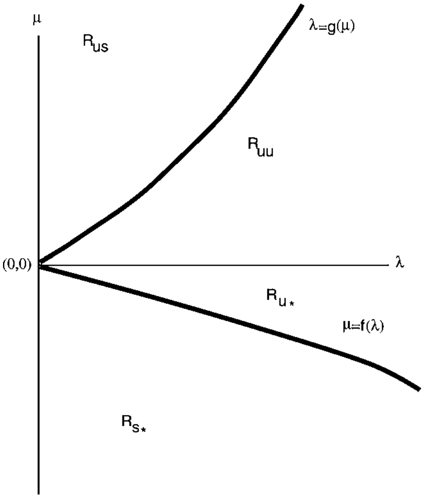


Fig. 5. Curves of neutral stability in the predator-prey model.

The configuration shown by Fig. 5 is reminiscent of the classical one for the autonomous model, where the curves of change of stability of the semi-trivial states $(\lambda, 0)$ and $(0, \mu)$ are $\mu = c\lambda$ and $\lambda = b\mu$, respectively. Actually, if $\ell = a$ and $m = d$, then $\mu = \hat{c}\lambda$ and $\lambda = \hat{b}\mu$ provide us with the tangents at $(0, 0)$ and at infinity to the curves $\mu = f(\lambda)$ and $\lambda = g(\mu)$.

4.3. *The species u and v cooperate (b < 0 and c < 0)*

In this case, $\lambda \mapsto f(\lambda)$ and $\mu \mapsto g(\mu)$ are decreasing and, in general, they look like shows Fig. 6. Now, the regions defined in (4.1)–(4.4) are given by the following identities:

$$\mathrm{R}_{\mathrm{uu}} = (0, \infty) \times (0, \infty), \quad \mathrm{R}_{\mathrm{ss}} = \emptyset, \quad \mathrm{R}_{\mathrm{us}} = \emptyset, \quad \mathrm{R}_{\mathrm{su}} = \emptyset,$$
$$\mathrm{R}_{\mathrm{u}*} = \bigl\{(\lambda, \mu)\colon \lambda > 0,\, f(\lambda) < \mu \leqslant 0\bigr\}, \quad \mathrm{R}_{\mathrm{s}*} = \bigl\{(\lambda, \mu)\colon \lambda > 0,\, \mu < f(\lambda)\bigr\},$$
$$\mathrm{R}_{*\mathrm{u}} = \bigl\{(\lambda, \mu)\colon \mu > 0,\ g(\mu) < \lambda \leqslant 0\}, \quad \mathrm{R}_{*\mathrm{s}} = \{(\lambda, \mu)\colon \mu > 0,\ \lambda < g(\mu)\}.$$

As in the predator-prey model, if we further assume $\ell \gg 0$, $m \gg 0$, $a \gg 0$ and $d \gg 0$, then, besides the fact that $f(\lambda)$ and $g(\mu)$ are monotone, the tangents to the curves of neutral stability $\mu = f(\lambda)$ and $\lambda = g(\mu)$ at $(0, 0)$ and at infinity are given by the straight lines (4.12) and (4.13), respectively. Similarly, the configuration shown by Fig. 6 is reminiscent of the classical one for the autonomous model, where the curves of change of stability of the semi-trivial states $(\lambda, 0)$ and $(0, \mu)$ are $\mu = c\lambda$ and $\lambda = b\mu$, respectively. Actually, if $\ell = a$ and $m = d$, then $\mu = \hat{c}\lambda$ and $\lambda = \hat{b}\lambda$ provide us with the tangents at $(0, 0)$ and at infinity to the curves $\mu = f(\lambda)$ and $\lambda = g(\mu)$.

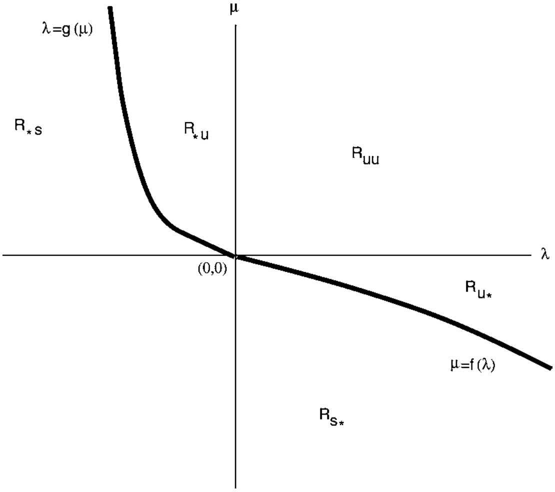


Fig. 6. Curves of neutral stability for the symbiotic model.

4.4. *The general case when b and c change of sign*

As for the competing species model, in order to simplify our analysis, besides (1.7) in this section we assume that

$$a \gg 0, \quad d \gg 0, \quad \ell \gg 0, \quad m \gg 0. \tag{4.26}$$

Then, as in Section 4.1, (4.12) and (4.13) provide us with the tangents at $(0,0)$ and at infinity of the curves of neutral stability of the semi-trivial states. As we are assuming that b and c change of sign in $[0,T]$, we have

$$\begin{aligned} I_+^b &:= \operatorname{Int}\operatorname{supp} b^+ \neq \emptyset, \quad I_-^b := \operatorname{Int}\operatorname{supp} b^- \neq \emptyset, \\ I_+^c &:= \operatorname{Int}\operatorname{supp} c^+ \neq \emptyset, \quad I_-^c := \operatorname{Int}\operatorname{supp} c^- \neq \emptyset, \end{aligned}$$

and, hence, $f(\lambda)$ and $g(\mu)$ fail to be monotone. Nevertheless, since

$$\begin{aligned} \lim_{(c^-)_M \downarrow 0} f(\lambda) &= \frac{1}{T}\int_0^T c^+(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t, \\ \lim_{(b^-)_M \downarrow 0} g(\mu) &= \frac{1}{T}\int_0^T b^+(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t, \end{aligned} \tag{4.27}$$

uniformly on compact subsets of (λ,μ), the curves of change of stability of the semi-trivial states approximate the corresponding curves of the competition model with interaction function coefficients b^+ and c^+, as $b^- \downarrow 0$ and $c^- \downarrow 0$. Therefore, on compact subsets of (λ,μ), the regions defined by (4.1) and (4.2) approximate the corresponding regions of the associated competition model. Similarly, since

$$\begin{aligned} \lim_{(c^+)_M \downarrow 0} f(\lambda) &= -\frac{1}{T}\int_0^T c^-(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t, \\ \lim_{(b^+)_M \downarrow 0} g(\mu) &= -\frac{1}{T}\int_0^T b^-(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t, \end{aligned} \tag{4.28}$$

uniformly on compact subsets of (λ,μ), the curves of change of stability of the semi-trivial states approximate the corresponding curves of the symbiotic model with interaction function coefficients $-b^-$ and $-c^-$ as $b^+ \downarrow 0$ and $c^+ \downarrow 0$. Consequently, for (λ,μ) varying in any compact subset of $(0,\infty)^2$, the regions defined by (4.1) and (4.2) approximate the corresponding regions of the symbiotic model with interactions $-b^-$ and $-c^-$. Therefore, the general case when b and c change of sign can be regarded as a sort of intermediate situation between a competitive and a symbiotic model. Actually, the map $(B,C):[0,1] \to \mathcal{C}_T \times \mathcal{C}_T$ defined, for each $\varepsilon \in [0,1]$, by

$$\begin{aligned} B(\varepsilon) &:= \varepsilon b^+ - (1-\varepsilon)b^- + 2\varepsilon(1-\varepsilon)\bigl(b^+ - b^-\bigr), \\ C(\varepsilon) &:= \varepsilon c^+ - (1-\varepsilon)c^- + 2\varepsilon(1-\varepsilon)\bigl(c^+ - c^-\bigr), \end{aligned}$$

establishes a homotopy between (b^+, c^+), $(-b^-, -c^-)$ and (b, c). Indeed,

$$\big(B(0), C(0)\big) = \big(-b^-, -c^-\big), \qquad \big(B(1), C(1)\big) = \big(b^+, c^+\big),$$

and

$$\big(B(1/2), C(1/2)\big) = (b, c).$$

Consequently, (1.2) should exhibit an intermediate behavior between those exhibited by a competition and a symbiotic model.

In Fig. 7, we have represented two admissible curves $\mu = f(\lambda)$ and $\lambda = g(\mu)$ in the special case when conditions

$$\hat{b} > 0, \quad \hat{c} > 0, \quad \hat{b}\hat{c} < 1, \tag{4.29}$$

and

$$\alpha := \frac{1}{T}\int_0^T c(t)\frac{\ell(t)}{a(t)}\,\mathrm{d}t < 0, \quad \beta := \frac{1}{T}\int_0^T b(t)\frac{m(t)}{d(t)}\,\mathrm{d}t < 0 \tag{4.30}$$

are satisfied. Conditions (4.29) can be reached by taking b^+ and c^+ with $\hat{b}^+\hat{c}^+ < 1$ and choosing sufficiently small b^- and c^-. In order to get (4.30) it suffices to choose ℓ sufficiently small in I_+^c and m sufficiently small in I_+^b. This entails the birth rate of each of the species must be relatively slow during the time periods where the intensity of its aggressions on the other species reaches its highest level. Consequently, although the species compete, the model exhibits some sort of cooperative behaviour. Nevertheless, one should be very careful in choosing the adequate ranges of values of the parameters where this phenomenology occurs.

For the situation illustrated in Fig. 7, the first quadrant is divided into there regions: the region $\mathrm{R_{uu}}$, where both semi-trivial states are linearly unstable, which is the region where $\mu > f(\lambda)$ and $\lambda > g(\mu)$, the region $\mathrm{R_{su}}$, where $(\theta_{[\lambda\ell,a]}, 0)$ is linearly stable and $(0, \theta_{[\mu m,d]})$ is linearly unstable, which is the region where $\lambda > 0$ and $0 < \mu < f(\lambda)$, and the region $\mathrm{R_{us}}$, where $(\theta_{[\lambda\ell,a]}, 0)$ is linearly unstable and $(0, \theta_{[\mu m,d]})$ is linearly stable, which is given by $\mu > 0$ and $0 < \lambda < g(\mu)$. In the second quadrant ($\lambda \leqslant 0$ and $\mu > 0$), $(0, \theta_{[\mu m,d]})$ is the unique semi-trivial state; it is linearly stable if $\lambda < g(\mu)$ and linearly unstable if $\lambda > g(\mu)$. In the fourth quadrant ($\lambda > 0$ and $\mu \leqslant 0$), $(\theta_{[\lambda\ell,a]})$ is the unique semi-trivial state; it is linearly stable if $\mu < f(\lambda)$ and linearly unstable if $\mu > f(\lambda)$. It should be noted that under conditions (4.29) and (4.30) the curves $\mu = f(\lambda)$ and $\lambda = g(\mu)$ might cross. In such case, their total number of crossing points, counting contact orders, must be even, because $\mu = f(\lambda)$ and $\lambda = g(\mu)$ are real analytic. Also, note that Fig. 7 shows a low-competitive-intensity behaviour around $(\lambda, \mu) = (0, 0)$ and a symbiotic type of behavior for sufficiently large λ and μ.

Now, by choosing b^+ and c^+ such that $\hat{b}^+\hat{c}^+ > 1$, $b^- \sim 0$, $c^- \sim 0$, $\ell \sim 0$ in I_+^c and $m \sim 0$ in I_+^b, one gets

$$\hat{b} > 0, \quad \hat{c} > 0, \quad \hat{b}\hat{c} > 1, \quad \alpha < 0, \quad \beta < 0. \tag{4.31}$$

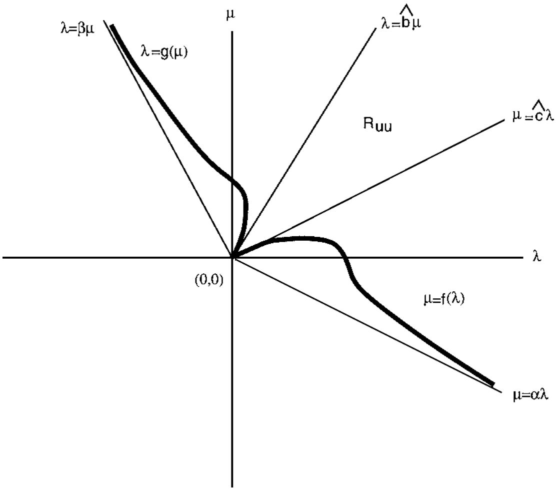


Fig. 7. Curves of neutral stability under conditions (4.29) and (4.30).

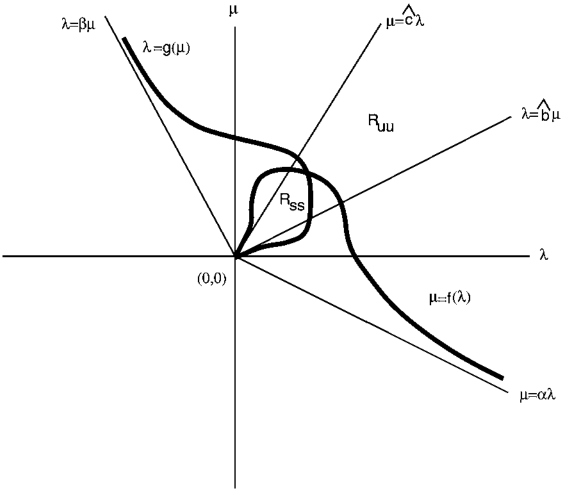


Fig. 8. Curves of neutral stability under condition (4.31).

Fig. 8 shows two admissible curves of neutral stability of semi-trivial states adjusted to (4.31). Now, these curves must cross at an odd number of points, counting contact orders, in such a way that both semi-trivial states are linearly unstable in the unbounded component of the enclosed region between them. Rather naturally, the model exhibits a high-intensity-competition behavior around $(0, 0)$, while it exhibits a genuine cooperative behavior for sufficiently large λ and μ; as far as to the structure of the regions defined in (4.1) and (4.2) concerns. It seems the first occasion that such a kind of behaviour has been observed had been in [22] within the context of reaction diffusion models.

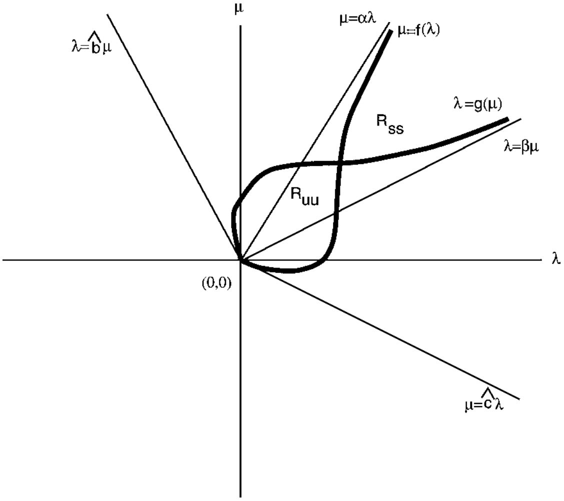


Fig. 9. Curves of neutral stability under condition (4.32).

Now, by choosing sufficiently small b^+ and c^+, $\ell \sim 0$ in I_-^c and sufficiently large in I_+^c, and $m \sim 0$ in I_-^b and sufficiently large in I_+^b, the following conditions can be reached

$$\hat{b} < 0, \quad \hat{c} < 0, \quad \alpha > 0, \quad \beta > 0, \quad \alpha\beta > 1. \tag{4.32}$$

Fig. 9 shows two admissible curves of neutral stability according to (4.32). As in case (4.31), the total number of crossing points between $\mu = f(\lambda)$ and $\lambda = g(\mu)$ must be odd, though, in case (4.32), $\mathrm{R_{ss}}$ includes the unbounded component enclosed by the two curves, while the bounded component containing $(0,0)$ is a subset of $\mathrm{R_{uu}}$, in strong contrast with the situation described by Fig. 8. Actually, in case (4.32), the model exhibits a genuine symbiotic behavior around $(0,0)$, while it possesses a high-intensity-competitive behavior for sufficiently large values of λ and μ.

Finally, by choosing sufficiently small b^- and c^+, one can easily reach

$$\hat{b} > 0, \quad \hat{c} < 0, \quad \alpha < 0, \quad \beta > 0, \tag{4.33}$$

for instance, by taking $\ell = m = a = d = 1$. In such cases the curves $\mu = g(\lambda)$ and $\lambda = g(\mu)$ exhibit a behaviour quite reminiscent of the case when v preys on u (cf. Fig. 10); rather naturally, since these models can be regarded as perturbations from the case when $b^- = 0$ and $c^+ = 0$, where $b > 0$ and $c < 0$—v preys on u—though the symbiotic effects of the interaction might play a significant role for large values of λ and μ. In Fig. 10 we have represented the curves of change of stability of the semi-trivial states of (1.2) in one of the several admissible situation cases.

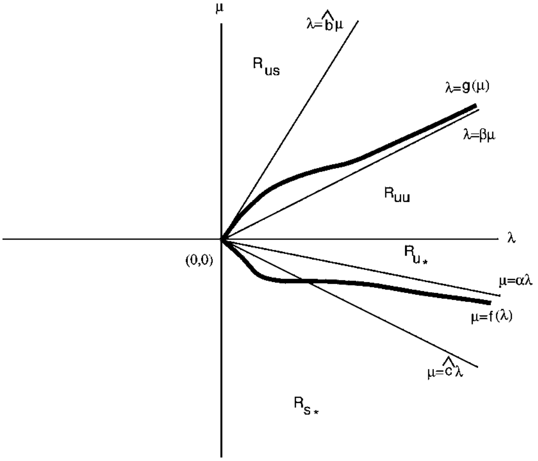


Fig. 10. Curves of neutral stability under condition (4.33).

5. The existence of coexistence states

In this section we are going to give a very general result about the existence of coexistence states of (1.2) for the general case where b and c change of sign. As a matter of fact, our result covers all admissible situations, as we are not imposing any sign restriction on b and c. The general methodology adopted in this section goes back to the pioneering chapter by J.M. Cushing [7], where one of the parameters, λ or μ, was fixed while the other was considered as a bifurcation parameter in order to get the coexistence states as emanating from one of the semi-trivial states in the context of competing species models. This idea has been extremely fruitful in a number of different contexts [19, Chapter 7].

Most precisely, in Section 5.1 we use the abstract unilateral theorem of [19, Chapter 6] to get an abstract global unilateral theorem for systems, which is the basic abstract result to get—as an easy consequence—our theorem about the existence of coexistence states, which will be given in Section 5.2. Our abstract result admits a number of applications in rather different contexts, but we refrain of giving more details here in. It should be noted that the well known, and very celebrated, unilateral theorems of P.H. Rabinowitz [31] are wrong as they are stated. Actually, there is an available—very recent—counterexample by E.N. Dancer [9]. In the subsequent subsections we shall consider each of the possible interactions separately, as in each special circumstance one can obtain sharper consequences from our abstract theorem.

5.1. *Abstract unilateral bifurcation from semi-trivial states*

Let X be an ordered Banach space whose positive cone P is normal and it has nonempty interior, and consider a nonlinear abstract equation of the form

$$\mathfrak{F}(\rho, x, y) := \mathfrak{L}(\rho)(x - x_0, y) + \mathfrak{N}(\rho, x, y) = 0, \quad (\rho, x, y) \in \mathbb{R} \times X^2, \tag{5.1}$$

for some $x_0 \in P \setminus \{0\}$, where

- (HL) $\mathfrak{K}(\rho) := I_{X^2} - \mathfrak{L}(\rho) \in \mathcal{L}(X^2)$, $\rho \in \mathbb{R}$, is a compact and continuous operator pencil with a discrete set of *singular values*, denoted by Σ. Since $\mathfrak{L}(\rho)$ is Fredlhom of index zero, $\sigma \in \Sigma$ if, and only if, σ is an *eigenvalue* of $\mathfrak{L}(\rho)$, i.e., if $\dim N[\mathfrak{L}(\sigma)] \geqslant 1$.
- (HR) $\mathfrak{R} \in \mathcal{C}(\mathbb{R} \times X^2; X^2)$ is compact on bounded sets and

$$\lim_{(x,y)\to(x_0,0)} \frac{\mathfrak{R}(\rho,x,y)}{\|(x-x_0,y)\|} = 0$$

uniformly on compact intervals of $\mathbb{R}$.
- (HP) The solutions of (5.1) satisfy the *strong maximum principle* in the sense that

$$(\rho,x,y) \in \mathbb{R} \times \bigl(P \setminus \{0\}\bigr) \times X \quad \text{and} \quad \mathfrak{F}(\rho,x,y) = 0 \quad \text{imply} \quad x \in \operatorname{Int} P,$$

and

$$(\rho,x,y) \in \mathbb{R} \times X \times \bigl(P \setminus \{0\}\bigr) \quad \text{and} \quad \mathfrak{F}(\rho,x,y) = 0 \quad \text{imply} \quad y \in \operatorname{Int} P,$$

where $\operatorname{Int} P$ stands for the interior of the cone P.

Subsequently, the space X^2 is viewed as an ordered Banach space with positive cone P^2, and, given any ordered Banach space E with positive cone W, and $e_1, e_2 \in E$, we write $e_1 > e_2$ if $e_1 - e_2 \in W \setminus \{0\}$, and $e_1 \gg e_2$ if $e_1 - e_2 \in \operatorname{Int} W$. As

$$\operatorname{Int} P^2 = \operatorname{Int} P \times \operatorname{Int} P,$$

it is said that (ρ,x,y) is a *positive solution* of (5.1), if (ρ,x,y) is a solution of (5.1) with $(x,y) > 0$ ($x \geqslant 0$, $y \geqslant 0$ and $(x,y) \neq (0,0)$). Thanks to Hypothesis (HP), any positive solution (ρ,x,y) of (5.1) must have one of the following forms: either $x \gg 0$ and $y \gg 0$—*strongly positive*, or $x \gg 0$ and $y = 0$, or $x = 0$ and $y \gg 0$—*semi-trivial positive solutions*.

Under Hypotheses (HL) and (HR),

$$\mathfrak{F}(\rho,x_0,0) = 0 \quad \text{for each } \rho \in \mathbb{R}, \tag{5.2}$$

and, due to Hypothesis (HP), $x_0 \gg 0$, since $x_0 > 0$. Moreover, since $\mathfrak{L}(\rho)$ is a linear isomorphism for any $\rho \in \mathbb{R} \setminus \Sigma$, the *fixed point index*

$$\operatorname{Ind}\bigl(0, \mathfrak{K}(\rho)\bigr)$$

is well defined—the topological degree of $\mathfrak{L}(\rho)$ in any bounded open set containing 0. Thus, the following *crossing number* $C : \Sigma \to \{-1,0,1\}$ can be introduced

$$C(\sigma) := \frac{1}{2} \lim_{\varepsilon \downarrow 0} \bigl[\operatorname{Ind}\bigl(0, \mathfrak{K}(\sigma+\varepsilon)\bigr) - \operatorname{Ind}\bigl(0, \mathfrak{K}(\sigma-\varepsilon)\bigr)\bigr], \quad \sigma \in \Sigma. \tag{5.3}$$

Note that $C(\sigma) \neq 0$ if, and only if, $\operatorname{Ind}(0, \mathfrak{K}(\rho))$ changes as ρ crosses σ, and that this occurs if, and only if, $\operatorname{Ind}((x_0, 0), \mathfrak{F}(\rho, \cdot, \cdot))$ changes as ρ crosses σ.

Our main result concerns the bounded components of strongly positive solutions of (5.1) emanating from the curve $(\rho, x, y) = (\rho, x_0, 0)$ at a singular value $\rho_0 \in \Sigma$ with—classical—geometric multiplicity one and $C(\rho_0) \in \{-1, 1\}$. By a *component* of strongly positive solutions of (5.1) it is meant a maximal (for the inclusion) relatively closed and connected subset of the set of strongly positive solutions of (5.1)—in $\mathbb{R} \times \operatorname{Int} P^2$. Thanks to [19, Theorem 6.2.1], Eq. (5.1) possesses a component emanating from $(\rho, x, y) = (\rho, x_0, 0)$ at ρ_0 if $C(\rho_0) \in \{-1, 1\}$. Such a component will be subsequently denoted by $\mathfrak{C}_0$. Throughout the remaining of this section, we shall denote by P_x and P_y the projection operators

$$P_x(x, y) = x, \quad P_y(x, y) = y, \quad (x, y) \in X^2. \tag{5.4}$$

The main abstract result of this section can now be stated as follows.

THEOREM 5.1. *Suppose $\rho_0 \in \Sigma$ satisfies $C(\rho_0) \neq 0$,*

$$N\big[\mathfrak{L}(\rho_0)\big] = \operatorname{span}\big[(\varphi_{0x}, \varphi_{0y})\big], \quad \varphi_{0y} \in P \setminus \{0\}, \tag{5.5}$$

and, for each $x \in X$, $P_y\mathfrak{K}(\rho_0)(x, \cdot)$ is strongly positive, in the sense that

$$P_y\mathfrak{K}(\rho_0)(x, \cdot)\big(P \setminus \{0\}\big) \subset \operatorname{Int} P. \tag{5.6}$$

Then,

$$\operatorname{spr} P_y\mathfrak{K}(\rho_0)(\varphi_{0x}, \cdot) = 1 \tag{5.7}$$

and there exists a component $\mathfrak{C}_0^P$ of $\mathfrak{C}_0 \cap [\mathbb{R} \times (\operatorname{Int} P)^2]$ such that

$$(\rho_0, x_0, 0) \in \bar{\mathfrak{C}}_0^P.$$

Suppose, in addition, that $\mathfrak{C}_0^P$ is bounded in $\mathbb{R} \times X^2$ and that

$$\operatorname{spr} P_y\mathfrak{K}(\rho_0)(x, \cdot) = 1 \quad \textit{for each } x \in X. \tag{5.8}$$

Then, some of the following alternatives occurs:

A1. *There exists $(\rho_1, x_1) \in (\mathbb{R} \times P) \setminus \{(\rho_0, x_0)\}$ such that $(\rho_1, x_1, 0) \in \bar{\mathfrak{C}}_0^P$. In such case, thanks to* (HP), *$x_1 \gg 0$ if $x_1 > 0$.*

A2. *There exists $(\rho_1, y_1) \in \mathbb{R} \times P$ such that $(\rho_1, 0, y_1) \in \bar{\mathfrak{C}}_0^P$. In such case, thanks to* (HP), *$y_1 \gg 0$ if $y_1 > 0$.*

Furthermore, alternative A2 *occurs if $(\rho, x, 0) \in \mathfrak{F}^{-1}(0)$ implies $x = x_0$, and ρ_0 is the unique value of ρ for which one is an eigenvalue of $P_y\mathfrak{K}(\rho)(x_0, \cdot)$ to a positive eigenvector.*

PROOF. The set of *non-trivial solutions* of (5.1) is defined through

$$\mathfrak{S} := \mathfrak{F}^{-1}\big((0,0)\big) \setminus \big[(\mathbb{R} \setminus \Sigma) \times \big\{(x_0, 0)\big\}\big]. \tag{5.9}$$

Note that $(\rho, x, y) \in \mathfrak{S}$ if $\mathfrak{F}(\rho, x, y) = 0$ with $(x, y) \neq (x_0, 0)$, or $(\rho, x, y) = (\rho, x_0, 0)$ with $\rho \in \Sigma$. Since $C(\rho_0) \neq 0$, thanks to [19, Theorem 6.2.1] there exists a component of $\mathfrak{S}$, denoted by $\mathfrak{C}_0$, such that $(\rho_0, x_0, 0) \in \mathfrak{C}_0$. Subsequently, we suppose that

$$\varphi_0 := (\varphi_{0x}, \varphi_{0y}) \in X \times \operatorname{Int} P$$

has been normalized so that

$$\|\varphi_0\|_{X^2} = 1.$$

It should be noted that, thanks to (HL),

$$\mathfrak{K}(\rho_0)\varphi_0 = \varphi_0$$

and, hence,

$$P_y \mathfrak{K}(\rho_0)(\varphi_{0x}, \varphi_{0y}) = \varphi_{0y}.$$

Thus, it follows from (5.5) and (5.6) that $\varphi_{0y} \gg 0$. Moreover, thanks to (5.6), (5.7) is a consequence from Krein–Rutman theorem, since the spectral radius is the unique real eigenvalue associated with it there is a positive eigenvector.

Now, let Z be a closed subspace of X^2 such that

$$X^2 = N\big[\mathfrak{L}(\rho_0)\big] \oplus Z.$$

Thanks to Hahn–Banach's theorem, there exists

$$\varphi_0^* := \big(\varphi_{0x}^*, \varphi_{0y}^*\big) \in X' \times X'$$

for which

$$Z = \big\{(x, y) \in X^2 \colon \big\langle \varphi_0^*, (x, y)\big\rangle = 0\big\}, \quad \big\langle \varphi_0^*, \varphi_0\big\rangle = 1,$$

where $\langle \cdot, \cdot \rangle$ stands for the duality between X^2 and $X' \times X'$. Now, adapting the theory developed by P.H. Rabinowitz [31], for each $\eta \in (0, 1)$ and sufficiently small $\varepsilon > 0$ we set

$$\begin{aligned} \mathcal{Q}_{\varepsilon,\eta} := \big\{&(\rho, x, y) \in \mathbb{R} \times X^2 \colon |\rho - \rho_0| < \varepsilon, \\ &\big|\big\langle \varphi_0^*, (x - x_0, y)\big\rangle\big| > \eta \big\|(x - x_0, y)\big\|\big\}. \end{aligned}$$

Since the mapping

$$(x, y) \mapsto \big|\big\langle \varphi_0^*, (x - x_0, y)\big\rangle\big| - \eta \big\|(x - x_0, y)\big\|$$

is continuous, $\mathcal{Q}_{\varepsilon,\eta}$ is an open subset of $\mathbb{R}\times X^2$ consisting of the two disjoint components $\mathcal{Q}^+_{\varepsilon,\eta}$ and $\mathcal{Q}^-_{\varepsilon,\eta}$ defined through

$$\begin{aligned}\mathcal{Q}^+_{\varepsilon,\eta} &:= \big\{(\rho,x,y)\in\mathbb{R}\times X^2\colon\ |\rho-\rho_0|<\varepsilon,\\ &\qquad \big\langle\varphi_0^*,(x-x_0,y)\big\rangle>\eta\big\|(x-x_0,y)\big\|\big\},\\ \mathcal{Q}^-_{\varepsilon,\eta} &:= \big\{(\rho,x,y)\in\mathbb{R}\times X^2\colon\ |\rho-\rho_0|<\varepsilon,\\ &\qquad \big\langle\varphi_0^*,(x-x_0,y)\big\rangle<-\eta\big\|(x-x_0,y)\big\|\big\}.\end{aligned}$$

The following result collects the main consequences from [19, Theorem 6.2.1, Lemma 6.4.1, Proposition 6.4.2]; it is entirely attributable to P.H. Rabinowitz [31].

THEOREM 5.2. *For each sufficiently small* $\delta>0$,

$$\mathfrak{C}_0\cap B_\delta(\rho_0,x_0,0)\subset\mathcal{Q}_{\varepsilon,\eta}\cup\big\{(\rho_0,x_0,0)\big\}$$

and each of the sets

$$\mathfrak{S}\setminus\big[\mathcal{Q}^-_{\varepsilon,\eta}\cap B_\delta(\rho_0,x_0,0)\big]\quad\textit{and}\quad\mathfrak{S}\setminus\big[\mathcal{Q}^+_{\varepsilon,\eta}\cap B_\delta(\rho_0,x_0,0)\big]$$

contains a component, denoted by $\mathfrak{C}_0^+$ *and* $\mathfrak{C}_0^-$, *respectively, such that*

$$(\rho_0,x_0,0)\in\mathfrak{C}_0^+\cap\mathfrak{C}_0^-$$

and

$$\mathfrak{C}_0\cap B_\delta(\rho_0,x_0,0)=\big(\mathfrak{C}_0^+\cup\mathfrak{C}_0^-\big)\cap B_\delta(\rho_0,x_0,0).\tag{5.10}$$

Moreover, for each

$$(\rho,x,y)\in\big[\mathfrak{C}_0\setminus\big\{(\rho_0,x_0,0)\big\}\big]\cap B_\delta(\rho_0,x_0,0),$$

there exists a unique $(s,\tilde{x},\tilde{y})\in\mathbb{R}\times Z$ *such that* $\rho=\rho(s)$ *and*

$$(x,y)=(x_0,0)+s\varphi_0+\big(\tilde{x},\tilde{y}\big)=\big(x_0+s\varphi_{0x}+\tilde{x},s\varphi_{0y}+\tilde{y}\big)\tag{5.11}$$

with $|s|>\eta\,\|(x-x_0,y)\|$. *Actually,*

$$\rho=\rho_0+\mathrm{o}(1)\quad\textit{and}\quad\big(\tilde{x},\tilde{y}\big)=\mathrm{o}(s),\quad\textit{as } s\to 0.\tag{5.12}$$

Note that if

$$(\rho,x,y)\in\mathfrak{C}_0^+\cap B_\delta(\rho_0,x_0,0),\quad (x,y)\neq(x_0,0),$$

then (5.11) holds with $s > \eta \,\|(x - x_0, y)\| > 0$ and, hence,

$$s^{-1} y = \varphi_{0y} + s^{-1}\tilde{y}.$$

Thus, due to (5.12), for sufficiently small $s > 0$ we have that $s^{-1}y \in \operatorname{Int} P$, since $\varphi_{0y} \gg 0$, and, hence, $y \gg 0$. The fact that $x \gg 0$ is an easy consequence from (5.12) and the fact that $x_0 \gg 0$. Therefore, for any sufficiently small $\delta > 0$ we have that

$$\big[\mathfrak{C}_0^+ \setminus \big\{(\rho_0, x_0, 0)\big\}\big] \cap B_\delta(\rho_0, x_0, 0) \subset \mathbb{R} \times \operatorname{Int} P^2. \tag{5.13}$$

This shows the existence of the component $\mathfrak{C}_0^P$. Actually, $\mathfrak{C}_0^P$ is the maximal sub-continuum of $\mathfrak{C}_0^+$ in $\mathbb{R} \times \operatorname{Int} P^2$ such that

$$(\rho_0, x_0, 0) \in \bar{\mathfrak{C}}_0^P.$$

If $\mathfrak{C}_0^P$ is unbounded in $\mathbb{R} \times X^2$ the proof is concluded. So, for the remaining of the proof we will assume that $\mathfrak{C}_0^P$ is bounded.

The following result, which is [19, Theorem 6.4.3], provides us with an updated version of the very celebrated unilateral theorem of P.H. Rabinowitz [31, Theorem 1.27] which is necessary to conclude the proof of Theorem 5.1. It should be noted that [31, Theorem 1.27] is wrong as originally stated (cf. the detailed discussion carried out in [19, p. 180] and the counterexample of E.N. Dancer [9]).

THEOREM 5.3. *The component $\mathfrak{C}_0^+$ satisfies some of the following alternatives*:

1. $\mathfrak{C}_0^+$ *is unbounded in* $\mathbb{R} \times X^2$.
2. *There exists* $\rho_1 \in \Sigma \setminus \{\rho_0\}$ *such that* $(\rho_1, x_0, 0) \in \mathfrak{C}_0^+$.
3. *There exists* $(\rho, x, y) \in \mathfrak{C}_0^+$ *such that*

$$(\rho, x - x_0, y) \in \mathbb{R} \times \big(Z \setminus \{0\}\big).$$

Now, we shall prove that

$$X^2 = N\big[\mathfrak{L}(\rho_0)\big] \oplus R\big[\mathfrak{L}(\rho_0)\big]. \tag{5.14}$$

Since $\mathfrak{L}(\rho_0)$ is Fredholm of index zero, to show (5.14) it suffices to prove that $\varphi_0 \notin R[\mathfrak{L}(\rho_0)]$. This will be accomplished by contradiction. Suppose, there exists $(x, y) \in X^2$ such that

$$\mathfrak{L}(\rho_0)(x, y) = (x, y) - \mathfrak{K}(\rho_0)(x, y) = \varphi_0 = (\varphi_{0x}, \varphi_{0y}).$$

Then,

$$y - P_y \mathfrak{K}(\rho_0)(x, y) = \varphi_{0y} \gg 0. \tag{5.15}$$

Note that, thanks to (HL) and (5.6), the map

$$y \mapsto P_y\mathfrak{K}(\rho_0)(x, y)$$

defines a linear compact strongly order preserving operator. Moreover, by construction,

$$P_y\mathfrak{K}(\rho_0)(\varphi_{0x}, \varphi_{0y}) = \varphi_{0y}$$

and, due to (5.8),

$$\operatorname{spr} P_y\mathfrak{K}(\rho_0)(x, \cdot) = 1.$$

Let $\psi_x \gg 0$ denote the principal eigenfunction associated to $P_y\mathfrak{K}(\rho_0)(x, \cdot)$. Then, thanks to (5.15), for each $\beta > 0$ we have that

$$y + \beta\psi_x - P_y\mathfrak{K}(\rho_0)(x, y + \beta\psi_x) = \varphi_{0y}.$$

On the other hand, since

$$\lim_{\beta\uparrow\infty} \left(\beta^{-1}y + \psi_x\right) = \psi_x \gg 0,$$

it is apparent that, for each sufficiently large $\beta > 0$, one has that

$$y + \beta\psi_x \gg 0$$

provides us with a positive solution of (5.15). Thanks to H. Amann [4, Theorem 3.2], this is impossible. This contradiction concludes the proof of (5.14). Consequently, in order to apply Theorem 5.3, we can make the special choice

$$Z = R\left[\mathfrak{L}(\rho_0)\right], \tag{5.16}$$

which will be maintained throughout the rest of the proof.

Since $\mathfrak{C}_0^P \subset \mathfrak{C}_0^+$, some of the following alternatives occurs: either

$$\mathfrak{C}_0^P = \mathfrak{C}_0^+ \setminus \left\{(\rho_0, x_0, 0)\right\}, \tag{5.17}$$

or

$$\mathfrak{C}_0^P \text{ is a proper subset of } \mathfrak{C}_0^+ \setminus \left\{(\rho_0, x_0, 0)\right\}. \tag{5.18}$$

Suppose (5.17). Then, $\bar{\mathfrak{C}}_0^P = \mathfrak{C}_0^+$ satisfies some of the alternatives of Theorem 5.3. Alternative 1 cannot be satisfied, since $\bar{\mathfrak{C}}_0^P$ is compact. Suppose alternative 3 occurs. Then, thanks to the choice (5.16), there exists

$$(\rho, x, y) \in \mathfrak{C}_0^+ = \bar{\mathfrak{C}}_0^P \subset \mathbb{R} \times P^2$$

such that

$$(\rho, x - x_0, y) \in \mathbb{R} \times R\big[\mathfrak{L}(\rho_0)\big], \quad y \neq 0.$$

Since $(\rho, x, y) \in \mathbb{R} \times P^2$ with $y \neq 0$, necessarily $y > 0$ and, consequently, $y \gg 0$, by (HP). Thus, there exists $(\tilde{x}, \tilde{y}) \in X^2$ such that

$$\mathfrak{L}(\rho_0)(\tilde{x}, \tilde{y}) = (\tilde{x}, \tilde{y}) - \mathfrak{K}(\rho_0)(\tilde{x}, \tilde{y}) = (x - x_0, y)$$

and, taking projections,

$$\tilde{y} - P_y\mathfrak{K}(\rho_0)\big(\tilde{x}, \tilde{y}\big) = y \gg 0.$$

Arguing as above, this is impossible, because, due to (5.8),

$$\operatorname{spr} P_y\mathfrak{K}(\rho_0)(\tilde{x}, \cdot) = 1.$$

Thus, alternative 2 of Theorem 5.3 must be satisfied. Consequently, there exists $\rho_1 \in \Sigma \setminus \{\rho_0\}$ such that

$$(\rho_1, x_0, 0) \in \mathfrak{C}_0^+ = \bar{\mathfrak{C}}_0^P$$

and, therefore, alternative A1 of the theorem is satisfied.

Now, suppose (5.18), instead of (5.17). Then, since

$$\mathfrak{C}_0^+ \cap B_\delta(\rho_0, x_0, 0) = \big[\mathfrak{C}_0^P \cap B_\delta(\rho_0, x_0, 0)\big] \cup \big\{(\rho_0, x_0, 0)\big\}$$

for each sufficiently small $\delta > 0$, fixing one of these δ's, there exists

$$(\rho_1, x_1, y_1) \notin B_\delta(\rho_0, x_0, 0) \tag{5.19}$$

such that

$$(\rho_1, x_1, y_1) \in \mathfrak{C}_0^+ \cap \big(\mathbb{R} \times \partial P^2\big) \cap \partial\mathfrak{C}_0^P,$$

where, given any set U, ∂U stands for the boundary of U. Then,

$$\mathfrak{F}(\rho_1, x_1, y_1) = 0 \quad \text{and} \quad (x_1, y_1) \in P^2.$$

If $x_1 > 0$ and $y_1 > 0$, then, due to (HP), $x_1 \gg 0$ and $y_1 \gg 0$, which is impossible, since $(x_1, y_1) \in \partial P^2$. Thus, either $x_1 = 0$, or $y_1 = 0$. Suppose $y_1 = 0$. Then, thanks to (5.19), $(\rho_1, x_1) \neq (\rho_0, x_0)$, and, therefore, alternative A1 is satisfied again. Obviously, if $x_1 = 0$, then the alternative A2 is satisfied.

Finally, suppose that $(\rho, x, 0) \in \mathfrak{F}^{-1}(0)$ implies $x = x_0$, and ρ_0 is the unique value of ρ for which one is an eigenvalue of $P_y\mathfrak{K}(\rho)(x_0, \cdot)$ to a positive eigenvector. To conclude the proof of the theorem we proceed by contradiction. Assume alternative A1 is satisfied. Then, there exists $(\rho_1, x_1) \in (\mathbb{R} \times P) \setminus \{(\rho_0, x_0)\}$ such that $(\rho_1, x_1, 0) \in \bar{\mathfrak{C}}_0^P$. Thus, $x_1 = x_0$, since $(\rho_1, x_1, 0) \in \mathfrak{F}^{-1}(0)$, and, hence,

$$\rho_1 \neq \rho_0. \tag{5.20}$$

Let $\{(\gamma_n, u_n, v_n)\}_{n\geqslant 1}$ be any subsequence of $\mathfrak{C}_0^P$ such that

$$\lim_{n\to\infty} (\gamma_n, u_n, v_n) = (\rho_1, x_0, 0). \tag{5.21}$$

Then, for each $n \geqslant 1$, $v_n \gg 0$, and, setting,

$$\psi_n := \frac{v_n}{\|v_n\|},$$

we find from $\mathfrak{F}(\gamma_n, u_n, v_n) = 0$ that

$$\begin{aligned}\psi_n &= P_y\mathfrak{K}(\gamma_n)(u_n, \psi_n) - \frac{P_y\mathfrak{R}(\gamma_n, u_n, v_n)}{\|v_n\|}\\ &= P_y\mathfrak{K}(\rho_1)(u_n, \psi_n) + P_y\big[\mathfrak{K}(\gamma_n) - \mathfrak{K}(\rho_1)\big](u_n, \psi_n) - \frac{P_y\mathfrak{R}(\gamma_n, u_n, v_n)}{\|v_n\|}.\end{aligned}$$

As $\{(u_n, \psi_n)\}_{n\geqslant 1}$ is bounded, since we are assuming that $\mathfrak{C}_0^P$ is bounded, we have

$$\lim_{n\to\infty} P_y\big[\mathfrak{K}(\gamma_n) - \mathfrak{K}(\rho_1)\big](u_n, \psi_n) = 0,$$

by the continuity of $\mathfrak{K}(\rho)$ in ρ. Moreover, thanks to (5.21), it follows from (HR) that

$$\lim_{n\to\infty} \frac{P_y\mathfrak{R}(\gamma_n, u_n, v_n)}{\|v_n\|} = 0.$$

Thus, for each $n \geqslant 1$, there exists $w_n \in X$ such that

$$\psi_n = P_y\mathfrak{K}(\rho_1)(x_0, \psi_n) + w_n$$

with $\lim_{n\to\infty} w_n = 0$. Consequently, by compactness, there exists a subsequence of $\{\psi_n\}_{n\geqslant 1}$, again labeled by n, for which the limit

$$\Psi := \lim_{n\to\infty} \psi_n \in X$$

exists and it satisfies

$$\Psi = P_y\mathfrak{K}(\rho_1)(x_0, \Psi). \tag{5.22}$$

Necessarily $\Psi \in P$ and $\|\Psi\| = 1$. Thus, $\Psi > 0$ and, hence, Ψ provides us with a positive eigenvector of $P_y\mathfrak{K}(\rho_1)(x_0, \cdot)$ associated with the eigenvalue 1. By our assumptions, this is impossible, unless $\rho_1 = \rho_0$, which contradicts (5.20). This concludes the proof. □

5.2. *The existence of coexistence states in the general case*

Subsequently, we fix $\lambda > 0$ and consider μ as the main bifurcation parameter. Also, $X := \mathcal{C}_T$ will be regarded as an ordered Banach space by its cone of non-negative functions, P. Note that P is normal, since every order interval is bounded. Also, note that, given any constant $M > 0$—to be chosen later—$(u, v) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ is a solution of (1.2) if and only if,

$$(u, v) \in \mathcal{C}_T \times \mathcal{C}_T \quad \text{and} \quad \mathfrak{F}(\mu, u, v) = 0, \tag{5.23}$$

where

$$\mathfrak{F} : \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T \to \mathcal{C}_T \times \mathcal{C}_T$$

is the operator defined by

$$\mathfrak{F}(\mu, u, v) := \begin{pmatrix} u - \mathfrak{R}_M\big[(\lambda\ell + M)u - au^2 - buv\big] \\ v - \mathfrak{R}_M\big[(\mu m + M)v - dv^2 - cuv\big] \end{pmatrix} \tag{5.24}$$

and $\mathfrak{R}_M$ stands for the operator introduced in (2.10),

$$\mathfrak{R}_M := J\left(\frac{\mathrm{d}}{\mathrm{d}t} + M\right)^{-1} : \mathcal{C}_T \to \mathcal{C}_T.$$

Thanks to Proposition 2.1, $\mathfrak{R}_M$ is compact and strongly order preserving. Moreover, for each $\mu \in \mathbb{R}$ we have that

$$\mathfrak{F}(\mu, \theta_{[\lambda\ell, a]}, 0) = 0$$

and, hence, since $\mathfrak{F}$ is real analytic, the general assumptions (HL) and (HR) of Section 5.1 are satisfied by choosing

$$(\rho, x_0, 0) := (\mu, \theta_{[\lambda\ell, a]}, 0), \qquad \mu \in \mathbb{R}. \tag{5.25}$$

Moreover, if $\mathfrak{F}(\mu, u, v) = 0$ with $u(t_0) > 0$ for some $t_0 \in \mathbb{R}$, then

$$u(t) = \mathrm{e}^{\int_{t_0}^{t} (\lambda\ell - au - bv)} u(t_0) > 0$$

for each $t \in \mathbb{R}$, since

$$u' = (\lambda\ell - au - bv)u,$$

and, hence, $u \gg 0$. Similarly, $v \gg 0$ if it is somewhere positive. Therefore, $\mathfrak{F}$ satisfies Hypothesis (HP) as well.

Maintaining the notations introduced in Section 5.1, for the choice (5.25) we have that the linearization

$$\mathfrak{L}(\mu) := D_{(u,v)}\mathfrak{F}(\mu, \theta_{[\lambda\ell,a]}, 0)$$

is given, for each $(u, v) \in \mathcal{C}_T \times \mathcal{C}_T$, through

$$\mathfrak{L}(\mu)\begin{pmatrix} u \\ v \end{pmatrix} := \begin{pmatrix} u - \mathfrak{R}_M\big[(\lambda\ell + M - 2a\theta_{[\lambda\ell,a]})u - b\theta_{[\lambda\ell,a]}v\big] \\ v - \mathfrak{R}_M\big[(\mu m + M - c\theta_{[\lambda\ell,a]})v\big] \end{pmatrix} \tag{5.26}$$

and, therefore,

$$\mathfrak{K}(\mu)\begin{pmatrix} u \\ v \end{pmatrix} := \begin{pmatrix} \mathfrak{R}_M\big[(\lambda\ell + M - 2a\theta_{[\lambda\ell,a]})u - b\theta_{[\lambda\ell,a]}v\big] \\ \mathfrak{R}_M\big[(\mu m + M - c\theta_{[\lambda\ell,a]})v\big] \end{pmatrix}.$$

Suppose $(u, v) \in N[\mathfrak{L}(\mu)]$ for some $\mu \in \mathbb{R}$ and $(u, v) \neq (0, 0)$. Then, $u, v \in \mathcal{C}^1_T$, and

$$\begin{cases} u' = (\lambda\ell - 2a\theta_{[\lambda\ell,a]})u - b\theta_{[\lambda\ell,a]}v, \\ v' = (\mu m - c\theta_{[\lambda\ell,a]})v. \end{cases} \tag{5.27}$$

Suppose $v = 0$. Then, $u \neq 0$ and the second equation of (5.27) holds true. Moreover, the first one reduces to

$$u' = (\lambda\ell - 2a\theta_{[\lambda\ell,a]})u. \tag{5.28}$$

Thanks to (2.28), we have that

$$\int_0^T (\lambda\ell - a\theta_{[\lambda\ell,a]}) = 0$$

and, hence,

$$\int_0^T (\lambda\ell - 2a\theta_{[\lambda\ell,a]}) = -\int_0^T a(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t < 0. \tag{5.29}$$

Therefore, (5.28) cannot admit a non-trivial T-periodic solution. This contradiction shows that, necessarily, $v \neq 0$, which implies

$$\mu = \mu_0 := \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t, \tag{5.30}$$

since $\hat{m} = 1$. Actually, in this case, there exists a constant $A \neq 0$ such that

$$v = A\varphi_v, \quad \varphi_v(t) := \mathrm{e}^{\int_0^t (\mu_0 m - c\theta_{[\lambda\ell,a]})}, \ t \in \mathbb{R}. \tag{5.31}$$

Thanks to (5.29), it follows from Proposition 2.1 that the problem

$$u' = (\lambda\ell - 2a\theta_{[\lambda\ell,a]})u - b\theta_{[\lambda\ell,a]}\varphi_v, \quad u \in \mathcal{C}_T^1, \tag{5.32}$$

possesses a unique solution. Namely,

$$\varphi_u := -\mathfrak{R}_{2a\theta_{[\lambda\ell,a]}-\lambda\ell}(b\theta_{[\lambda\ell,a]}\varphi_v). \tag{5.33}$$

Consequently, $\mathfrak{L}(\mu)$ is an isomorphism if $\mu \neq \mu_0$, whereas

$$N\big[\mathfrak{L}(\mu_0)\big] = \operatorname{span}\big[(\varphi_u, \varphi_v)\big]. \tag{5.34}$$

In particular, using the notations introduced in Section 5.1, this entails $\Sigma = \{\mu_0\}$. Note that, since $\varphi_v \gg 0$, condition (5.5) is satisfied. Also, for each $u, v \in X = \mathcal{C}_T$, we have that

$$P_v\mathfrak{K}(\mu_0)(u, v) = \mathfrak{R}_M\big[(\mu_0 m + M - c\theta_{[\lambda\ell,a]})v\big], \tag{5.35}$$

which is independent of u. Now, choosing M sufficiently large so that

$$\mu_0 m + M - c\theta_{[\lambda\ell,a]} \gg 0,$$

it follows from Proposition 2.1 that $P_v\mathfrak{K}(\mu_0)(u, v) \gg 0$ if $v > 0$. Therefore, (5.6) and (5.7) are satisfied. Actually, since $P_v\mathfrak{K}(\mu_0)(u, \cdot)$ is constant in u, condition (5.8) is as well satisfied. Furthermore, since $(\theta_{[\lambda\ell,a]}, 0)$ is the unique semi-trivial solution of (1.2) for each $\mu \in \mathbb{R}$ and 1 cannot be an eigenvalue of $P_v\mathfrak{K}(\mu)(0, \cdot)$, unless $\mu = \mu_0$—we have already seen it in the proof of (5.34)—to show that all assumptions of Theorem 5.1 are satisfied it remains to prove that $C(\mu_0) \neq 0$. To prove it, one can argue as follows. Note that $\mathfrak{L}(\mu)$ is analytic in μ and that, for each $u, v \in \mathcal{C}_T$, we have

$$\frac{\mathrm{d}\mathfrak{L}}{\mathrm{d}\mu}(\mu)(u, v) = \big(0, -\mathfrak{R}_M(mv)\big).$$

We now show that the following holds:

$$\frac{\mathrm{d}\mathfrak{L}}{\mathrm{d}\mu}(\mu_0)(\varphi_u, \varphi_v) = \big(0, -\mathfrak{R}_M(m\varphi_v)\big) \notin R\big[\mathfrak{L}(\mu_0)\big]. \tag{5.36}$$

Indeed, if (5.36) is not true, then there exists $v \in \mathcal{C}_T$ such that

$$-\mathfrak{R}_M(m\varphi_v) = v - \mathfrak{R}_M\big[(\mu_0 m + M - c\theta_{[\lambda\ell,a]})v\big].$$

Equivalently,

$$v' = (\mu_0 m - \theta_{[\lambda\ell,a]})v - m\varphi_v, \quad v \in \mathcal{C}_T. \tag{5.37}$$

Thanks to (5.30), it follows from Proposition 2.1(d) that (5.37) does not admit a solution, since $m\varphi_v > 0$. This contradiction shows (5.36). It should be noted that (5.36) is the classical transversality condition of M.G. Crandall and P.H. Rabinowitz [6]. Since $\mathfrak{L}(\mu_0)$ is Fredholm of index zero, because it is a compact perturbation of the identity, it is apparent from (5.36) that

$$\frac{\mathrm{d}\mathfrak{L}}{\mathrm{d}\mu}(\mu_0)\big(N\big[\mathfrak{L}(\mu_0)\big]\big) \oplus R\big[\mathfrak{L}(\mu_0)\big] = \mathcal{C}_T \times \mathcal{C}_T. \tag{5.38}$$

Thus, μ_0 is a 1-transversal eigenvalue of the analytic family $\mathfrak{L}(\mu)$, $\mu \in \mathbb{R}$, in the sense of [19, Definition 4.2.1]. Hence, the algebraic multiplicity defined there in equals one, i.e.,

$$\chi\big[\mathfrak{L}(\mu); \mu_0\big] = 1,$$

and, therefore, thanks to [19, Theorem 5.3.1], μ_0 is a pole of order one of the resolvent operator $\mathfrak{L}^{-1}(\mu)$, $\mu \neq \mu_0$. Consequently, thanks to [19, Theorem 5.6.2], the crossing number $C(\mu_0)$ defined in (5.3) satisfies

$$C(\mu_0) \in \{-1, 1\}.$$

Therefore, Theorem 5.1 provides us with the following existence result.

THEOREM 5.4. *Fix* $\lambda > 0$ *and consider* μ *as the main bifurcation parameter. Then, Problem* (1.2) *possesses a component of coexistence states*

$$\mathfrak{C}^\lambda_\mu \subset \mathbb{R} \times \operatorname{Int} P \times \operatorname{Int} P$$

such that

$$(\mu_0, \theta_{[\lambda\ell,a]}, 0) \in \bar{\mathfrak{C}}^\lambda_\mu, \tag{5.39}$$

where μ_0 *is defined through* (5.30). *Moreover, if* $\mathfrak{C}^\lambda_\mu$ *is bounded, then there exists* $\mu_* > 0$ *such that*

$$\lambda = \frac{1}{T}\int_0^T b(t)\theta_{[\mu_* m,d]}(t)\,\mathrm{d}t \tag{5.40}$$

and

$$(\mu_*, 0, \theta_{[\mu_* m,d]}) \in \bar{\mathfrak{C}}^\lambda_\mu. \tag{5.41}$$

Therefore, $\mathfrak{C}^\lambda_\mu$ *links the semi-trivial states* $(\mu_0, \theta_{[\lambda\ell,a]}, 0)$ *and* $(\mu_*, 0, \theta_{[\mu_* m,d]})$ *if it is bounded.*

PROOF. The existence of $\mathfrak{C}_\mu^\lambda$ bifurcating from the curve $(\mu, u, v) = (\mu, \theta_{[\lambda\ell, a]}, 0)$ at $\mu = \mu_0$ is an immediate consequence from Theorem 5.1.

Suppose $\mathfrak{C}_\mu^\lambda$ is bounded. Then, it satisfies alternative A2 of Theorem 5.1 and, hence, there exists $(\mu_*, v_*) \in \mathbb{R} \times P$ such that

$$(\mu_*, 0, v_*) \in \bar{\mathfrak{C}}_\mu^\lambda.$$

Thus, there exists a sequence of coexistence states of (1.2), say $\{(\mu_n, u_n, v_n)\}_{n \geqslant 1}$, such that

$$\lim_{n \to \infty} (\mu_n, u_n, v_n) = (\mu_*, 0, v_*).$$

For each $n \geqslant 1$, we have that

$$u_n = \mathfrak{R}_M\big[(\lambda\ell + M)u_n - a u_n^2 - b u_n v_n\big].$$

Thus, setting

$$U_n := \frac{u_n}{\|u_n\|_{\mathcal{C}_T}}, \quad n \geqslant 1,$$

gives

$$U_n = \mathfrak{R}_M\big[(\lambda\ell + M)U_n - a u_n U_n - b v_n U_n\big], \quad n \geqslant 1. \tag{5.42}$$

Since $\{(\lambda\ell + M)U_n - a u_n U_n - b v_n U_n\}_{n \geqslant 1}$ is bounded in $\mathcal{C}_T$ and $\mathfrak{R}_M$ is compact, there exists a subsequence of U_n, labeled again by n, such that

$$\lim_{n \to \infty} U_n = \Phi.$$

Necessarily, $\|\Phi\|_{\mathcal{C}_T} = 1$, and, hence, $\Phi > 0$, since $u_n \gg 0$, $n \geqslant 1$. Moreover, passing to the limit as $n \to \infty$ in (5.42), we find that

$$\Phi = \mathfrak{R}_M\big[(\lambda\ell + M - b v_*)\Phi\big].$$

Thus, $\Phi \in \mathcal{C}_T^1$, $\Phi > 0$, and

$$\Phi' = (\lambda\ell - b v_*)\Phi. \tag{5.43}$$

Suppose $v_* = 0$, then, we find from (5.43) that $\lambda\hat{\ell} = \lambda = 0$, which is impossible, since we are assuming $\lambda > 0$. Thus, $v_* > 0$. Actually, since $(\mu_*, 0, v_*)$ solves (1.2), necessarily

$$\mu_* > 0 \quad \text{and} \quad v_* = \theta_{[\mu_* m, d]} \gg 0.$$

Consequently, (5.43) becomes into

$$\Phi' = (\lambda\ell - b\theta_{[\mu_* m,d]})\Phi, \tag{5.44}$$

and, therefore,

$$\lambda = \frac{1}{T}\int_0^T b(t)\theta_{[\mu_* m,d]}(t)\,\mathrm{d}t,$$

which concludes the proof of the theorem. □

By symmetry, if $\mu > 0$ is fixed and $\lambda \in \mathbb{R}$ is regarded as the main bifurcation parameter, then the following result is satisfied.

THEOREM 5.5. *Fix $\mu > 0$ and consider λ as the main bifurcation parameter. Then, Problem* (1.2) *possesses a component of coexistence states*

$$\mathfrak{C}_\lambda^\mu \subset \mathbb{R} \times \operatorname{Int} P \times \operatorname{Int} P$$

such that

$$(\lambda_0, 0, \theta_{[\mu m,d]}) \in \bar{\mathfrak{C}}_\lambda^\mu, \quad \lambda_0 := \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t. \tag{5.45}$$

Moreover, if $\mathfrak{C}_\lambda^\mu$ is bounded, then there exists $\lambda_* > 0$ such that

$$\mu = \frac{1}{T}\int_0^T c(t)\theta_{[\lambda_*\ell,a]}(t)\,\mathrm{d}t \tag{5.46}$$

and

$$(\lambda_*, \theta_{[\lambda_*\ell,a]}, 0) \in \bar{\mathfrak{C}}_\lambda^\mu. \tag{5.47}$$

Therefore, $\mathfrak{C}_\lambda^\mu$ links the semi-trivial states $(\lambda_0, 0, \theta_{[\mu m,d]})$ and $(\lambda_*, \theta_{[\lambda_*\ell,a]}, 0)$ if it is bounded.

5.3. *Local structure of $\mathfrak{C}_\mu^\lambda$ and $\mathfrak{C}_\lambda^\mu$ at $(\mu_0, \theta_{[\lambda\ell,a]}, 0)$ and $(\lambda_0, 0, \theta_{[\mu m,d]})$*

The following result shows that, for each $\lambda > 0$, $\mathfrak{C}_\mu^\lambda$ consists of a real analytic curve of coexistence states of (1.2) in a neighborhood of

$$(\mu, u, v) = (\mu_0, \theta_{[\lambda\ell,a]}, 0).$$

Although this is a consequence from the celebrated theorem of M.G. Crandall and P.H. Rabinowitz [6], we shall give a self-contained proof—whose uniqueness part is based on Theorem 5.2—in order to calculate the bifurcation direction of the curve of coexistence states. Throughout this section, all notation introduced in Section 5.2 will be maintained.

THEOREM 5.6. *Fix* $\lambda > 0$ *and regard to* $\mu \in \mathbb{R}$ *as the main bifurcation parameter. Then, there exist* $\varepsilon > 0$ *and two real analytic maps*

$$\mu : (-\varepsilon, \varepsilon) \to \mathbb{R}, \qquad (u_1, v_1) : (-\varepsilon, \varepsilon) \to R\big[\mathfrak{L}(\mu_0)\big] \times R\big[\mathfrak{L}(\mu_0)\big], \tag{5.48}$$

such that

$$\mu(0) = \mu_0 = \frac{1}{T}\int_0^T c(t)\,\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t, \quad \big(u_1(0), v_1(0)\big) = (0,0), \tag{5.49}$$

and, for each $s \in (-\varepsilon, \varepsilon)$,

$$\mathfrak{F}\big(\mu(s), \theta_{[\lambda\ell,a]} + s\big[\varphi_u + u_1(s)\big], s\big[\varphi_v + v_1(s)\big]\big) = 0, \tag{5.50}$$

where φ_v *and* φ_u *are the functions defined by* (5.31) *and* (5.33), *respectively. Moreover, there exists* $\delta > 0$ *such that, if* B_δ^0 *stands for the ball of radius* δ *centered at* $(\mu_0, \theta_{[\lambda\ell,a]}, 0) \in \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ *and, for each* $s \in (-\varepsilon, \varepsilon)$, *we set*

$$\big(u(s), v(s)\big) := \big(\theta_{[\lambda\ell,a]} + s\big[\varphi_u + u_1(s)\big], s\big[\varphi_v + v_1(s)\big]\big), \tag{5.51}$$

then

$$\mathfrak{C}_\mu^\lambda \cap B_\delta^0 = \big\{\big(\mu(s), u(s), v(s)\big) : \ 0 < s < \varepsilon\big\} \cap B_\delta^0. \tag{5.52}$$

Furthermore,

$$\frac{\mathrm{d}\mu}{\mathrm{d}s}(0) = \frac{1}{T}\int_0^T \big[c(t)\varphi_u(t) + d(t)\varphi_v(t)\big]\,\mathrm{d}t. \tag{5.53}$$

PROOF. Consider the operator

$$\mathfrak{G} : \mathbb{R}^2 \times R\big[\mathfrak{L}(\mu_0)\big] \to \mathcal{C}_T \times \mathcal{C}_T$$

defined, for each $(s, \mu, u_1, v_1) \in \mathbb{R}^2 \times R[\mathfrak{L}(\mu_0)]$, by

$$\mathfrak{G}(s, \mu, u_1, v_1) := \begin{cases} s^{-1}\mathfrak{F}\big(\mu, \theta_{[\lambda\ell,a]} + s(\varphi_u + u_1), s(\varphi_v + v_1)\big), & \text{if } s \neq 0, \\ D_{(u,v)}\mathfrak{F}(\mu, \theta_{[\lambda\ell,a]}, 0)(\varphi_u + u_1, \varphi_v + v_1), & \text{if } s = 0. \end{cases}$$

The operator $\mathfrak{G}$ is real analytic and it satisfies

$$\mathfrak{G}(0, \mu_0, 0, 0) = D_{(u,v)}\mathfrak{F}(\mu_0, \theta_{[\lambda\ell,a]}, 0)(\varphi_u, \varphi_v) = \mathfrak{L}(\mu_0)(\varphi_u, \varphi_v) = (0, 0).$$

Moreover, for each $(\mu, u_1, v_1) \in \mathbb{R} \times R[\mathfrak{L}(\mu_0)]$,

$$\begin{aligned}
D_{(\mu,u_1,v_1)}\mathfrak{G}(0,\mu_0,0,0)(\mu,u_1,v_1) &= \lim_{h\to 0}\left[h^{-1}\mathfrak{G}(0,\mu_0+h\mu,hu_1,hv_1)\right]\\
&= \lim_{h\to 0}\left[h^{-1}D_{(u,v)}\mathfrak{F}(\mu_0+h\mu,\theta_{[\lambda\ell,a]},0)(\varphi_u+hu_1,\varphi_v+hv_1)\right]\\
&= \lim_{h\to 0}\left[h^{-1}\mathfrak{L}(\mu_0+h\mu)(\varphi_u+hu_1,\varphi_v+hv_1)\right]\\
&= \mathfrak{L}(\mu_0)(u_1,v_1)+\mu\frac{\mathrm{d}\mathfrak{L}}{\mathrm{d}\mu}(\mu_0)(\varphi_u,\varphi_v).
\end{aligned}$$

Thanks to the proof of Theorem 5.1, we already know that

$$N\big[\mathfrak{L}(\mu_0)\big] \oplus R\big[\mathfrak{L}(\mu_0)\big] = \mathcal{C}_T \times \mathcal{C}_T,$$

though this easily follows from the fact that $(\varphi_u, \varphi_v) \notin R[\mathfrak{L}(\mu_0)]$. Therefore,

$$\mathfrak{L}(\mu_0)|_{R[\mathfrak{L}(\mu_0)]} : R\big[\mathfrak{L}(\mu_0)\big] \to R\big[\mathfrak{L}(\mu_0)\big]$$

is a linear isomorphism and, hence, it follows from (5.36) that

$$D_{(\mu,u_1,v_1)}\mathfrak{G}(0,\mu_0,0,0) : \mathbb{R} \times R\big[\mathfrak{L}(\mu_0)\big] \to \mathcal{C}_T \times \mathcal{C}_T$$

is a linear isomorphism as well. Consequently, thanks to the implicit function theorem, there exist $\varepsilon > 0$ and the real analytic maps (5.48), satisfying (5.49), such that, for each $s \in (-\varepsilon, \varepsilon)$,

$$\mathfrak{G}\big(s, \mu(s), u_1(s), v_1(s)\big) = 0.$$

Moreover, if

$$\mathfrak{G}(s, \mu, u_1, v_1) = 0$$

for some (s, μ, u_1, v_1) sufficiently close to $(0, \mu_0, 0, 0)$, then

$$(s, \mu, u_1, v_1) = \big(s, \mu(s), u_1(s), v_1(s)\big).$$

It should be noted that, by the definition of $\mathfrak{G}$ itself, for each $s \in (-\varepsilon, \varepsilon)$, (5.50) is satisfied. Now, we will prove the uniqueness result.

Thanks to Theorem 5.2, there exists $\delta > 0$ such that, for each

$$(\mu, u, v) \in \big[\mathfrak{C}^\lambda_\mu \setminus \big\{(\mu_0, \theta_{[\lambda\ell,a]}, 0)\big\}\big] \cap B_\delta,$$

there exists a unique $(s, \tilde{u}, \tilde{v}) \in \mathbb{R} \times R[\mathfrak{L}(\mu_0)]$ such that

$$\mu = \mu(s), \quad (u, v) = (\theta_{[\lambda\ell,a]}, 0) + s(\varphi_u, \varphi_v) + \big(\tilde{u}, \tilde{v}\big), \tag{5.54}$$

with $s > \eta\,\|(u-\theta_{[\lambda\ell,a]}, v)\| > 0$. Moreover,

$$\mu = \mu_0 + \mathrm{o}(1) \quad \text{and} \quad \big(\tilde{u}, \tilde{v}\big) = \mathrm{o}(s) \quad \text{as } s \to 0. \tag{5.55}$$

So, setting

$$U(s) := s^{-1}\tilde{u}(s) = \mathrm{o}(1), \qquad V(s) := s^{-1}\tilde{v}(s) = \mathrm{o}(1),$$

we have that $(U(s), V(s)) \in R[\mathfrak{L}(\mu_0)]$ and

$$\mathfrak{F}\big(\mu(s), \theta_{[\lambda\ell,a]} + s\big(\varphi_u + U(s)\big), s\big(\varphi_s + V(s)\big)\big) = 0.$$

Thus,

$$\mathfrak{G}\big(s, \mu(s), U(s), V(s)\big) = 0$$

and, therefore, by the uniqueness obtained as an application of the implicit function theorem to $\mathfrak{G}$, if δ is taken sufficiently small, we have that

$$\big(U(s), V(s)\big) = \big(u_1(s), v_1(s)\big)$$

and, consequently,

$$(\mu, u, v) = \big(\mu(s), u(s), v(s)\big),$$

which concludes the proof of (5.52).

To prove (5.53), consider the second equation of (5.50). Then, after division by s, gives

$$\begin{aligned}
&\varphi_v + s\frac{\mathrm{d}v_1}{\mathrm{d}s}(0) + \mathrm{O}\big(s^2\big) \\
&\quad = \mathfrak{R}_M\bigg\{\bigg[\bigg(\mu_0 + s\frac{\mathrm{d}\mu}{\mathrm{d}s}(0) + \mathrm{O}\big(s^2\big)\bigg)m + M\bigg]\bigg(\varphi_v + s\frac{\mathrm{d}v_1}{\mathrm{d}s}(0) + \mathrm{O}\big(s^2\big)\bigg) \\
&\qquad - \mathrm{d}s\bigg(\varphi_v + s\frac{\mathrm{d}v_1}{\mathrm{d}s}(0) + \mathrm{O}\big(s^2\big)\bigg)^2 - c\big(\theta_{[\lambda\ell,a]} + s\varphi_u + \mathrm{O}\big(s^2\big)\big) \\
&\qquad \times \bigg(\varphi_v + s\frac{\mathrm{d}v_1}{\mathrm{d}s}(0) + \mathrm{O}\big(s^2\big)\bigg)\bigg\}.
\end{aligned}$$

By definition,

$$\varphi_v = \mathfrak{R}_M\big\{(\mu_0 m + M - c\theta_{[\lambda\ell,a]})\varphi_v\big\}.$$

Thus, dividing by s the previous identity, passing to the limit as $s \to 0$, and rearranging terms yields to

$$\frac{\mathrm{d}v_1}{\mathrm{d}s}(0) = \mathfrak{R}_M\bigg[(\mu_0 m + M - c\theta_{[\lambda\ell,a]})\frac{\mathrm{d}v_1}{\mathrm{d}s}(0) + \frac{\mathrm{d}\mu}{\mathrm{d}s}(0)m\varphi_v - d\varphi_v^2 - c\varphi_u\varphi_v\bigg],$$

or, equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}v_1}{\mathrm{d}s}(0) = (\mu_0 m - c\theta_{[\lambda\ell,a]})\frac{\mathrm{d}v_1}{\mathrm{d}s}(0) + \frac{\mathrm{d}\mu}{\mathrm{d}s}(0)m\varphi_v - d\varphi_v^2 - c\varphi_u\varphi_v.$$

Thus, thanks to Proposition 2.1(d),

$$\int_0^T \mathrm{e}^{-\int_0^t(\mu_0 m - c\theta_{[\lambda\ell,a]})}\left(\frac{\mathrm{d}\mu}{\mathrm{d}s}(0)m(t)\varphi_v(t) - d(t)\varphi_v^2(t) - c(t)\varphi_u(t)\varphi_v(t)\right)\mathrm{d}t = 0,$$

and, hence, (5.31) entails

$$\int_0^T \left(\frac{\mathrm{d}\mu}{\mathrm{d}s}(0)m(t) - d(t)\varphi_v(t) - c(t)\varphi_u(t)\right)\mathrm{d}t = 0.$$

Therefore,

$$\frac{\mathrm{d}\mu}{\mathrm{d}s}(0) = \frac{1}{T}\int_0^T \big[c(t)\varphi_u(t) + d(t)\varphi_v(t)\big]\,\mathrm{d}t,$$

since $\hat{m} = 1$. This concludes the proof. □

By symmetry, for each $\mu > 0$, $\mathfrak{C}_\lambda^\mu$ consists of a real analytic curve of coexistence states of (1.2) in a neighborhood of $(\lambda, u, v) = (\lambda_0, 0, \theta_{[\mu m,d]})$. Actually, setting

$$\lambda_0 := \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t, \qquad \tilde{\varphi}_u(t) := \mathrm{e}^{\int_0^t(\lambda_0\ell - b\theta_{[\mu m,d]})}, \quad t \in \mathbb{R}, \tag{5.56}$$

and

$$\tilde{\varphi}_v := -\mathfrak{R}_{2d\theta_{[\mu m,d]}-\mu m}(c\theta_{[\mu m,d]}\tilde{\varphi}_u), \tag{5.57}$$

there exists $\delta > 0$ such that in the ball of radius $\delta > 0$ centered at $(\lambda_0, 0, \theta_{[\mu m,d]})$ the component $\mathfrak{C}_\lambda^\mu$ consists of a real analytic curve of the form

$$(\lambda, u, v) = \big(\lambda(s), s\big(\tilde{\varphi}_u + \mathrm{O}(s)\big), \theta_{[\mu m,d]} + s\big(\tilde{\varphi}_v + \mathrm{O}(s)\big)\big), \quad s \sim 0,\ s > 0,$$

where $\lambda(0) = \lambda_0$ and

$$\frac{\mathrm{d}\lambda}{\mathrm{d}s}(0) = \frac{1}{T}\int_0^T \big[b(t)\tilde{\varphi}_v(t) + a(t)\tilde{\varphi}_u(t)\big]\,\mathrm{d}t. \tag{5.58}$$

6. The symbiotic model ($b < 0$ and $c < 0$)

In this section we analyze the structure of the set of coexistence states of (1.2) in the special case when $b < 0$ and $c < 0$. Our main result shows that, for each $\lambda > 0$ and $\mu > 0$, the components of coexistence states of (1.2) constructed in Section 5, $\mathfrak{C}_\lambda^\mu$ and $\mathfrak{C}_\mu^\lambda$, are real analytic curves. These curves fiber a real analytic surface entirely formed by coexistence states and linking the corresponding surfaces of semi-trivial solutions along their curves of neutral stability. Also, the stability of the coexistence states along these curves is completely ascertained. It turns out that the stability of the coexistence states along these curves changes at any turning point along the curve, while it remains unchanged when an hysteresis point is passed by. Some of the results of this section are substantial improvements of some results of M.A. Krasnosel'skii [16]. The theory developed in this section can be viewed as a development of many ideas and results going back to [17].

6.1. *Periodic systems of cooperative type*

In this section we study the linear periodic system

$$\begin{cases} u' = \alpha u + \beta v, \\ v' = \gamma u + \rho v, \end{cases} \tag{6.1}$$

where $\alpha, \beta, \gamma, \rho \in \mathcal{C}_T$ satisfy

$$\int_0^T \alpha < 0, \qquad \int_0^T \rho < 0, \qquad \beta > 0, \qquad \gamma > 0, \tag{6.2}$$

as well as its associated Cauchy problem

$$\begin{cases} u' = \alpha u + \beta v, \\ v' = \gamma u + \rho v, \\ u(0) = u_0, \quad v(0) = v_0. \end{cases} \tag{6.3}$$

The next result is the key theorem from which most of the results of this section will follow.

THEOREM 6.1. *Suppose* (6.1) *possesses a solution* $(u, v) \in \mathcal{C}_T^1 \times \mathcal{C}_T^1 \setminus \{(0,0)\}$. *Then, either* $u \gg 0$ *and* $v \gg 0$, *or else* $u \ll 0$ *and* $v \ll 0$. *Therefore, all the coexistence states of* (1.2) *must be ordered, in the sense that, for any pair of coexistence states,* $(u_1, v_1) \neq (u_2, v_2)$, *either* $u_1 \gg u_2$ *and* $v_1 \gg v_2$, *or* $u_1 \ll u_2$ *and* $v_1 \ll v_2$.

PROOF. If $u = 0$, then $v' = \rho v$, and it follows from Proposition 2.1 that $v = 0$. Thus, $u \neq 0$, since $(u, v) \neq (0, 0)$. As the system is linear, without lost of generality we can assume that $u(t) > 0$ for some $t \in \mathbb{R}$. If $u \gg 0$, then $\gamma u > 0$ and, therefore,

$$v = \mathfrak{R}_{-\rho}(\gamma u) \gg 0,$$

which concludes the proof of the first thesis of the theorem. To show that $u \gg 0$ we will argue by contradiction. Suppose there are $t_0 < t_1$ such that

$$u(t_0) = u(t_1) = 0 \quad \text{and} \quad u(t) > 0 \quad \text{for each } t \in (t_0, t_1). \tag{6.4}$$

Then, for each $t \in \mathbb{R}$, we have that

$$u(t) = \int_{t_0}^{t} \mathrm{e}^{\int_\tau^t \alpha} \beta(\tau) v(\tau)\, \mathrm{d}\tau \tag{6.5}$$

and

$$v(t) = \mathrm{e}^{\int_{t_0}^t \rho} v(t_0) + \int_{t_0}^{t} \mathrm{e}^{\int_\tau^t \rho} \gamma(\tau) u(\tau)\, \mathrm{d}\tau. \tag{6.6}$$

Suppose $v(t_0) \geqslant 0$. Then, thanks to (6.4) and (6.6), we find that, for each $t \in [t_0, t_1]$,

$$v(t) \geqslant \mathrm{e}^{\int_{t_0}^t \rho} v(t_0) \geqslant 0,$$

since $\gamma u \geqslant 0$. Thus, (6.4) and (6.5) imply

$$\begin{aligned} 0 = u(t_1) &= \int_{t_0}^{t_1} \mathrm{e}^{\int_\tau^{t_1} \alpha} \beta(\tau) v(\tau)\, \mathrm{d}\tau \\ &\geqslant \int_{t_0}^{\frac{t_0+t_1}{2}} \mathrm{e}^{\int_\tau^{t_1} \alpha} \beta(\tau) v(\tau)\, \mathrm{d}\tau \\ &= \exp\biggl(\int_{\frac{t_0+t_1}{2}}^{t_1} \alpha\biggr) \int_{t_0}^{\frac{t_0+t_1}{2}} \exp\biggl(\int_\tau^{\frac{t_0+t_1}{2}} \alpha\biggr) \beta(\tau) v(\tau)\, \mathrm{d}\tau \\ &= \exp\biggl(\int_{\frac{t_0+t_1}{2}}^{t_1} \alpha\biggr) u\biggl(\frac{t_0+t_1}{2}\biggr) > 0, \end{aligned}$$

which is impossible. Hence, $v(t_0) < 0$. So, there is $\delta > 0$ such that $t_0 < t_0 + \delta < t_1$, for which $v(t) < 0$ if $t \in [t_0, t_0 + \delta]$. Thus, we find from (6.5) that $u(t) \leqslant 0$ for each $t \in [t_0, t_0 + \delta]$, which contradicts (6.4). Therefore, necessarily $u \gg 0$, which concludes the proof of the first statement.

Now, suppose $(u_1, v_1) \neq (u_2, v_2)$ are two coexistence states of (1.2). Then,

$$\begin{aligned} (u_2 - u_1)' &= \lambda \ell (u_2 - u_1) - a(u_2 + u_1)(u_2 - u_1) - b u_2 v_2 + b u_1 v_1 \\ &= \bigl[\lambda \ell - a(u_2 + u_1) - b v_2\bigr](u_2 - u_1) - b u_1 (v_2 - v_1). \end{aligned}$$

By symmetry,

$$(v_2 - v_1)' = \bigl[\mu m - d(v_2 + v_1) - c u_2\bigr](v_2 - v_1) - c v_1 (u_2 - u_1).$$

Thus, the pair

$$(U,V):=(u_2-u_1,v_2-v_1)\in\mathcal{C}^1_T\times\mathcal{C}^1_T$$

satisfies

$$\begin{cases} U'=\bigl[\lambda\ell-a(u_2+u_1)-bv_2\bigr]U-bu_1V,\\ V'=\bigl[\mu m-d(v_2+v_1)-cu_2\bigr]V-cv_1U. \end{cases}$$

Since $u_1\gg 0$, $v_1\gg 0$, $b<0$ and $c<0$, necessarily $-bu_1>0$ and $-cv_1>0$. Moreover, since $u_2\gg 0$ and $v_2\gg 0$ satisfy

$$\frac{u_2'}{u_2}=\lambda\ell-au_2-bv_2,\qquad \frac{v_2'}{v_2}=\mu m-dv_2-cu_2,$$

integrating in $[0,T]$ gives

$$\int_0^T(\lambda\ell-au_2-bv_2)=0,\qquad \int_0^T(\mu m-dv_2-cu_2)=0. \tag{6.7}$$

Thus,

$$\int_0^T\bigl[\lambda\ell-a(u_2+u_1)-bv_2\bigr]=-\int_0^T a(t)u_1(t)\,\mathrm{d}t<0$$

and

$$\int_0^T\bigl[\mu m-d(v_2+v_1)-cu_2\bigr]=-\int_0^T d(t)v_1(t)\,\mathrm{d}t<0.$$

Therefore, by the first part of the theorem, either $U\gg 0$ and $V\gg 0$, or $U\ll 0$ and $V\ll 0$, which concludes the proof. □

PROPOSITION 6.2. *Suppose* $(\bar u,\bar v)\in\mathcal{C}^1(\mathbb{R}_+)\times\mathcal{C}^1(\mathbb{R}_+)$, $\mathbb{R}_+:=[0,\infty)$, *satisfy*

$$\begin{cases} \bar u'\geqslant\alpha\bar u+\beta\bar v,\\ \bar v'\geqslant\gamma\bar u+\rho\bar v,\\ \bar u(0)>0,\quad \bar v(0)>0. \end{cases} \tag{6.8}$$

Then, $\bar u(t)>0$ *and* $\bar v(t)>0$ *for each* $t\geqslant 0$.

If $u_0+v_0>0$ *and* $u_0v_0=0$, *then there exists* $t_0>0$ *such that the solution* (u,v) *of* (6.3) *satisfies* $u(t)>0$ *and* $v(t)>0$ *for each* $t\geqslant t_0$.

PROOF. Setting

$$d:=\bar u'-\alpha\bar u-\beta\bar v\geqslant 0,\qquad \delta:=\bar v'-\gamma\bar u-\rho\bar v\geqslant 0,$$

and integrating the resulting system show that, for each $t \in \mathbb{R}_+$,

$$\bar{u}(t) = \mathrm{e}^{\int_0^t \alpha} \bar{u}(0) + \int_0^t \mathrm{e}^{\int_\tau^t \alpha} \bigl[\beta(\tau)\bar{v}(\tau) + d(\tau)\bigr] \, \mathrm{d}\tau \tag{6.9}$$

and

$$\bar{v}(t) = \mathrm{e}^{\int_0^t \rho} \bar{v}(0) + \int_0^t \mathrm{e}^{\int_\tau^t \rho} \bigl[\gamma(\tau)\bar{u}(\tau) + \delta(\tau)\bigr] \, \mathrm{d}\tau. \tag{6.10}$$

Now, suppose $\bar{u}(0) > 0$ and $\bar{v}(0) > 0$. Then, either $\bar{u}(t) > 0$ for each $t \in \mathbb{R}_+$, or there exists t^* such that

$$\bar{u}\bigl(t^*\bigr) = 0 \quad \text{and} \quad \bar{u}(t) > 0 \quad \text{for each } t \in \bigl[0, t^*\bigr). \tag{6.11}$$

Suppose (6.11) occurs. Then, since $\bar{u}(0) > 0$, $\bar{v}(0) > 0$, $\beta > 0$ and $d \geqslant 0$, it follows from (6.9) that there exists $\tilde{t} \in (0, t^*)$ such that

$$\bar{v}\bigl(\tilde{t}\bigr) = 0 \quad \text{and} \quad \bar{v}(t) > 0 \quad \text{for each } t \in \bigl[0, \tilde{t}\bigr),$$

which contradicts (6.10), since $\bar{v}(0) > 0$ and $\gamma \geqslant 0$, $\delta \geqslant 0$ and $\bar{u} \geqslant 0$ in $[0, \tilde{t}]$. Therefore, $\bar{u}(t) > 0$ for each $t \in \mathbb{R}_+$ and, thanks to (6.10), $\bar{v}(t) > 0$ for each $t \in \mathbb{R}_+$. This concludes the proof of the first part of the proposition.

Now, suppose $u_0 > 0$ and $v_0 = 0$, and let (u, v) denote the unique solution of (6.3). Thanks to the first part of the proposition, for each $\varepsilon > 0$ the solution $(u_\varepsilon, v_\varepsilon)$ of (6.1) with initial data $(u_\varepsilon(0), v_\varepsilon(0)) = (u_0, v_0 + \varepsilon)$ satisfies $u_\varepsilon(t) > 0$ and $v_\varepsilon(t) > 0$ for each $t \geqslant 0$. Thus, by continuous dependence with respect to the initial data, passing to the limit as $\varepsilon \downarrow 0$ we find that $u > 0$ and $v \geqslant 0$. Hence, for each $t \geqslant 0$,

$$u(t) = \mathrm{e}^{\int_0^t \alpha} u_0 + \int_0^t \mathrm{e}^{\int_\tau^t \alpha} \beta(\tau) v(\tau) \, \mathrm{d}\tau \geqslant \mathrm{e}^{\int_0^t \alpha} u_0 > 0.$$

If $v = 0$ in $\mathbb{R}_+$, then we find from the v-equation that $\gamma u = 0$ in $\mathbb{R}_+$, which is impossible since $\gamma > 0$ and $u \gg 0$. Thus, there exists $t_0 > 0$ such that $v(t_0) > 0$. Therefore, since $u(t_0) > 0$, by applying the first part of this theorem, we find that $u(t) > 0$ and $v(t) > 0$ for each $t \geqslant t_0$, which concludes the proof. □

Subsequently, we denote by $U(t)$ the fundamental matrix of solutions of (6.1) such that $U(0) = I := I_{\mathbb{R}^2}$. Then, $U(T)$ provides us with the Poincaré map associated to (6.1). Thanks to Liouville's formula,

$$0 < \det U(T) = \mathrm{e}^{\int_0^T (\alpha + \rho)} < 1,$$

since $\int_0^T (\alpha + \rho) < 0$. Thus, if we denote by ν_1 and ν_2 the two eigenvalues of $U(T)$, counting algebraic multiplicities, then

$$0 < \nu_1 \nu_2 < 1. \tag{6.12}$$

Actually, the following result is satisfied.

THEOREM 6.3. *The characteristic multipliers of* (6.1), ν_1 *and* ν_2, *are real and distinct, and satisfy* (6.12). *Moreover, if we order them so that* $0 < \nu_1 < \nu_2$, *then,* $0 < \nu_1 < 1$,

$$N\big[U(T) - \nu_1 I\big] = \operatorname{span}\big[(x_0, y_0)\big] \quad \textit{with } x_0 < 0 \textit{ and } y_0 > 0,$$

and

$$N\big[U(T) - \nu_2 I\big] = \operatorname{span}\big[(x_0, y_0)\big] \quad \textit{with } x_0 > 0 \textit{ and } y_0 > 0.$$

In particular, (6.1) *is asymptotically stable if* $\nu_2 < 1$, *neutrally stable if* $\nu_2 = 1$, *and unstable if* $\nu_2 > 1$. *Furthermore,*

$$\lambda_0 := -\frac{1}{T}\operatorname{Log}\nu_2 \tag{6.13}$$

provides us with the unique value of $\lambda \in \mathbb{R}$ *for which the eigenvalue problem*

$$\begin{cases} u' = \alpha u + \beta v + \lambda u, \\ v' = \gamma u + \rho v + \lambda v, \end{cases} \tag{6.14}$$

admits a solution $(u, v) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ *with* $u \gg 0$ *and* $v \gg 0$.

PROOF. Suppose $u_0 > 0$, $v_0 > 0$, and let $(u(t), v(t))$ denote the unique solution of (6.3). Then, thanks to Proposition 6.2, $u(T) > 0$ and $v(T) > 0$, and, hence, $U(T)$ sends the interior of the first quadrant of the plane into itself. Setting

$$U(T) = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix},$$

the invariance of the first quadrant entails $u_{ij} \geqslant 0$ for each $i, j \in \{1, 2\}$. On the other hand, the eigenvalues of $U(T)$ (ν_1 and ν_2) are the roots of

$$z^2 - (u_{11} + u_{22})z + u_{11}u_{22} - u_{12}u_{21} = 0,$$

which are given through

$$\frac{1}{2}\big(u_{11} + u_{22} \pm \sqrt{\Delta}\big),$$

where

$$\Delta := (u_{11} + u_{22})^2 - 4(u_{11}u_{22} - u_{12}u_{21}) = (u_{11} - u_{22})^2 + 4u_{12}u_{21} \geqslant 0.$$

Thus, both are real. Moreover, since

$$\det U(T) = u_{11}u_{22} - u_{12}u_{21} = \nu_1\nu_2 > 0,$$

necessarily

$$u_{11} > 0, \quad u_{22} > 0, \quad \nu_1 + \nu_2 = u_{11} + u_{22}.$$

We now show that $u_{12} > 0$ and $u_{21} > 0$. Indeed, if $u_{12} = 0$, then

$$\{\nu_1, \nu_2\} = \{u_{11}, u_{22}\}$$

and, hence,

$$N\big[U(T) - u_{22}I\big] = \operatorname{span}\big[(0, 1)\big].$$

Consequently, the solution of (6.3) with initial data $(0, 1)$, say (u, v), satisfies $u(nT) = 0$ for each $n \geqslant 1$, which contradicts the second part of Proposition 6.2. This contradiction shows that $u_{12} > 0$. Similarly, $u_{21} > 0$ follows. Thus, $\Delta > 0$ and, therefore, $\nu_1 \neq \nu_2$, say $0 < \nu_1 < \nu_2$. Now, note that

$$\min\{u_{11}, u_{22}\} < \nu_2 = \frac{1}{2}\big(u_{11} + u_{22} + \sqrt{\Delta}\,\big).$$

In case $u_{11} = \min\{u_{11}, u_{22}\}$, $(u_{12}/(\nu_2 - u_{11}), 1)$ provides us with an eigenvector of $U(T)$ associated to ν_2, while, in case $u_{22} = \min\{u_{11}, u_{22}\}$, $(1, u_{21}/(\nu_2 - u_{22}))$ provides us with an eigenvector of $U(T)$ associated to ν_2. In both situations the two components of the eigenvector are positive. Similarly, in case $u_{11} = \min\{u_{11}, u_{22}\}$, $(1, u_{21}/(\nu_1 - u_{22}))$ is an eigenvector of $U(T)$ associated to ν_1 whose second component is negative, since

$$\nu_1 = u_{11} + u_{22} - \nu_2 < u_{22}.$$

By symmetry, in case $u_{22} = \min\{u_{11}, u_{22}\}$, $(u_{12}/(\nu_1 - u_{11}), 1)$ is an eigenvector of $U(T)$ associated to ν_1 whose second component is negative. Therefore, the eigenvector of $U(T)$ associated to ν_2 can be chosen in the interior of the first quadrant, while the eigenvector associated to ν_1 can be chosen in the interior of the second quadrant.

Now, note that the change of variable

$$(u, v) = \mathrm{e}^{\lambda t}(x, y)$$

transforms (6.14) into

$$\begin{cases} x' = \alpha x + \beta y, \\ y' = \gamma x + \rho y. \end{cases}$$

Thus, for any $(u_0, v_0) \in \mathbb{R}^2$, the unique solution of (6.14) satisfying

$$\big(u(0), v(0)\big) = (u_0, v_0)$$

is given through

$$\big(u(t), v(t)\big) = \mathrm{e}^{\lambda t} U(t)(u_0, v_0).$$

Therefore, the Poincaré map of (6.14) is given by

$$P_T := \mathrm{e}^{\lambda T} U(T)$$

and, consequently, the multipliers of (6.14) are $\mathrm{e}^{\lambda T}\nu_j$, $j \in \{1, 2\}$. Thus, the eigenvalue problem (6.14) possesses a T-periodic solution if, and only, if

$$\lambda \in \left\{ -\frac{1}{T}\operatorname{Log}\nu_1, -\frac{1}{T}\operatorname{Log}\nu_2 \right\}.$$

By the analysis already done, the unique value of λ providing us with a T-periodic solution having both components positive is the one given by (6.13), as the other eigenvalue provides us with a T-periodic solution whose components have contrary sign. This concludes the proof. □

Subsequently, the value λ_0 defined by (6.13) will be called the *principal eigenvalue* of (6.14). It should be noted that (6.1) is asymptotically stable if, and only if $\lambda_0 > 0$, whereas it is unstable if, and only if, $\lambda_0 < 0$. The following result provides us with an extremely useful criterion for ascertaining the stability of (6.1).

THEOREM 6.4. *The following properties are satisfied*:

(a) *System* (6.1) *is stable* ($\nu_2 < 1$) *if, and only if, there exists* $(\bar{u}, \bar{v}) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ *with* $\bar{u} \gg 0$ *and* $\bar{v} \gg 0$ *such that*

$$\begin{cases} \bar{u}' \geqslant \alpha\bar{u} + \beta\bar{v}, \\ \bar{v}' \geqslant \gamma\bar{u} + \rho\bar{v}, \end{cases} \tag{6.15}$$

with some of these inequalities strict.

(b) *System* (6.1) *is unstable* ($\nu_2 > 1$) *if, and only if, there exits* $(\underline{u}, \underline{v}) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ *with* $\underline{u} \gg 0$ *and* $\underline{v} \gg 0$ *such that*

$$\begin{cases} \underline{u}' \leqslant \alpha\underline{u} + \beta\underline{v}, \\ \underline{v}' \leqslant \gamma\underline{u} + \rho\underline{v}, \end{cases} \tag{6.16}$$

with some of these inequalities strict.

PROOF. First we will prove part (a). If $\nu_2 < 1$, then the principal eigenvalue of (6.14) is positive. Indeed,

$$\lambda_0 = -\frac{1}{T}\operatorname{Log}\nu_2 > 0$$

and, hence, any T-periodic strongly positive solution associated to λ_0 provides us with a strict supersolution of (6.1). This shows the necessity of the existence of the supersolution. To show the sufficiency of this condition, suppose there exists $(\bar{u}, \bar{v}) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ with $\bar{u} \gg 0$ and $\bar{v} \gg 0$ satisfying (6.15) with some of the inequalities strict. Now, pick

$$(u_0, v_0) \in N\big[U(T) - \nu_2 I\big]$$

with $u_0 > 0$ and $v_0 > 0$, and let $(u(t), v(t))$ denote the unique solution of (6.3). Then,

$$\big(u(nT), v(nT)\big) = \nu_2^n (u_0, v_0), \quad n \geqslant 0.$$

Now, pick $R > 0$ satisfying

$$0 < u_0 < R\bar{u}(0), \qquad 0 < v_0 < R\bar{v}(0),$$

and consider the auxiliary function

$$x := R\bar{u} - u, \qquad y := R\bar{v} - v.$$

Then, by construction,

$$\begin{cases} x' \geqslant \alpha x + \beta y, \\ y' \geqslant \gamma x + \rho y, \\ x(0) > 0, \quad y(0) > 0, \end{cases} \tag{6.17}$$

and, hence, thanks to Proposition 6.2,

$$x(t) > 0 \quad \text{and} \quad y(t) > 0 \quad \text{for each } t \in \mathbb{R}_+.$$

Equivalently,

$$0 < u(t) < R\bar{u}(t) \quad \text{and} \quad 0 < v(t) < R\bar{v}(t) \quad \text{for each } t \in \mathbb{R}_+,$$

since $u_0 > 0$ and $v_0 > 0$. In particular, the sequence

$$\big(u(nT), v(nT)\big) = \nu_2^n (u_0, v_0), \quad n \geqslant 0,$$

is bounded above and, therefore, $\nu_2 \leqslant 1$. Suppose $\nu_2 = 1$. Then, $\lambda_0 = 0$ and, thanks to Theorem 6.3, (6.1) possesses a T-solution (p, q) with $p \gg 0$ and $q \gg 0$. On the other hand, setting

$$d := \bar{u}' - \alpha\bar{u} - \beta\bar{v}, \qquad \delta := \bar{v}' - \gamma\bar{u} - \rho\bar{v},$$

we have that $d \geqslant 0$, $\delta \geqslant 0$, $d + \delta > 0$, and

$$\begin{cases} \bar{u}' = \alpha\bar{u} + \beta\bar{v} + d, \\ \bar{v}' = \gamma\bar{u} + \rho\bar{v} + \delta. \end{cases}$$

Consequently, for each $\xi \in \mathbb{R}$, the pair

$$(u_\xi, v_\xi) := (\bar{u} - \xi p, \bar{v} - \xi q)$$

provides us with a T-periodic solution of

$$\begin{cases} u' = \alpha u + \beta v + d, \\ v' = \gamma u + \rho v + \delta. \end{cases} \tag{6.18}$$

Since $\bar{u} \gg 0$ and $\bar{v} \gg 0$, we have that $u_\xi \gg 0$ and $v_\xi \gg 0$ for sufficiently small $\xi > 0$. Moreover, since $p \gg 0$ and $q \gg 0$, we have that $u_\xi \ll 0$ and $v_\xi \ll 0$ for sufficiently large $\xi > 0$. Thus, there exists $\xi > 0$ such that $u_\xi \geqslant 0$, $v_\xi \geqslant 0$ and, for some $t_0 \in \mathbb{R}$, either $u_\xi(t_0) = 0$, or $v_\xi(t_0) = 0$. Without lost of generality we can assume that $u_\xi(t_0) = 0$. Then, for each $t \geqslant t_0$, we have that

$$u_\xi(t) = \int_{t_0}^{t} \mathrm{e}^{\int_\tau^t \alpha} \big[\beta v_\xi(\tau) + d(\tau)\big] \, \mathrm{d}\tau. \tag{6.19}$$

If $\beta v_\xi + d > 0$, then, we find from (6.19) that $u_\xi(t_0 + nT) > 0$ for each $n \geqslant 2$, which contradicts the periodicity of u_ξ, since $u_\xi(t_0) = 0$. Thus,

$$\beta v_\xi = d = 0, \tag{6.20}$$

and, thanks again to (6.19),

$$u_\xi = 0.$$

Thus, substituting into the v-equation of (6.18) shows that

$$v'_\xi = \rho v_\xi + \delta.$$

Necessarily,

$$v_\xi = \mathfrak{R}_{-\rho}(\delta) \gg 0,$$

since $\delta > 0$, and, therefore, $\beta v_\xi > 0$, which contradicts (6.20) and shows $\nu_2 < 1$. This concludes the proof of part (a).

Now, we will prove part (b). Suppose $\nu_2 > 1$. Then,

$$\lambda_0 = -\frac{1}{T} \operatorname{Log} \nu_2 < 0$$

and, hence, any T-periodic strongly positive solution of (6.14) associated to λ_0 provides us with a strict subsolution of (6.1). To show the converse, let $(\underline{u}, \underline{v}) \in \mathcal{C}_T^1 \times \mathcal{C}_T^1$ with $\underline{u} \gg 0$ and $\underline{v} \gg 0$ satisfying (6.16) with some of the inequalities strict. Pick

$$(u_0, v_0) \in N\big[U(T) - \nu_2 I\big]$$

with $u_0 > 0$ and $v_0 > 0$, and let $(u(t), v(t))$ denote the unique solution of (6.3). Then,

$$\big(u(nT), v(nT)\big) = \nu_2^n(u_0, v_0), \quad n \geqslant 0. \tag{6.21}$$

Now, for each $\varepsilon > 0$ we consider the auxiliary function

$$(U_\varepsilon, V_\varepsilon) := (u - \varepsilon \underline{u}, v - \varepsilon \underline{v}).$$

For any $\varepsilon > 0$, $(U_\varepsilon, V_\varepsilon)$ is a strict supersolution of (6.1). Moreover, for sufficiently small $\varepsilon > 0$ we have that

$$U_\varepsilon(0) = u_0 - \varepsilon \underline{u}(0) > 0 \quad \text{and} \quad V_\varepsilon(0) = v_0 - \varepsilon \underline{v}(0) > 0.$$

Choose one of these ε's. Then, thanks to Proposition 6.2, $U_\varepsilon(t) > 0$ and $V_\varepsilon(t) > 0$ for each $t \geqslant 0$. In particular, for each $n \geqslant 0$,

$$\varepsilon \underline{u}(0) = \varepsilon \underline{u}(nT) < u(nT) = \nu_2^n u_0$$

and, therefore, $\nu_2 \geqslant 1$.

Suppose $\nu_2 = 1$. Then, $\lambda_0 = 0$ and, thanks to Theorem 6.3, (6.1) possesses a T-solution (p, q) with $p \gg 0$ and $q \gg 0$. On the other hand, setting

$$d := \underline{u}' - \alpha \underline{u} - \beta \underline{v}, \qquad \delta := \underline{v}' - \gamma \underline{u} - \rho \underline{v},$$

we have that $d \leqslant 0$, $\delta \leqslant 0$, $d + \delta < 0$, and

$$\begin{cases} \underline{u}' = \alpha \underline{u} + \beta \underline{v} + d, \\ \underline{v}' = \gamma \underline{u} + \rho \underline{v} + \delta. \end{cases}$$

Consequently, for each $\xi \in \mathbb{R}$, the pair

$$(u_\xi, v_\xi) := (p - \xi \underline{u}, q - \xi \underline{v})$$

provides us with a T-periodic solution of

$$\begin{cases} u' = \alpha u + \beta v - d, \\ v' = \gamma u + \rho v - \delta. \end{cases} \tag{6.22}$$

Since $p \gg 0$ and $q \gg 0$, we have that $u_\xi \gg 0$ and $v_\xi \gg 0$ for sufficiently small $\xi > 0$. Moreover, since $\underline{u} \gg 0$ and $\underline{v} \gg 0$, we have that $u_\xi \ll 0$ and $v_\xi \ll 0$ for sufficiently large $\xi > 0$. Thus, there exists $\xi > 0$ such that $u_\xi \geqslant 0$, $v_\xi \geqslant 0$ and, for some $t_0 \in \mathbb{R}$, either $u_\xi(t_0) = 0$, or $v_\xi(t_0) = 0$. The result obtained in ending the proof of part (a) shows that this is impossible. It should be noted that $-d \geqslant 0$, $-\delta \geqslant 0$ and $-(\delta + d) > 0$. Therefore, $\nu_2 > 1$, which concludes the proof of the theorem. □

6.2. *Structure of the set of coexistence states*

In this section, the coexistence states of (1.2) will be regarded as zeroes of the operator

$$\mathfrak{F} : \mathbb{R}^2 \times \mathcal{C}_T \times \mathcal{C}_T \to \mathcal{C}_T \times \mathcal{C}_T$$

defined by

$$\mathfrak{F}(\lambda, \mu, u, v) := \begin{pmatrix} u - \mathfrak{R}_M\big[(\lambda\ell + M)u - au^2 - buv\big] \\ v - \mathfrak{R}_M\big[(\mu m + M)v - dv^2 - cuv\big] \end{pmatrix} \tag{6.23}$$

where $M > 0$ is fixed. The following result will provide us with the sharp structure of the whole surface, in terms of the parameters (λ, μ), of the coexistence states of (1.2) linking the surfaces of semi-trivial states of (1.2) along their curves of neutral stability.

PROPOSITION 6.5. *Suppose* $(u_0, v_0) \in \mathcal{C}_T \times \mathcal{C}_T$ *satisfies* $u_0 \gg 0$, $v_0 \gg 0$, *and*

$$\mathfrak{F}(\lambda, \mu, u_0, v_0) = 0.$$

Then, each of the operators

$$D_{(\lambda,u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0) : \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T \to \mathcal{C}_T \times \mathcal{C}_T$$

and

$$D_{(\mu,u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0) : \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T \to \mathcal{C}_T \times \mathcal{C}_T$$

is surjective. In other words, $(0, 0)$ *is a regular value of* $\mathfrak{F}$ *with respect to* $\mathbb{R}^2 \times \operatorname{Int} P \times \operatorname{Int} P$.

PROOF. By differentiating, we have that

$$D_{(u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0) = I_{\mathcal{C}_T \times \mathcal{C}_T} - \mathfrak{R}_M \begin{pmatrix} \alpha_M & -bu_0 \\ -cv_0 & \beta_M \end{pmatrix}, \tag{6.24}$$

$$D_\lambda\mathfrak{F}(\lambda, \mu, u_0, v_0) = -\mathfrak{R}_M \begin{pmatrix} \ell u_0 \\ 0 \end{pmatrix}, \tag{6.25}$$

and

$$D_\mu\mathfrak{F}(\lambda, \mu, u_0, v_0) = -\mathfrak{R}_M \begin{pmatrix} 0 \\ mv_0 \end{pmatrix}, \tag{6.26}$$

where, for any $\xi \geqslant 0$, we are denoting

$$\alpha_\xi := \lambda\ell + \xi - 2au_0 - bv_0, \quad \beta_\xi := \mu m + \xi - 2dv_0 - cu_0. \tag{6.27}$$

Subsequently, it will be said that (λ,μ,u_0,v_0) is a non-degenerate coexistence state of (1.2) if $D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)$ is an isomorphism. In the contrary case, it will be said that (λ,μ,u_0,v_0) is degenerate.

Obviously, $D_{(\lambda,u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)$ and $D_{(\mu,u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)$ are surjective if (λ,μ,u_0,v_0) is non-degenerate. So, suppose (λ,μ,u_0,v_0) is degenerate. Then, since $D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)$ is a compact perturbation of the identity, it is Fredholm of index zero and, hence,

$$N := \dim N\big[D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)\big] = \operatorname{codim} R\big[D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)\big] \geqslant 1.$$

Now, note that $N[D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)]$ is the set of T-periodic solutions (u,v) of the linear system

$$\begin{cases} u' = \alpha_0 u - b u_0 v, \\ v' = -c v_0 u + \beta_0 v. \end{cases} \tag{6.28}$$

Since $b<0$ and $c<0$, we have that

$$-bu_0 > 0 \quad \text{and} \quad -cv_0 > 0.$$

Moreover, since (λ,μ,u_0,v_0) is a coexistence state of (1.2), relations (6.7) hold with $(u_2,v_2)=(u_0,v_0)$ and, hence,

$$\int_0^T \alpha_0 = -\int_0^T a(t)u_0(t)\,\mathrm{d}t < 0 \quad \text{and} \quad \int_0^T \beta_0 = -\int_0^T d(t)v_0(t)\,\mathrm{d}t < 0.$$

Therefore, (6.28) fits into the abstract setting of Section 6.1. Consequently, thanks to Theorems 6.1 and 6.3, the biggest multiplier of (6.28), denoted by ν_2, equals one and any T-periodic solution of (6.28) must be a multiple of a fixed T-periodic solution (φ,ψ) such that $\varphi \gg 0$ and $\psi \gg 0$. In particular, $N=1$. So, thanks to (6.25) and (6.26), to conclude the proof it suffices to show the following:

$$\begin{aligned} &-\mathfrak{R}_M(\ell u_0, 0) \notin R\big[D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)\big], \\ &-\mathfrak{R}_M(0, m v_0) \notin R\big[D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)\big]. \end{aligned} \tag{6.29}$$

On the contrary, suppose that, e.g.,

$$-\mathfrak{R}_M(\ell u_0, 0) \in R\big[D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)\big].$$

Then, the system

$$\begin{cases} u' = \alpha_0 u - b u_0 v + \ell u_0, \\ v' = -c v_0 u + \beta_0 v, \end{cases} \tag{6.30}$$

possesses a T-periodic solution, say (u, v). Actually, for each $R > 0$, the pair

$$(u_R, v_R) := (u + R\varphi, v + R\psi)$$

provides us with a T-periodic solution of (6.30). Since $\ell u_0 > 0$, for sufficiently large $R > 0$, (u_R, v_R) provides us with a T-periodic strongly positive strict supersolution of (6.28) and, therefore, thanks to Theorem 6.4(a), $\nu_2 < 1$, which is a contradiction, since $\nu_2 = 1$. Similarly, the second relation of (6.29) holds. This concludes the proof. □

The following result, which is an easy consequence from Proposition 6.5, ascertains the global structure of the components of coexistence states $\mathfrak{C}_\mu^\lambda$ and $\mathfrak{C}_\lambda^\mu$ of (1.2) constructed in Theorems 5.4 and 5.5.

THEOREM 6.6. *The following properties are satisfied*:

(a) *For each* $\lambda > 0$, *the component of coexistence states* $\mathfrak{C}_\mu^\lambda$ *is unbounded in* $\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ *and it consists of a real analytic curve. More precisely, there exits a real analytic map* $(\mu, u, v) : [0, \infty) \to \mathbb{R} \times \operatorname{Int} P \times \operatorname{Int} P$ *such that*

$$\big(\mu(0), u(0), v(0)\big) = (\mu_0, \theta_{[\lambda\ell,a]}, 0)$$

and

$$\mathfrak{C}_\mu^\lambda = \big\{\big(\mu(s), u(s), v(s)\big) \colon\ s \in (0, \infty)\big\}.$$

(b) *For each* $\mu > 0$, *the component of coexistence states* $\mathfrak{C}_\lambda^\mu$ *is unbounded in* $\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ *and it consists of a real analytic curve. More precisely, there exists a real analytic map* $(\lambda, u, v) : [0, \infty) \to \mathbb{R} \times \operatorname{Int} P \times \operatorname{Int} P$ *such that*

$$\big(\lambda(0), u(0), v(0)\big) = (\lambda_0, 0, \theta_{[\mu m,d]})$$

and

$$\mathfrak{C}_\lambda^\mu = \big\{\big(\lambda(s), u(s), v(s)\big) \colon\ s \in (0, \infty)\big\}.$$

(c) *The set of coexistence states of* (1.2) $(\lambda, \mu, u, v) \in \mathbb{R}^2 \times \operatorname{Int} P \times \operatorname{Int} P$ *such that* $\lambda > 0$ *and, for some* $\mu \in \mathbb{R}$, $(\mu, u, v) \in \mathfrak{C}_\mu^\lambda$, *or* $\mu > 0$ *and, for some* $\lambda \in \mathbb{R}$, $(\lambda, u, v) \in \mathfrak{C}_\lambda^\mu$ *is a real analytic surface linking the surfaces of semi-trivial states*

$$\big\{(\lambda, \mu, \theta_{[\lambda\ell,a]}, 0) \colon\ \lambda > 0,\ \mu \in \mathbb{R}\big\}$$

and

$$\big\{(\lambda, \mu, 0, \theta_{[\mu m,d]}) \colon\ \lambda \in \mathbb{R},\ \mu > 0\big\}$$

along their respective curves of neutral stability

$$\mu = \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t \quad \textit{and} \quad \lambda = \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t.$$

PROOF. Suppose (λ, μ, u_0, v_0) is a coexistence state of (1.2) with, e.g., $\lambda > 0$. Subsequently, we fix λ and regard to μ as the main path-following parameter. Suppose (λ, μ, u_0, v_0) is non-degenerate. Then, the implicit function theorem shows that (μ, u_0, v_0) lies in an analytic curve through it that can be locally parameterized by μ, or ... by some length, or pseudo-length of arc of curve, as it is usual in numerical analysis of bifurcation problems. Now, suppose that (λ, μ, u_0, v_0) is degenerate and let $\omega_0 \in \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ such that

$$N\big[D_{(\mu,u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0)\big] = \operatorname{span}[\omega_0].$$

Let Z be any closed supplement of ω_0 in $\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$, i.e.,

$$\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T = \operatorname{span}[\omega_0] \oplus Z.$$

Then, any element $(\mu, u, v) \in \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ admits a unique decomposition as

$$(\mu, u, v) = s\omega_0 + z, \qquad (s, z) \in \mathbb{R} \times Z.$$

So, the problem of solving $\mathfrak{F}(\lambda, \mu, u, v) = 0$, with λ fixed, around (μ, u_0, v_0) is equivalent to the problem of solving

$$0 = \mathfrak{G}(s, z) := \mathfrak{F}\big(\lambda, (\mu, u_0, v_0) + s\omega_0 + z\big)$$

around $(s, z) = (0, 0)$. The map $\mathfrak{G} : \mathbb{R} \times Z \to \mathcal{C}_T \times \mathcal{C}_T$ is real analytic and it satisfies $\mathfrak{G}(0, 0) = 0$. Moreover, thanks to Proposition 6.5, the linearized operator

$$D_z\mathfrak{G}(0, 0) = D_{(\mu,u,v)}\mathfrak{F}(\mu, u_0, v_0)|_Z$$

is a linear isomorphism. Therefore, thanks to the implicit function, the set of coexistence states of (1.2) around (λ, μ, u_0, v_0) consists of an analytic curve. Actually, the curve can be parameterized by the projection of the coexistence states on ω_0.

Now, fix $\lambda > 0$. Since $b < 0$, for each $\mu > 0$ we have that

$$\lambda > 0 > \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t$$

and, therefore, thanks to Theorem 5.4, $\mathfrak{C}_\mu^\lambda$ is unbounded. Moreover, thanks to Theorem 5.6, there exists $L > 0$ and a real analytic map

$$(\mu, u, v) : [0, L] \to \mathbb{R} \times \operatorname{Int} P \times \operatorname{Int} P$$

such that

$$\big(\mu(0), u(0), v(0)\big) = (\mu_0, \theta_{[\lambda\ell,a]}, 0)$$

and, in a neighborhood of $(\mu_0, \theta_{[\lambda\ell,a]}, 0)$, $\mathfrak{C}^\lambda_\mu$ consists of

$$\big\{\big(\mu(s), u(s), v(s)\big)\colon\ 0 < s \leqslant L\big\}.$$

By a global continuation argument based on the implicit function theorem, the *local curve* $(\mu(s), u(s), v(s))$, $s \in (0, L]$, admits a *maximal continuation* (non-prolonging) to a real analytic curve, $(\mu(s), u(s), v(s))$, $s \in (0, \infty)$, of coexistence states of (1.2). Actually, the parameter s can be taken as the length of arc of curve. Clearly,

$$\varGamma := \big\{\big(\mu(s), u(s), v(s)\big)\colon\ s > 0\big\} \subset \mathfrak{C}^\lambda_\mu,$$

since $\mathfrak{C}^\lambda_\mu$ is connected. Moreover, the uniqueness obtained as an application of the implicit function theorem shows that there exists an open subset $\mathcal{O} \subset \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ such that

$$\varGamma \subset \mathcal{O} \quad \text{and} \quad \mathfrak{F}^{-1}(0) \cap (\mathbb{R} \times \operatorname{Int} P \times \operatorname{Int} P) \cap \mathcal{O} = \varGamma.$$

Therefore, $\mathfrak{C}^\lambda_\mu = \varGamma$, since $\mathfrak{C}^\lambda_\mu$ is connected. Part (b) follows from the symmetry of the problem. part (c) is a two-parameter re-interpretation of parts (a) and (b) that can be easily obtained from the fact that $\mathfrak{F}$ is real analytic in all its arguments. This concludes the proof. □

For the autonomous counterpart of (1.2) with $bc = 1$, is very simple to check that, for each $\lambda > 0$, the curve

$$\big(\mu(s), u(s), v(s)\big) = (c\lambda, \lambda - bs, s), \qquad s \geqslant 0,$$

provides us with the component $\mathfrak{C}^\lambda_\mu$ bifurcating from the semi-trivial solution $(\lambda, 0)$ at $\mu = c\lambda$. As this simple example shows, in general, condition $\mu' = 0$ cannot be avoided unless some additional assumptions are made, as, e.g., having a priori bounds for the coexistence states of the model at $\mu = \mu_0$. Note that in case $\mu' = 0$, one necessarily has $\mu(s) = \mu_0$ for each $s \geqslant 0$ and

$$\limsup_{s\uparrow\infty} \big\|\big(u(s), v(s)\big)\big\|_{\mathcal{C}_T\times\mathcal{C}_T} = \infty,$$

since $\mathfrak{C}^\lambda_\mu$ is unbounded. Actually, as an immediate consequence from the analysis carried out in the next section, in case $\mu' = 0$ one has that, for each $s \geqslant 0$, $\dot{u}(s) \gg 0$ and $\dot{v}(s) \gg 0$, where $\cdot := \frac{\mathrm{d}}{\mathrm{d}s}$. Therefore,

$$\lim_{s\uparrow\infty} \big\|\big(u(s), v(s)\big)\big\|_{\mathcal{C}_T\times\mathcal{C}_T} = \infty.$$

6.3. *Stability of the coexistence states along $\mathfrak{C}^{\lambda}_{\mu}$ and $\mathfrak{C}^{\mu}_{\lambda}$*

Fix $\lambda > 0$ and let $(\mu(s), u(s), v(s))$, $s \geqslant 0$, be the real analytic map constructed in Theorem 6.6, where s is assumed to be the length of arc of curve. Subsequently, we shall denote $\cdot := \frac{\mathrm{d}}{\mathrm{d}s}$. Since

$$\mathfrak{F}\big(\lambda, \mu(s), u(s), v(s)\big) = 0, \quad s \geqslant 0,$$

differentiating with respect to s gives

$$D_{(\mu,u,v)}\mathfrak{F}\big(\lambda, \mu(s), u(s), v(s)\big)\big(\dot{\mu}(s), \dot{u}(s), \dot{v}(s)\big) = 0, \quad s > 0, \tag{6.31}$$

and, hence,

$$N\big[D_{(\mu,u,v)}\mathfrak{F}\big(\lambda, \mu(s), u(s), v(s)\big)\big] = \operatorname{span}\big[\big(\dot{\mu}(s), \dot{u}(s), \dot{v}(s)\big)\big], \quad s > 0.$$

Moreover, since s has been taken as the length of arc of curve,

$$\big\|\big(\dot{\mu}(s), \dot{u}(s), \dot{v}(s)\big)\big\|_{\mathbb{R}\times\mathcal{C}_T\times\mathcal{C}_T} = 1, \quad s > 0. \tag{6.32}$$

On the other hand, since

$$\big(\mu(0), u(0), v(0)\big) = (\mu_0, \theta_{[\lambda\ell,a]}, 0),$$

it follows from Theorem 5.6 that

$$\dot{v}(s) = \varphi_v + \mathrm{O}(s) = \mathrm{e}^{\int_0^{\cdot}(\mu_0 m - c\theta_{[\lambda\ell,a]})} + \mathrm{O}(s) \gg 0, \quad s \downarrow 0. \tag{6.33}$$

Now, set

$$(p_s, q_s) := \big(\dot{u}(s), \dot{v}(s)\big), \quad s > 0,$$

and note that (6.31) can be equivalently rewritten in the form

$$\begin{cases} (p_s)' = \big[\lambda\ell - 2au(s) - bv(s)\big]p_s - bu(s)q_s, \\ (q_s)' = -cv(s)p_s + \big[\mu(s)m - 2dv(s) - cu(s)\big]q_s + \dot{\mu}(s)mv(s). \end{cases} \tag{6.34}$$

Note that, since $-bu(s) > 0$, $-cv(s) > 0$,

$$\int_0^T \big[\lambda\ell - 2au(s) - bv(s)\big] = -2\int_0^T \big[au(s)\big] < 0, \tag{6.35}$$

and

$$\int_0^T \big[\mu(s)m - 2dv(s) - cu(s)\big] = -2\int_0^T \big[dv(s)\big] < 0, \tag{6.36}$$

the linear system (6.34) fits within the general setting of Section 6.1. In particular, thanks to Proposition 2.1, we find from the first equation of (6.34) that, for each $s > 0$ sufficiently small,

$$p_s = \mathfrak{R}_{-[\lambda\ell - 2au(s) - bv(s)]}\bigl(-bu(s)q_s\bigr) \gg 0.$$

Therefore, there exists $\varepsilon > 0$ such that

$$p_s \gg 0 \quad \text{and} \quad q_s \gg 0 \quad \text{for each } s \in [0, \varepsilon]. \tag{6.37}$$

Suppose $\mu = \mu_0$ for each $s \geqslant 0$. Then, (6.34) becomes into

$$\begin{cases} (p_s)' = \bigl[\lambda\ell - 2au(s) - bv(s)\bigr]p_s - bu(s)q_s, \\ (q_s)' = -cv(s)p_s + \bigl[\mu(s)m - 2dv(s) - cu(s)\bigr]q_s. \end{cases} \tag{6.38}$$

Thanks to (6.32), $(p_s, q_s) \neq (0, 0)$ and, hence, thanks to Theorem 6.1, for each $s \geqslant 0$, either $p_s \gg 0$ and $q_s \gg 0$, or $p_s \ll 0$ and $q_s \ll 0$. Thanks to (6.37), necessarily

$$p_s \gg 0 \quad \text{and} \quad q_s \gg 0 \quad \text{for each } s \in [0, \infty). \tag{6.39}$$

Indeed, since $s \to (p_s, q_s)$ is real analytic, if (6.39) is not satisfied, then, due to (6.37), there is some value of s, say $\tilde{s} > 0$, where $p_{\tilde{s}} > 0$, $q_{\tilde{s}} > 0$, and, e.g., $p_{\tilde{s}} \gg 0$ fails. Thanks to Theorem 6.1, this would imply $p_{\tilde{s}} \gg 0$ and $q_{\tilde{s}} \gg 0$, which is a contradiction. Thus, (6.39) is satisfied, as it was claimed in the last paragraph of Section 6.2, and, consequently, thanks to Theorem 6.3, the principal characteristic multiplier of $(\mu(s), u(s), v(s))$, which will be subsequently denoted by $\nu_2(s)$, satisfies $\nu_2(s) = 1$ for each $s \geqslant 0$. Therefore, if $\dot{\mu} = 0$, then $\dot{\nu}_2 = 0$ and, hence, all the coexistence states of $\mathfrak{C}^{\lambda}_{\mu}$ are neutrally stable. When $\dot{\mu} \neq 0$, then the following result is satisfied.

THEOREM 6.7. *Suppose $\dot{\mu} \neq 0$ in $[0, \infty)$. Then, the set*

$$\mathcal{S} := \bigl\{s \in (0, \infty)\colon\ \dot{\mu}(s) = 0\bigr\} \tag{6.40}$$

is discrete, each $s \in \mathcal{S}$ is a zero of finite order of $\dot{\mu}$, and

$$\mathcal{S} = \bigl\{s \in (0, \infty)\colon\ \nu_2(s) = 1\bigr\}. \tag{6.41}$$

Moreover,

$$\dot{\mu}(s)\bigl(1 - \nu_2(s)\bigr) > 0, \quad s \in (0, \infty) \setminus \mathcal{S}, \tag{6.42}$$

and

$$p_s \gg 0 \quad \textit{and} \quad q_s \gg 0 \quad \textit{if} \quad \dot{\mu}(s) \geqslant 0. \tag{6.43}$$

PROOF. The fact that $\mathcal{S}$ is discrete and that each $s \in \mathcal{S}$ has finite order, as a zero of $\dot{\mu}$, are consequences from the fact that $\dot{\mu}$ is real analytic. Now, we shall prove (6.41). Pick $s \in \mathcal{S}$. Then, (p_s, q_s) satisfies (6.38) and, due to (6.32), $(p_s, q_s) \neq (0, 0)$. Thus, thanks to Theorem 6.1, $p_s \gg 0$ and $q_s \gg 0$, or $p_s \ll 0$ and $q_s \ll 0$. In any of these cases, it follows from Theorem 6.3 that $\lambda = 0$ is the *principal characteristic exponent* of $(\mu(s), u(s), v(s))$ and, hence, $\nu_2(s) = 1$. To prove the converse inclusion, let $s \in (0, \infty)$ such that $\nu_2(s) = 1$. Then, thanks to Theorem 6.3, there exists a T-periodic pair (P, Q) with $P \gg 0$ and $Q \gg 0$ such that

$$\begin{cases} P' = \big[\lambda\ell - 2au(s) - bv(s)\big]P - bu(s)Q, \\ Q' = -cv(s)P + \big[\mu(s)m - 2dv(s) - cu(s)\big]P. \end{cases} \tag{6.44}$$

Now, choose a sufficiently small $\varepsilon > 0$ so that the auxiliary pair

$$(U, V) := (P - \varepsilon p_s, Q - \varepsilon q_s),$$

satisfy $U \gg 0$ and $V \gg 0$. Thanks to (6.34) and (6.44) we have that

$$\begin{cases} U' = \big[\lambda\ell - 2au(s) - bv(s)\big]U - bu(s)V, \\ V' = -cv(s)U + \big[\mu(s)m - 2dv(s) - cu(s)\big]V - \varepsilon\dot{\mu}(s)mv(s). \end{cases}$$

Assume $\dot{\mu}(s) > 0$. Then, since $\varepsilon mv(s) > 0$, (U, V) is a strict subsolution of

$$\begin{cases} x' = \big[\lambda\ell - 2au(s) - bv(s)\big]x - bu(s)y, \\ y' = -cv(s)x + \big[\mu(s)m - 2dv(s) - cu(s)\big]y, \end{cases} \tag{6.45}$$

and, hence, thanks to Theorem 6.4(b), $\nu_2(s) > 1$, which is a contradiction. Assume $\dot{\mu}(s) < 0$. Then, (U, V) is a strict supersolution of (6.45) and, due to Theorem 6.4(a), $\nu_2(s) < 1$, which is a contradiction again. Consequently, $\dot{\mu}(s) = 0$ and, therefore, (6.41) holds.

Now, pick $\tilde{s} \in \mathcal{S}$. Thanks to (6.32), it follows from Theorem 6.1 that some of the following alternatives occurs:

1. $p_{\tilde{s}} \gg 0$ and $q_{\tilde{s}} \gg 0$.
2. $p_{\tilde{s}} \ll 0$ and $q_{\tilde{s}} \ll 0$.

Assume alternative 1 occurs. As $\tilde{s}$ is an isolated zero of $\dot{\mu}(s)$, there exists $\varepsilon > 0$ such that $p_s = \dot{u}(s) \gg 0$ and $q_s = \dot{v}(s) \gg 0$ for each $s \in (\tilde{s} - \varepsilon, \tilde{s} + \varepsilon)$, and

$$\dot{\mu}^{-1}(0) \cap \big(\tilde{s} - \varepsilon, \tilde{s} + \varepsilon\big) \cap (0, \infty) = \big\{\tilde{s}\big\}.$$

Thus, for each

$$s \in J := (0, \infty) \cap \big[\big(\tilde{s} - \varepsilon, \tilde{s} + \varepsilon\big) \setminus \big\{\tilde{s}\big\}\big],$$

either $\dot{\mu}(s) > 0$, or $\dot{\mu}(s) < 0$. Suppose $\dot{\mu}(s) > 0$. Then, (p_s, q_s) provides us with a strongly positive strict supersolution of (6.45) and it follows from Theorem 6.4(a) that $\nu_2(s) < 1$.

Hence, $\dot{\mu}(s)(1-\nu_2(s))>0$. Now, suppose $\dot{\mu}(s)<0$. Then, (p_s,q_s) provides us with a strongly positive strict subsolution of (6.45) and it follows from Theorem 6.4(b) that $\nu_2(s)>1$. Hence, $\dot{\mu}(s)(1-\nu_2(s))>0$. Therefore, condition (6.42) is satisfied for each $\tilde{s}\in\mathcal{S}$ satisfying alternative 1.

Now, suppose $\tilde{s}\in\mathcal{S}$ satisfies alternative 2. As $\tilde{s}$ is an isolated zero of $\dot{\mu}(s)$, there exists $\varepsilon>0$ such that $p_s=\dot{u}(s)\ll 0$ and $q_s=\dot{v}(s)\ll 0$ for each $s\in(\tilde{s}-\varepsilon,\tilde{s}+\varepsilon)$, and

$$\dot{\mu}^{-1}(0)\cap\big(\tilde{s}-\varepsilon,\tilde{s}+\varepsilon\big)\cap(0,\infty)=\big\{\tilde{s}\big\}.$$

Thus, for each

$$s\in J:=(0,\infty)\cap\big[\big(\tilde{s}-\varepsilon,\tilde{s}+\varepsilon\big)\setminus\big\{\tilde{s}\big\}\big],$$

either $\dot{\mu}(s)>0$, or $\dot{\mu}(s)<0$. Suppose $\dot{\mu}(s)>0$. Then, $(-p_s,-q_s)$ provides us with a strongly positive strict subsolution of (6.45) and it follows from Theorem 6.4(b) that $\nu_2(s)>1$. Hence, $\dot{\mu}(s)(1-\nu_2(s))<0$. Similarly, if $\dot{\mu}(s)<0$, then, $(-p_s,-q_s)$ provides us with a strongly positive strict supersolution of (6.45) and Theorem 6.4(a) implies $\nu_2(s)<1$. Therefore, the following condition is satisfied

$$\dot{\mu}(s)\big(1-\nu_2(s)\big)<0,\quad s\in(0,\infty)\cap\big[\big(\tilde{s}-\varepsilon,\tilde{s}+\varepsilon\big)\setminus\big\{\tilde{s}\big\}\big],\tag{6.46}$$

for each $\tilde{s}\in\mathcal{S}$ satisfying alternative 2.

Now, suppose

$$\mathcal{S}=\{s_n\colon n\geqslant 1\},\quad 0<s_1<s_n<s_{n+1},\ n\geqslant 2;$$

the subsequent argument can be easily adapted to cover the case when $\mathcal{S}$ is finite, possibly empty. Thanks to (6.37), for sufficiently small $s>0$, $p_s\gg 0$ and $q_s\gg 0$. Moreover, either $\dot{\mu}(s)>0$ for each $s\in(0,s_1)$, or $\dot{\mu}(s)<0$ for each $s\in(0,s_1)$.

Suppose $\dot{\mu}(s)>0$ for each $s\in(0,s_1)$. Then, for sufficiently small $s>0$, (p_s,q_s) provides us with a strongly positive strict supersolution of (6.45) and, hence, thanks to Theorem 6.4(a), $\nu_2(s)<1$. Actually, this is the inter-exchange stability principle between the semi-trivial state $(\mu,\theta_{[\lambda\ell,a]},0)$ and the bifurcating coexistence state $(\mu(s),u(s),v(s))$ as μ crosses μ_0. Thanks to (6.41), necessarily $\nu_2(s)<1$ for every $s\in(0,s_1)$ and, therefore,

$$\dot{\mu}(s)\big(1-\nu_2(s)\big)>0,\quad s\in(0,s_1).\tag{6.47}$$

Necessarily $p_{s_1}\gg 0$ and $q_{s_1}\gg 0$. Indeed, if $p_{s_1}\ll 0$ and $q_{s_1}\ll 0$, then (6.46) holds, which contradicts (6.47). Therefore,

$$\dot{\mu}(s)\big(1-\nu_2(s)\big)>0,\quad s\in(0,s_2)\setminus\{s_2\}.\tag{6.48}$$

Reiterating this argument, (6.42) holds readily.

Now, suppose $\dot{\mu}(s)<0$ for each $s\in(0,s_1)$. Then, (p_s,q_s) provides us with a strongly positive strict subsolution of (6.45) and, hence, thanks to Theorem 6.4(b), $\nu_2(s)>1$. As in the previous case, this provides us with the inter-exchange stability principle between the semi-trivial state $(\mu, \theta_{[\lambda\ell,a]}, 0)$ and the bifurcating coexistence state $(\mu(s), u(s), v(s))$. Thanks to (6.41), $\nu_2(s) > 1$ for every $s \in (0, s_1)$ and, therefore, (6.47) is as well satisfied. Arguing as above, the proof of (6.42) is concluded.

Note that, as a result from the previous analysis, we have that

$$p_{\tilde{s}} \gg 0 \quad \text{and} \quad q_{\tilde{s}} \gg 0 \quad \text{for each } \tilde{s} \in \mathcal{S} \cup \{0\}. \tag{6.49}$$

Let $s_1^*, s_2^* \in \mathcal{S} \cup \{0\}$ such that

$$\dot{\mu}(s) > 0, \quad s \in \left(s_1^*, s_2^*\right).$$

Thanks to (6.49), $p_s \gg 0$ and $q_s \gg 0$ for each $s > s_1^*$ sufficiently close to s_1^*. Thus, either $p_s \gg 0$ and $q_s \gg 0$ for each $s \in (s_1^*, s_2^*)$, or there exist $s_0 \in (s_1^*, s_2^*)$ and $t \in \mathbb{R}$ such that $p_{s_0} \geqslant 0$ and $q_{s_0} \geqslant 0$ and either $p_{s_0}(t) = 0$, or $q_{s_0}(t) = 0$. Since

$$-cv(s_0)p_{s_0} + \dot{\mu}(s_0)mv(s_0) > 0,$$

thanks to Proposition 2.1, it follows from the second equation of (6.34) that $q_{s_0} \gg 0$. Thus,

$$-bu(s_0)q_{s_0} > 0$$

and, thanks again to Proposition 2.1, the first equation of (6.34) gives $p_{s_0} \gg 0$, which is impossible. Therefore, for each $s \in (s_1^*, s_2^*)$, $p_s \gg 0$ and $q_s \gg 0$, which concludes the proof of (6.43). The proof is completed. □

It should be noted that, since the dependence of the Poincaré map of (6.45) on s is analytic, the map $s \to \nu_2(s)$ is as well real analytic. Moreover,

$$\nu_2(0) = 1,$$

because this is the critical value of the parameter where the stability of the semi-trivial state changes. As far as to $\dot{\mu}(0)$ concerns, it might take any real value, according with the value of the integral of (5.53).

Subsequently, we shall assume

$$\dot{\mu} \neq 0$$

in order to use Theorem 6.7 for ascertaining the attractive character of all the coexistence states along the curve $\mathfrak{C}_\mu^\lambda$. In this case, setting

$$s_1 = \min \mathcal{S},$$

some of the following alternatives occurs. Either

$$\dot{\mu}(s) > 0, \quad s \in (0, s_1) \tag{6.50}$$

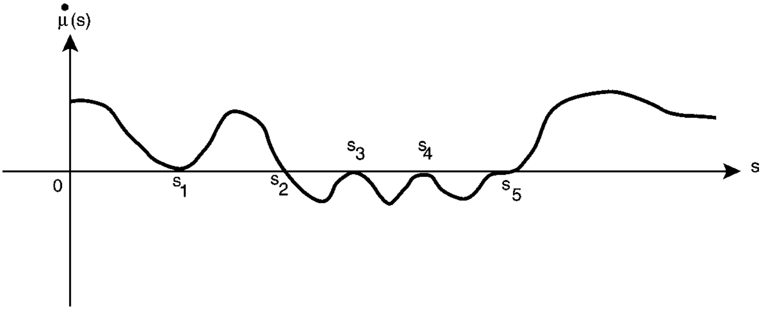


Fig. 11. An admissible graph of $\dot\mu(s)$ in case (6.50).

or

$$\dot\mu(s) < 0, \quad s \in (0, s_1). \tag{6.51}$$

If $\mathcal{S} = \emptyset$, then either $\dot\mu(s) > 0$ for each $s > 0$, or $\dot\mu(s) < 0$ for each $s > 0$. The bifurcation of $\mathfrak{C}^\lambda_\mu$ from $(\mu_0, \theta_{[\lambda\ell,a]}, 0)$ is said to be *super-critical* if (6.50) occurs, while it is said to be *sub-critical* if (6.51) is satisfied.

The following concepts will be extremely useful in the forthcoming discussion. Given $\tilde s \in \mathcal{S}$, we shall denote by $\operatorname{ord}(\tilde s)$ the order of $\tilde s$ as a zero of $\dot\mu(s)$.

- A value $\tilde s \in \mathcal{S}$ is said to be a *hysteresis point* of $\mathfrak{C}^\lambda_\mu$ if $\operatorname{ord}(\tilde s)$ is even.
- A value $\tilde s \in \mathcal{S}$ is said to be a *super-critical turning point* of $\mathfrak{C}^\lambda_\mu$ if $\operatorname{ord}(\tilde s)$ is odd and $\mu(s) > \mu(\tilde s)$ in a perforated neighborhood of $\tilde s$.
- A value $\tilde s \in \mathcal{S}$ is said to be a *sub-critical turning point* of $\mathfrak{C}^\lambda_\mu$ if $\operatorname{ord}(\tilde s)$ is odd and $\mu(s) < \mu(\tilde s)$ in a perforated neighborhood of $\tilde s$.

Suppose (6.50) and, e.g., the graph of $\dot\mu(s)$ is of the type sketched in Fig. 11. Although far from necessary we shall assume that $\mathcal{S} = \{s_1, s_2, s_3, s_4, s_5\}$. Hence, $\dot\mu(s) > 0$ for each $s > s_5$. For this special configuration, s_1, s_3 and s_4 are hysteresis points, while s_2 is a sub-critical turning point and s_5 is a super-critical turning point. Thanks to Theorem 6.7, the principal characteristic multiplier $\nu_2(s)$ of the coexistence state $(\mu(s), u(s), v(s))$ must adjust to the pattern shown in Fig. 12. Therefore, the coexistence state $(\mu(s), u(s), v(s))$ is asymptotically stable if $s \in (0, s_1) \cup (s_1, s_2) \cup (s_5, \infty)$, whereas it is unstable if $s \in (s_2, s_5) \setminus \{s_3, s_4\}$. In Fig. 13 we have represented the corresponding bifurcation diagram of coexistence states, where we are representing the value of $u(t)$ for a given time $t \in \mathbb{R}$ versus the parameter μ, not s. Except for $(\mu, 0, 0)$, continuous lines represent stable solutions, while dashed lines represent unstable solutions. The horizontal lines represent the curves of states $(\mu, 0, 0)$ and $(\mu, \theta_{[\lambda\ell,0]}, 0)$. The curve $\mathfrak{C}^\lambda_\mu$ bifurcates from $(\mu, \theta_{[\lambda\ell,a]}, 0)$ at

$$\mu = \mu_0 = \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t < 0,$$

since $c < 0$. The wiggled thick line represent $\mathfrak{C}^\lambda_\mu$, whose bifurcation from the semi-trivial state is super-critical. At the sub-critical turning point $(\mu(s_2), u(s_2), v(s_2))$ the stability of

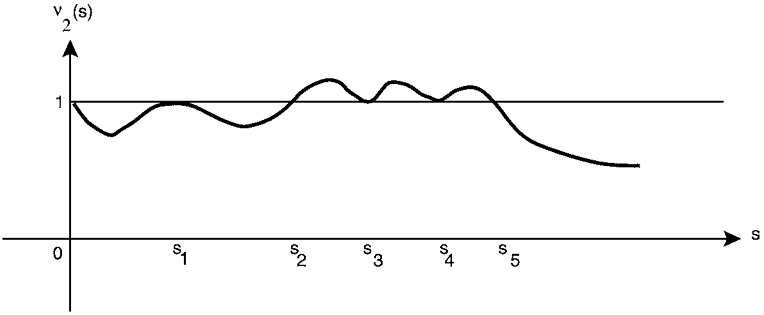


Fig. 12. The corresponding principal characteristic multiplier $\nu_2(s)$.

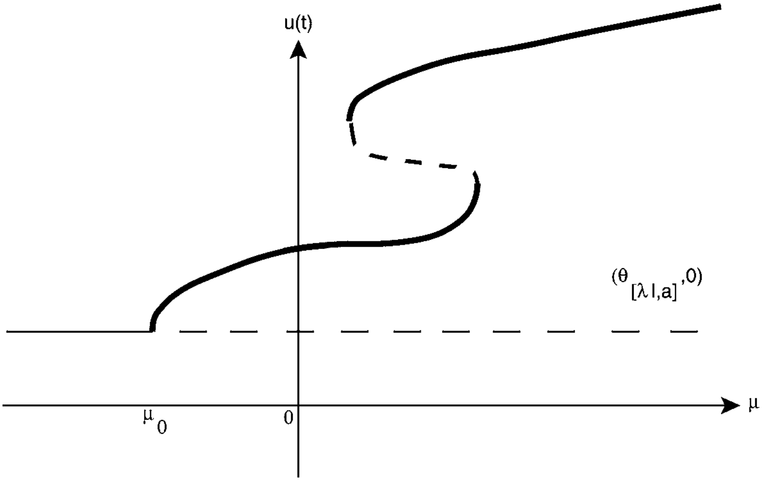


Fig. 13. The corresponding bifurcation diagram.

the solutions along the lower arc of curve of $\mathfrak{C}^\lambda_\mu$ is lost. Then, the solutions stay unstable until the second turning point, $(\mu(s_5), u(s_5), v(s_5))$, which is super-critical, is passed by. Once crossed the super-critical turning point, all remaining solutions for further values of s stay stable. It should be noted that, thanks to Theorem 6.7, $u(s)$ and $v(s)$ are strongly increasing along the pieces of the curve $\mathfrak{C}^\lambda_\mu$ where $\dot\mu(s) \geqslant 0$.

Now, suppose (6.51), instead of (6.50), and, e.g., the graph of $\dot\mu(s)$ adjust to the profile shown in Fig. 14. According to Theorem 6.7, the principal characteristic multiplier $\nu_2(s)$ of the coexistence state $(\mu(s), u(s), v(s))$ must adjust to the pattern shown in Fig. 15. In Fig. 16 we have represented an admissible bifurcation diagram corresponding to the graph of $\dot\mu(s)$ shown in Fig. 14. Now the bifurcation of $\mathfrak{C}^\lambda_\mu$ from $(\mu, \theta_{[\lambda\ell,a]}, 0)$ at μ_0 is sub-critical, the solutions along the lower curve being unstable until they reach the super-critical turning point $(\mu(s_2), u(s_2), v(s_2))$, where they become stable until reaching the second turning point, $(\mu(s_5), u(s_5), v(s_5))$, where they become unstable and stay unstable for any further value of s, $s > s_5$.

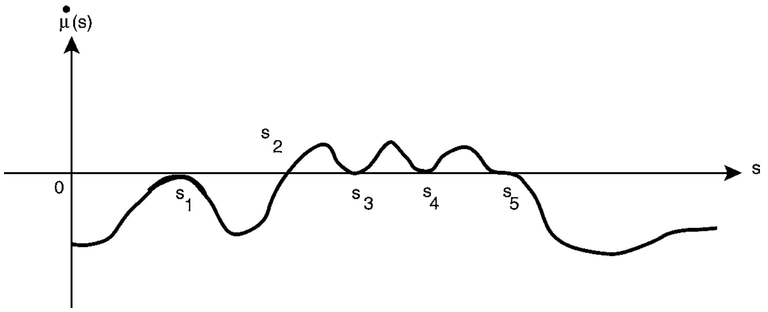


Fig. 14. An admissible graph for $\dot{\mu}(s)$ in case (6.51).

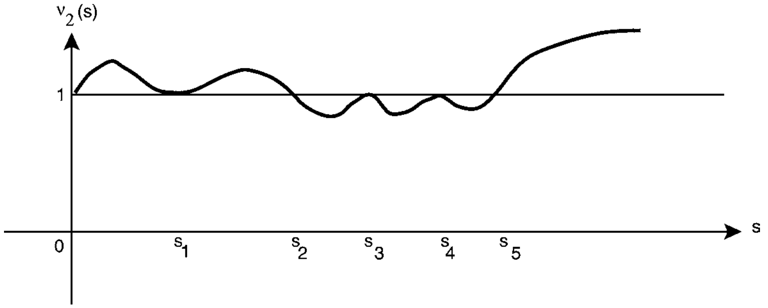


Fig. 15. The corresponding principal characteristic multiplier $\nu_2(s)$.

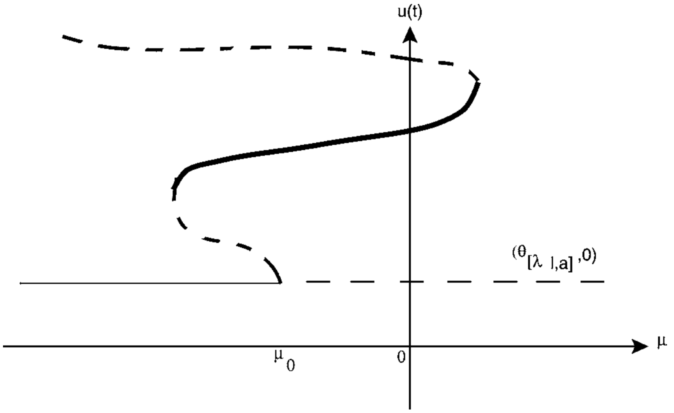


Fig. 16. The corresponding bifurcation diagram.

As in the previous example, as a result of Theorem 6.7, $s \mapsto (u(s), v(s))$ is strongly increasing along any arc of curve filled in by stable solutions, though some of the components might decrease along some pieces of the arcs of curve where the coexistence states are unstable.

Generically, when some additional parameter of the model is varied, each hysteresis point will either disappear, or it will generate an even number of turning points; half among them sub-critical and the other half super-critical. This might provide with a mechanism to generate models with an arbitrarily large number of coexistence states.

It should be noted that (1.2) might possess some additional coexistence state not lying in $\mathfrak{C}_\mu^\lambda$. If this is the case, i.e., if (1.2) possesses a coexistence state $(\mu, u_0, v_0) \notin \mathfrak{C}_\mu^\lambda$, then the global continuation argument used to ascertain the analytic structure of $\mathfrak{C}_\mu^\lambda$ also shows that there is an analytic curve

$$(\mu, u, v) : (-\infty, \infty) \to \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$$

of coexistence states of (1.2) such that

$$\big(\mu(0), u(0), v(0)\big) = (\mu, u_0, v_0)$$

and $\dot{\mu} \neq 0$. Indeed, if $\dot{\mu} = 0$, then $\mu(s) = \mu$ for each s, and $s \mapsto (u(s), v(s))$ must be strongly monotone, and, as a result, it should bifurcate from $(\mu_0, \theta_{[\lambda\ell, a]}, 0)$, which is not possible, because $(\mu, u_0, v_0) \notin \mathfrak{C}_\mu^\lambda$. Theorem 6.7 also applies along these curves, showing that $\nu_2(s) - 1$ must change of sign at any turning point along them. The corresponding curves should be either isolated bounded components, or unbounded components, but a sharper analysis of this and other related problems escapes from the general scope of this chapter.

By symmetry, all the results that we have obtained for $\mathfrak{C}_\mu^\lambda$ are also valid for the components $\mathfrak{C}_\lambda^\mu$ obtained by fixing $\mu > 0$ and using λ as the main bifurcation parameter.

6.4. *Low and strong symbiosis effects*

Throughout this section we will assume that

$$\ell \gg 0, \quad m \gg 0, \quad a \gg 0, \quad d \gg 0, \quad -b \gg 0, \quad -c \gg 0. \tag{6.52}$$

Then, it is said that (1.2) is lowly symbiotic if

$$\left(\frac{-b}{a}\right)_M \left(\frac{-c}{d}\right)_M < 1, \tag{6.53}$$

while it is said that (1.2) is highly symbiotic if

$$\left(\frac{-b}{a}\right)_L \left(\frac{-c}{d}\right)_L > 1. \tag{6.54}$$

Although in the intermediate situation cases where conditions (6.53) and (6.54) fail, the curves of coexistence states $\mathfrak{C}_\mu^\lambda$ and $\mathfrak{C}_\lambda^\mu$ might exhibit a bounded μ projection and a bounded λ projection, respectively, because of the lost of uniform priori bounds in $\mathcal{C}_T \times \mathcal{C}_T$, in the cases when some of these conditions is satisfied, the μ and λ-projections are unbounded. Actually, the following result is satisfied for the case of low symbiosis.

THEOREM 6.8. *Assume* (6.52), (6.53), *and* (1.2) *possesses a coexistence state, say* (λ, μ, u, v). *Then,*

$$\begin{aligned} &\lambda \geqslant -(-b/\ell)_M (m/d)_M \mu \quad \text{if } \mu > 0, \\ &\mu \geqslant -(-c/m)_M (\ell/a)_M \lambda \quad \text{if } \lambda > 0. \end{aligned} \tag{6.55}$$

Moreover,

$$\begin{aligned} \|u\|_{\mathcal{C}_T} &\leqslant \frac{(\ell/a)_M |\lambda| + (-b/a)_M (m/d)_M |\mu|}{1 - (-b/a)_M (-c/d)_M}, \\ \|v\|_{\mathcal{C}_T} &\leqslant \frac{(m/d)_M |\mu| + (-c/d)_M (\ell/a)_M |\lambda|}{1 - (-b/a)_M (-c/d)_M}, \end{aligned} \tag{6.56}$$

and, therefore, for each $\lambda > 0$ *and* $\mu > 0$ *there exist*

$$\mu_1 = \mu_1(\lambda) \in (-\infty, \mu_0] \quad \text{and} \quad \lambda_1 = \lambda_1(\mu) \in (-\infty, \lambda_0]$$

such that

$$P_\mu \mathfrak{C}^\lambda_\mu \in \{[\mu_1, \infty), (\mu_0, \infty)\} \quad \text{and} \quad P_\lambda \mathfrak{C}^\mu_\lambda \in \{[\lambda_1, \infty), (\lambda_0, \infty)\},$$

where P_γ *stands for the* γ*-projection operator,* $\gamma \in \{\lambda, \mu\}$.

PROOF. Pick $t_0, t_1 \in \mathbb{R}$ such that

$$u(t_0) = \|u\|_{\mathcal{C}_T} \quad \text{and} \quad v(t_1) = \|v\|_{\mathcal{C}_T}.$$

Then, $u'(t_0) = v'(t_1) = 0$ and, hence,

$$\lambda \ell(t_0) - a(t_0)u(t_0) - b(t_0)v(t_0) = 0, \qquad \mu m(t_1) - d(t_1)v(t_1) - c(t_1)u(t_1) = 0.$$

Thus,

$$u(t_0) = \lambda \frac{\ell(t_0)}{a(t_0)} - \frac{b(t_0)}{a(t_0)} v(t_0), \qquad v(t_1) = \mu \frac{m(t_1)}{d(t_1)} - \frac{c(t_1)}{d(t_1)} u(t_1). \tag{6.57}$$

Now, since

$$v(t_0) \leqslant v(t_1) = \mu \frac{m(t_1)}{d(t_1)} - \frac{c(t_1)}{d(t_1)} u(t_1) \leqslant \mu \frac{m(t_1)}{d(t_1)} - \frac{c(t_1)}{d(t_1)} u(t_0),$$

the first identity of (6.57) implies

$$u(t_0) \leqslant \frac{\ell(t_0)}{a(t_0)}\lambda - \frac{b(t_0)}{a(t_0)}\frac{m(t_1)}{d(t_1)}\mu + \frac{b(t_0)}{a(t_0)}\frac{c(t_1)}{d(t_1)}u(t_0)$$

and, hence,

$$\left[1 - \frac{b(t_0)}{a(t_0)}\frac{c(t_1)}{d(t_1)}\right]u(t_0) \leqslant \frac{\ell(t_0)}{a(t_0)}\lambda - \frac{b(t_0)}{a(t_0)}\frac{m(t_1)}{d(t_1)}\mu. \tag{6.58}$$

Similarly,

$$\left[1 - \frac{b(t_0)}{a(t_0)}\frac{c(t_1)}{d(t_1)}\right]v(t_1) \leqslant \frac{m(t_1)}{d(t_1)}\mu - \frac{c(t_1)}{d(t_1)}\frac{\ell(t_0)}{a(t_0)}\lambda. \tag{6.59}$$

Moreover, thanks to (6.53),

$$0 < 1 - \left(\frac{-b}{a}\right)_M \left(\frac{-c}{d}\right)_M \leqslant 1 - \frac{b(t_0)}{a(t_0)}\frac{c(t_1)}{d(t_1)} \tag{6.60}$$

and, consequently, it follows from (6.58) and (6.59) that, necessarily,

$$\lambda \geqslant -\frac{-b(t_0)}{\ell(t_0)}\frac{m(t_1)}{d(t_1)}\mu \quad \text{and} \quad \mu \geqslant -\frac{-c(t_1)}{m(t_1)}\frac{\ell(t_0)}{a(t_0)}\lambda.$$

Therefore, (6.55) are necessary conditions for the existence of a coexistence state. Also, substituting (6.60) into (6.58) and (6.59), gives (6.56).

Now, fix $\lambda > 0$ and consider the analytic curve $\mathfrak{C}^{\lambda}_{\mu}$. Thanks to Theorem 6.6(a), $\mathfrak{C}^{\lambda}_{\mu}$ is unbounded in $\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$. Thus, thanks to the uniform a priori estimates (6.56), $P_\mu \mathfrak{C}^{\lambda}_{\mu}$ must be unbounded. Note that it is connected, since $\mathfrak{C}^{\lambda}_{\mu}$ is a curve. Therefore, the conclusion follows from the second necessary condition of (6.55). Similarly, one can obtain the corresponding assertion for $\mathfrak{C}^{\mu}_{\lambda}$. □

As a consequence from Theorem 6.8, for each $\lambda > 0$, the curve $\mathfrak{C}^{\lambda}_{\mu}$ looks like the one shown in Fig. 13, whereas, for the case of strong symbiosis, the following result shows that $\mathfrak{C}^{\lambda}_{\mu}$ looks like shows Fig. 16.

THEOREM 6.9. *Assume* (6.52), (6.54),

$$(a/d)_L (d/a)_L > \left[(-b/a)_L (-c/d)_L\right]^{-1}, \tag{6.61}$$

and (1.2) *possesses a coexistence state, say* (λ, μ, u, v). *Then,*

$$\lambda \leqslant -(-b/a)_L (a/d)_L \mu, \qquad \mu \leqslant -(-c/d)_L (d/a)_L \lambda, \tag{6.62}$$

and

$$\begin{aligned}\int_0^T u &\leqslant \frac{-T(\lambda + (-b/a)_L(a/d)_L\mu)}{a_L[(-b/a)_L(-c/d)_L(a/d)_L(d/a)_L - 1]},\\ \int_0^T v &\leqslant \frac{-T(\mu + (-c/d)_L(d/a)_L\lambda)}{d_L[(-b/a)_L(-c/d)_L(a/d)_L(d/a)_L - 1]}.\end{aligned} \tag{6.63}$$

Therefore, for each $\lambda > 0$ *and* $\mu > 0$ *there exist*

$$\mu_1 = \mu_1(\lambda) \in [\mu_0, 0) \quad \textit{and} \quad \lambda_1 = \lambda_1(\mu) \in [\lambda_0, 0)$$

such that

$$P_\mu \mathfrak{C}^\lambda_\mu \in \{(-\infty, \mu_1], (-\infty, \mu_0)\} \quad \textit{and} \quad P_\lambda \mathfrak{C}^\mu_\lambda \in \{(-\infty, \lambda_1], (-\infty, \lambda_0)\}.$$

PROOF. Dividing by u the u-equation, by v the v-equation, and integrating in $[0, T]$ gives

$$\lambda = \frac{1}{T}\int_0^T (au + bv), \qquad \mu = \frac{1}{T}\int_0^T (dv + cu). \tag{6.64}$$

Thus,

$$\begin{aligned}\int_0^T (au) &= \lambda T - \int_0^T (bv) \geqslant \lambda T + (-b/a)_L \int_0^T (av)\\ &\geqslant \lambda T + (-b/a)_L(a/d)_L \int_0^T (dv)\\ &= T\lambda + (-b/a)_L(a/d)_L T\mu + (-b/a)_L(a/d)_L \int_0^T (-cu)\\ &\geqslant T\lambda + (-b/a)_L(a/d)_L T\mu\\ &\quad + (-b/a)_L(-c/d)_L(a/d)_L(d/a)_L \int_0^T (au)\end{aligned}$$

and, hence,

$$\bigl[1 - (-b/a)_L(-c/d)_L(a/d)_L(d/a)_L\bigr]\int_0^T (au) \geqslant T\bigl[\lambda + (-b/a)_L(a/d)_L\mu\bigr].$$

Similarly,

$$\bigl[1 - (-b/a)_L(-c/d)_L(a/d)_L(d/a)_L\bigr]\int_0^T (dv) \geqslant T\bigl[\mu + (-c/d)_L(d/a)_L\lambda\bigr].$$

On the other hand, thanks to (6.61), we have that

$$(-b/a)_L(-c/d)_L(a/d)_L(d/a)_L > 1,$$

and, hence, (6.62) holds true. Therefore,

$$\begin{aligned} a_L \int_0^T u &\leqslant \int_0^T (au) \leqslant -T\frac{\lambda + (-b/a)_L(a/d)_L\mu}{(-b/a)_L(-c/d)_L(a/d)_L(d/a)_L - 1}, \\ d_L \int_0^T v &\leqslant \int_0^T (dv) \leqslant -T\frac{\mu + (-c/d)_L(d/a)_L\lambda}{(-b/a)_L(-c/d)_L(a/d)_L(d/a)_L - 1}, \end{aligned} \tag{6.65}$$

and, consequently, (6.63) is as well true.

Now, fix $\lambda > 0$ and consider $\mathfrak{C}_\mu^\lambda$. Thanks to Theorem 6.6(a), $\mathfrak{C}_\mu^\lambda$ is unbounded in $\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$. Moreover, due to second estimate of (6.62),

$$P_\mu \mathfrak{C}_\mu^\lambda \cap [0, \infty) = \emptyset.$$

On the other hand, thanks to (6.63), u and v possess uniform a priori bounds in $\mathcal{C}_T$ on compact subsets of (λ, μ). Indeed, it is rather clear that (6.63) shows that u_L and v_L are uniformly bounded. Now, pick $t_u, t_v \in \mathbb{R}$ such that

$$u_L = u(t_u), \qquad v_L = v(t_v).$$

Then,

$$u(t) = \mathrm{e}^{\int_{t_u}^t (\lambda\ell - au - bv)} u(t_u), \qquad v(t) = \mathrm{e}^{\int_{t_v}^t (\mu m - dv - cu)} v(t_v),$$

must be uniformly bounded, as well. This concludes the proof. □

Even in the case when the model possesses a stable coexistence state, the solutions of (1.1) might blow up in finite time for sufficiently large initial data. In the case of high symbiosis, where according to Theorem 6.9 Problem (1.2) does not admit a coexistence state if $\lambda > 0$ and $\mu > 0$, all solutions of (1.1) will blow-up in finite time independently of the size of the initial populations. Actually, this is the reason why (1.2) cannot admit a coexistence state. We refrain of giving more details here in, though we send to the interested reader to [11] for further technical details concerning the elliptic counterpart of the model studied here in.

7. The competing species model ($b > 0$ and $c > 0$)

Undoubtedly, this is the case that has attracted more attention in the mathematical literature. Perhaps, because in a certainly pioneer chapter, P. de Mottoni and A. Schiaffino [29] showed that any solution of (1.1) converges in the large to a T-periodic component-wise non-negative solution of (1.2), which extraordinary facilitates the analysis of the dynamics of (1.1). Among the most important works related to this model one should include the most pioneering chapter of J.M. Cushing [7], where the average of the birth rates, λ and μ, were used by the first time as the main bifurcation parameters to show the existence of a continuum, not necessarily a curve, of coexistence states of (1.2) linking the two semi-trivial states of the model. Then, a huge industry grew from these pioneering chapters. Among the most important posterior works one should quote those of S. Ahmad [1], S. Ahmad and A.C. Lazer [2], C. Alvárez and A.C. Lazer [3], J.C. Eilbeck and J. López-Gómez [12], J.K. Hale and A. Somolinos [14], P. Hess and A.C. Lazer [15], J. López-Gómez [17], R. Ortega and A. Tineo [30], H. Smith [32,33], A. Tineo [34,35], and F. Zanolin [36], among many others that are omitted here; many of them authorized by the previous authors.

Rather naturally, in this section we will focus our attention into the results obtained in [17] and [12], where it was proved that the components $\mathfrak{C}^{\lambda}_{\mu}$ and $\mathfrak{C}^{\mu}_{\lambda}$ are real analytic curves consisting of coexistence states whose attractive character changes when a turning point is crossed. This theorem is a very sharp improvement of a previous result by R. Ortega and A. Tineo [30], were it was shown that either the model possess a segment of coexistence states, or it exhibits a finite number of coexistence states.

Although in the previous references some uniqueness results were given, the problem of the uniqueness remains, in its wide amplitude, entirely open, and it will be treated elsewhere.

7.1. *Periodic systems of quasi-cooperative type*

In this section we study the periodic linear system

$$\begin{cases} u' = \alpha u - \beta v, \\ v' = -\gamma u + \rho v, \end{cases} \tag{7.1}$$

where $\alpha, \beta, \gamma, \rho \in \mathcal{C}_T$ satisfy

$$\int_0^T \alpha < 0, \qquad \int_0^T \rho < 0, \qquad \beta > 0, \qquad \gamma > 0, \tag{7.2}$$

and its associated Cauchy problem

$$\begin{cases} u' = \alpha u - \beta v, \\ v' = -\gamma u + \rho v, \\ u(0) = u_0, \quad v(0) = v_0. \end{cases} \tag{7.3}$$

As an immediate consequence from Theorem 6.1, it follows the next result.

THEOREM 7.1. *Suppose* (7.1) *possesses a solution* $(u, v) \in \mathcal{C}^1_T \times \mathcal{C}^1_T \setminus \{(0, 0)\}$. *Then, either* $u \gg 0$ *and* $v \ll 0$, *or else* $u \ll 0$ *and* $v \gg 0$. *Therefore, all the coexistence states* *of* (1.2) *must be ordered, in the sense that for any pair of coexistence states,* $(u_1, v_1) \neq (u_2, v_2)$, *either* $u_1 \gg u_2$ *and* $v_1 \ll v_2$, *or* $u_1 \ll u_2$ *and* $v_1 \gg v_2$.

PROOF. The change of variable

$$(x, y) = (u, -v) \tag{7.4}$$

transforms (7.1) into

$$\begin{cases} x' = \alpha x + \beta y, \\ y' = \gamma x + \rho y, \end{cases} \tag{7.5}$$

which is of the same type as (6.1). Therefore, the result is a corollary of Theorem 6.1. □

Similarly, the following quasi-cooperative counterpart of Proposition 6.2 holds.

PROPOSITION 7.2. *Suppose* $(\bar{u}, \bar{v}) \in \mathcal{C}^1(\mathbb{R}_+) \times \mathcal{C}^1(\mathbb{R}_+)$, $\mathbb{R}_+ := [0, \infty)$, *satisfy*

$$\begin{cases} \bar{u}' \geqslant \alpha \bar{u} - \beta \bar{v}, \\ \bar{v}' \leqslant -\gamma \bar{u} + \rho \bar{v}, \\ \bar{u}(0) > 0, \quad \bar{v}(0) < 0. \end{cases} \tag{7.6}$$

Then, $\bar{u}(t) > 0$ *and* $\bar{v}(t) < 0$ *for each* $t \geqslant 0$.

If $u_0 - v_0 > 0$ *and* $u_0 v_0 = 0$, *then there exists* $t_0 > 0$ *such that the unique solution* (u, v) *of* (7.3) *satisfies* $u(t) > 0$ *and* $v(t) < 0$ *for each* $t \geqslant t_0$.

PROOF. The auxiliary pair

$$(\bar{x}, \bar{y}) := (\bar{u}, -\bar{v})$$

satisfies

$$\begin{cases} \bar{x}' \geqslant \alpha \bar{x} + \beta \bar{y}, \\ \bar{y}' \geqslant \gamma \bar{x} + \rho \bar{y}, \\ \bar{x}(0) > 0, \quad \bar{y}(0) > 0, \end{cases}$$

and, hence, thanks to Proposition 6.2, $\bar{x}(t) > 0$ and $\bar{y}(t) > 0$ for each $t \geqslant 0$, which concludes the proof of the first claim.

Now, suppose $u_0 - v_0 > 0$ and $u_0 v_0 = 0$, and let $(u(t), v(t))$ denote the unique solution of (7.3). Then,

$$(x, y) := (u, -v)$$

satisfies

$$\begin{cases} x' = \alpha x + \beta y, \\ y' = \gamma x + \rho y, \\ x(0) = u_0, \quad y(0) = -v_0, \end{cases}$$

and

$$x(0) + y(0) = u_0 - v_0 > 0, \qquad x(0)y(0) = 0.$$

Thus, it follows from Proposition 6.2 that there exists $t_0 > 0$ such that $x(t) > 0$ and $y(t) > 0$ for each $t \geqslant t_0$. Therefore, $u(t) > 0$ and $v(t) < 0$ for each $t \geqslant t_0$, which concludes the proof. ☐

Also, the following counterpart of Theorem 6.3 holds. Subsequently, we shall denote by $\Phi(T)$ the Poincaré map of (7.1).

THEOREM 7.3. *The characteristic multipliers of* (7.1), ν_1 *and* ν_2, *are real and distinct, and satisfy* $0 < \nu_1\nu_2 < 1$. *Moreover, if we order them so that* $0 < \nu_1 < \nu_2$, *then,* $0 < \nu_1 < 1$,

$$N\big[\Phi(T) - \nu_1 I\big] = \operatorname{span}\big[(x_0, y_0)\big] \quad \textit{with } x_0 > 0 \textit{ and } y_0 > 0,$$

and

$$N\big[\Phi(T) - \nu_2 I\big] = \operatorname{span}\big[(x_0, y_0)\big] \quad \textit{with } x_0 > 0 \textit{ and } y_0 < 0.$$

In particular, (7.1) *is asymptotically stable if* $\nu_2 < 1$, *neutrally stable if* $\nu_2 = 1$, *and unstable if* $\nu_2 > 1$. *Furthermore, the value*

$$\lambda_0 := -\frac{1}{T}\operatorname{Log}\nu_2 \tag{7.7}$$

provides us with the unique value of $\lambda \in \mathbb{R}$ *for which the eigenvalue problem*

$$\begin{cases} u' = \alpha u - \beta v + \lambda u, \\ v' = -\gamma u + \rho v + \lambda v \end{cases} \tag{7.8}$$

admits a solution $(u, v) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ *such that* $u \gg 0$ *and* $v \ll 0$.

PROOF. As the change of variable (7.4) transforms (7.1) into (7.5), if

$$U(T) = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}$$

denotes the Poincaré map of (7.5), necessarily

$$\Phi(T)=\begin{pmatrix} u_{11} & -u_{12} \\ -u_{21} & u_{22} \end{pmatrix}. \tag{7.9}$$

Indeed, if we denote by $(u(t), v(t))$ the unique solution of (7.3) and $(x(t), y(t))$ stands for the unique solution of (7.5) satisfying $(x(0), y(0)) = (u_0, -v_0)$, then

$$\begin{pmatrix} u(T) \\ v(T) \end{pmatrix} = \begin{pmatrix} x(T) \\ -y(T) \end{pmatrix} = \begin{pmatrix} u_{11} & -u_{12} \\ -u_{21} & u_{22} \end{pmatrix} \begin{pmatrix} x(0) \\ -y(0) \end{pmatrix} = \Phi(T) \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}.$$

Clearly, the eigenvalues of $U(T)$ and $\Phi(T)$ are the same. Moreover, thanks to Theorem 6.3,

$$N\big[U(T)-\nu_1 I\big] = \operatorname{span}\big[(\tilde{x}_0, \tilde{y}_0)\big] \quad \text{with } \tilde{x}_0 < 0 \text{ and } \tilde{y}_0 > 0,$$
$$N\big[U(T)-\nu_2 I\big] = \operatorname{span}\big[(\tilde{x}_0, \tilde{y}_0)\big] \quad \text{with } \tilde{x}_0 > 0 \text{ and } \tilde{y}_0 > 0,$$

and a direct calculation shows that, for each $j \in \{1, 2\}$, (x, y) is an eigenvector of $U(T)$ associated to the eigenvalue ν_j if, and only if, $(x, -y)$ is an eigenvector of $\Phi(T)$ associated to ν_j. Therefore,

$$N\big[\Phi(T)-\nu_1 I\big] = \operatorname{span}\big[(x_0, y_0)\big] \quad \text{with } x_0 > 0 \text{ and } y_0 > 0,$$

and

$$N\big[\Phi(T)-\nu_2 I\big] = \operatorname{span}\big[(x_0, y_0)\big] \quad \text{with } x_0 > 0 \text{ and } y_0 < 0,$$

which concludes the proof of the first block of statements.

Finally, note that the change of variable (7.4) transforms (7.8) into the equivalent system

$$\begin{cases} x' = \alpha x + \beta y + \lambda x, \\ y' = \gamma x + \rho y + \lambda y, \end{cases}$$

for which the last assertion of Theorem 6.3 is valid. This concludes the proof. □

Finally, by performing the change of variable (7.4), it is easy to see that the following counterpart of Theorem 6.4 is satisfied.

THEOREM 7.4. *The following properties are satisfied*:

(a) *System* (7.1) *is stable* ($\nu_2 < 1$) *if, and only if, there exists* $(\bar{u}, \bar{v}) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ *with* $\bar{u} \gg 0$ *and* $\bar{v} \ll 0$ *such that*

$$\begin{cases} \bar{u}' \geqslant \alpha \bar{u} - \beta \bar{v}, \\ \bar{v}' \leqslant -\gamma \bar{u} + \rho \bar{v}, \end{cases} \tag{7.10}$$

with some of these inequalities strict.

(b) *System* (6.1) *is unstable* ($\nu_2 > 1$) *if, and only if, there exits* $(\underline{u}, \underline{v}) \in \mathcal{C}^1_T \times \mathcal{C}^1_T$ *with* $\underline{u} \gg 0$ *and* $\underline{v} \ll 0$ *such that*

$$\begin{cases} \underline{u}' \leqslant \alpha \underline{u} - \beta \underline{v}, \\ \underline{v}' \geqslant -\gamma \underline{u} + \rho \underline{v}, \end{cases} \tag{7.11}$$

with some of these inequalities strict.

7.2. *Structure of the components* $\mathfrak{C}^\lambda_\mu$ *and* $\mathfrak{C}^\mu_\lambda$

As in Section 6.2, the coexistence states of (1.2) will be regarded as zeroes of the operator $\mathfrak{F} : \mathbb{R}^2 \times \mathcal{C}_T \times \mathcal{C}_T \to \mathcal{C}_T \times \mathcal{C}_T$ defined by (6.23), where $M > 0$ is fixed. The following counterpart of Proposition 6.5 is satisfied.

PROPOSITION 7.5. *Suppose* $(u_0, v_0) \in \mathcal{C}_T \times \mathcal{C}_T$ *satisfies* $u_0 \gg 0$, $v_0 \gg 0$, *and*

$$\mathfrak{F}(\lambda, \mu, u_0, v_0) = 0.$$

Then, each of the operators

$$D_{(\lambda,u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0) : \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T \to \mathcal{C}_T \times \mathcal{C}_T$$

and

$$D_{(\mu,u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0) : \mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T \to \mathcal{C}_T \times \mathcal{C}_T$$

is surjective. In other words, $(0, 0)$ *is a regular value of* $\mathfrak{F}$ *with respect to* $\mathbb{R}^2 \times \operatorname{Int} P \times \operatorname{Int} P$.

PROOF. As in the proof of Proposition 6.5, differentiating gives (6.24), (6.25) and (6.26). Similarly, $D_{(\lambda,u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0)$ and $D_{(\mu,u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0)$ are surjective if (λ, μ, u_0, v_0) is non-degenerate. So, suppose (λ, μ, u_0, v_0) is degenerate. Then, since $D_{(u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0)$ is a compact perturbation of the identity,

$$N := \dim N\big[D_{(u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0)\big] = \operatorname{codim} R\left[D_{(u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0)\right] \geqslant 1,$$

and $N[D_{(u,v)}\mathfrak{F}(\lambda, \mu, u_0, v_0)]$ is the set of T-periodic solutions (u, v) of the linear system

$$\begin{cases} u' = \alpha_0 u - b u_0 v, \\ v' = -c v_0 u + \beta_0 v. \end{cases} \tag{7.12}$$

Since $b > 0$ and $c > 0$, we have that

$$-b u_0 < 0 \quad \text{and} \quad -c v_0 < 0.$$

Moreover, since (λ,μ,u_0,v_0) is a coexistence state of (1.2),

$$\int_0^T \alpha_0 = -\int_0^T a(t)u_0(t)\,\mathrm{d}t < 0 \quad \text{and} \quad \int_0^T \beta_0 = -\int_0^T d(t)v_0(t)\,\mathrm{d}t < 0.$$

Therefore, (7.12) fits into the abstract setting of Section 7.1. Thus, thanks to Theorem 7.3, the biggest multiplier of (7.12), denoted by ν_2, equals one and any T-periodic solution of (7.12) must be a multiple of a fixed T-periodic solution (φ,ψ) such that $\varphi \gg 0$ and $\psi \ll 0$. In particular, $N = 1$. Now, due to (6.25) and (6.26), to conclude the proof it suffices to prove (6.29). On the contrary, suppose that, e.g.,

$$-\mathfrak{R}_M(\ell u_0, 0) \in R\big[D_{(u,v)}\mathfrak{F}(\lambda,\mu,u_0,v_0)\big].$$

Then, the system

$$\begin{cases} u' = \alpha_0 u - b u_0 v + \ell u_0, \\ v' = -c v_0 u + \beta_0 v, \end{cases} \tag{7.13}$$

possesses a T-periodic solution (u,v). Actually, for each $R > 0$, the pair

$$(u_R, v_R) := (u + R\varphi, v + R\psi)$$

provides us with a T-periodic solution of (7.13) such that, for sufficiently large $R > 0$,

$$u_R \gg 0 \quad \text{and} \quad v_R \ll 0.$$

Moreover, it follows from $\ell u_0 > 0$ that

$$\begin{cases} u_R' > \alpha_0 u_R - b u_0 v_R, \\ v_R' = -c v_0 u_R + \beta_0 v_R, \end{cases}$$

and, therefore, thanks to Theorem 7.4(a), $\nu_2 < 1$, which is a contradiction, since $\nu_2 = 1$. Similarly, the second relation of (6.29) holds. This concludes the proof. □

The following result ascertains the global structure of the coexistence states of (1.2) for the competing species model. To reduce the complexity of the number of cases to be considered in our analysis we shall throughout assume that conditions (4.10) are satisfied, though most of the results are still valid in the absence of (4.10)—such restrictions are exclusively needed in order to get some non-existence results entailing the relative compactness of the components $\mathfrak{C}_\mu^\lambda$ and $\mathfrak{C}_\lambda^\mu$.

THEOREM 7.6. *Assume* $b > 0$, $c > 0$, *and* (1.2) *exhibits a coexistence state, say* (λ,μ,u_0,v_0). *Then,* $\lambda > 0$, $\mu > 0$, *and*

$$u_0 \ll \theta_{[\lambda\ell,a]}, \qquad v_0 \ll \theta_{[\mu m,d]}. \tag{7.14}$$

Moreover, the following properties are satisfied:

(a) *For each* $\lambda > 0$ *there exists* $\mu_1(\lambda) > 0$ *such that* (1.2) *does not admit a coexistence state if* $\mu \geqslant \mu_1(\lambda)$. *Moreover, the component of coexistence states* $\mathfrak{C}_\mu^\lambda$ *constructed in Theorem* 5.4 *is bounded in* $\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ *and it consists of a real analytic curve. More precisely, there exits a real analytic map* $(\mu, u, v) : [0, 1] \to \mathbb{R} \times \operatorname{Int} P \times \operatorname{Int} P$ *such that*

$$\big(\mu(0), u(0), v(0)\big) = (\mu_0, \theta_{[\lambda\ell,a]}, 0),$$
$$\big(\mu(1), u(1), v(1)\big) = (\mu_*, 0, \theta_{[\mu_* m,d]}),$$

and

$$\mathfrak{C}_\mu^\lambda = \big\{\big(\mu(s), u(s), v(s)\big) \colon\ s \in (0, 1)\big\},$$

where $\mu_* > 0$ *is the unique value of* μ *satisfying* (5.40).

(b) *For each* $\mu > 0$ *there exists* $\lambda_1(\mu) > 0$ *such that* (1.2) *does not admit a coexistence state if* $\lambda \geqslant \lambda_1(\mu)$. *Moreover, the component of coexistence states* $\mathfrak{C}_\lambda^\mu$ *constructed in Theorem* 5.5 *is bounded in* $\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ *and it consists of a real analytic curve. More precisely, there exists a real analytic map* $(\lambda, u, v) : [0, 1] \to \mathbb{R} \times \operatorname{Int} P \times \operatorname{Int} P$ *such that*

$$\big(\lambda(0), u(0), v(0)\big) = (\lambda_0, 0, \theta_{[\mu m,d]}),$$
$$\big(\lambda(1), u(1), v(1)\big) = (\lambda_*, \theta_{[\lambda_*\ell,a]}, 0),$$

and

$$\mathfrak{C}_\lambda^\mu = \big\{(\lambda(s), u(s), v(s)\big) \colon\ s \in (0, 1)\big\},$$

where $\lambda_* > 0$ *is the unique value of* λ *satisfying* (5.46).

(c) *The set of coexistence states of* (1.2) $(\lambda, \mu, u, v) \in \mathbb{R}^2 \times \operatorname{Int} P \times \operatorname{Int} P$ *such that, for some* $\mu \in \mathbb{R}$, $(\mu, u, v) \in \mathfrak{C}_\mu^\lambda$, *or, for some* $\lambda \in \mathbb{R}$, $(\lambda, u, v) \in \mathfrak{C}_\lambda^\mu$, *is a real analytic surface linking the surfaces of semi-trivial states*

$$\big\{(\lambda, \mu, \theta_{[\lambda\ell,a]}, 0) : \lambda > 0,\ \mu \in \mathbb{R}\big\}$$

and

$$\big\{(\lambda, \mu, 0, \theta_{[\mu m,d]}) \colon\ \lambda \in \mathbb{R},\ \mu > 0\big\}$$

along their respective curves of neutral stability

$$\mu = \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t \quad \textit{and} \quad \lambda = \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t.$$

Therefore, (1.2) *possesses a coexistence state if both semi-trivial states are, simultaneously, linearly unstable, or linearly stable.*

PROOF. The fact that the components $\mathfrak{C}_\mu^\lambda$ and $\mathfrak{C}_\lambda^\mu$ are real analytic curves follows from Proposition 7.5 adapting the argument of the proof of Theorem 6.6. Now, we shall show that both components are bounded. The remaining assertions of the theorem follow straight ahead from Theorems 5.4 and 5.5.

Suppose (λ, μ, u_0, v_0) is a coexistence state of (1.2). Then, dividing by u_0 the u-equation, by v_0 the v-equation, and integrating in $[0, T]$ gives

$$\lambda = \frac{1}{T}\int_0^T (au_0 + bv_0) > 0 \quad \text{and} \quad \mu = \frac{1}{T}\int_0^T (dv_0 + cu_0) > 0.$$

Thus, $\lambda > 0$ and $\mu > 0$ are necessary for the existence of a coexistence state. So, throughout the remaining of the proof we will assume that $\lambda > 0$ and $\mu > 0$.

Fix $\lambda > 0$. Then, $u_0 \gg 0$ is a positive strict subsolution of the problem

$$u' = \lambda\ell - au^2, \quad u \in \mathcal{C}_T^1,$$

since $bu_0v_0 > 0$, and, hence, thanks to Proposition 2.6, $u_0 \ll \theta_{[\lambda\ell,a]}$. Similarly, $v_0 \ll \theta_{[\mu m,d]}$. Thus,

$$v_0 = \mu m v_0 - dv_0^2 - cu_0v_0 > (\mu m - c\theta_{[\lambda\ell,a]})v_0 - dv_0^2.$$

Hence, if

$$\mu > \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t,$$

then, thanks again to Proposition 2.6, we have that

$$v_0 \gg \theta_{[\mu m - c\theta_{[\lambda\ell,a]},d]} := \Theta_\mu,$$

and, so,

$$u_0 < (\lambda\ell - b\Theta_\mu)u_0 - au_0^2.$$

Consequently, dividing by u_0 and integrating in $[0, T]$ gives

$$\lambda > \frac{1}{T}\int_0^T b(t)\Theta_\mu(t)\,\mathrm{d}t. \tag{7.15}$$

To show the existence of $\mu_1(\lambda)$ such that (1.2) cannot admit a coexistence state if $\mu \geqslant \mu_1$ it suffices to show that (7.15) cannot be satisfied for sufficiently large μ. To show this, note that the function Ψ_μ defined through $\Theta_\mu := \mu\Psi_\mu$ satisfies

$$\frac{1}{\mu}\Psi_\mu' = \left(m - \frac{c\theta_{[\lambda\ell,a]}}{\mu}\right)\Psi_\mu - \mathrm{d}\Psi_\mu^2.$$

Adapting the argument of the proof of (2.20) it is apparent that $\lim_{\mu\uparrow\infty}\Psi_\mu = \frac{m}{d}$ uniformly in $\mathbb{R}$, since $\lim_{\mu\uparrow\infty}\frac{c\theta_{[\lambda\ell,a]}}{\mu} = 0$. Therefore,

$$\lim_{\mu\uparrow\infty}\Theta_\mu = \infty$$

uniformly in $[0,T]$, and, consequently, (7.15) fails to be true for sufficiently large μ. By symmetry, for any $\mu > 0$ there exists $\lambda_1(\mu) > 0$ such that (1.2) does not admit a coexistence state if $\lambda \geqslant \lambda_1(\mu)$.

The previous bounds show that $\mathfrak{C}^\lambda_\mu$ and $\mathfrak{C}^\mu_\lambda$ must be bounded for each $\lambda > 0$ and $\mu > 0$. The remaining assertions of the theorem follow readily from Theorems 5.4, 5.5 and 5.6. □

For the autonomous counterpart of (1.2) with $bc = 1$, is easy to check that, for any $\lambda > 0$, the curve

$$\big(\mu(s), u(s), v(s)\big) = (c\lambda, \lambda - bs, s), \quad 0 \leqslant s \leqslant \frac{\lambda}{b},$$

provides us with the component $\mathfrak{C}^\lambda_\mu$ bifurcating from the semi-trivial solution $(\lambda, 0)$ at $\mu = c\lambda$. As this simple example shows, condition $\dot{\mu} = 0$ cannot be avoided in general, unless some additional assumptions are imposed; e.g., that the semi-trivial states cannot be simultaneously neutrally stable.

As in the symbiotic model, fixing λ, or μ, and due to Proposition 7.5, any coexistence state of (1.2), say (λ, μ, u_0, v_0), must lay into a real analytic curve of coexistence states of (1.2). Fix, e.g., λ and let Γ denote the trajectory of the μ–curve of coexistence states containing (λ, μ, u_0, v_0). As a consequence from the existence of uniform a priori bounds in $\mathcal{C}_T$ for the coexistence states of (1.2) (cf. Theorem 7.6), Γ must be bounded. If $\Gamma \cap \mathfrak{C}^\lambda_\mu \neq \emptyset$, then $\Gamma = \mathfrak{C}^\lambda_\mu$. Thus, if $\Gamma \neq \mathfrak{C}^\lambda_\mu$, necessarily $\Gamma \cap \mathfrak{C}^\lambda_\mu = \emptyset$, and, consequently, Γ must be an isolated component with respect of the surfaces of semi-trivial positive solutions of (1.2). Suppose such a component exists. Then, by varying an additional parameter of the model, e.g., the amplitude of b, or c, it is possible to show that when b, or c are sufficiently small, then (1.2) possesses a unique coexistence state, necessarily belonging to $\mathfrak{C}^\lambda_\mu$. Therefore, at some intermediate value of the *unfolding parameter*, the isolated component Γ must reduce to a single point. Thanks to Proposition 7.5 this is impossible. Therefore, the only coexistence states of (1.2) are those contained in the components $\mathfrak{C}^\lambda_\mu$. The technical details of this analysis will appear elsewhere.

As an easy consequence from Theorem 7.6, except in the case when both semi-trivial states are neutrally stable, (1.2) possesses, at most, a finite number of coexistence states and, according to Theorem 7.1, all of them must be ordered, which provides us with the main result of R. Ortega and A. Tineo [30].

In the case when both semi-trivial states are neutrally stable, either there is an analytic curve of coexistence states linking both states, or the model possesses a finite number of coexistence states; its cardinal depending on the values of the several parameter functions involved in the setting of the model.

7.3. *Stability of the coexistence states along $\mathfrak{C}^{\lambda}_{\mu}$ and $\mathfrak{C}^{\mu}_{\lambda}$*

Fix $\lambda > 0$ and let $(\mu(s), u(s), v(s))$, $s \geqslant 0$, be the real analytic map constructed in Theorem 7.6, where $s \in (0, L)$ is assumed to be the length of arc of curve. Subsequently, we shall denote $\cdot := \frac{\mathrm{d}}{\mathrm{d}s}$. Repeating the argument of Section 6.3, we have (6.31) and (6.32). Moreover, since

$$\big(\mu(0), u(0), v(0)\big) = (\mu_0, \theta_{[\lambda\ell,a]}, 0),$$

it follows from Theorem 5.6 that

$$q_s := \dot{v}(s) = \varphi_v + \mathrm{O}(s) = \mathrm{e}^{\int_0^{\cdot}(\mu_0 m - c\theta_{[\lambda\ell,a]})} + \mathrm{O}(s) \gg 0, \quad s \downarrow 0. \tag{7.16}$$

Similarly, setting

$$(p_s, q_s) := \big(\dot{u}(s), \dot{v}(s)\big), \quad s > 0,$$

(6.31) can be rewritten in the form

$$\begin{cases} (p_s)' = \big[\lambda\ell - 2au(s) - bv(s)\big]p_s - bu(s)q_s, \\ (q_s)' = -cv(s)p_s + \big[\mu(s)m - 2dv(s) - cu(s)\big]q_s + \dot{\mu}(s)mv(s). \end{cases} \tag{7.17}$$

In the competing species model, we have that $-bu(s) < 0$, $-cv(s) < 0$, (6.35) and (6.36). Thus, (7.17) fits within the general setting of Section 7.1. In particular, thanks to Proposition 2.1, we find from the first equation of (7.17) that, for each sufficiently small $s > 0$,

$$p_s = \mathfrak{R}_{-[\lambda\ell - 2au(s) - bv(s)]}\big(-bu(s)q_s\big) \ll 0.$$

Therefore, there exists $\varepsilon > 0$ such that

$$p_s \ll 0 \quad \text{and} \quad q_s \gg 0 \quad \text{for each } s \in [0, \varepsilon]. \tag{7.18}$$

Suppose $\mu(s) = \mu_0$ for each $s \in [0, L]$. Then, (7.17) becomes into (6.38) and, due to (6.32), $(p_s, q_s) \neq (0, 0)$. Thus, thanks to Theorem 7.1, for each $s \in [0, L]$, either $p_s \gg 0$ and $q_s \ll 0$, or $p_s \ll 0$ and $q_s \gg 0$. Thanks to (7.18), necessarily

$$p_s \ll 0 \quad \text{and} \quad q_s \gg 0 \quad \text{for each } s \in [0, L]. \tag{7.19}$$

Indeed, since $s \to (p_s, q_s)$ is analytic, if (7.19) is not satisfied, then, due to (6.37), there is some value of s, say $\tilde{s} \in (0, L]$ where $p_{\tilde{s}} < 0$, $q_{\tilde{s}} > 0$, and, e.g., $p_{\tilde{s}} \ll 0$ fails. Due to Theorem 7.1, this implies $p_{\tilde{s}} \ll 0$ and $q_{\tilde{s}} \gg 0$, which is a contradiction. Thus, (7.19) holds true. Consequently, thanks to Theorem 7.3, the principal characteristic multiplier of $(\mu(s), u(s), v(s))$, $\nu_2(s)$, equals 1 for each $s \in [0, L]$. Therefore, if $\dot{\mu} = 0$, then $\dot{\nu}_2 = 0$ and, hence, all the coexistence states of $\mathfrak{C}^{\lambda}_{\mu}$ are neutrally stable. When $\dot{\mu} \neq 0$, the following result is satisfied.

THEOREM 7.7. *Suppose $\dot\mu \neq 0$ in $[0, L]$. Then, the set*

$$\mathcal{S} := \{s \in (0, L)\colon\ \dot\mu(s) = 0\} \tag{7.20}$$

is finite (possibly empty), each $s \in \mathcal{S}$ is a zero of finite order of $\dot\mu$, and

$$\mathcal{S} = \{s \in (0, L)\colon\ \nu_2(s) = 1\}. \tag{7.21}$$

Moreover,

$$\dot\mu(s)\bigl(1 - \nu_2(s)\bigr) > 0, \quad s \in (0, L) \setminus \mathcal{S}, \tag{7.22}$$

and

$$p_s \ll 0 \quad \textit{and} \quad q_s \gg 0 \quad \textit{whenever} \quad \dot\mu(s) \geqslant 0. \tag{7.23}$$

PROOF. The fact that $\mathcal{S}$ is, at most, finite and that each $s \in \mathcal{S}$ has finite order—as a zero of $\dot\mu$—are consequences from the fact that $\dot\mu$ is real analytic. Now, we shall prove (7.21). Pick $s \in \mathcal{S}$. Then, (p_s, q_s) satisfies (6.38) and, due to (6.32), $(p_s, q_s) \neq (0, 0)$. Thus, thanks to Theorem 7.1, $p_s \ll 0$ and $q_s \gg 0$, or $p_s \gg 0$ and $q_s \ll 0$. In any of these cases, we find from Theorem 7.3 that $\lambda = 0$ is the *principal characteristic exponent* of $(\mu(s), u(s), v(s))$ and, hence, $\nu_2(s) = 1$. To prove the converse inclusion, let $s \in (0, L)$ such that $\nu_2(s) = 1$. Then, thanks to Theorem 7.3, there exists a T-periodic pair (P, Q) with $P \gg 0$ and $Q \ll 0$ such that

$$\begin{cases} P' = [\lambda\ell - 2au(s) - bv(s)]P - bu(s)Q, \\ Q' = -cv(s)P + [\mu(s)m - 2dv(s) - cu(s)]P. \end{cases} \tag{7.24}$$

Now, choose a sufficiently small $\varepsilon > 0$ so that the auxiliary pair

$$(U, V) := (P - \varepsilon p_s, Q - \varepsilon q_s),$$

satisfy $U \gg 0$ and $V \ll 0$. Thanks to (7.17) and (7.24) we have that

$$\begin{cases} U' = [\lambda\ell - 2au(s) - bv(s)]U - bu(s)V, \\ V' = -cv(s)U + [\mu(s)m - 2dv(s) - cu(s)]V - \varepsilon\dot\mu(s)mv(s). \end{cases}$$

Assume $\dot\mu(s) > 0$. Then, since $\varepsilon mv(s) > 0$, the pair (U, V) satisfies

$$\begin{cases} U' = [\lambda\ell - 2au(s) - bv(s)]U - bu(s)V, \\ V' < -cv(s)U + [\mu(s)m - 2dv(s) - cu(s)]V, \end{cases}$$

and, hence, thanks to Theorem 7.4(a), $\nu_2(s) < 1$, which is a contradiction. Assume $\dot{\mu}(s) < 0$. Then, (U, V) satisfies

$$\begin{cases} U' = \big[\lambda\ell - 2au(s) - bv(s)\big]U - bu(s)V, \\ V' > -cv(s)U + \big[\mu(s)m - 2dv(s) - cu(s)\big]V, \end{cases}$$

and, due to Theorem 7.4(b), $\nu_2(s) > 1$, which is a contradiction again. Consequently, $\dot{\mu}(s) = 0$ and, therefore, (7.21) holds true.

Now, pick $\tilde{s} \in \mathcal{S}$. Thanks to (6.32), it follows from Theorem 7.1 that some of the following alternatives occurs:

1. $p_{\tilde{s}} \gg 0$ and $q_{\tilde{s}} \ll 0$.
2. $p_{\tilde{s}} \ll 0$ and $q_{\tilde{s}} \gg 0$.

Assume alternative 1 occurs. As $\tilde{s}$ is an isolated zero of $\dot{\mu}(s)$, there exists $\varepsilon > 0$ such that $p_s = \dot{u}(s) \gg 0$ and $q_s = \dot{v}(s) \ll 0$ for each $s \in (\tilde{s} - \varepsilon, \tilde{s} + \varepsilon)$, and

$$\dot{\mu}^{-1}(0) \cap (\tilde{s} - \varepsilon, \tilde{s} + \varepsilon) \cap (0, L) = \{\tilde{s}\}.$$

Thus, for each

$$s \in J := (0, L) \cap \big[(\tilde{s} - \varepsilon, \tilde{s} + \varepsilon) \setminus \{\tilde{s}\}\big],$$

either $\dot{\mu}(s) > 0$, or $\dot{\mu}(s) < 0$. Suppose $\dot{\mu}(s) > 0$. Then,

$$\begin{cases} (p_s)' = \big[\lambda\ell - 2au(s) - bv(s)\big]p_s - bu(s)q_s, \\ (q_s)' > -cv(s)p_s + \big[\mu(s)m - 2dv(s) - cu(s)\big]q_s, \end{cases} \tag{7.25}$$

and it follows from Theorem 7.4(b) that $\nu_2(s) > 1$. Hence, $\dot{\mu}(s)(1 - \nu_2(s)) < 0$. Now, suppose $\dot{\mu}(s) < 0$. Then,

$$\begin{cases} (p_s)' = \big[\lambda\ell - 2au(s) - bv(s)\big]p_s - bu(s)q_s, \\ (q_s)' < -cv(s)p_s + \big[\mu(s)m - 2dv(s) - cu(s)\big]q_s, \end{cases} \tag{7.26}$$

and it follows from Theorem 7.4(a) that $\nu_2(s) < 1$. Hence, $\dot{\mu}(s)(1 - \nu_2(s)) < 0$. Therefore, the following condition is satisfied

$$\dot{\mu}(s)\big(1 - \nu_2(s)\big) < 0, \quad s \in (0, L) \cap \big[(\tilde{s} - \varepsilon, \tilde{s} + \varepsilon) \setminus \{\tilde{s}\}\big], \tag{7.27}$$

for each $\tilde{s} \in \mathcal{S}$ satisfying alternative 1.

Now, suppose $\tilde{s} \in \mathcal{S}$ satisfies alternative 2. As $\tilde{s}$ is an isolated zero of $\dot{\mu}(s)$, there exists $\varepsilon > 0$ such that $p_s = \dot{u}(s) \ll 0$ and $q_s = \dot{v}(s) \gg 0$ for each $s \in (\tilde{s} - \varepsilon, \tilde{s} + \varepsilon)$, and

$$\dot{\mu}^{-1}(0) \cap (\tilde{s} - \varepsilon, \tilde{s} + \varepsilon) \cap (0, L) = \{\tilde{s}\}.$$

Thus, for each

$$s \in J := (0, L) \cap \big[(\tilde{s} - \varepsilon, \tilde{s} + \varepsilon) \setminus \{\tilde{s}\}\big],$$

either $\dot{\mu}(s) > 0$, or $\dot{\mu}(s) < 0$. Suppose $\dot{\mu}(s) > 0$. Then,

$$\begin{cases} (-p_s)' = \big[\lambda\ell - 2au(s) - bv(s)\big](-p_s) - bu(s)(-q_s), \\ (-q_s)' < -cv(s)(-p_s) + \big[\mu(s)m - 2dv(s) - cu(s)\big](-q_s), \end{cases} \tag{7.28}$$

with $-p_s \gg 0$ and $-q_s \ll 0$, and, hence, due to Theorem 7.4(a), $\nu_2(s) < 1$. Thus, $\dot{\mu}(s)(1 - \nu_2(s)) > 0$. Similarly, if $\dot{\mu}(s) < 0$, then, due to Theorem 7.4(b), $\nu_2(s) < 1$. Therefore, condition (7.22) is satisfied for each $\tilde{s} \in \mathcal{S}$ satisfying alternative 2.

Now, suppose

$$\mathcal{S} = \{s_n \colon 1 \leqslant n \leqslant N\}, \quad 0 < s_n < s_{n+1},\ 1 \leqslant n \leqslant N - 1;$$

the subsequent argument is easily adapted to cover the case when $\mathcal{S}$ is empty. Thanks to (7.19), for sufficiently small $s > 0$, $p_s \ll 0$ and $q_s \gg 0$. Moreover, either $\dot{\mu}(s) > 0$ for each $s \in (0, s_1)$, or $\dot{\mu}(s) < 0$ for each $s \in (0, s_1)$.

Suppose $\dot{\mu}(s) > 0$ for each $s \in (0, s_1)$. Then, for sufficiently small $s > 0$, $(-p_s, -q_s)$ satisfies (7.28) and, hence, thanks to Theorem 7.4(a), $\nu_2(s) < 1$. Actually, this is the inter-exchange stability principle between the semi-trivial state $(\mu, \theta_{[\lambda\ell,a]}, 0)$ and the bifurcating coexistence state $(\mu(s), u(s), v(s))$ as μ crosses the bifurcation value μ_0. Thanks to (7.21), necessarily $\nu_2(s) < 1$ for each $s \in (0, s_1)$ and, therefore,

$$\dot{\mu}(s)\big(1 - \nu_2(s)\big) > 0, \quad s \in (0, s_1). \tag{7.29}$$

Necessarily $p_{s_1} \ll 0$ and $q_{s_1} \gg 0$. Indeed, if $p_{s_1} \gg 0$ and $q_{s_1} \ll 0$, then (7.27) holds, which contradicts (7.29). Therefore,

$$\dot{\mu}(s)\big(1 - \nu_2(s)\big) > 0, \quad s \in (0, s_2) \setminus \{s_2\}. \tag{7.30}$$

Reiterating this argument, (7.22) holds true.

Now, suppose $\dot{\mu}(s) < 0$ for each $s \in (0, s_1)$. Then, $(-p_s, -q_s)$ satisfies

$$\begin{cases} (-p_s)' = \big[\lambda\ell - 2au(s) - bv(s)\big](-p_s) - bu(s)(-q_s), \\ (-q_s)' > -cv(s)(-p_s) + \big[\mu(s)m - 2dv(s) - cu(s)\big](-q_s), \end{cases} \tag{7.31}$$

and, hence, thanks to Theorem 7.4(b), $\nu_2(s) > 1$. As in the previous case, this provides us with the inter-exchange stability principle between the semi-trivial state $(\mu, \theta_{[\lambda\ell,a]}, 0)$ and the bifurcating coexistence state $(\mu(s), u(s), v(s))$. Thanks to (7.21), $\nu_2(s) > 1$ for each $s \in (0, s_1)$ and, therefore, (7.22) is as well satisfied. Arguing as above, the proof of (7.22) is concluded.

Note that, as a result from the previous analysis, we have that

$$p_{\tilde{s}} \ll 0 \quad \text{and} \quad q_{\tilde{s}} \gg 0 \quad \text{for each } \tilde{s} \in \mathcal{S} \cup \{0\}. \tag{7.32}$$

Finally, let $s_1^*, s_2^* \in \mathcal{S} \cup \{0\}$ such that

$$\dot{\mu}(s) > 0, \quad s \in \left(s_1^*, s_2^*\right).$$

Thanks to (7.32), $p_s \ll 0$ and $q_s \gg 0$ for each $s > s_1^*$ sufficiently close to s_1^*. Thus, either $p_s \ll 0$ and $q_s \gg 0$ for each $s \in (s_1^*, s_2^*)$, or there exist $s_0 \in (s_1^*, s_2^*)$ and $t \in \mathbb{R}$ such that $p_{s_0} \leqslant 0$ and $q_{s_0} \geqslant 0$ and either $p_{s_0}(t) = 0$, or $q_{s_0}(t) = 0$. Since

$$-cv(s_0)p_{s_0} + \dot{\mu}(s_0)mv(s_0) > 0,$$

we find from Proposition 2.1 applied to the second equation of (7.17) that $q_{s_0} \gg 0$. Thus, $-bu(s_0)q_{s_0} < 0$ and, thanks again to Proposition 2.1, the first equation of (7.17) gives $p_{s_0} \ll 0$, which is impossible. Therefore, for each $s \in (s_1^*, s_2^*)$, $p_s \ll 0$ and $q_s \gg 0$, which concludes the proof of (7.23). The proof of the theorem is completed. □

7.4. *Some bifurcation diagrams*

In order to study the curves $\mathfrak{C}_\lambda^\mu$ we did carried out some numerical investigations in [12], where we solved (1.2) using spectral collocation methods coupled with path-following techniques. To compute T-periodic solutions of (1.2), the most appropriate set of basis functions to use is trigonometric polynomials, and these were used to generate the bifurcation diagrams of Figs. 17 and 18. Fig. 17 shows the approximated components $\mathfrak{C}_\lambda^\mu$ for the following problems, with $T = 1$ and $a = d = 1$,

(a) $\mu = 3.32$, $m(t) = 1 + 0.7\sin(2\pi t)$, $\ell(t) = 1 + 0.7\cos(2\pi t)$, $b(t) = 0.9 + 0.7\sin(2\pi t + \pi/4)$, $c(t) = 0.9 + 0.7\cos(2\pi t + \pi/4)$.

(b) $\mu = 3.32$, $m(t) = 1 + 0.7\sin(2\pi t)$, $\ell(t) = 1 + 0.7\cos(2\pi t)$, $b(t) = 0.9 + 0.7\sin(2\pi t)$, $c(t) = 0.9 + 0.7\cos(2\pi t)$.

(c) $\mu = 5.00$, $\ell(t) = m(t) = 1 + 0.7\sin(2\pi t)$, $b(t) = c(t) = 0.9 + 0.7\sin(2\pi t)$.

(d) $\mu = 3.32$, $\ell(t) = m(t) = 1 - 0.7\sin(2\pi t)$, $b(t) = c(t) = 1.1 + 0.7\sin(2\pi t)$.

Given a $\mu > 0$, we have plotted the parameter λ along the horizontal axis, and along the vertical one we give the value $\frac{\|u\|_2}{2} + \|v\|_2$, where

$$\|f\|_2 = \sqrt{\sum_{j=1}^{\infty} c_j^2} \quad \text{if} \quad f = c_1 + \sum_{j=1}^{\infty} \left[c_{2j}\sin(2\pi jt) + c_{2j+1}\cos(2\pi jt)\right].$$

This particular choice of norm makes it easy to distinguish between the semi-trivial states $(\lambda, \mu, \theta_{[\lambda\ell,a]}, 0)$ and $(\lambda, \mu, 0, \theta_{[\mu m,d]})$, which would otherwise cross near the bifurcation points and confuse the picture. Stable solutions are represented by continuous lines, unstable ones by dashed lines. Several different cases are shown according to the nature of the

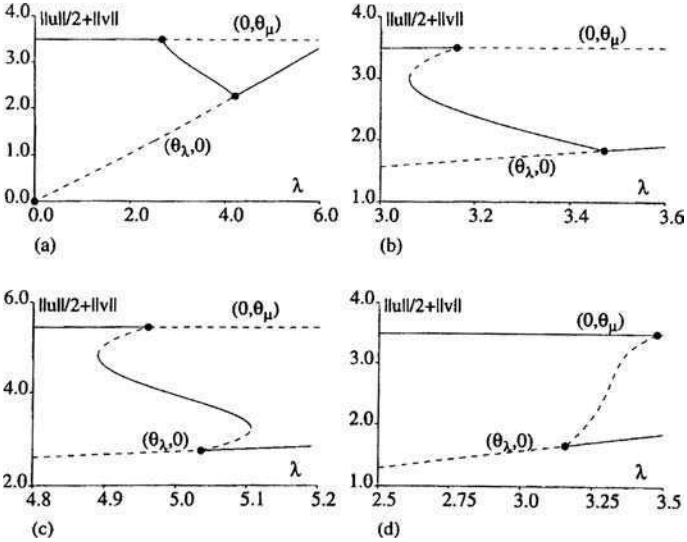


Fig. 17. Bifurcation diagrams for $\mu > 0$ constant.

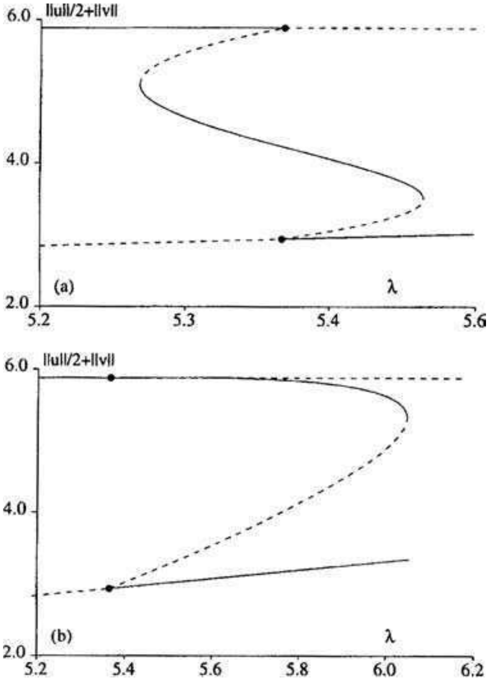


Fig. 18. Two examples with both semi-trivial states neutrally stable.

bifurcation to coexistence states from the semi-trivial solutions. In all cases, $\mathfrak{C}_\lambda^\mu$ possesses at most a finite number of hysteresis and turning points and the stability of the coexistence states change according to Theorem 7.7. Fig. 18 shows some of the possible situations when both semi-trivial states are neutrally stable. In some cases, (1.2) possesses a coexistence states, while in others it does not. Fig. 18 shows the following two cases, with $T=1$ and $a=d=1$:

(a) $\mu = 5.365$, $m(t) = 1 + 0.7\sin(2\pi t)$, $\ell(t) = 1 + 0.7\cos(2\pi t)$, $b(t) = c(t) = 0.9 + 0.7\sin(2\pi t)$.

(b) $\mu = 5.365$, $m(t) = 1 + 0.7\sin(2\pi t)$, $\ell(t) = 1 + 0.7\cos(2\pi t)$, $b(t) = 0.9 + 0.7\sin(2\pi t)$, $c(t) = 0.9 + 0.7\cos(2\pi t)$.

Case (a) provides us with an example where both semi-trivial states are neutrally stable and (1.2) does not admit a coexistence state. It should be noted that, thanks to the discussion closing Section 7.3, (1.2) cannot admit a further coexistence state outside $\mathfrak{C}_\lambda^\mu$. Case (b) provides us with an example where both semi-trivial states are neutrally stable and (1.2) possesses a unique coexistence state, which is stable.

7.5. *Dynamics of* (1.1)

Thanks to the theory developed by P. de Mottoni and A. Schiaffino [29], all solutions of (1.1) must approach a T-periodic solution of (1.2) as time passes by. Combining this information with the theory developed in this section, one can easily obtain the validity of the following assertions. Subsequently, we shall denote by $(u(t), v(t))$ the unique positive solution of (1.1).

1. If $\lambda \leqslant 0$ and $\mu \leqslant 0$, then $\lim_{t\uparrow\infty} (u(t), v(t)) = (0,0)$.
2. If $\lambda > 0$ and $\mu \leqslant 0$, then

$$\lim_{t\uparrow\infty} \big(u(t), v(t)\big) = (\theta_{[\lambda\ell,a]}, 0). \tag{7.33}$$

3. If $\lambda \leqslant 0$ and $\mu > 0$, then

$$\lim_{t\uparrow\infty} \big(u(t), v(t)\big) = (0, \theta_{[\mu m,d]}). \tag{7.34}$$

4. If

$$\lambda > 0, \quad 0 < \mu < \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t, \tag{7.35}$$

   and (1.2) does not admit a coexistence state, then, (7.33) holds true.
5. If (7.35) is satisfied, and (1.2) has a coexistence state, then, either (7.33) holds true, or there exists a coexistence state (u_0, v_0) of (1.2) such that

$$\lim_{t\uparrow\infty} \big(u(t), v(t)\big) = (u_0, v_0). \tag{7.36}$$

Moreover, in this case the number of stable coexistence states of (1.2) equals the number of unstable coexistence states, if there is some.

6. If

$$\mu > 0, \quad 0 < \lambda < \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t, \tag{7.37}$$

and (1.2) does not admit a coexistence state, then, (7.34) holds true.

7. If (7.37) is satisfied and (1.2) has a coexistence state, then, either (7.34) holds true, or there exists a coexistence state (u_0, v_0) of (1.2) satisfying (7.36). Moreover, in this case the number of stable coexistence states of (1.2) equals the number of unstable coexistence states, if there is some.
8. If

$$\lambda > \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t, \quad \mu > \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t, \tag{7.38}$$

then (1.2) possesses a stable coexistence state. Moreover, if $n \geqslant 1$ stands for its number of stable coexistence states, then (1.2) possesses $n-1$ unstable coexistence states and a finite number of neutrally stable coexistence states. Furthermore, there exists a coexistence state (u_0, v_0) for which (7.36) holds true.

9. If

$$0 < \lambda < \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t, \quad 0 < \mu < \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t,$$

then (1.2) possesses an unstable coexistence state. Moreover, if $n \geqslant 1$ stands for its number of unstable coexistence states, then (1.2) possesses $n-1$ stable coexistence states and a finite number of neutrally stable coexistence states. Furthermore, either there exists a coexistence state (u_0, v_0) for which (7.36) holds, or either (7.33), or (7.34), holds, according to the values of (x, y).

8. The species v preys on u ($b > 0$ and $c < 0$)

Now, the linearization of (1.2) around any coexistence state can be expressed in the form

$$\begin{cases} u' = \alpha u - \beta v, \\ v' = \gamma u + \rho v, \end{cases} \tag{8.1}$$

where $\alpha, \beta, \gamma, \rho \in \mathcal{C}_T$ satisfy

$$\int_0^T \alpha < 0, \quad \int_0^T \rho < 0, \quad \beta > 0, \quad \gamma > 0, \tag{8.2}$$

which is a non-cooperative periodic system. As a result, the nice monotonicity properties available for competing species and symbiotic models are lost, and, in general, the components of coexistence states $\mathfrak{C}^{\lambda}_{\mu}$ and $\mathfrak{C}^{\mu}_{\lambda}$ bifurcating from the semi-trivial states are far from being analytic curves, except in a neighborhood of their bifurcation points from the semi-trivial states. Nevertheless, in strong contrast with the classical cases of competition and symbiosis, in this model the existence of a coexistence state is characterized by the linear stability of the semi-trivial states.

8.1. *Characterizing the existence of coexistence states*

Subsequently, we suppose $m \gg 0$ and $d \gg 0$. Then,

$$\lim_{\mu\uparrow\infty} \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t = \infty,$$

since

$$\lim_{\mu\uparrow\infty} \frac{\theta_{[\mu m,d]}}{\mu} = \frac{m}{d},$$

and, hence, for each $\lambda > 0$ there exists a unique value of μ, $\mu_* > 0$, such that

$$\lambda = \frac{1}{T}\int_0^T b(t)\theta_{[\mu_* m,d]}(t)\,\mathrm{d}t. \tag{8.3}$$

The following result is satisfied.

THEOREM 8.1. *The following conditions are equivalent*:

(a) *Problem* (1.2) *possesses a coexistence state.*

(b) $\lambda > 0$ *and any semi-trivial state of* (1.2) *is linearly unstable.*

(c) *The parameters* (λ, μ) *satisfy*

$$\lambda > \max\left\{0, \frac{1}{T}\int_0^T (b\theta_{[\mu m,d]})\right\}, \quad \mu > \frac{1}{T}\int_0^T (c\theta_{[\lambda\ell,a]}), \tag{8.4}$$

where, for each $\mu \in \mathbb{R}$, $\theta_{[\mu m,d]} \in \mathcal{C}^1_T$ *stands for the maximal non-negative solution of* $u' = \lambda\ell u - au^2$. *In other words,* $\theta_{[\mu m,d]} \gg 0$ *if* $\mu > 0$, *while* $\theta_{[\mu m,d]} = 0$ *if* $\mu \leqslant 0$.

Moreover, for each $\lambda > 0$, *the component* $\mathfrak{C}^{\lambda}_{\mu}$ *of coexistence states constructed in Theorem* 5.4 *is bounded in* $\mathbb{R} \times \mathcal{C}_T \times \mathcal{C}_T$ *and it satisfies*

$$(\lambda, \mu_0, \theta_{[\lambda\ell,a]}, 0), (\lambda, \mu_*, 0, \theta_{[\mu_* m,d]}) \in \bar{\mathfrak{C}}^{\lambda}_{\mu}, \tag{8.5}$$

where μ_* *is the unique value of* μ *satisfying* (8.3). *Therefore, the* μ*-projection of* $\mathfrak{C}^{\lambda}_{\mu}$ *is the interval* (μ_0, μ_*).

PROOF. Suppose (λ, μ, u_0, v_0) is a coexistence state of (1.2). Then, dividing by u_0 the u-equation and integrating in $[0, T]$ gives

$$\lambda = \frac{1}{T}\int_0^T (au_0 + bv_0) > 0.$$

So, $\lambda > 0$ is necessary for the existence. Subsequently, we will assume that $\lambda > 0$. Then, thanks again to the u-equation,

$$u_0' = \lambda \ell u_0 - au_0^2 - bu_0 v_0 < \lambda \ell u_0 - au_0^2,$$

and, due to Proposition 2.6,

$$u_0 \ll \theta_{[\lambda\ell, a]}. \tag{8.6}$$

So, dividing by v_0 the v-equation, integrating in $[0, T]$, and using (8.6) shows that

$$\mu = \frac{1}{T}\int_0^T (cu_0 + dv_0) > \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t,$$

since $c < 0$. Thus, the second estimate of (8.4) is necessary for the existence.

Now, suppose $\mu > 0$. Then, $\theta_{[\mu m, d]} \gg 0$,

$$v_0' = \mu m v_0 - dv_0^2 - cu_0 v_0 > \mu m v_0 - dv_0^2,$$

and, it is apparent from Proposition 2.6 that

$$v_0 \gg \theta_{[\mu m, d]}. \tag{8.7}$$

Consequently,

$$\lambda = \frac{1}{T}\int_0^T (au_0 + bv_0) > \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t$$

and, therefore, (8.4) is necessary for the existence of a coexistence state. It should be noted that (8.4) provides us with the set of values of (λ, μ) with $\lambda > 0$ for which any semi-trivial state of (1.2) is linearly unstable, i.e., (b) and (c) are equivalent and we have just seen that (a) implies (c).

Now, fix $\lambda > 0$ and pick $(\mu, u_0, v_0) \in \mathfrak{C}_\mu^\lambda$. Then, (8.7) is satisfied. Moreover, at any time t_0 where $v_0(t_0) = (v_0)_M$, we have that $v_0'(t_0) = 0$ and, hence, it follows from the v-equation that

$$\mu m(t_0) - d(t_0)v_0(t_0) - c(t_0)u_0(t_0) = 0.$$

Thus, thanks to (8.6),

$$\begin{aligned}\|v_0\|_{\mathcal{C}_T} = v_0(t_0) &= \frac{m(t_0)}{d(t_0)}\mu + \frac{-c(t_0)}{d(t_0)}u(t_0)\\ &\leqslant \left(\frac{m}{d}\right)_M |\mu| + \left(\frac{-c}{d}\right)_M \theta_{[\lambda\ell,a]},\end{aligned}$$

and, consequently, $\mathfrak{C}^{\lambda}_{\mu}$ is bounded. Now, it is apparent that Theorem 5.4 concludes the proof. Indeed, the μ-projection of $\mathfrak{C}^{\lambda}_{\mu}$ must be the interval (μ_0, μ_*) and, therefore, (8.4) is sufficient for the existence of a coexistence state of (1.2). □

8.2. *Dynamics of* (1.1)

Thanks to the theory developed in [25], the following results, whose proofs are omitted here, are satisfied. Subsequently, we shall denote by $(u(t), v(t))$ the unique positive solution of (1.1).

1. If $(0,0)$ is linearly stable, i.e., if $\lambda \leqslant 0$ and $\mu \leqslant 0$, then $\lim_{t\uparrow\infty}(u(t), v(t)) = (0,0)$.
2. If $\lambda > 0$ and $(\theta_{[\lambda\ell,a]}, 0)$ is linearly stable, i.e., $\mu \leqslant \frac{1}{T}\int_0^T c(t)\theta_{[\lambda\ell,a]}(t)\,\mathrm{d}t$, then $\lim_{t\uparrow\infty}(u(t), v(t)) = (\theta_{[\lambda\ell,a]}, 0)$.
3. If $\mu > 0$ and $(0, \theta_{[\mu m,d]})$ is linearly stable, i.e., $\lambda \leqslant \frac{1}{T}\int_0^T b(t)\theta_{[\mu m,d]}(t)\,\mathrm{d}t$, then $\lim_{t\uparrow\infty}(u(t), v(t)) = (0, \theta_{[\mu m,d]})$.

As far as to the problem of the uniqueness, multiplicity and stability of the coexistence states things are much more difficult to deal with than in the competing species and symbiotic prototype counterparts, since the linearization of (1.2) at any coexistence state has a *non-cooperative structure*—the off-diagonal entries of the coupling matrix of the system have contrary signs—and for such problems a general comparison principle within the spirit of Theorems 6.1 and 7.1 is far from being available. Therefore, the problem of ascertaining the signs of the Floquet exponents of a given coexistence state is extremely difficult to treat with.

Quite strikingly, although the one-dimensional predator–prey model with diffusion under homogeneous Dirichlet boundary conditions is known to admit at most a coexistence state [26,10,18], there are choices of the several coefficient functions in the setting of (1.2) for which the predator–prey model admits two $2T$-periodic coexistence states and one unstable T-periodic coexistence state [5,8,18]. Hence, the uniqueness result for the one-dimensional elliptic model does not extend in general to the periodic prototype models, however both prototypes, viewed as abstract non-cooperative systems, possess the same operator structure [18]. In [25] were obtained some persistence results in the presence of coexistence states, and in [5] it was shown that if the product of the interactions b and c is sufficiently small, then the model admits at most a coexistence state, though we refrain of giving more details here. Actually, this uniqueness principle holds true in the general case when b and c are arbitrary. The details will appear elsewhere.

References

[1] S. Ahmad, *On the non-autonomous Volterra–Lotka competition equations*, Proc. Amer. Math. Soc. **117** (1993), 199–204.

[2] S. Ahmad and A.C. Lazer, *Asymptotic behaviour of solutions of periodic competition diffusion systems*, Nonlinear Anal. **13** (1989), 263–284.

[3] C. Alvárez and A.C. Lazer, *An application of topological degree to the periodic competing species problem*, J. Austral. Math. Soc. Ser. B **28** (1986), 202–219.

[4] H. Amann, *Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces*, SIAM Rev. **18** (1976), 620–709.

[5] Z. Amine and R. Ortega, *A periodic prey-predator system*, J. Math. Anal. Appl. **185** (1994), 477–489.

[6] M.G. Crandall and P.H. Rabinowitz, *Bifurcation from simple eigenvalues*, J. Funct. Anal. **8** (1971), 321–340.

[7] J.M. Cushing, *Two species competition in a periodic environment*, J. Math. Biol. **10** (1980), 385–390.

[8] E.N. Dancer, *Turing instabilities for systems of two equations with periodic coefficients*, Differential Integral Equations **7** (1994), 1253–1264.

[9] E.N. Dancer, *Bifurcation from simple eigenvalues and eigenvalues of geometric multiplicity one*, Bull. London Math. Soc. **34** (2002), 533–538.

[10] E.N. Dancer, J. López-Gómez and R. Ortega, *On the spectrum of some linear noncooperative elliptic systems with radial symmetry*, Differential Integral Equations **8** (1995), 515–523.

[11] M. Delgado, J. López-Gómez and A. Suárez, *On the symbiotic Lotka–Volterra model with diffusion and transport effects*, J. Differential Equations **160** (2000), 175–262.

[12] J.C. Eilbeck and J. López-Gómez, *On the periodic Lotka–Volterra competition model*, J. Math. Anal. Appl. **210** (1997), 58–87.

[13] J.E. Furter, J.C. Eilbeck and J. López-Gómez, *Coexistence in the competition model with diffusion*, J. Differential Equations **107** (1994), 96–139.

[14] J.K. Hale and A. Somolinos, *Competition for fluctuating nutrient*, J. Math. Biol. **18** (1983), 255–280.

[15] P. Hess and A.C. Lazer, *On an abstract competition model and applications*, Nonlinear Anal. **16** (1991), 917–940.

[16] M.A. Krasnosel'skii, *Translations Along Trajectories of Differential Equations*, Math. Monogr. of the AMS, Amer. Math. Soc., Providence, RI (1968).

[17] J. López-Gómez, *Structure of positive solutions in the periodic Lotka–Volterra competition model when a parameter varies*, Nonlinear Anal. **27** (1996), 739–768.

[18] J. López-Gómez, *A bridge between Operator Theory and Mathematical Biology*, Fields Inst. Commun. **25** (2000), 383–397.

[19] J. López-Gómez, *Spectral Theory and Nonlinear Functional Analysis*, Pitman Res. Notes Math. Ser., Vol. 426, Chapman & Hall/CRC, Boca Raton, FL (2001).

[20] J. López-Gómez, *Classifying smooth supersolutions in a general class of elliptic boundary value problems*, Adv. Differential Equations **8** (2003), 1025–1042.

[21] J. López-Gómez and M. Molina-Meyer, *The maximum principle for cooperative weakly coupled elliptic systems and some applications*, Differential Integral Equations **7** (1994), 383–398.

[22] J. López-Gómez and M. Molina-Meyer, *Singular perturbations in Economy and Ecology. The effect of strategic symbiosis in random competitive environments*, Adv. Math. Sci. Appl. **14** (2004), 87–107.

[23] J. López-Gómez and M. Molina-Meyer, *Symbiosis in Competitive environments*, Res. Inst. Math. Sci. Kyoto **12** (2003), 118–141.

[24] J. López-Gómez and M. Molina-Meyer, *Bounded components of positive solutions of abstract fixed point equations: mushrooms, loops and isolas*, J. Differential Equations **209** (2005) 416–441.

[25] J. López-Gómez, R. Ortega and A. Tineo, *The periodic predator–prey Lotka–Volterra model*, Adv. Differential Equations **1** (1996), 403–423.

[26] J. López-Gómez and R. Pardo, *Existence and uniqueness for the predator-prey model with diffusion: The scalar case*, Differential Integral Equations **6** (1993), 1025–1031.

[27] J. Mawhin, *Les héritiers de Pierre-François Verhulst: une population dynamique*, Acad. Roy. Belg. Bull. Cl. Sci. (6) **13** (2003), 349–378.

[28] J. Mawhin, *The legacy of P.F. Verhulst and V. Volterra in population dynamics*, The First 60 Years of Nonlinear Analysis of Jean Mawhin, M. Delgado, J. López-Gómez, R. Ortega and A. Suárez, eds., World-Scientific (2004).

[29] P. de Mottoni and A. Schiaffino, *Competition systems with periodic coefficients: a geometric approach*, J. Math. Biol. **11** (1981), 319–335.

[30] R. Ortega and A. Tineo, *On the number of positive solutions for planar competing Lotka–Volterra systems*, J. Math. Anal. Appl. **193** (19945), 975–978.

[31] P.H. Rabinowitz, *Some global results for nonlinear eigenvalue problems*, J. Funct. Anal. **7** (1971), 487–513.

[32] H. Smith, *Periodic competitive differential equations and the discrete dynamics of competitive maps*, J. Differential Equations **64** (1986), 165–194.

[33] H. Smith, *Periodic solutions of periodic competitive and cooperative systems*, SIAM J. Math. Anal. **17** (1986), 1289–1238.

[34] A. Tineo, *Existence of global coexistence state for periodic competition diffusion systems*, Nonlinear Anal. **19** 81992), 335–344.

[35] A. Tineo, *An iterative scheme for the N-competing species problem*, J. Differential Equations (1995), 1–15.

[36] F. Zanolin, *Permanence and positive periodic solutions of Kolmogorov competing species models*, Results Math. **21** (1992), 224–250.

CHAPTER 6

Nonlocal Initial and Boundary Value Problems: A Survey

S.K. Ntouyas

Department of Mathematics, University of Ioannina, 451 10 Ioannina, Greece
E-mail: *sntouyas@cc.uoi.gr*

Contents

1. Introduction . . . 463
   1.1. Examples of nonlocal conditions . . . 463
2. Nonlocal initial value problems . . . 465
   2.1. Existence and uniqueness . . . 465
3. Nonlocal boundary value problems . . . 469
   3.1. First-order three-point boundary value problems . . . 470
   3.2. Second-order three-point boundary value problems . . . 475
   3.3. Third-order three-point boundary value problems . . . 484
   3.4. m-point boundary value problems reduced to three-point boundary value problems . . . 489
   3.5. Nonresonance results . . . 490
   3.6. Results at resonance . . . 498
   3.7. m-point boundary value problems reduced to four-point boundary value problems . . . 505
   3.8. The general case . . . 513
   3.9. Positive solutions of some nonlocal boundary value problems . . . 518
   3.10. Positive solutions of nonlocal boundary value problems with dependence on the first-order derivative . . . 542
   3.11. Positive solutions of nonlocal problems for p-Laplacian . . . 547
References . . . 554

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 2
Edited by A. Cañada, P. Drábek and A. Fonda
© 2005 Elsevier B.V. All rights reserved

1. Introduction

The evolution of a physical system in time is described by an initial value problem, i.e. a differential equation (ordinary or partial) and an initial condition. In many cases it is better to have more initial information. The local condition is replaced then by a nonlocal condition, which gives better effect than the local initial condition, since the measurement given by a nonlocal condition is usually more precise than the only one measurement given by a local condition.

The study of initial value problems (IVP for short) with nonlocal conditions is of significance, since they have applications in problems in physics and other areas of applied mathematics. Conditions of this type can be applied in the theory of elasticity with better effect than the initial or Darboux conditions.

Consider the differential equation

$$x'(t) = f\big(t, x(t)\big), \quad x(t_0) = x_0, \quad t \geqslant 0,\ x(t) \in \mathbb{R}^n. \tag{1.1}$$

The fundamental assumption used when modeling a system using a differential equation is that the time rate at time t, given as $x'(t)$, depends only on the current status at time t, given as $f(t, x(t))$. Moreover, the initial condition is given in the form $x(t_0) = x_0$. In applications, this assumption and the initial condition should be improved so that we can model the situations more accurately and therefore derive better results.

One improvement of Eq. (1.1) is to assume that the time rate depends not only on the current status, but also on the status in the past; that is, the past history will contribute to the future development, or, there is a *time-delay* effect.

Another improvement of Eq. (1.1) is the so-called *differential equations with nonlocal conditions*. That is, we extend the initial condition (also called *local condition*)

$$x(t_0) = x_0$$

to the following *nonlocal condition*

$$x(t_0) + g\big(x(\cdot)\big) = x_0, \tag{1.2}$$

where $x(\cdot)$ denotes a solution (that is, $x(\cdot)$ is a function) and g is a mapping defined on some space consisting of certain functions. (Of course, g may be identically zero, in which case it reduces to the local condition or initial condition $x(t_0) = x_0$.) The advantage of using nonlocal conditions is that measurements at more places can be incorporated to get better models.

1.1. *Examples of nonlocal conditions*

Let $p \in \mathbb{N}$ and let $t_1, t_2, \ldots, t_p$ be given real numbers such that

$$t_0 < t_1 < t_2 < \cdots < t_p \leqslant t_0 + \alpha.$$

Then g can be defined by the formula

$$g(x)=\sum_{i=1}^{p} c_i x(t_i), \tag{1.3}$$

where c_i, $i=1,2,\dots,p$ are given constants or by the relation

$$g(x)=\sum_{i=1}^{p} \frac{c_i}{\epsilon_i}\int_{t_i-\epsilon_i}^{t_i} x(t)\,\mathrm{d}t, \tag{1.4}$$

where ϵ_i, $i=1,2,\dots,p$ are given positive constants such that: $t_{i-1}<t_i-\epsilon_i<t_i$, $i=1,2,\dots,p$. Particularly, if $x_0=0$, $p=1$, $t_1=t_0+\alpha$ then

$$u(t_0)=-c_1 x(t_0+\alpha) \quad \text{and} \quad x(t_0)=\frac{-c_1}{\epsilon_1}\int_{t_0+\alpha-\epsilon_1}^{t_0+\alpha} x(t)\,\mathrm{d}t$$

respectively.

The constants c_i, $i=1,2,\dots,p$ in the nonlocal condition (1.3) can satisfy the inequalities $|c_i|>1$, $i=1,2,\dots,p$. It is remarkable that if $c_i\neq 0$, $i=1,2,\dots,p$ the results can be applied to kinematics to determine the evolution $t\to x(t)$ of the location of a physical object for which we do not know the positions $x(0), x(t_1),\dots,x(t_p)$, but we know that the nonlocal condition of (1.3) holds. Consequently, to describe some physical phenomena, the nonlocal condition can be more useful than the standard initial condition $x(0)=x_0$. From the nonlocal condition of (1.3) it is clear that when $c_i=0$, $i=1,2,\dots,p$ we have the classical initial condition.

Let us have one more example for partial differential equations. In the theory of diffusion and heat conduction one can encounter a mathematical model of the form:

$$\begin{aligned}
&Lx+c(w,t)x=f(w,t), \quad t\in\Omega,\ 0<t<T,\\
&x(w,t)=\phi(w,t), \quad w\in\partial\Omega,\ 0<t<T,\\
&x(w,0)+\sum_{k=1}^{N}\beta_k(w)x(w,t_k)=\psi(w), \quad w\in\Omega \text{ and } t_k\in(0,T],\ k=1,\dots,N,
\end{aligned}$$

where Ω is a bounded region in $\mathbb{R}^n$ and L is a uniformly parabolic operator with continuous and bounded coefficients. It represents the diffusion phenomenon of a small amount of gas in a transparent tube. If there is very little gas at the initial time, the measurement $x(w,0)$ of the amount of the gas in this instant may be less precise that the measurement $x(w,0)+\sum_{k=1}^{N}\beta_k(x)x(w,t_k)$ of the sum of the amount of this gas.

2. Nonlocal initial value problems

Here, we consider the following initial value problem

$$x' = f(t,x), \quad t \in I := [t_0, T],\ x \in \mathbb{R}^n, \tag{2.1}$$

$$x(t_0) + g\big(x(\cdot)\big) = x_0, \tag{2.2}$$

where $x(\cdot)$ denotes a solution and g is a mapping acting on some space of functions defined on $[0, T]$.

Since $g(x(\cdot))$ in Eq. (2.2) is defined on the interval $[t_0, T]$ rather than a single point, the IVP (2.1)–(2.2) is called an "*initial value problem with nonlocal conditions*" or "*nonlocal initial value problem*", and can be applied with better effect than just using the classical initial value problem with $x(t_0) = x_0$, because new measurements at more places are allowed, thus more information is available.

For the IVP (2.1)–(2.2), observe that if x is a solution, then

$$x(t) = x(t_0) + \int_{t_0}^{t} f\big(s, x(s)\big)\,\mathrm{d}s = \big[x_0 - g\big(x(\cdot)\big)\big] + \int_{t_0}^{t} f\big(s, x(s)\big)\,\mathrm{d}s. \tag{2.3}$$

On the other hand, assume that a function x on the interval $[t_0, T]$ satisfies (2.3). For this fixed function x, $g(x(\cdot))$ is a fixed element in $\mathbb{R}^n$, hence $\frac{\mathrm{d}}{\mathrm{d}t} g(x(\cdot)) = 0$. Therefore, if we take a derivative in t, then x is a solution of IVP (2.1)–(2.2). That is, we conclude that a continuous function x is a solution of the IVP (2.1)–(2.2) if and only if

$$x(t) = \big[x_0 - g\big(x(\cdot)\big)\big] + \int_{t_0}^{t} f\big(s, x(s)\big)\,\mathrm{d}s.$$

This leads to a mapping P on $C([t_0, T], \mathbb{R}^n)$ such that

$$(Px)(t) = \big[x_0 - g\big(x(\cdot)\big)\big] + \int_{t_0}^{t} f\big(s, x(s)\big)\,\mathrm{d}s.$$

2.1. *Existence and uniqueness*

Motivated by physical problems, Byszewski and Lakshmikantham [12] considered the following nonlocal IVP

$$x' = f(t,x), \quad t \in I := [t_0, b], \tag{2.4}$$

$$x(t_0) + g\big(t_1, t_2, \ldots, t_p, x(\cdot)\big) = x_0, \tag{2.5}$$

where $t_0 < t_1 < \cdots < t_p < b$, $p \in \mathbb{N}$, $x = (x_1, x_2, \ldots, x_n) \in \Omega$, $x_0 = (x_{10}, \ldots, x_{n0}) \in \Omega$, $f = (f_1, \ldots, f_n) \in C(I \times \Omega, E)$, $g = (g_1, \ldots, g_n) : I^p \times \Omega \to E$, $g(t_1, \ldots, t_p, \cdot) \in$

$C(\Omega, E)$ and $\Omega \subset E$ is a suitable subset. The symbol $g(t_1, \ldots, t_p, x(\cdot))$ is used in the sense that in the place of $\cdot$ we can substitute only elements of the set $\{t_1, \ldots, t_p\}$.

Let $E = E_1 \times E_2 \times \cdots \times E_n$, where E_i, $i = 1, 2, \ldots, n$ are Banach spaces with norm $\|\cdot\|$.

THEOREM 2.1. *Assume that*:

(i) *E is a Banach space with norm $\|\cdot\|$, $x_0 \in E$, and $\Omega := \overline{B(x_0, r)} = \{y\colon \|y - x_0\| \leqslant r\} \subset E$.*

(ii) *There exist constants $L_i > 0$, $i = 1, \ldots, n$ such that*

$$\bigl\| f_i(s, y) - f_i(s, \bar{y}) \bigr\| \leqslant L_i \|y - \bar{y}\|$$

for (s, y), $(s, \bar{y}) \in I \times \Omega$, $i = 1, \ldots, n$.

(iii) *There exist constants $K_i > 0$, $i = 1, \ldots, n$ such that*

$$\bigl\| g_i(t_1, \ldots, t_p, z) - g_i(t_1, \ldots, t_p, \bar{z}) \bigr\| \leqslant K_i \|z - \bar{z}\|, \quad z, \bar{z} \in \Omega,\ i = 1, \ldots, n.$$

(iv) $$(Md + N)\sqrt{n} \leqslant r, \qquad (Ld + K)\sqrt{n} < 1,$$

where

$$M_i = \sup_{(s,y) \in I \times \Omega} \bigl\| f_i(s, y) \bigr\|, \quad N_i = \sup_{z \in \Omega} \bigl\| g(t_1, \ldots, t_p, z) \bigr\|, \quad i = 1, \ldots, n,$$

and

$$d = b - t_0, \quad L = \max_{i=1,\ldots,n} L_i, \quad K = \max_{i=1,\ldots,n} K_i,$$

$$M = \max_{i=1,\ldots,n} M_i, \quad N = \max_{i=1,\ldots,n} N_i.$$

Then there exists a unique solution on the nonlocal IVP (2.4)–(2.5).

PROOF. It is easy to see, that problem (2.4)–(2.5) is equivalent to the integral equation

$$x(t) = x_0 - g\bigl(t_1, \ldots, t_p, x(\cdot)\bigr) + \int_{t_0}^{t} f\bigl(s, x(s)\bigr)\,\mathrm{d}s, \quad t \in I,$$

which is also equivalent to the following system of the integral equations

$$x_i(t) = x_{i0} - g_i\bigl(t_1, \ldots, t_p, x(\cdot)\bigr) + \int_{t_0}^{t} f_i\bigl(s, x(s)\bigr)\,\mathrm{d}s, \quad t \in I,\ i = 1, \ldots, n.$$

We introduce the operator T given by

$$(Ty)(t) := \bigl((T_1 y)(t), \ldots, (T_n y)(t)\bigr), \quad t \in I,$$

where

$$(T_i y)(t) = x_{i0} - g_i\bigl(t_1, \ldots, t_p, x(\cdot)\bigr) + \int_{t_0}^{t} f_i\bigl(s, x(s)\bigr)\,\mathrm{d}s, \quad t \in I,\ i = 1, \ldots, n.$$

Let $X := C(I, \Omega)$. We will show that the operator T maps X into X. We observe that

$$\begin{aligned}
\|Ty - x_0\| &= \left(\sum_{i=1}^{n} \|T_i y - x_{i0}\|^2\right)^{1/2} \\
&= \left(\sum_{i=1}^{n} \Bigl[\sup_{t\in I} \bigl\|(T_i y)(t) - x_{i0}\bigr\|\Bigr]^2\right)^{1/2} \\
&\leqslant \left(\sum_{i=1}^{n} \biggl[\sup_{t\in I} \int_{t_0}^{t} \bigl\|f_i\bigl(s, y(s)\bigr)\bigr\|\,\mathrm{d}s + \bigl\|g_i\bigl(t_1, \ldots, t_p, y(\cdot)\bigr)\bigr\|\biggr]^2\right)^{1/2} \\
&\leqslant \left(\sum_{i=1}^{n} \Bigl[\sup_{t\in I} \bigl(|t - t_0| M_i + N_i\bigr)\Bigr]^2\right)^{1/2} \\
&\leqslant (Md + N)\sqrt{n} \leqslant r,
\end{aligned}$$

for $y \in X$. Therefore $T : X \to X$.

Next, we will show that T is a contraction on X. We have:

$$\begin{aligned}
\|Ty - T\bar{y}\| &= \left(\sum_{i=1}^{n} \|T_i y - T_i \bar{y}\|^2\right)^{1/2} \\
&= \left(\sum_{i=1}^{n} \Bigl[\sup_{t\in I} \bigl\|(T_i y)(t) - (T_i \bar{y})(t)\bigr\|\Bigr]^2\right)^{1/2} \\
&\leqslant \left(\sum_{i=1}^{n} \biggl[\sup_{t\in I} \int_{t_0}^{t} \bigl\|f_i\bigl(s, y(s)\bigr) - f_i\bigl(s, \bar{y}(s)\bigr)\bigr\|\,\mathrm{d}s\right. \\
&\qquad \left. + \bigl\|g_i\bigl(t_1, \ldots, t_p, y(\cdot)\bigr) - g_i\bigl(t_1, \ldots, t_p, \bar{y}(\cdot)\bigr)\bigr\|\biggr]^2\right)^{1/2} \\
&\leqslant \left(\sum_{i=1}^{n} \Bigl[\sup_{t\in I} \bigl(|t - t_0| L_i \bigl\|y(s) - \bar{y}(s)\bigr\| + K_i \bigl\|y(\cdot) - \bar{y}(\cdot)\bigr\|\bigr)\Bigr]^2\right)^{1/2} \\
&\leqslant (Ld + K)\sqrt{n}\|y - \bar{y}\|
\end{aligned}$$

for arbitrary $y, \bar{y} \in X$ which proves that T is a contraction. Hence, by Banach fixed point theorem, in space X there is the only one fixed point of T and this point is the solution of the nonlocal problem (2.4)–(2.5). So, the proof of the theorem is complete. $\square$

Consider now a nonlocal IVP with a special form of the nonlocal initial condition,

$$x'(t) = f\bigl(t, x(t)\bigr), \quad t \in J := [0, T], \tag{2.6}$$

$$x(0) + \sum_{k=1}^{p} c_k x(t_k) = x_0, \tag{2.7}$$

where $0 < t_1 < \cdots < t_p \leqslant T$, $f : J \times E \to E$ is a given function and E is a Banach space with norm $\|\cdot\|$, $x_0 \in E$, $c_k \neq 0$ and $p \in \mathbb{N}$.

By X we denote the Banach space $C(J, E)$ with the standard norm $\|\cdot\|_X$.

DEFINITION 2.2. Assume that $\sum_{k=1}^{p} c_k \neq -1$. A function $x \in X$ is called a mild solution of (2.6)–(2.7) if satisfies the integral equation

$$x(t) = A\left(x_0 - \sum_{k=1}^{p} c_k \int_0^{t_k} f\bigl(s, x(s)\bigr)\,\mathrm{d}s\right) + \int_0^t f\bigl(s, x(s)\bigr)\,\mathrm{d}s,$$

where $A = (1 + \sum_{k=1}^{p} c_k)^{-1}$.

DEFINITION 2.3. A function $x : J \to E$ is said to be a classical solution of the nonlocal problem (2.6)–(2.7) if

(i) x is continuous on J and continuously differentiable on J,
(ii) $x'(t) = f(t, x(t))$ for $t \in J$, and
(iii) $x(0) + \sum_{k=1}^{p} c_k x(t_k) = x_0$.

It is easy to see that if $f \in C(J \times E, E)$ and $\sum_{k=1}^{p} c_k \neq -1$, then x is the unique classical solution of the nonlocal problem (2.6)–(2.7) if and only if x is the unique mild solution of this problem.

THEOREM 2.4. *Assume that*:

(1) $f : J \times E \to E$ *is continuous with respect to the first variable on* J *and there is* $L > 0$ *such that*

$$\bigl\| f(t, z) - f(t, \bar{z}) \bigr\| \leqslant L \sum_{i=1}^{2} \|z - \bar{z}\|$$

for $t \in J$, $z, \bar{z} \in E$.
(2) $\sum_{k=1}^{p} c_k \neq -1$.
(3) $2LT(1 + |A \sum_{k=1}^{p} c_k|) < 1$.

Then the nonlocal IVP (2.6)–(2.7) *has a unique classical solution.*

PROOF. Introduce the operator P by the formula

$$(Px)(t) = A\left(x_0 - \sum_{k=1}^{p} c_k \int_0^{t_k} f\bigl(s, x(s)\bigr)\,\mathrm{d}s\right) + \int_0^t f\bigl(s, x(s)\bigr)\,\mathrm{d}s, \quad t \in J,\ x \in X.$$

It is easy to see that $P : X \to X$. We will prove that P is a contraction on X. We have

$$\begin{aligned}\bigl\|(Px)(t) - (P\bar{x})(t)\bigr\| &= \left\| A\left(-\sum_{k=1}^{p} c_k \int_0^{t_k} \bigl[f\bigl(s, x(s)\bigr) - f\bigl(s, \bar{x}(s)\bigr)\bigr]\,\mathrm{d}s\right)\right.\\ &\qquad \left. + \int_0^t \bigl[f\bigl(s, x(s)\bigr) - f\bigl(s, \bar{x}(s)\bigr)\bigr]\,\mathrm{d}s \right\|\\ &\leqslant 2LT\left(1 + \left|A \sum_{k=1}^{p} c_k\right|\right) \|x - \bar{x}\|_X.\end{aligned}$$

By (3) operator P is a contraction. Consequently by Banach fixed point theorem in space X there is only one fixed point of P and this fixed point is the unique classical solution of the nonlocal IVP (2.6)–(2.7). The proof of the theorem is complete. □

Theorem 2.4 is a special case of Theorem 2.4 of [11], where the functional differential equation $x'(t) = f(t, x(t), x(a(t)))$, $t \in J$ was considered. We point out that when we model some situations from physics and other applied sciences, we typically end up with partial differential equations. In recent years there has been increasing interest in studying evolution differential equations with nonlocal conditions. The interested reader is referred to [10] and [86] and the references cited therein.

3. Nonlocal boundary value problems

Let $f : [0, 1] \times \mathbb{R}^2 \to \mathbb{R}$ be a continuous function and let $e : [0, 1] \to \mathbb{R}$ be a function in $L^1[0, 1]$, $a_i \in \mathbb{R}$, $\xi_i \in (0, 1)$, $i = 1, 2, \ldots, m-2$, $0 < \xi_1 < \xi_2 < \cdots < \xi_{m-2} < 1$. We consider the following second-order ordinary differential equation

$$x''(t) = f\bigl(t, x(t), x'(t)\bigr) + e(t), \quad t \in (0, 1) \tag{3.1}$$

subject to the following boundary value conditions:

$$x(0) = 0, \qquad x(1) = \sum_{i=1}^{m-2} a_i x(\xi_i). \tag{3.2}$$

It is well known, that if a function $x \in C^1[0,1]$ satisfies the boundary conditions (3.2) and all α_i have the same sign then there exists an $\eta \in [\xi_1, \xi_{m-2}]$ such that

$$x(0) = 0, \qquad x(1) = \alpha x(\eta),$$

with $\alpha = \sum_{k=1}^{m-2} a_i$. Accordingly, the problem of the existence of a solution for the boundary value problem (BVP for short) (3.1)–(3.2) can be studied via the existence of a solution of the following three-point BVP

$$x(0) = 0, \qquad x(1) = \alpha x(\eta). \tag{3.3}$$

The above so-called "*nonlocal*" or "*multi-point*" or "*m-point*" BVP was initiated by Il'in and Moiseev in [43,44], motivated by the work of Bitsadze and Samarskii on nonlocal linear elliptic BVPs [7,8]. The nonlocal BVP for ordinary differential equations arise in a variety of different areas of applied mathematics and physics, and describe many phenomena in the applied mathematical sciences. For example, the vibrations of a guy wire of a uniform cross-section and composed on N parts of different densities can be set up as a multi-point BVP (see [83]); many problems in the theory of elastic stability can be handled by the method of multi-point problems (see [90]).

We will study nonlocal BVP for first, second or higher order ordinary differential equations.

3.1. *First-order three-point boundary value problems*

Given $a, b, c \in \mathbb{R}$ with $a < b < c$. Let $J := [a, c]$ and $M_{n\times n}$ the Banach space of all constant matrices of order n with the norm

$$\|B\| = \max_{1\leqslant i,j\leqslant n} |b_{i,j}|.$$

Let $C(J), L^1(J)$, and $AC(J)$ the usual Banach spaces of all continuous functions, all Lebesque integrable functions and all absolutely continuous functions with the norms

$$\|x\|_\infty = \max\bigl\{\bigl|x(t)\bigr|: t \in J\bigr\}, \qquad \|x\|_{L^1} = \int_J \bigl|x(t)\bigr|\,\mathrm{d}t,$$
$$\|x\|_{AC} = \|x\|_\infty + \|x'\|_{L^1},$$

respectively. For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, define

$$\|\alpha\| = \max\bigl\{|\alpha_i|: i = 1, \dots, n\bigr\}.$$

Denote by X, Y and Z the Banach spaces $C[J, \mathbb{R}^n] = \{(x_1, \dots, x_n): x_i \in C(J)\}$, $L^1[J, \mathbb{R}^n] = \{(x_1, \dots, x_n): x_i \in L^1C(J)\}$, and $AC[J, \mathbb{R}^n] = \{(x_1, \dots, x_n): x_i \in AC(J)\}$ with the norms $\|x\|_X = \max\{\|x_i\|_\infty: i = 2, \dots, n\}$, $\|x\|_Y = \sum_{i=1}^n \|x_i\|_{L^1}$ and $\|x\|_Z = \max\{\|x_i\|_{AC}: i = 2, \dots, n\}$ respectively.

DEFINITION 3.1. We say that $f : J \times \mathbb{R}^n \to \mathbb{R}^n$ is a Carathéodory function if

(i) $f(\cdot, u)$ is Lebesgue measurable on J for each $u \in \mathbb{R}^n$,
(ii) $f(t, \cdot)$ is continuous on $\mathbb{R}^n$ for a.e. $t \in J$,
(iii) for each $r \in (0, \infty)$, there exists an $h_r \in L^1(J, \mathbb{R}^n)$ such that

$$\left|f_i(t, u)\right| \leqslant (h_r)_i(t), \quad \text{for a.e. } t \in J,\ \|u\| \leqslant r,\ i = 1, \dots, n.$$

We study the problem of existence of solutions for the three-point BVP

$$x'(t) = f\bigl(t, x(t)\bigr), \quad \text{for a.e. } t \in J, \tag{3.4}$$

$$Mx(a) + Nx(b) + Rx(c) = \alpha, \tag{3.5}$$

where $f : [a, c] \times \mathbb{R}^n \to \mathbb{R}^n$ is a Carathéodory function, $M, N, R \in M_{n\times n}$, and $\alpha \in \mathbb{R}^n$ are given.

THEOREM 3.2. *Let $f : [a, c] \times \mathbb{R}^n \to \mathbb{R}^n$ be a Carathéodory function and $\alpha \in \mathbb{R}^n$. Assume that*

(A1) *M, N, R are constant square matrices of order n such that*

$$\det(M + N + R) \neq 0.$$

(A2) *There exist functions $p, r \in L^1(J)$ such that*

$$\left\|f(t, u)\right\| \leqslant p(t)\|u\| + r(t),$$

for a.e. $t \in J$ and $u \in \mathbb{R}^n$.

Then the three-point BVP (3.4)–(3.5) *has at least one solution in $C(J, \mathbb{R}^n)$ provided*

$$\Gamma \|p\|_1 < 1,$$

where

$$\Gamma = \max\Bigl\{\left\|(M + N + R)^{-1}R\right\|, \left\|(M + N + R)^{-1}M\right\|, \left\|(M + N + R)^{-1}(N + R)\right\|, \left\|(M + N + R)^{-1}(N + M)\right\|\Bigr\}.$$

PROOF. Let $e(\cdot) \in L^1(J, \mathbb{R}^n)$ and $\alpha \in \mathbb{R}^n$, $x \in AC(J)$ be such that

$$x'(t) = e(t), \quad \text{for a.e. } t \in J,$$

$$Mx(a) + Nx(b) + Rx(c) = \alpha.$$

Then

$$x(t) = \int_a^t e(s)\,\mathrm{d}s + (M + N + R)^{-1}\left[\alpha - N\int_a^b e(s)\,\mathrm{d}s - R\int_a^c e(s)\,\mathrm{d}s\right].$$

If E is the identity matrix, for $a \leqslant t \leqslant b < c$, we have

$$\begin{aligned} x(t) = & \int_a^t \bigl[E - (M+N+R)^{-1}N - (M+N+R)^{-1}R\bigr]e(s)\,\mathrm{d}s \\ & - \int_t^b \bigl[(M+N+R)^{-1}N + (M+N+R)^{-1}R\bigr]e(s)\,\mathrm{d}s \\ & - \int_b^c \bigl[(M+N+R)^{-1}R\bigr]e(s)\,\mathrm{d}s \\ = & \int_a^t (M+N+R)^{-1}Me(s)\,\mathrm{d}s \\ & - \int_t^b (M+N+R)^{-1}(N+R)e(s)\,\mathrm{d}s \\ & - \int_b^c (M+N+R)^{-1}Re(s)\,\mathrm{d}s, \end{aligned}$$

and hence

$$\begin{aligned} \bigl\|x(t)\bigr\| \leqslant \max\bigl\{ & \bigl\|(M+N+R)^{-1}M\bigr\|, \bigl\|(M+N+R)^{-1}(N+R)\bigr\|, \\ & \bigl\|(M+N+R)^{-1}R\bigr\|\bigr\}\|e\|_Y. \end{aligned}$$

For $a \leqslant b \leqslant t \leqslant c$ we have

$$\begin{aligned} x(t) = & \int_a^b \bigl[E - (M+N+R)^{-1}N - (M+N+R)^{-1}R\bigr]e(s)\,\mathrm{d}s \\ & - \int_b^t \bigl[-E + (M+N+R)^{-1}R\bigr]e(s)\,\mathrm{d}s \\ & - \int_t^c \bigl[(M+N+R)^{-1}R\bigr]e(s)\,\mathrm{d}s \\ = & \int_a^b (M+N+R)^{-1}Me(s)\,\mathrm{d}s \\ & - \int_b^t (M+N+R)^{-1}(-M-N)e(s)\,\mathrm{d}s \\ & - \int_t^c (M+N+R)^{-1}Re(s)\,\mathrm{d}s, \end{aligned}$$

and hence

$$\begin{aligned} \bigl\|x(t)\bigr\| \leqslant \max\bigl\{ & \bigl\|(M+N+R)^{-1}M\bigr\|, \bigl\|(M+N+R)^{-1}(M+N)\bigr\|, \\ & \bigl\|(M+N+R)^{-1}R\bigr\|\bigr\}\|e\|_Y. \end{aligned}$$

By combining the above relations we get

$$\left|x(t)\right| \leqslant \Gamma \|e\|_1$$

where

$$\Gamma = \max\bigl\{ \bigl\|(M+N+R)^{-1}R\bigr\|, \bigl\|(M+N+R)^{-1}M\bigr\|, \\ \bigl\|(M+N+R)^{-1}(N+R)\bigr\|, \bigl\|(M+N+R)^{-1}(N+M)\bigr\| \bigr\}.$$

It is obvious that, if $e = 0$, then the problem

$$x'(t) = 0, \quad \text{for a.e. } t \in J,$$
$$Mx(a) + Nx(b) + Rx(c) = \alpha,$$

has the unique solution $x(t) = w$, where

$$w = (M+M+R)^{-1}\alpha.$$

Moreover the problem

$$x'(t) = f(t,x), \quad \text{for a.e. } t \in J,$$
$$Mx(a) + Nx(b) + Rx(c) = \alpha,$$

has a solution $x = y_1 + w$ if and only if y_1 is a solution of

$$y'(t) = f(t, y+w), \quad \text{for a.e. } t \in J, \tag{3.6}$$
$$My(a) + Ny(b) + Ry(c) = 0. \tag{3.7}$$

Now we define $L : D(L) \subset X \to Y$ by

$$D(L) = \bigl\{ y \in AC(J)\colon\ My(a) + Ny(b) + Ry(c) = 0 \bigr\},$$

and for $y \in D(L)$,

$$Ly = y'.$$

We also define a nonlinear mapping $G : X \to X$ by

$$Gy = f\bigl(t, y(t) + w\bigr), \quad t \in J.$$

We note that G is a bounded operator from X into Y. Next it is easy to see from (A1) that the linear mapping $L : D(L) \subset X \to Y$, is a one-to-one mapping. Let the linear mapping $K : Y \to X$ be defined for $e \in Y$ by

$$(Ke)(t) = \int_a^t e(s)\,\mathrm{d}s + (M+M+R)^{-1}\biggl[\alpha - N\int_a^b e(s)\,\mathrm{d}s - R\int_a^c e(s)\,\mathrm{d}s\biggr].$$

Then for $e \in Y$, we have $Ke \in D(L)$ and $LKe = e$. For $y \in D(L)$, we have $KLy = y$. Furthermore, it follows easily using the Arzelá–Ascoli theorem that KG maps a bounded subset of X into a relatively compact subset of X. Hence, $KG: X \to X$ is a compact mapping.

We next note that $y \in C(J, \mathbb{R}^n)$ is a solution of BVP (3.6)–(3.7) if and only if y is a solution to operator equation

$$Ly - Gy,$$

which is equivalent to the equation

$$y = KGy.$$

We apply the Leray–Schauder continuation theorem (see, e.g., [82, Corollary IV.7]) to obtain existence of a solution to $y = KGy$ or equivalently to the BVP (3.6)–(3.7).

To do this, it suffices to verify that the set of all possible solutions of the family of equations

$$y'(t) = \lambda f(t, y + w), \quad \text{for a.e. } t \in J, \tag{3.8}$$

$$My(a) + Ny(b) + Ry(c) = 0, \tag{3.9}$$

is bounded in $X = C(J, \mathbb{R}^n)$ by a constant independent of $\lambda \in [0, 1]$.

Let $y(t)$ be a solution of (3.8)–(3.9) for some $\lambda \in [0, 1]$. It follows from our assumptions that $f(t, y(t) + w(t)) \in L^1(J, \mathbb{R}^n)$. We have

$$\begin{aligned} \|y\|_X &\leqslant \lambda \Gamma \big\| f\big(t, y(t) + w\big)\big\|_Y \\ &\leqslant \Gamma\big[\|p\|_{L^1}\|y + w\|_X + \|r\|_{L^1}\big] \\ &\leqslant \Gamma\big[\|p\|_{L^1}\big(\|y\|_X + \|w\|_X\big) + \|r\|_{L^1}\big]. \end{aligned}$$

It follows from $\Gamma\|p\|_{L^1} < 1$ that there exists a constant c_0, independent of $\lambda \in [0, 1]$, such that

$$\|y\|_X \leqslant c_0.$$

This completes the proof of the theorem. □

THEOREM 3.3. *Let* $f : [a, c] \times \mathbb{R}^n \to \mathbb{R}^n$ *be a Carathéodory function and* $\alpha \in \mathbb{R}^n$. *Assume that* (A1) *holds. Moreover we suppose that*

(A3) *There exists function* $p \in L^1(J)$ *such that*

$$\big\| f(t, u) - f(t, v)\big\| \leqslant p(t)\|u - v\|$$

for a.e. $t \in J$ *and all* $u, v \in \mathbb{R}^n$.

Then the three-point BVP (3.4)–(3.5) *has a unique solution in* $C(J, \mathbb{R}^n)$ *provided*

$$\Gamma \|p\|_1 < 1.$$

PROOF. We note that (A3) implies

$$\|f(t,u)\| \leqslant \|f(t,0)\| + p(t)\|u\|,$$

for a.e. $t \in J$ and all $u \in \mathbb{R}^n$. Hence the BVP (3.4)–(3.5) has at least one solution in $C(J, \mathbb{R}^n)$ by Theorem 3.2. Suppose that u_1, u_2 are two solutions of (3.4)–(3.5) in $C(J, \mathbb{R}^n)$. Setting $z = u_1 - u_2$, we have

$$z' = f\bigl(t, u_1(t)\bigr) - f\bigl(t, u_2(t)\bigr),$$
$$Mz(a) + Nz(b) + Rz(c) = 0.$$

Then we have

$$\begin{aligned}\|z\|_X &\leqslant \lambda\Gamma \bigl\| f\bigl(t, u_1(t)\bigr) - f\bigl(t, u_2(t)\bigr)\bigr\|_Y \\ &\leqslant \Gamma\bigl(\|p\|_{L^1}\|u_1 - u_2\|_X\bigr) \\ &\leqslant \Gamma\|p\|_{L^1}\|z\|_X,\end{aligned}$$

which implies that $z(t) = 0$, for $t \in J$ and hence $u_1 = u_2$. This completes the proof of the theorem. □

Theorems 3.2 and 3.3 were taken from [74]. Existence and uniqueness of solutions of the BVP (3.4)–(3.5) were studied in [84] by using the successive over relaxation iteration and the Banach contraction mapping principle.

3.2. *Second-order three-point boundary value problems*

In this section we consider the following three-point BVP

$$x''(t) = f\bigl(t, x(t), x'(t)\bigr) + e(t), \quad t \in (0,1) \tag{3.10}$$

$$x(0) = 0, \qquad x(1) = x(\eta), \tag{3.11}$$

where $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$, $e : [0,1] \to \mathbb{R}$ are given functions and $\eta \in (0,1)$.

We will study the existence and uniqueness of the BVP (3.10)–(3.11). We use the classical spaces $C[0,1]$, $C^k[0,1]$, $L^k[0,1]$, and $L^\infty[0,1]$ of continuous, k-times continuously differentiable, measurable real valued functions whose kth power of the absolute value is Lebesgue integrable on $[0,1]$, or measurable functions that are essentially bounded on $[0,1]$. We also use the Sobolev space $W^{2,k}(0,1)$, $k = 1, 2$ defined by

$$W^{2,k}(0,1) = \bigl\{x : [0,1] \to \mathbb{R}\colon\ x, x' \text{are absolutely continuous on } [0,1] \text{ with } x \in L^k[0,1]\bigr\}$$

with its usual norm. We denote the norm in $L^k[0,1]$ by $\|\cdot\|_k$, and the norm in $L^\infty[0,1]$ by $\|\cdot\|_\infty$.

The BVP (3.10)–(3.11) can be put in the form of an operator equation

$$Lx + Nx = w,$$

where $L : D(L) \subset X \to Y$ is a linear operator, $N : X \to Y$ is a nonlinear operator, and X, Y are suitable spaces in duality. Clearly the linear operator L in (3.10) is given by

$$Lx = -x''$$

where the boundary conditions (3.11) are used to define the domain, $D(L)$, of L.

We begin by giving the following definition.

DEFINITION 3.4. A function $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ satisfies Carathéodory's conditions if:

(i) for each $(x,y) \in \mathbb{R}^2$, the function $t \in [0,1] \to f(t,x,y) \in \mathbb{R}$ is measurable on $[0,1]$,

(ii) for a.e. $t \in [0,1]$, the function $(x,y) \in \mathbb{R}^2 \to f(t,x,y) \in \mathbb{R}$ is continuous on $\mathbb{R}^2$, and

(iii) for each $r > 0$, there exists $g_r \in L^1[0,1]$ such that $|f(t,x,y)| \leqslant g_r(t)$ for a.e. $t \in [0,1]$ and $(x,y) \in \mathbb{R}^2$ with $\sqrt{x^2+y^2} \leqslant r$.

In the next existence theorems, we obtain appropriate *a priori* bounds by applying degree theory. We use Wirtinger-type inequalities to obtain the necessary a priori bounds needed to apply the Leray–Schauder continuation theorem [82, Corollary IV.7].

THEOREM 3.5. *Let $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ satisfy Carathéodory's conditions. Assume that there exist functions $p, q, r \in L^1[0,1]$ such that*

$$\bigl|f(t,u,v)\bigr| \leqslant p(t)|u| + q(t)|v| + r(t), \tag{3.12}$$

for a.e. $t \in [0,1]$ and all $(u,v) \in \mathbb{R}^2$. Then for every given function $e \in L^1[0,1]$, the BVP (3.10)–(3.11)*, has at least one solution in $C^1[0,1]$ provided*

$$\|q\|_1 \leqslant 1,$$

and

$$\sqrt{\eta}\,\|p\|_{L^1[\eta,1]} \frac{1 - \|q\|_{L^1[0,1]}}{1 - \|q\|_{L^1[\eta,1]}} + \|p\|_{L^1[0,1]} + \|q\|_{L^1[0,1]} < 1.$$

PROOF. Let X be the Banach space $C^1[0,1]$ and Y denote the Banach space $L^1(0,1)$ with their usual norms. We define a linear mapping $L : D(L) \subset X \to Y$ by setting

$$D(L) = \bigl\{x \in W^{2,1}(0,1):\ x(0) = 0,\ x(1) = x(\eta)\bigr\},$$

and for $x \in D(L)$,

$$Lx = -x''.$$

We also define a nonlinear mapping $N : X \to Y$ by setting

$$(Nx)(t) = f\bigl(t, x(t), x'(t)\bigr), \quad t \in [0, 1].$$

We note that N is a bounded mapping from X into Y. Next, it is easy to see that the linear mapping $L : D(L) \subset X \to Y$, is one-to-one mapping. Next, the linear mapping $K : Y \to X$, defined for $y \in Y$ by

$$(Ky)(t) = -\int_0^t (t-s)y(s)\,\mathrm{d}s + t\int_0^\eta y(s)\,\mathrm{d}s + \frac{t}{1-\eta}\int_0^1 (1-s)y(s)\,\mathrm{d}s$$

is such that for $y \in Y$, $Ky \in D(L)$ and $LKy = y$; and for $u \in D(L)$, $KLu = u$. Furthermore, it follows easily using the Arzelá–Ascoli theorem that KN maps a bounded subset of X into a relatively compact subset of X. Hence $KN : X \to X$ is a compact mapping.

We next note that $x \in C^1[0, 1]$ is a solution of the BVP (3.10)–(3.11) if and only if x is a solution to the operator equation

$$Lx = Nx + e.$$

Now, the operator equation $Lx = Nx + e$ is equivalent to the equation

$$x = KNx + Ke.$$

We apply the Leray–Schauder continuation theorem (see, e.g., [82, Corollary IV.7]) to obtain the existence of a solution for $x = KNx + Ke$ or equivalently to the BVP (3.10)–(3.11).

To do this, it suffices to verify that the set of all possible solutions of the family of equations

$$x''(t) = \lambda f\bigl(t, x(t), x'(t)\bigr) + \lambda e(t), \quad t \in (0, 1), \tag{3.13}$$

$$x(0) = 0, \qquad x(1) = x(\eta), \tag{3.14}$$

is, a priori, bounded in $C^1[0, 1]$ by a constant independent of $\lambda \in [0, 1]$.

We observe first that for $x \in W^{2,1}(0, 1)$ with $x(0) = 0, x(1) = x(\eta)$ there exists a $\zeta \in (0, 1)$, $\eta < \zeta < 1$, such that $x'(\zeta) = 0$. It follows that

$$\|x\|_2 \leqslant \frac{2}{\pi}\|x'\|_2, \qquad \|x\|_\infty \leqslant \|x'\|_2,$$

$$\bigl|x(1)\bigr| = \bigl|x(\eta)\bigr| = \left|\int_0^\eta x'(t)\,\mathrm{d}t\right| \leqslant \sqrt{\eta}\|x'\|_2,$$

$$\left|x'(1)\right| = \left|\int_\zeta^1 x''(t)\,\mathrm{d}t\right| \leqslant \int_\zeta^1 \left|x''(t)\right|\mathrm{d}t.$$

We multiply Eq. (3.13) by x and integrate from 0 to 1 to get

$$\begin{aligned}
0 &= -\int_0^1 x''(t)x(t)\,\mathrm{d}t + \lambda\int_0^1 f\big(t,x(t),x'(t)\big)x(t)\,\mathrm{d}t - \lambda\int_0^1 e(t)x(t)\,\mathrm{d}t \\
&\geqslant \|x'\|_2^2 - \left|x'(1)\right|\left|x(1)\right| - \int_0^1 \left|f\big(t,x(t),x'(t)\big)\right|\left|x(t)\right|\mathrm{d}t - \|e\|_1\|x\|_\infty \\
&\geqslant \|x'\|_2^2 - \left|x'(1)\right|\left|x(1)\right| \\
&\quad - \int_0^1 \left[p(t)\left|x(t)\right|^2 + q(t)\left|x(t)\right|\left|x'(t)\right| + r(t)\left|x(t)\right|\right]\mathrm{d}t - \|e\|_1\|x\|_\infty \\
&\geqslant \|x'\|_2^2 - \sqrt{\eta}\|x'\|_2\|x''\|_{L^1[\eta,1]} \\
&\quad - \left[\|p\|_1\|x\|_2^2 + \|q\|_1\|x\|_\infty\|x'\|_\infty + \|r\|_1\|x\|_\infty\right] - \|e\|_1\|x\|_\infty \\
&\geqslant \|x'\|_2^2 - \sqrt{\eta}\|x'\|_2\|x''\|_{L^1[\eta,1]} - \left[\|p\|_1\|x'\|_2^2 + \|q\|_1\|x'\|_2\|x''\|_1\right] \\
&\quad - \big(\|r\|_1 + \|e\|_1\big)\|x'\|_2.
\end{aligned}$$

From Eq. (3.13) we have

$$\begin{aligned}
\|x''\|_1 &\leqslant \|p\|_1\|x\|_\infty + \|q\|_1\|x'\|_\infty + \|r\|_1 \\
&\leqslant \|p\|_1\|x'\|_2 + \|q\|_1\|x''\|_1 + \|r\|_1,
\end{aligned}$$

so that

$$\|x''\|_1 \leqslant \frac{\|p\|_1\|x'\|_2}{1-\|q\|_1} + \frac{\|r\|_1}{1-\|q\|_1}.$$

Similarly we have

$$\|x''\|_{L^1[\eta,1]} \leqslant \frac{\|p\|_{L^1[\eta,1]}\|x'\|_2}{1-\|q\|_{L^1[\eta,1]}} + \frac{\|r\|_1}{1-\|q\|_{L^1[\eta,1]}}.$$

Consequently

$$\begin{aligned}
&\left\{1 - \|q\|_1 - \sqrt{\eta}\|p\|_{L^1[\eta,1]}\frac{1-\|q\|_1}{1-\|q\|_{L^1[\eta,1]}} - \|p\|_1\right\}\|x'\|_2 \\
&\quad\leqslant \|q\|_1\|r\|_1 + \sqrt{\eta}\|r\|_1\frac{1-\|q\|_1}{1-\|q\|_{L^1[\eta,1]}} + \big(\|r\|_1 + \|e\|_1\big)\big(1-\|q\|_1\big).
\end{aligned}$$

Hence, there exists a constant C, independent of $\lambda \in [0,1]$ such that $\|x'\|_2 \leqslant C$. It follows that there is a constant, still denoted by C, independent of $\lambda \in [0,1]$ such that

$$\|x\|_{C^1[0,1]} \leqslant C.$$

This completes the proof of the theorem. □

THEOREM 3.6. *Let* $f:[0,1]\times\mathbb{R}^2 \to \mathbb{R}$ *satisfy Carathéodory's conditions. Assume that* $a \geqslant 0, b \geqslant 0, \alpha(t) \in L^1[0,1]$ *are such that*

$$\left|f(t,u,v)\right| \leqslant a|u| + b|v| + \alpha(t),$$

for a.e. $t \in [0,1]$ *and all* $(u,v) \in \mathbb{R}^2$*. Then for every given function* $e(t) \in L^1[0,1]$*, the BVP* (3.10)–(3.11) *has at least one solution in* $C^1[0,1]$ *provided*

$$\left(\frac{2}{\pi}a + b\right)\left(\frac{2}{\pi} + \sqrt{\eta(1-\eta)}\right) < 1.$$

PROOF. As in the proof of Theorem 3.5 it suffices to verify that the set of all possible solutions of the family of equations (3.13)–(3.14) is a priori bounded in $C^1[0,1]$ by a constant independent of $\lambda \in [0,1]$.

Letting x to be a solution of (3.13)–(3.14) for some $\lambda \in [0,1]$ we get, as in Theorem 3.5, that

$$\begin{aligned}
0 &= -\int_0^1 x''(t)x(t)\,\mathrm{d}t + \lambda\int_0^1 f\bigl(t,x(t),x'(t)\bigr)x(t)\,\mathrm{d}t - \lambda\int_0^1 e(t)x(t)\,\mathrm{d}t \\
&\geqslant \|x'\|_2^2 - \left|x'(1)\right|\left|x(1)\right| - \int_0^1 \bigl[a|x|^2 + b|x||x'| + \alpha(t)|x|\bigr]\,\mathrm{d}t - \|e\|_1\|x\|_\infty \\
&\geqslant \|x'\|_2^2 - \sqrt{\eta}\|x'\|_2\|x''\|_{L^1[\eta,1]} - a\|x\|_2^2 - b\|x\|_2\|x'\|_2 \\
&\quad - \bigl(\|\alpha\|_1 + \|e\|_1\bigr)\|x\|_\infty \\
&\geqslant \|x'\|_2^2 - \sqrt{\eta}\|x'\|_2\|x''\|_{L^1[\eta,1]} - \left(\frac{4}{\pi^2}a + \frac{2}{\pi}b\right)\|x'\|_2^2 - \bigl(\|\alpha\|_1 + \|e\|_1\bigr)\|x'\|_2.
\end{aligned}$$

Also

$$\begin{aligned}
\|x''\|_{L^1[\eta,1]} &\leqslant \left\|f\bigl(t,x(t),x'(t)\bigr)\right\|_{L^1[\eta,1]} + \|e\|_1 \\
&\leqslant a\int_\eta^1 |x|\,\mathrm{d}x + b\int_\eta^1 |x'|\,\mathrm{d}x + \|\alpha\|_1 + \|e\|_1 \\
&\leqslant \sqrt{1-\eta}\left(\frac{2}{\pi}a + b\right)\|x'\|_2 + \|\alpha\|_1 + \|e\|_1.
\end{aligned}$$

Consequently

$$\|x'\|_2 \leqslant \frac{(\|\alpha\|_1 + \|e\|_1)(1+\sqrt{\eta})}{1-((2/\pi)a+b)(2/\pi+\sqrt{\eta(1-\eta)})}.$$

It then follows as in proof of Theorem 3.5 that there exists a constant C independent of $\lambda \in [0,1]$ such that $\|x\|_{C^1[0,1]} \leqslant C$, which proves the theorem. □

We give now a uniqueness result for the BVP (3.10)–(3.11).

THEOREM 3.7. *Let* $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ *satisfy Carathéodory's conditions. Assume that* $a \geqslant 0$, $b \geqslant 0$, *are such that*

$$\bigl|f(t,u_1,v_1) - f(t,u_2,v_2)\bigr| \leqslant a|u_1-u_2| + b|v_1-v_2|, \tag{3.15}$$

for a.e. $t \in [0,1]$ *and all* $(u_i,v_i) \in \mathbb{R}^2, i = 1,2$. *Then for every given function* $e(t) \in L^1[0,1]$, *the BVP* (3.10)–(3.11) *has unique solution in* $C^1[0,1]$ *provided*

$$\left(\frac{2}{\pi}a + b\right)\left(\frac{2}{\pi} + \sqrt{\eta(1-\eta)}\right) < 1.$$

PROOF. We note that (3.15) implies that

$$\bigl|f(t,u,v)\bigr| < a|u| + b|v| + \bigl|f(t,0,0)\bigr|$$

for all $t \in [0,1]$ and all $(u,v) \in \mathbb{R}^2$. Hence the BVP (3.10)–(3.11) has a solution by Theorem 3.6. Let now u_1, u_2 be two solutions of (3.10)–(3.11) in $C^1[0,1]$. Setting $w = u_1 - u_2$ we have

$$w'' = f(t,u_1,u_1') - f(t,u_2,u_2'), \quad t \in (0,1), \tag{3.16}$$

$$w(0) = 0, \qquad w(1) = w(\eta). \tag{3.17}$$

Multiplying Eq. (3.16) by w and integrating over $[0,1]$ we get

$$\begin{aligned}
0 = {} & -\int_0^1 w''w\,\mathrm{d}t + \int_0^1 \bigl[f\bigl(t,u_1(t),u_1'(t)\bigr) - f\bigl(t,u_2(t),u_2'(t)\bigr)\bigr]w(t)\,\mathrm{d}t \\
\geqslant {} & \|w'\|_2^2 - \bigl|w(1)\bigr|\bigl|w'(1)\bigr| \\
& - \int_0^1 \bigl|f\bigl(t,u_1(t),u_1'(t)\bigr) - f\bigl(t,u_2(t),u_2'(t)\bigr)\bigr|\bigl|w(t)\bigr|\,\mathrm{d}t \\
\geqslant {} & \|w'\|_2^2 - \sqrt{\eta}\|w'\|_2\|w''\|_{L^1[\eta,1]} \\
& - \int_0^1 \bigl[a|u_1-u_2|^2 + b|u_1'-u_2'||u_1-u_2|_2\bigr]\,\mathrm{d}t
\end{aligned}$$

$$\geqslant \|w'\|_2^2 - \sqrt{\eta}\|w'\|_2\|w''\|_{L^1[\eta,1]} - \left(\frac{4}{\pi^2}a + \frac{2}{\pi}b\right)\|w'\|_2^2.$$

But

$$\begin{aligned}\|w''\|_{L^1[\eta,1]} &= \left\| f\big(t, u_1(t), u_1'(t)\big) - f\big(t, u_2(t), u_2'(t)\big)\right\|_{L^1[\eta,1]} \\ &\leqslant \sqrt{1-\eta}\left(\frac{2}{\pi}a + b\right)\|w'\|_2^2,\end{aligned}$$

and consequently

$$\left(1 - \left(\sqrt{\eta(1-\eta)} + \frac{2}{\pi}\right)\left(\frac{2}{\pi}a + b\right)\right)\|w'\|_2^2 \leqslant 0,$$

so that $\|w'\|_2 = 0$. Since $w(0) = 0$ we obtain using $w(t) = \int_0^t w'(t)\,\mathrm{d}t$ that $\|w\|_\infty \leqslant \|w'\|_2 = 0$. Hence $w = 0$ a.e. in $[0,1]$ and thus $w(t) = 0$ for every $t \in [0,1]$ because w is continuous. □

Theorems 3.5–3.7 were proved by Gupta in [25]. By means of rather different approach based on an existence theorem for operator inclusions, Marano proved in [80], that if p, q, r are in $L^1[0,1]$ and (3.12) holds, then the BVP (3.10)–(3.11) has at least one generalized solution $x \in W^{2,1}[0,1]$ provided

$$\|p\|_1 + \|q\|_1 < 1,$$

which is an improvement of Theorem 3.5. For this result of Marano, Gupta in [26] gave a simple proof by using degree theory and Wirtinger-type inequalities. Marano in [80] proved also that if p, q, r are in $L^2[0,1]$ and (3.12) holds, then the BVP (3.10)–(3.11) has at least one generalized solution $x \in W^{2,1}[0,1]$ provided

$$\frac{\sqrt{\eta}+1}{\pi}\|p\|_2 + \|q\|_2 < 1.$$

In the case when p, q, r are in $L^2(0,1)$, and (3.12) holds, Gupta in [32] proved that the BVP (3.10)–(3.11) has at least one generalized solution $x \in C^1[0,1]$ provided

$$\frac{2}{\pi}\|p\|_2 + \|q\|_2 < 1.$$

The Leray–Schauder continuation theorem can be used without the use of Wirtinger-type inequalities and using instead Green's functions. The Green's functions for various BVPs are easily computed. Then one uses Green's functions to reduce the BVPs to integral equations. Thus finding a priori estimates can be realized by estimating the norms of integral operators, which can be solved by Green's functions and its derivatives. In this direction Zhang and Han [93] proved that if p, q, r are in $L^1[0,1]$ and (3.12) holds, then the BVP (3.10)–(3.11) has at least one solution in $C^1[0,1]$ provided

$$\max\biggl\{\int_0^\eta \bigl(sp(s)+q(s)\bigr)\,\mathrm{d}s + \int_\eta^1 \frac{1-s}{1-\eta}\bigl(sp(s)+q(s)\bigr)\,\mathrm{d}s, \\ \int_\eta^1 \frac{s-\eta}{1-\eta}\bigl(sp(s)+q(s)\bigr)\,\mathrm{d}s\biggr\} < 1.$$

In [9] Boucherif and Bouguima proved the existence of solutions of the BVP (3.10)–(3.11) by considering two types of nonlinearities, continuous and Carathéodory's nonlinearities. The methods used in the first case rely on the topological transversality theorem of Granas [14], and in the second case on the technique of cut-off functions. They motivated their results by the remark that if $f(t,u,v) = u + v + R(t,u,v)$ with $R(t,u,v) \leqslant r(t)$ for a.e. $t \in [0,1]$ then f satisfies condition (3.12) i.e. $|f(t,u,v)| \leqslant p_0(t)|u| + q_0|v| + r(t)$ but $\|p_0\|_1 + \|q_0\|_1 = 2 > 1$.

THEOREM 3.8. *Let* $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ *be a continuous function satisfying the following conditions*:

(f1) *there exists* $r > 0$ *such that*

$$\int_0^t f\bigl(s, x(s), x'(s)\bigr)x(s)\,\mathrm{d}s > 0 \quad \textit{whenever} \quad \bigl|x(t)\bigr| > r.$$

(f2) $|f(t,u,v)| \leqslant \psi(|v|)$ *for* $(t,u) \in [0,1] \times [-r,r]$, $v \in \mathbb{R}$ *where* $\psi : [0,\infty) \to (0,\infty)$ *is such that* $1/\psi$ *is integrable over bounded intervals in* $[0,\infty)$ *and* $\int_0^{+\infty} \frac{s}{\psi(s)}\,\mathrm{d}s > 2r$.

Then the BVP (3.10)–(3.11) *has at least one solution*.

PROOF. We begin by examining the linear problem

$$x'' = z, \qquad x(0) = 0, \qquad x(1) = x(\eta)$$

where $z \in C([0,1],\mathbb{R})$. This problem has a unique solution $x \in C^2[0,1]$, given by

$$x(t) = \int_0^t (t-s)z(s)\,\mathrm{d}s + t\int_0^\eta z(s)\,\mathrm{d}s + \frac{t}{1-\eta}\int_\eta^1 (1-s)z(s)\,\mathrm{d}s.$$

This relation defines a linear operator $K : C[0,1] \to C^2[0,1]$ such that $x = Kz$ is the unique solution. Moreover $L = K^{-1}$ is defined by

$$D(L) = \bigl\{x \in C^2[0,1]\colon\ x(0) = 0,\ x(1) = x(\eta)\bigr\}, \qquad Lx = x''.$$

For $\lambda \in [0,1]$ consider the one-parameter family of problems

$$x''(t) = \lambda f\bigl(t, x(t), x'(t)\bigr), \quad t \in (0,1), \tag{3.18}$$

$$x(0) = 0, \qquad x(1) = x(\eta). \tag{3.19}$$

We shall prove that any solution x of (3.18)–(3.19) is a priori bounded, independently of λ.

Let x be a solution of (3.18)–(3.19). If $\lambda = 0$ then $x = 0$ and so $\|x\|_0 = 0 < r$. Assume $0 < \lambda \leqslant 1$. Then λf satisfies (f1). Suppose x takes on a positive maximum at $t_0 \in [0,1]$. Then $x'(t_0) = 0$ and $x''(t_0) \leqslant 0$. Now, if $x(t_0) > r$ it follows from (f1) that

$$\int_0^{t_0} x''(s)x(s)\,\mathrm{d}s = x(t_0)x'(t_0) - \int_0^{t_0} x'(s)^2\,\mathrm{d}s > 0.$$

Since $x'(t_0) = 0$ we have a contradiction. Hence $x(t_0) \leqslant r$ and therefore $\|x\|_0 \leqslant r$. Note that $t_0 \neq 0$ since $x(0) = 0$. Also, the case $t_0 = \eta$ or $t_0 = 1$ leads to a contradiction.

Now we want to show that there exists a constant m_0, such that $|x'(t)| \leqslant m_0$ for all $t \in [0,1]$. Suppose on the contrary that for all positive constants m, there exists $t_0 \in (0,1]$ such that $x'(t_0) > m$ or $x'(t_0) < -m$. We will assume only the case $x'(t_0) > m$, since the other case can be handled in a similar way. Let $\tau \in (0,1)$ be such that $x'(\tau) = 0$ and $x'(t) > 0$ for all $t \in (\tau, t_0]$. This implies that

$$x''(t) \leqslant \psi\bigl(x'(t)\bigr) \quad \text{for all } t \in [\tau, t_0].$$

Hence

$$\frac{x''(t)x'(t)}{\psi(x'(t))} \leqslant x'(t) \quad \text{for all } t \in [\tau, t_0].$$

Integrating this inequality from τ to t_0 we get

$$\int_0^{x'(t_0)} \frac{s}{\psi(s)}\,\mathrm{d}s \leqslant 2r.$$

It follows from the assumptions on ψ that there exists a constant $\tilde{m}$ such that $x'(t_0) \leqslant \tilde{m}$, and so we get a contradiction.

Since f is continuous, we have $|x''(t)| \leqslant m_1 := \sup\{|f(t,u,v)|\colon\ 0 \leqslant t \leqslant 1, \|u\|_0 \leqslant r, \|v\|_0 \leqslant m_0\}$. Hence $\|x''\|_0 \leqslant m_1$. Finally, let $M = \max(r, m_0, m_1)$. Then we get, from $\|x\| = \max(\|x\|_0, \|x'\|_0, \|x''\|_0)$ that $\|x\| \leqslant M$, and M is independent of λ.

We apply the topological transversality of Granas [14]. Let $U = \{x \in D(L)\colon\ \|x\| < M + 1\}$, and $j : D(L) \to C^1[0,1]$ with $j(x) = x$, the natural embedding. The Nemitskii operator $F : C^1[0,1] \to C[0,1]$ is defined by $F(x)(t) := f(t, x(t), x'(t))$ for all $t \in [0,1]$.

For $0 \leqslant \lambda \leqslant 1$ define a family of mappings

$$H_\lambda : \overline{U} \to D(L) \quad \text{by } H_\lambda(x) = \lambda KFj(x).$$

H_λ is a compact homotopy which has no fixed points on the boundary ∂U of U. Since the constant map H_0 is essential, the topological transversality theorem implies that H_1 is essential i.e. has a fixed point. This completes the proof. $\square$

In the case of Carathéodory's nonlinearities we state the following result whose proof is given in [9].

THEOREM 3.9. *Let* $f:[0,1]\times\mathbb{R}^2\to\mathbb{R}$ *be a function satisfying Carathéodory's conditions. Assume that*

(f) $|f(t,u,v)|\leqslant\omega(t,|u|+|v|)$ *for a.e.* $t\in[0,1]$ *and all* $(u,v)\in\mathbb{R}^2$ *where* $\omega:[0,1]\times\mathbb{R}^+\to\mathbb{R}^+$ *is nondecreasing in its second argument and*

$$\lim_{\rho\to\infty}\frac{1}{\rho}\int_0^1\omega(t,\rho)\,\mathrm{d}t=0.$$

Then the BVP (3.10)–(3.11) *has at least one solution.*

3.3. *Third-order three-point boundary value problems*

In this section we consider third-order three-point BVPs. The existence and uniqueness of third-order BVPs deserve a good deal of attention, since they occur in a wide variety of applications. For example, a three-layer beam is formed by parallel layers of different materials. For an equally loaded beam of this type, has shown that the deflection ψ is governed by an ordinary third-order linear differential equation

$$\psi'''-K^2\psi'+a=0,$$

where K^2 and a are physical parameters depending on the elasticity of the layers. The condition of zero moment at the free ends implies the boundary conditions

$$\psi'(0)=\psi'(1)=0,$$

and the symmetry yields the third boundary condition

$$\psi(1/2)=0.$$

We will study existence and uniqueness of the following three-point BVP

$$x'''(t)=f\big(t,x(t),x'(t),x''(t)\big)+e(t),\quad t\in(0,1),\tag{3.20}$$

$$x(0)=x(\eta)=x(1)=0\tag{3.21}$$

where $e:[0,1]\to\mathbb{R}$ is a function in $L^1[0,1]$, $\eta\in(0,1)$, and $f:[0,1]\times\mathbb{R}^3\to\mathbb{R}$ is a function satisfying Carathéodory's conditions, that is

(i) for each $(x,y,z)\in\mathbb{R}^3$, the function $t\in[0,1]\to f(t,x,y,z)\in\mathbb{R}$ is measurable on $[0,1]$,

(ii) for a.e. $t\in[0,1]$, the function $(x,y,z)\in\mathbb{R}^3\to f(t,x,y,z)\in\mathbb{R}$ is continuous on $\mathbb{R}^3$, and

(iii) for each $r > 0$, there exists $g_r \in L^1[0,1]$ such that $|f(t,x,y,z)| \leqslant g_r(t)$ for a.e. $t \in [0,1]$ and $(x,y,z) \in \mathbb{R}^3$ with $\sqrt{x^2+y^2+z^2} \leqslant r$.

THEOREM 3.10. *Let $f : [0,1] \times \mathbb{R}^3 \to \mathbb{R}$ satisfy Carathéodory's conditions. Assume that*

(i) *there exist $a, b, c \in \mathbb{R}$ and $\alpha(t) \in L^1[0,1]$ such that*

$$f(t,x,y,z)y \geqslant ay^2 + b|y||x| + c|y||z| + \alpha(x)y,$$

for a.e. $t \in [0,1]$ and all $(x,y,z) \in \mathbb{R}^3$;

(ii) *there exist functions p, q, r in $L^2[0,1]$ and a function $S(t) \in L^1[0,1]$ such that*

$$\bigl|f(t,x,y,z)\bigr| \leqslant p(t)|x| + q(t)|y| + r(t)|z| + S(t),$$

for a.e. $t \in [0,1]$ and all $(x,y,z) \in \mathbb{R}^3$.

Let $\eta \in (0,1)$ be given. Then for any given function $e \in L^1[0,1]$, the tree-point BVP (3.20)–(3.21) *has at least one solution in $C^2[0,1]$ provided*

$$\frac{4}{\pi^2}|a| + \frac{4}{\pi^3}|b| + \frac{2}{\pi}|c| + \frac{2}{\pi^2}\|p\|_2 + \frac{2}{\pi}\|q\|_2 + \|r\|_2 < 1.$$

PROOF. Let X be the Banach space $C^2[0,1]$ and Y denote the Banach space $L^1(0,1)$ with their usual norms. We define a linear mapping $L : D(L) \subset X \to Y$ by setting

$$D(L) = \bigl\{x \in W^{3,1}(0,1)\colon\ x(0) = x(\eta) = x(1) = 0\bigr\},$$

and for $x \in D(L)$,

$$Lx = -x'''.$$

We also define a nonlinear mapping $N : X \to Y$ by setting

$$(Nx)(t) = f\bigl(t, x(t), x'(t), x''(t)\bigr), \quad t \in [0,1].$$

We note that N is a bounded mapping from X into Y. Next, it is easy to see that the linear mapping $L : D(L) \subset X \to Y$, is one-to-one mapping. Next, the linear mapping $K : Y \to X$, defined for $y \in Y$ by

$$(Ky)(t) = \frac{1}{2}\int_0^t (t-s)^2 y(s)\,\mathrm{d}s + Bt + Ct^2,$$

where

$$B = \frac{1}{2\eta(\eta-1)}\int_0^\eta (t-s)^2 y(s)\,\mathrm{d}s + \frac{\eta}{2(1-\eta)}\int_0^1 (t-s)^2 y(s)\,\mathrm{d}s,$$
$$C = \frac{1}{2\eta(1-\eta)}\int_0^\eta (t-s)^2 y(s)\,\mathrm{d}s - \frac{1}{2(1-\eta)}\int_0^1 (t-s)^2 y(s)\,\mathrm{d}s,$$

is such that for $y \in Y$, $Ky \in D(L)$ and $LKy = y$; and for $u \in D(L)$, $KLu = u$. Furthermore, it follows easily using the Arzelá–Ascoli theorem that KN maps a bounded subset of X into a relatively compact subset of X. Hence $KN : X \to X$ is a compact mapping.

We next note that $x \in C^1[0,1]$ is a solution of the BVP (3.20)–(3.21) if and only if x is a solution to the operator equation

$$Lx = Nx + e.$$

Now, the operator equation $Lx = Nx + e$ is equivalent to the equation

$$x = KNx + Ke.$$

We apply the Leray–Schauder continuation theorem (see, e.g., [82, Corollary IV.7]) to obtain the existence of a solution for $x = KNx + Ke$ or equivalently to the BVP (3.20)–(3.21).

To do this, it suffices to verify that the set of all possible solutions of the family of equations

$$x'''(t) = \lambda f\bigl(t, x(t), x'(t), x''(t)\bigr) + \lambda e(t), \quad t \in (0,1), \tag{3.22}$$

$$x(0) = x(\eta) = x(1) = 0, \tag{3.23}$$

is, a priori, bounded in $C^2[0,1]$ by a constant independent of $\lambda \in [0,1]$.

We observe that for $x \in W^{3,1}(0,1)$ with $x(0) = x(\eta) = x(1) = 0$, there exist ξ_1, ξ_2, ξ_3 in $(0,1)$ with $\xi_1 < \eta < \xi_2$, $x'(\xi_1) = x'(\xi_2) = 0$, $x''(\xi_3) = 0$. This gives

$$\|x\|_2 \leqslant \frac{1}{\pi}\|x'\|_2, \qquad \|x'\|_2 \leqslant \frac{2}{\pi}\|x''\|_2,$$

$$\bigl|x'(0)\bigr| = \left|\int_0^{\xi_1} x''(t)\,\mathrm{d}t\right| \leqslant \int_0^{\eta} \bigl|x''(t)\bigr|\,\mathrm{d}t,$$

$$\bigl|x'(1)\bigr| = \left|\int_{\xi_2}^{1} x''(t)\,\mathrm{d}t\right| \leqslant \int_{\eta}^{1} \bigl|x''(t)\bigr|\,\mathrm{d}t,$$

$$\bigl|x''(0)\bigr| = \left|\int_0^{\xi_3} x''(t)\,\mathrm{d}t\right| \leqslant \|x'''\|_1, \qquad \bigl|x''(1)\bigr| = \left|\int_{\xi_3}^{1} x''(t)\,\mathrm{d}t\right| \leqslant \|x'''\|_1.$$

We multiply Eq. (3.22) by x' and integrate from 0 to 1 to get

$$0 = -\int_0^1 x'''(t)x'(t)\,\mathrm{d}t + \lambda\int_0^1 f\bigl(t, x(t), x'(t), x''(t)\bigr)x'(t)\,\mathrm{d}t$$
$$- \lambda\int_0^1 e(t)x'(t)\,\mathrm{d}t$$

$$
\begin{aligned}
&\geqslant \|x''\|_2^2 - |x''(1)||x'(1)| - |x''(0)||x'(0)| \\
&\quad + \lambda \int_0^1 \left\{a x'^2(t) + b|x(t)||x'(t)| + c|x'(t)||x''(t)| + \alpha(t)x'(t)\right\} \mathrm{d}t \\
&\quad - \lambda \|e\|_1 \|x'\|_\infty \\
&\geqslant \|x''\|_2^2 - \left(|x'(0)| + |x'(1)|\right)\|x''\|_1 - |a|\|x'\|_2^2 - |b|\|x\|_2\|x'\|_2 \\
&\quad - |c|\|x'\|_2\|x''\|_2 - \left(\|\alpha\|_1 + \|e\|_1\right)\|x'\|_\infty \\
&\geqslant \|x''\|_2^2 - \left(\int_0^\eta |x''(t)|\,\mathrm{d}t + \int_\eta^1 |x''(t)|\,\mathrm{d}t\right)\|x'''\|_1 \\
&\quad - \left(\frac{4}{\pi^2}|a| + \frac{4}{\pi^3}|b| + \frac{2}{\pi}|c|\right)\|x''\|_2^2 - \left(\|\alpha\|_1 + \|e\|_1\right)\|x''\|_2 \\
&\geqslant \|x''\|_2^2 - \|x''\|_1\|x'''\|_1 - \left(\frac{4}{\pi^2}|a| + \frac{4}{\pi^3}|b| + \frac{2}{\pi}|c|\right)\|x''\|_2^2 \\
&\quad - \left(\|\alpha\|_1 + \|e\|_1\right)\|x''\|_2.
\end{aligned}
$$

From Eq. (3.22) we have

$$
\begin{aligned}
\|x'''\|_1 &\leqslant \|p\|_2\|x\|_2 + \|q\|_2\|x'\|_2 + \|r\|_2\|x''\|_2 + \|S\|_1 + \|e\|_1 \\
&\leqslant \left(\frac{2}{\pi^2}\|p\|_2 + \frac{2}{\pi}\|q\|_2 + \|r\|_2\right)\|x''\|_2 + \|S\|_1 + \|e\|_1.
\end{aligned}
$$

Putting this in the previous inequality we find

$$
\|x''\|_2 \leqslant \frac{\|S\|_1 + 2\|e\|_1 + \|\alpha\|_1}{1 - \left(\frac{4}{\pi^2}|a| + \frac{4}{\pi^3}|b| + \frac{2}{\pi}|c| + \frac{2}{\pi^2}\|p\|_2 + \frac{2}{\pi}\|q\|_2 + \|r\|_2\right)}.
$$

Hence, there exists a constant C, independent of $\lambda \in [0,1]$ such that $\|x'''\|_1 \leqslant C$. It follows that there is a constant, still denoted by C, independent of $\lambda \in [0,1]$ such that

$$
\|x\|_{C^2[0,1]} \leqslant C.
$$

This completes the proof of the theorem. □

In Theorem 3.10 we can do without assumption (i), if we observe that

$$
\begin{aligned}
f(t,x,y,z)y &\geqslant |f(t,x,y,z)||y| \\
&\geqslant -|p(t)||y||x| - |q(t)|y^2 - |r(t)||y||z| - |S(t)||y|.
\end{aligned}
$$

It is easy to see from the proof of Theorem 3.10 that the BVP (3.20)–(3.21) has at least one solution in $C^2[0,1]$ provided

$$\frac{4}{\pi^2}\|p\|_2 + \frac{4}{\pi}\|q\|_2 + 2\|r\|_2 < 1.$$

THEOREM 3.11. *Let $f:[0,1]\times\mathbb{R}^3 \to \mathbb{R}$ satisfy Carathéodory's conditions. Suppose that*

(iii) *there exist functions $p, q, r \in L^2[0,1]$ such that*

$$\begin{aligned}&\left|f(t,x_1,y_1,z_1) - f(t,x_2,y_2,z_2)\right|\\ &\quad \leqslant p(t)|x_1-x_2| + q(t)|y_1-y_2| + r(t)|z_1-z_2|,\end{aligned}$$

for a.e. $t \in [0,1]$ and all $(x_i,y_i,z_i) \in \mathbb{R}^3$, $i=1,2$.

Let $\eta \in (0,1)$ be given. Then for every $e \in L^1[0,1]$ the BVP (3.20)–(3.21) *has exactly one solution in $C^2[0,1]$ provided*

$$\frac{4}{\pi^2}\|p\|_2 + \frac{4}{\pi}\|q\|_2 + 2\|r\|_2 < 1.$$

PROOF. We note that (iii) implies that

$$\left|f(t,x,y,z)\right| \leqslant p(t)|x| + q(t)|y| + r(t)|z| + \left|f(t,0,0,0)\right|$$

for a.e. $t \in [0,1]$ and all $(x,y,z) \in \mathbb{R}^3$. Accordingly the BVP (3.20)–(3.21) has at least one solution in $C^2[0,1]$.

Now, to prove the uniqueness, let x_1, x_2 are two solutions of the BVP (3.20)–(3.21). Setting $u(t) = x_(t) - x_2(t)$ we then get

$$u'''(t) = f\big(t,x_1(t),y_1(t),z_1(t)\big) - f\big(t,x_2(t),y_2(t),z_2(t)\big), \tag{3.24}$$

$$u(0) = u(\eta) = u(1) = 0. \tag{3.25}$$

We multiply Eq. (3.24) by u' and integrate over $[0,1]$ to get

$$\begin{aligned}0 &= -\int_0^1 u'''(t)u'(t)\,\mathrm{d}t\\ &\quad + \int_0^1 \big[f\big(t,x_1(t),y_1(t),z_1(t)\big) - f\big(t,x_2(t),y_2(t),z_2(t)\big)\big]u'(t)\,\mathrm{d}t\\ &\geqslant \|u''\|_2^2 - \|u''\|_2\|u'''\|_1 - \|p\|_2\|u\|_2\|u'\|_\infty - \|q\|_2\|u'\|_2\|u'\|_\infty\\ &\quad - \|r\|_2\|u'\|_\infty\|u''\|_2\\ &\geqslant \|u''\|_2^2 - \|u''\|_2\big(\|p\|_2\|u\|_2 + \|q\|_2\|u'\|_2 + \|r\|_2\|u''\|_2\big)\\ &\quad - \left(\frac{2}{\pi^2}\|p\|_2 + \frac{2}{\pi}\|q\|_2 + \|r\|_2\right)\|u''\|_2^2\end{aligned}$$

$$\geqslant \left(1 - \left[\frac{4}{\pi^2}\|p\|_2 + \frac{4}{\pi}\|q\|_2 + 2\|r\|_2\right]\right)\|u''\|_2^2,$$

hence $\|u''\|_2 = 0$. It now follows from the estimate $\|u\|_2 \leqslant (2/\pi^2)\|u''\|_2$ and the continuity of u that $u(t) = 0$ for every $t \in [0,1]$, i.e. $x_1(t) = x_2(t)$ for every $t \in [0,1]$. This completes the proof of the theorem. □

Theorems 3.10 and 3.11 were proved in [31]. For other results for third-order three-point BVPs the interested reader is referred to [1,24,81] and the references cited therein.

3.4. *m-point boundary value problems reduced to three-point boundary value problems*

Let $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ be either a continuous or a Carathéodory's function and let $e : [0,1] \to \mathbb{R}$ be a function in $L^1[0,1]$, $a_i \in \mathbb{R}$, with all of the a_i's having the same sign, $\xi_i \in (0,1), i = 1,2,\ldots,m-2$, $0 < \xi_1 < \xi_2 < \cdots < \xi_{m-2} < 1$. We consider the following second-order ordinary differential equation

$$x''(t) = f\big(t, x(t), x'(t)\big) + e(t), \quad t \in (0,1) \tag{3.26}$$

subject to one of the following boundary value conditions:

$$x(0) = 0, \qquad x(1) = \sum_{i=1}^{m-2} a_i x(\xi_i), \tag{3.27}$$

$$x'(0) = 0, \qquad x(1) = \sum_{i=1}^{m-2} a_i x(\xi_i), \tag{3.28}$$

$$x(0) = 0, \qquad x'(1) = \sum_{i=1}^{m-2} a_i x'(\xi_i). \tag{3.29}$$

It is well known, that the existence of a solution for these BVPs can be studied via the existence of a solution for Eq. (3.26) subject to one of the following three-point boundary value conditions:

$$x(0) = 0, \qquad x(1) = \alpha x(\eta), \tag{3.30}$$

$$x'(0) = 0, \qquad x(1) = \alpha x(\eta), \tag{3.31}$$

$$x(0) = 0, \qquad x'(1) = \alpha x'(\eta), \tag{3.32}$$

where $\alpha = \sum_{i=1}^{m-2} a_i$ and $\eta \in [\xi_1, \xi_{m-2}]$.

For certain boundary condition case such that the linear operator $Lx = x''$, defined in a suitable Banach space, is invertible, is the so-called *nonresonance* case. Otherwise, the so-called *resonance* case. In the next sections we will give existence results for both, nonresonance and resonance cases.

3.5. *Nonresonance results*

Using the Leray–Schauder continuation theorem (see [82]) we first give an existence result for the three-point BVP (3.26), (3.30) in the case when $\alpha \neq 1$.

THEOREM 3.12. *Let* $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ *be a function satisfying Carathéodory's conditions. Assume that* (3.12) *holds. Also let* $\alpha \in \mathbb{R}$, *and* $\eta \in (0,1)$ *be given. Then the BVP* (3.26), (3.30) *has at least one solution in* $C^1[0,1]$ *provided*

$$\begin{cases} \|p\|_1 + \|q\|_1 < 1, & \text{if } \alpha \leqslant 1, \\ \|p\|_1 + \|q\|_1 < \dfrac{1-\alpha\eta}{\alpha(1-\eta)}, & \text{if } 1 < \alpha < 1/\eta. \end{cases}$$

PROOF. Let X be the Banach space $C^1[0,1]$ and Y denote the Banach space $L^1(0,1)$ with their usual norms. We denote a linear mapping $L : D(L) \subset X \to Y$ by setting

$$D(L) = \{x \in W^{2,1}(0,1)\colon\ x(0) = 0, x(1) = \alpha x(\eta)\},$$

and for $x \in D(L)$,

$$Lx = x''.$$

We also define a nonlinear mapping $N : X \to Y$ by setting

$$(Nx)(t) = f\bigl(t, x(t), x'(t)\bigr), \quad t \in [0,1].$$

We note that N is a bounded mapping from X into Y. Next, it is easy to see that the linear mapping $L : D(L) \subset X \to Y$, is one-to-one mapping. We note that since $\alpha \leqslant 1$, $\alpha \neq 1/\eta$. Next, the linear mapping $K : Y \to X$, defined for $y \in Y$ by

$$(Ky)(t) = \int_0^t (t-s)y(s)\,\mathrm{d}s + \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta - s)y(s)\,\mathrm{d}s \\ - \frac{t}{1-\alpha\eta}\int_0^1 (1-s)y(s)\,\mathrm{d}s$$

is such that for $y \in Y$, $Ky \in D(L)$ and $LKy = y$; and for $u \in D(L)$, $KLu = u$. Furthermore, it follows easily using the Arzelá–Ascoli theorem that KN maps a bounded subset of X into a relatively compact subset of X. Hence $KN : X \to X$ is a compact mapping.

We next note that $x \in C^1[0,1]$ is a solution of the BVP (3.26), (3.30) if and only if x is a solution to the operator equation

$$Lx = Nx + e.$$

Now, the operator equation $Lx = Nx + e$ is equivalent to the equation

$$x = KNx + Ke.$$

We apply the Leray–Schauder continuation theorem (see, e.g., [82, Corollary IV.7]) to obtain the existence of a solution for $x = KNx + Ke$ or equivalently to the BVP (3.26), (3.30).

To do this, it suffices to verify that the set of all possible solutions of the family of equations

$$x''(t) = \lambda f\big(t, x(t), x'(t)\big) + \lambda e(t), \quad t \in (0,1), \tag{3.33}$$

$$x(0) = 0, \qquad x(1) = \alpha x(\eta), \tag{3.34}$$

is, a priori, bounded in $C^1[0,1]$ by a constant independent of $\lambda \in [0,1]$.

We observe that if $x \in W^{2,1}(0,1)$, with $x(0) = 0$, $x(1) = \alpha x(\eta)$, with $\alpha \leqslant 1$, there exists a $\zeta \in (0,1)$ with $x'(\zeta) = 0$. Indeed, suppose that $\alpha \leqslant 1$ and $x'(t) > 0$ for every $t \in (0,1)$. Then $x(t) > 0$ for every $t \in (0,1)$ and $x(t)$ is strictly increasing on $[0,1]$. We then get that $0 < x(1) - x(\eta) = (\alpha - 1)x(\eta) \leqslant 0$, a contradiction. Then we have

$$\|x\|_\infty \leqslant \|x'\|_\infty \leqslant \|x''\|_1. \tag{3.35}$$

Now let $x(t)$ be a solution of (3.33)–(3.34) for some $\lambda \in [0,1]$, so that $x \in W^{2,1}(0,1)$ with $x(0) = 0$, $x(1) = \alpha x(\eta)$. We then get from Eq. (3.33) that

$$\begin{aligned} \|x''\|_1 &= \lambda \big\| f\big(t, x(t), x'(t)\big) + e(t) \big\|_1 \\ &\leqslant \|p\|_1 \|x\|_\infty + \|q\|_1 \|x'\|_\infty + \|r\|_1 + \|e\|_1 \\ &\leqslant \big(\|p\|_1 + \|q\|_1\big)\|x''\|_1 + \|r\|_1 + \|e\|_1. \end{aligned}$$

It follows from the assumption that there is a constant c, independent of $\lambda \in [0,1]$, such that

$$\|x''\|_1 \leqslant c.$$

It is now immediate from (3.35) that the set of solutions of the family of equations (3.33)–(3.34) is, a priori, bounded in $C^1[0,1]$ by a constant independent of $\lambda \in [0,1]$. This completes the proof of the theorem when $\alpha \leqslant 1$.

Let now $\alpha > 1$ and $\eta \in (0,1)$ with $\alpha\eta < 1$. It suffices to verify that the set of all possible solutions of the family of equations (3.33)–(3.34) is, a priori, bounded in $C^1[0,1]$ by a constant independent of $\lambda \in [0,1]$.

For $x \in W^{2,1}(0,1)$, with $x(0) = 0$, $x(1) = \alpha x(\eta)$, if $\alpha > 1$, there exists a $\zeta \in (\eta, 1)$ such that $(\alpha - 1)x(\eta) = x(1) - x(\eta) = (1-\eta)x'(\zeta)$, or $x(\eta) = (1-\eta)/(\alpha - 1)x'(\zeta)$.

Then we have

$$\|x\|_\infty \leqslant \|x'\|_\infty \leqslant \frac{\alpha(1-\eta)}{1-\alpha\eta}.$$

We then see from Eq. (3.33) that

$$\|x''\|_1 = \lambda \big\| f\big(t, x(t), x'(t)\big) + e(t) \big\|_1$$

$$\leqslant \|p\|_1\|x\|_\infty + \|q\|_1\|x'\|_\infty + \|r\|_1 + \|e\|_1$$

$$\leqslant \big(\|p\|_1 + \|q\|_1\big)\left(\frac{\alpha(1-\eta)}{1-\alpha\eta}\right)\|x''\|_1 + \|r\|_1 + \|e\|_1$$

for a solution x of the family of equations (3.33)–(3.34) for some $\lambda \in [0,1]$. It is then immediate that the set of solutions of the family of equations (3.33)–(3.34) is, a priori, bounded in $C^1[0,1]$ by a constant, independent of $\lambda \in [0,1]$. This completes the proof of the theorem. □

Now we study the m-point BVP (3.26)–(3.27) using the a priori estimates that we obtained for the three-point BVP (3.26), (3.30). This is because for every solution $x(t)$ of the BVP (3.26)–(3.27) there exists $\eta \in [\xi_1, \xi_{m-2}]$, depending on $x(t)$, such that $x(t)$ is also a solution of the BVP (3.26), (3.30) with $\alpha = \sum_{i=1}^{m-2} a_i \neq 1$.

THEOREM 3.13. *Let $f:[0,1]\times\mathbb{R}^2 \to \mathbb{R}$ be a function satisfying Carathéodory's conditions. Assume that* (3.12) *holds. Also let $\alpha = \sum_{i=1}^{m-2} a_i$, and $\eta \in (0,1)$ be given. Then the BVP* (3.26)–(3.27) *has at least one solution in $C^1[0,1]$ provided*

$$\begin{cases} \|p\|_1 + \|q\|_1 < 1, & \text{if } \alpha \leqslant 1,\\ \|p\|_1 + \|q\|_1 < \dfrac{1-\alpha\xi_{m-2}}{\alpha(1-\xi_1)}, & \text{if } 1 < \alpha < 1/\xi_{m-2}. \end{cases}$$

PROOF. We will give only the proof for $\alpha = \sum_{i=1}^{m-2} a_i \leqslant 1$; the other case is similar. The proof is quite similar to that of Theorem 3.12 and uses the a priori bounds for the set of solutions of the family of equations (3.33)–(3.34).

Let X be the Banach space $C^1[0,1]$ and Y denote the Banach space $L^1(0,1)$ with their usual norms. We denote a linear mapping $L: D(L) \subset X \to Y$ by setting

$$D(L) = \left\{x \in W^{2,1}(0,1)\colon\ x(0)=0, x(1) = \sum_{i=1}^{m-2} a_i x(\xi_i)\right\},$$

and for $x \in D(L)$,

$$Lx = x''.$$

We also define a nonlinear mapping $N: X \to Y$ by setting

$$(Nx)(t) = f\big(t, x(t), x'(t)\big), \quad t \in [0,1].$$

We note that N is a bounded mapping from X into Y. Next, it is easy to see that the linear mapping $L: D(L) \subset X \to Y$, is one-to-one mapping. Next, the linear mapping $K: Y \to X$, defined for $y \in Y$ by

$$(Ky)(t) = \int_0^t (t-s)y(s)\,\mathrm{d}s + At$$

where A is given by

$$A\left(1-\sum_{i=1}^{m-2}a_i\xi_i\right)=\sum_{i=1}^{m-2}a_i\int_0^{\xi_1}(\xi_i-s)y(s)\,\mathrm{d}s-\int_0^1(1-s)y(s)\,\mathrm{d}s,$$

is such that for $y\in Y$, $Ky\in D(L)$ and $LKy=y$; and for $u\in D(L)$, $KLu=u$. Furthermore, it follows easily using the Arzelá–Ascoli theorem that KN maps bounded subsets of X into a relatively compact subsets of X. Hence $KN:X\to X$ is a compact mapping.

We, next, note that $x\in C^1[0,1]$ is a solution of the BVP (3.26)–(3.27) if and only if x is a solution to the operator equation

$$Lx=Nx+e.$$

Now, the operator equation $Lx=Nx+e$ is equivalent to the equation

$$x=KNx+Ke.$$

We apply the Leray–Schauder continuation theorem to obtain the existence of a solution for $x=KNx+Ke$ or equivalently to the BVP (3.26)–(3.27).

To do this, it suffices to verify that the set of all possible solutions of the family of equations

$$x''(t)=\lambda f\big(t,x(t),x'(t)\big)+\lambda e(t),\quad t\in(0,1), \tag{3.36}$$

$$x(0)=0,\qquad x(1)=\sum_{i=1}^{m-2}a_ix(\xi_i), \tag{3.37}$$

is, a priori, bounded in $C^1[0,1]$ by a constant independent of $\lambda\in[0,1]$.

Let, now, $x(t)$ be a solution of (3.36)–(3.37) for some $\lambda\in[0,1]$, so that $x\in W^{2,1}(0,1)$ with $x(0)=0$, $x(1)=\sum_{i=1}^{m-2}a_ix(\xi_i)$. Accordingly, there exists, an $\eta\in[\xi_1,\xi_{m-2}]$, depending on x, such that $x(t)$ is a solution of the three-point BVP

$$x''(t)=\lambda f\big(t,x(t),x'(t)\big)+\lambda e(t),\quad t\in(0,1),$$

$$x(0)=0,\qquad x(1)=\alpha x(\eta).$$

It then follows, as in the proof of Theorem 3.12 that there is a constant c, independent of $\lambda\in[0,1]$, and $\eta\in[\xi_1,\xi_{m-2}]$, such that

$$\|x\|_\infty\leqslant\|x'\|_\infty\leqslant\|x''\|_1\leqslant c.$$

Thus the set of solutions of the family of equations (3.36)–(3.37) is, a priori, bounded in $C^1[0,1]$ by a constant, independent of $\lambda\in[0,1]$.

This completes the proof of the theorem. □

Theorems 3.12 and 3.13 were proved in [33]. In Theorem 3.12 the functions p, q, r in the condition (3.12) are in $L^1(0,1)$. If p, q are in $L^\infty(0,1)$ and r, e in $L^2(0,1)$ then the BVP (3.26), (3.30) has at least one solution in $C^1[0,1]$ provided

$$\begin{cases} \frac{2}{\pi}\left(\frac{2}{\pi}\|p\|_\infty + \|q\|_\infty\right) < 1, & \text{if } \alpha \leqslant 1, \\ \frac{2}{\pi}\|p\|_\infty + \|q\|_\infty < \frac{\sqrt{2}(1-\alpha\eta)}{\alpha(1-\eta)}, & \text{if } 1 < \alpha < 1/\eta. \end{cases}$$

In the case when $\alpha \geqslant 1$ and $\alpha\eta \neq 1$ we have a sharper condition for the solvability of BVP (3.26), (3.30). Thus, if tp, q, r are in $L^1[0,1]$ and (3.12) holds, then the BVP (3.26), (3.30) has at least one solution in $C^1[0,1]$ provided

$$\begin{cases} \|tp(t)\|_1 + \|q(t)\|_1 < \frac{1-\alpha\eta}{1-\eta}, & \text{if } \alpha\eta \leqslant 1, \\ \|tp(t)\|_1 + \|q(t)\|_1 < \frac{\alpha\eta - 1}{(\alpha-1)\eta} & \text{if } \alpha\eta > 1. \end{cases}$$

The above conditions were proved in [33] and based on a priori estimates obtained by a different method than that was used in Theorem 3.12.

In the case when $\alpha \in \mathbb{R}$, $\alpha \leqslant 1$ Gupta and Trofimchuk in [38] gave constructive conditions for the solvability of the BVP (3.26), (3.30) using the spectral radius of a compact linear operator, and in [39] new solvability conditions based on a new type Wirtinger inequality. For other results see [37].

In the case when $\alpha \in \mathbb{R}$ with $\alpha\eta \neq 1$ a better condition for the solvability of the BVP (3.26), (3.30) is given by Gupta and Trofimchuk in [38]. They proved that, if tp, q, $t[tp+q]$, r are in $L^1[0,1]$ and (3.12) holds, then the BVP (3.26), (3.30) has at least one solution in $C^1[0,1]$ provided

$$\begin{aligned} \max\Bigg\{ &\int_0^\eta \left|\frac{1-\alpha}{1-\alpha\eta} - 1\right| \big[sp(s) + q(s)\big]\,\mathrm{d}s \\ &+ \frac{1}{|1-\alpha\eta|}\int_\eta^1 (1-s)\big[sp(s)+q(s)\big]\,\mathrm{d}s, \\ &\int_\eta^1 \frac{|s-\alpha\eta|}{|1-\alpha\eta|}\big[sp(s)+q(s)\big]\,\mathrm{d}s + \frac{|\alpha-1|}{|1-\alpha\eta|}\int_0^\eta s\big[sp(s)+q(s)\big]\,\mathrm{d}s \Bigg\} < 1. \end{aligned}$$

Moreover, Gupta in [30], when $\alpha \neq 1$ and $\alpha\eta \neq 1$, obtained new a priori estimates which are sharper than those used in the previous theorems. More precisely, he proved that if p, q, r are in $L^1(0,1)$ and (3.12) holds, then the BVP (3.26), (3.30) has at least one solution in $C^1[0,1]$ provided

$$M\|p\|_1 + \frac{1}{1-\tau}\|q\|_1 < 1,$$

where M and τ are appropriate constants.

Finally, we note that in [40] Gupta and Trofimchuk studied examples of three-point BVP by making extensive use of computer algebra systems like Maple and MathCad.

Concerning the uniqueness of the solutions of the BVPs (3.26), (3.30) and (3.26)–(3.27) we have the following results.

THEOREM 3.14. *Let $f:[0,1]\times\mathbb{R}^2\to\mathbb{R}$ be a function satisfying Carathéodory's conditions. Assume that there exist nonnegative constants c, d such that*

$$\bigl|f(t,x_1,x_2)-f(t,y_1,y_2)\bigr|\leqslant c|x_1-x_2|+d|y_1-y_2|$$

for a.e. $t\in[0,1]$ and all $(x_i,x_i)\in\mathbb{R}^2, i=1,2$. Also let $\alpha\in\mathbb{R}$, and $\eta\in(0,1)$ be given. Then the BVP (3.26), (3.30) *has exactly one solution in $C^1[0,1]$ provided*

$$\begin{cases} \dfrac{2}{\pi}\left(\dfrac{2}{\pi}c+d\right)<1, & \text{if } \alpha\leqslant 1,\\ \dfrac{2}{\pi}c+d<\dfrac{\sqrt{2}(1-\alpha\eta)}{\alpha(1-\eta)}, & \text{if } 1<\alpha<1/\eta. \end{cases}$$

PROOF. The existence of a solution for the BVP (3.26), (3.30) follows from Theorem 3.12. Let now $x_1(t), x_2(t)$ be two solutions of the BVP (3.26), (3.30). We then get

$$\begin{aligned} &x_1''(t)-x_2''(t)=f\bigl(t,x_1(t),x_1'(t)\bigr)-f\bigl(t,y_1(t),y_1'(t)\bigr), \quad 0<t<1,\\ &(x_1-x_2)(0)=0, \qquad (x_1-x_2)(1)=\alpha(x_1-x_2)(\eta). \end{aligned}$$

It follows that

$$\begin{aligned} \bigl\|x_1''-x_2''\bigr\|_2 &= \bigl\|f\bigl(t,x_1(t),x_1'(t)\bigr)-f\bigl(t,y_1(t),y_1'(t)\bigr)\bigr\|_2\\ &\leqslant c\|x_1-x_2\|_2+\bigl\|x_1'-x_2'\bigr\|_2\\ &\leqslant \frac{2}{\pi}\left(\frac{2}{\pi}c+d\right)\bigl\|x_1''-x_2''\bigr\|_2. \end{aligned}$$

Hence, we get that $\|x_1''-x_2''\|_2=0$. It now follows from the inequalities

$$\|x_1-x_2\|_2\leqslant\frac{2}{\pi}\bigl\|x_1'-x_2'\bigr\|_2\leqslant\frac{4}{\pi^2}\bigl\|x_1''-x_2''\bigr\|_2=0,$$

that $x_1(t)=x_2(t)$ for a.e. $t\in[0,1]$ and hence for every $t\in[0,1]$ because of the continuity of $x_1(t)$ and $x_2(t)$ on $[0,1]$. This completes the proof of the theorem in the case when $\alpha\leqslant 1$. The proof for the case $1<\alpha<1/\eta$ is similar. □

THEOREM 3.15. *Let* $f:[0,1]\times\mathbb{R}^2\to\mathbb{R}$ *be as in Theorem* 3.14. *Then for any given* $e(t)$ *in* $L^2(0,1)$, *the* m*-point BVP* (3.26)–(3.27) *has exactly one solution in* $C^1[0,1]$ *provided*

$$\begin{cases} \|p\|_1+\|q\|_1<1, & \text{if } \alpha\leqslant 1,\\ \|p\|_1+\|q\|_1<\dfrac{1-\alpha\xi_{m-2}}{\alpha(1-\xi_1)}, & \text{if } 1<\alpha<1/\eta. \end{cases}$$

We now give a theorem which allows f to have nonlinear growth. We do this by imposing a decomposition condition for f. We need the following simple lemma.

LEMMA 3.16. *Let* $x\in C^1[0,1]$ *satisfy* $x(0)=0, x(1)=\alpha x(\eta)$, *where* $\eta\in(0,1)$, *and* $\alpha>1$, $\alpha\neq 1/\eta$, *then there exist* ζ *and* $C_0\in(0,1)$ *such that* $x'(\zeta)=C_0x(1)$.

PROOF. If $\alpha\eta>1$, let $C_0=1/(\alpha\eta)$. There exists $\zeta\in(0,1)$ with

$$x'(\zeta)=\frac{x(\eta)-x(0)}{\eta}=C_0x(1).$$

If $\alpha\eta<1$, let $C_0=\frac{\alpha-1}{\alpha(1-\eta)}$. There exists $\zeta\in(\eta,1)$ with

$$x'(\zeta)=\frac{x(1)-x(\eta)}{1-\eta}=\frac{\alpha-1}{\alpha(1-\eta)}x(\eta)=C_0x(1). \qquad\square$$

THEOREM 3.17. *Assume that* $f:[0,1]\times\mathbb{R}^2\to\mathbb{R}$ *is continuous and has the decomposition*

$$f(t,x,p)=g(t,x,p)+h(t,x,p)$$

such that

(1) $pg(t,x,p)\leqslant 0$ *for all* $(t,x,p)\in[0,1]\times\mathbb{R}^2$;

(2) $h(t,x,p)|\leqslant a(t)|x|+b(t)|p|+u(t)|x|^r+v(t)|p|^k+c(t)$ *for all* $(t,x,p)\in[0,1]\times\mathbb{R}^2$ *where* a,b,u,v *are in* $L^1[0,1]$ *and* $0\leqslant r,k<1$.

Then, for $\alpha\neq 1/\eta$, *there exists a solution* $x\in C^1[0,1]$ *to BVP* (3.26), (3.30) *provided that*

$$\begin{cases} \|a\|_1+\|b\|_1<\dfrac{1}{2}, & \text{if } \alpha\leqslant 1,\\ \|a\|_1+\|b\|_1<\dfrac{1}{2}\left(1-\dfrac{(\alpha-1)^2}{\alpha^2(1-\eta)^2}\right), & \text{if } 1<\alpha<\dfrac{1}{\eta},\\ \|a\|_1+\|b\|_1<\dfrac{1}{2}\left(1-\dfrac{1}{\alpha^2\eta^2}\right), & \text{if } \dfrac{1}{\eta}<\alpha. \end{cases}$$

PROOF. By the same argument as in the proof of Theorem 3.12, it suffices to show that all possible solutions of the following family of equations (3.33)–(3.34) is bounded in $C^1[0,1]$ by a constant independent of $\lambda\in[0,1]$.

Suppose that x is a solution of (3.33) and let ζ be as in Lemma 3.16 and write $C = \|c\|_1 + \|e\|_1$. Multiplying both sides of Eq. (3.33) with x' and integrating, we obtain

$$\begin{aligned}\frac{1}{2}x'^2(t) &\leqslant \frac{1}{2}x'^2(\zeta) + \|x'\|_\infty\big(\|a\|_1\|x\|_\infty + \|b\|_1\|x'\|_\infty \\ &\quad + \|u\|_1\|x\|_\infty^r + \|v\|_1\|x'\|_\infty^k + C\big) \\ &= \frac{1}{2}C_0^2 x'^2(1) + \|x'\|_\infty\big(\|a\|_1\|x\|_\infty + \|b\|_1\|x'\|_\infty \\ &\quad + \|u\|_1\|x\|_\infty^r + \|v\|_1\|x'\|_\infty^k + C\big).\end{aligned}$$

Since $\|x\|_\infty \leqslant \|x'\|_\infty$ and by assumption

$$\|a\|_1 + \|b\|_1 < \frac{1}{2}\big(1 - C_0^2\big).$$

If $\|x'\|_\infty \neq 0$, we have

$$\left(\frac{1}{2}\big(1 - C_0^2\big) - \big(\|a\|_1 + \|b\|_1\big)\right)\|x'\|_\infty \leqslant \|u\|_1\|x\|_\infty^r + \|v\|_1\|x'\|_\infty^k + C.$$

This implies that $\|x'\|_\infty$ is bounded since $0 \leqslant r, k < 1$, that is, there is $M > 0$ such that $\|x\|_\infty \leqslant \|x'\|_\infty \leqslant M$. This completes the proof. □

Theorem 3.17 is taken over from Feng and Webb [15] and is more general than Theorem 3.12, since the assumptions of Theorem 3.12 are special cases of Theorem 3.17, when $g(t, x, p) \equiv 0$, $u(t) \equiv 0$ and $v(t) \equiv 0$. But in Theorem 3.12 we need $\|a\|_1 + \|b\|_1 < 1$ and in Theorem 3.17 in these special cases we need $\|a\|_1 + \|b\|_1 < 1/2$. In [15] also interesting examples are given to compare their results with other existence results.

We now give existence results for the BVPs (3.26), (3.31) and (3.26), (3.32). By using the same argument as in Theorem 3.12 we can prove the following existence results (see [32,34]).

THEOREM 3.18. *Let $f : [0, 1] \times \mathbb{R}^2 \to \mathbb{R}$ be a function satisfying Carathéodory's conditions. Assume that* (3.12) *holds. Also let $\alpha \in \mathbb{R}$, and $\eta \in (0, 1)$ be given. Then the BVP* (3.26), (3.31) *has at least one solution in $C^1[0, 1]$ provided*

$$\begin{cases} \|p\|_1 + \|q\|_1 < 1, & \textit{if}\ \alpha \leqslant 0, \\ \left(1 + \max\left\{\dfrac{\alpha(1-\eta)}{|\alpha - 1|}, \dfrac{1-\eta}{|\alpha - 1|}\right\}\right)\|p\|_1 + \|q\|_1 < 1, & \textit{if}\ 0 < \alpha \neq 1. \end{cases}$$

In Theorem 3.18 if p, q, r are in $L^2(0,1)$ then the BVP (3.26), (3.31) has at least one solution in $C^1[0,1]$ provided

$$\begin{cases} \dfrac{2}{\pi}\left(\dfrac{2}{\pi}\|p\|_2 + \|q\|_2\right) < 1, & \text{if } \alpha \leqslant 0, \\ \left(\dfrac{\sqrt{2}}{\pi} + \max\left\{\dfrac{\alpha(1-\eta)}{|\alpha-1|}, \dfrac{1-\eta}{|\alpha-1|}\right\}\right)\|p\|_1 + \dfrac{2}{\pi}\|q\|_1 < 1, & \text{if } 0 < \alpha \neq 1. \end{cases}$$

THEOREM 3.19. *Let $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ be a function satisfying Carathéodory's conditions. Assume that* (3.12) *holds. Also let $\alpha \in \mathbb{R}$, $\alpha \neq 1$ and $\eta \in (0,1)$ be given. Then the BVP* (3.26), (3.32) *has at least one solution in $C^1[0,1]$ provided*

$$\left(\max\left\{\frac{1+|\alpha-1|}{|\alpha-1|}, \frac{|\alpha|+|\alpha-1|}{|\alpha-1|}\right\}\right)\left(\|p\|_1 + \|q\|_1\right) < 1.$$

As in Theorem 3.13 we prove existence results for the m-point BVPs (3.26), (3.28) and (3.29), (3.31), by using the a priori bounds obtained for the BVPs (3.26), (3.31) and (3.26), (3.32).

3.6. *Results at resonance*

In the following we shall give existence results for BVP (3.26), (3.30) when $\alpha\eta = 1$. In this case, the linear operator L is noninvertible and the Leray–Schauder continuation theorem cannot be used. We shall apply the continuation theorem of Mawhin [82]. For the convenience of the reader, we recall this theorem.

Let X and Z be Banach spaces and $L : D(L) \subset X \to Z$ be a linear operator which is Fredholm of index zero (that is $\operatorname{im}(L)$ (the image of L) is closed in Z, and $\ker(L)$ (the kernel of L) and $Z/\operatorname{im}(L)$ (the cokernel of L) are finite-dimensional with equal dimension). Let $P : X \to \ker(L)$ and $Q : Z \to Z_1$, where $X = \ker(L) \oplus X_1$ and $Z = \operatorname{im}(L) \oplus Z_1$, be continuous projections. Let L_1 denote L restricted to $D(L) \cap X_1$, an invertible operator into $\operatorname{im}(L)$, and write $K = L_1^{-1}$. Let Ω be a bounded, open subset of X such that $D(L) \cap \Omega \neq \emptyset$ and let $N : \overline{\Omega} \to Z$ be an L-compact mapping, that is, the maps $QN : \overline{\Omega} \to Z$ and $K(I-Q)N : \overline{\Omega} \to X$ are compact.

THEOREM 3.20. *Let L be Fredholm of index zero and let N be L-compact on $\overline{\Omega}$. Assume that the following conditions are satisfied.*

(1) $Lx + \lambda Nx \neq 0$ *for each* $(x, \lambda) \in [D(L) \setminus \ker L \cap \partial\Omega] \times (0,1)$.
(2) $Nx \notin \operatorname{im}(L)$ *for each* $x \in \ker(L) \cap \partial\Omega$.
(3) $\deg(QN|_{\ker L}, \Omega \cap \ker(L), 0) \neq 0$, *where $Q : Z \to Z$ is a continuous projection as above.*

Then the equation $Lx + Nx = 0$ has at least one solution in $D(L) \cap \overline{\Omega}$.

The following existence theorems hold.

THEOREM 3.21. *Let* $f:[0,1]\times\mathbb{R}^2\to\mathbb{R}$ *be a continuous function. Assume that*:

(1) *There exist functions* p,q,r *in* $L^1[0,1]$ *such that*

$$\bigl|f(t,u,v)\bigr|\leqslant p(t)|u|+q(t)|v|+r(t),\quad \text{for } t\in[0,1] \text{ and } (u,v)\in\mathbb{R}^2.$$

(2) *There exists* $N>0$ *such that for* $v\in\mathbb{R}$ *with* $|v|>N$, *one has*

$$\bigl|f(t,u,v)\bigr|\geqslant -l|u|+n|v|-M\quad \text{for } t\in[0,1],\ u\in\mathbb{R}$$

where $n>l\geqslant 0$, $M\geqslant 0$.

(3) *There exists* $R>0$ *such that for* $|v|>R$ *one has either*

$$vf(t,vt,v)\leqslant 0\quad \text{for } t\in[0,1]$$

or else

$$vf(t,vt,v)\geqslant 0\quad \text{for } t\in[0,1].$$

Then, for every continuous function e, *the BVP* (3.26), (3.30) *with* $\alpha\eta=1$ *has at least one solution in* $C^1[0,1]$ *provided that*

$$2\bigl(\|p\|_1+\|q\|_1\bigr)+\frac{l}{n}<1.$$

We give the proof via several lemmas. We now let L be the linear operator from $D(L)\subset X\to Z=L^1(0,1)$ defined by

$$D(L)=\bigl\{x\in W^{2,1}(0,1):\ x(0)=0, x(1)=x(\eta)/\eta\bigr\}$$

and for $x\in D(L)$, $Lx=x''$. Let $X_1=\{x\in X:\ x(0)=0\}$.

LEMMA 3.22. *Suppose L is as above. Then* $L:D(L)\subset X\to Z$ *is Fredholm of index zero. Furthermore, the linear operator* $K:\operatorname{im}(L)\to D(L)\cap X_1$ *defined by*

$$(Ky)(t)=\int_0^t\int_0^\tau y(s)\,\mathrm{d}s\,\mathrm{d}\tau,\qquad y\in\operatorname{im}(L)$$

is such that $K=L_P^{-1}$, *where* $L_P=L|_{D(L)\cap X_1}$. *Also* $\|Ky\|\leqslant\|y\|_1$ *for all* $y\in\operatorname{im}(L)$.

PROOF. It is easy to see that $\ker(L)=\{ct:\ c\in\mathbb{R}\}$. We show that

$$\operatorname{im}(L)=\left\{y\in L^1[0,1]:\ \int_0^1 Y(t)\,\mathrm{d}t=\int_0^1 Y(\eta t)\,\mathrm{d}t,\ \text{where } Y(t)=\int_0^t y(s)\,\mathrm{d}s\right\}.$$

In fact, for $y \in \operatorname{im}(L)$, there is $x \in D(L)$ with $y(t) = x''(t)$. Therefore

$$\int_0^1 Y(t)\,\mathrm{d}t = \int_0^1 \big[x'(t) - x'(0)\big]\,\mathrm{d}t = x(1) - x'(0),$$

and

$$\int_0^1 Y(\eta t)\,\mathrm{d}t = \int_0^1 \big[x'(\eta t) - x'(0)\big]\,\mathrm{d}t = x(\eta)/\eta - x'(0).$$

Since $x(1) = \frac{1}{\eta}x(\eta)$, we have $\int_0^1 Y(t)\,\mathrm{d}t = \int_0^1 Y(\eta t)\,\mathrm{d}t$. On the other hand, suppose $y \in L^1[0,1]$ is such that $\int_0^1 Y(t)\,\mathrm{d}t = \int_0^1 Y(\eta t)\,\mathrm{d}t$. Let $x(t) = \int_0^t Y(s)\,\mathrm{d}s$, then $x \in D(L)$ and $x''(t) = y(t)$.

For $y \in L^1[0,1]$, let

$$Qy = \frac{2}{1-\eta}\int_0^1 \int_{\eta t}^t y(s)\,\mathrm{d}s\,\mathrm{d}t,$$

and let $y_1(t) = y(t) - Qy$. Then $Y_1(t) = \int_0^t y(s)\,\mathrm{d}s - (Qy)(t)$, and writing Qy in the form

$$Qy = \frac{2}{1-\eta}\left\{\int_0^1 \int_0^t y(s)\,\mathrm{d}s\,\mathrm{d}t - \int_0^1 \int_0^{\eta t} y(s)\,\mathrm{d}s\,\mathrm{d}t\right\},$$

we see that

$$\int_0^1 \int_0^t y(s)\,\mathrm{d}s\,\mathrm{d}t - \frac{Qy}{2} = \int_0^1 \int_0^{\eta t} y(s)\,\mathrm{d}s\,\mathrm{d}t - \eta\frac{Qy}{2},$$

that is

$$\int_0^1 Y_1(t)\,\mathrm{d}t = \int_0^1 Y_1(\eta t)\,\mathrm{d}t.$$

This shows that $y_1 \in \operatorname{im}(L)$. Hence $Z = \operatorname{im}(L) + \mathbb{R}$. Since $\operatorname{im}(L) \cap \mathbb{R} = \{0\}$, we have $Z = \operatorname{im}(L) \oplus \mathbb{R}$ and therefore L is Fredholm of index zero.

Now we define a projection from X onto $\ker(L)$ by setting $(Px)(t) = x'(0)t$ and let $X_1 = \{x \in X, x'(0) = 0\}$. Then, for $x \in D(L) \cap X_1$, we have

$$(KL_Px)(t) = Kx''(t) = \int_0^t \int_0^\tau x''(s)\,\mathrm{d}s\,\mathrm{d}\tau = \int_0^t \big(x'(\tau) - x'(0)\big)\,\mathrm{d}\tau = x(t),$$

and for $y \in \operatorname{im}(L)$, we have

$$(L_PKy)(t) = \left(\int_0^t \int_0^\tau y(s)\,\mathrm{d}s\,\mathrm{d}\tau\right)'' = y(t).$$

This shows that $K = L_P^{-1}$. Also we have

$$\left\|(Ky)(t)\right\|_\infty \leqslant \int_0^1 \int_0^1 \left|y(s)\right| \mathrm{d}s \, \mathrm{d}\tau \leqslant \left\|y(t)\right\|_1,$$

and

$$(Ky)'(t) = \int_0^t y(s)\,\mathrm{d}s, \quad \text{so} \quad \left\|(Ky)'(t)\right\|_\infty \leqslant \left\|y(t)\right\|_1,$$

and hence $\|Ky\| \leqslant \|y\|_1$ which completes the proof. □

LEMMA 3.23. *Let $U_1 = \{x \in D(L) \setminus \ker(L), Lx + \lambda Nx = 0$ for some $\lambda \in [0,1]\}$. Then U_1 is a bounded subset of X.*

PROOF. Suppose that $x \in U_1$, and $Lx = -\lambda Nx$, then $\lambda \neq 0$ and $QNx = 0$. Therefore

$$\int_0^1 \int_{\eta\tau}^{\tau} \left(f\left(t, x(t), x'(t)\right) + e(t)\right) \mathrm{d}t \, \mathrm{d}\tau = 0,$$

and hence there exists $\gamma \in (0,1)$ such that

$$\left|f\left(\gamma, x(\gamma), x'(\gamma)\right)\right| = \left|e(\gamma)\right| \leqslant \left\|e(t)\right\|_\infty.$$

Also for $x \in D(L) \setminus \ker(L)$, by Lemma 3.22 and condition (1)

$$\begin{aligned}\left\|(I-P)x\right\| &= \left\|KL(I-P)x\right\| \leqslant \left\|L(I-P)x\right\|_1 = \left\|L(x)\right\|_1 \\ &\leqslant \left\|N(x)\right\|_1 \leqslant \|p\|_1\|x\|_\infty + \|q\|_1\|x'\|_\infty + \|r\|_1 + \|e\|_1.\end{aligned}$$

If for some $t_0 \in [0,1]$, $|x'(t_0)| \leqslant N$ then we have

$$\left|x'(0)\right| = \left|x'(t_0) - \int_0^{t_0} x''(t)\,\mathrm{d}t\right| \leqslant N + \|x''\|_1.$$

Otherwise, if $|x'(t)| > N$ for all t, from condition (2) we obtain

$$\left|x'(\gamma)\right| \leqslant \frac{\|e\|_\infty + M}{n} + \frac{l}{n}\|x\|_\infty,$$

so that

$$\left|x'(0)\right| = \left|x'(\gamma) - \int_0^{\gamma} x''(t)\,\mathrm{d}t\right| \leqslant \frac{\|e\|_\infty + M}{n} + \frac{l}{n}\|x\|_\infty + \|x''\|_1.$$

Therefore we have in either case

$$\big\|P(x)\big\| = \big|x'(0)\big| \leqslant \max\left\{\frac{\|e\|_\infty + M}{n}, N\right\} + \frac{l}{n}\|x\|_\infty + \|x''\|_1.$$

Writing $x(t) = \int_0^t x'(s)\,\mathrm{d}s$, we obtain

$$\|x\|_\infty \leqslant \|x'\|_1 \leqslant \|x'\|_\infty.$$

Therefore

$$\begin{aligned}\|x'\|_\infty &\leqslant \|x\| \leqslant \big\|(I-P)(x)\big\| + \|Px\| \\ &\leqslant \left(\|p\|_1 + \|q\|_1 + \frac{l}{n}\right)\|x'\|_\infty + \|x''\|_1 + C,\end{aligned}$$

where $C = \|r\|_1 + \max\{\|e\|_\infty + M/n, N\}$. Let $C_1 = 1 - [\|p\|_1 + \|q\|_1 + \frac{l}{n}]$, so that $C_1 > 0$. Then we can write

$$\|x'\|_\infty \leqslant \frac{1}{C_1}\|x''\|_1 + \frac{C}{C_1},$$

and then

$$\begin{aligned}\|x''\|_1 = \big\|L(x)\big\|_1 &\leqslant \|Nx\|_1 \leqslant \|p\|_1\|x\|_\infty + \|q\|_1\|x'\|_\infty + \|r\|_1 + \|e\|_1 \\ &\leqslant \big(\|p\|_1 + \|q\|_1\big)\|x'\|_\infty + \|r\|_1 + \|e\|_1 \\ &\leqslant \frac{\|p\|_1 + \|q\|_1}{C_1}\|x''\|_1 + C_2,\end{aligned}$$

where $C_2 = \frac{\|p\|_1 + \|q\|_1}{C_1} + \|r\|_1 + \|e\|_1$. If $C_3 = \frac{\|p\|_1 + \|q\|_1}{C_1}$ then $C_3 < 1$ and hence

$$\|x\|_\infty \leqslant \|x'\|_\infty \leqslant \frac{C_2}{C_1(1-C_3)} + \frac{C}{C_1},$$

which proves that U_1 is bounded. □

LEMMA 3.24. *The set* $U_2 := \{x \in \ker(L)\colon\ N(x) \in \operatorname{im}(L)\}$ *is bounded.*

PROOF. Suppose that $x \in U_2$ so that $x(t) = ct$, where c is a constant. Since $QNx = 0$, we have

$$\int_0^1 \int_{\eta\tau}^{\tau} f(t, ct, c)\,\mathrm{d}t\,\mathrm{d}\tau = -\int_0^1 \int_{\eta\tau}^{\tau} e(t)\,\mathrm{d}t\,\mathrm{d}\tau.$$

Therefore there exists $\xi \in (0,1)$ such that $|f(\xi, c\xi, c)| \leqslant |e(\xi)| \leqslant \|e\|_\infty$. It follows that

$$|c| \leqslant \max\left\{N, \frac{M + \|e\|_\infty}{n-l}\right\}.$$

For, if $|c| > N$, then by condition (2) we obtain $\|e\|_\infty \geqslant -l|c\xi| + n|c| - M$ and hence $|c| \leqslant (M + \|e\|_\infty)/(n-l)$. Thus U_2 is bounded. □

LEMMA 3.25. *If in condition* (3) *we assume that there exists* $R > 0$ *such that for all* $|v| > R$, $vf(t, vt, v) \leqslant 0$, *for* $t \in [0,1]$, *then the set*

$$U_3 := \bigl\{x \in \ker(L)\colon\ H(x,\lambda) = \lambda Jx + (1-\lambda)QNx = 0,\ \lambda \in [0,1]\bigr\},$$

is bounded, where $J : \ker(L) \to Z_0$, *is the linear isomorphism given by* $J(ct) = c$.

PROOF. Assume that $x_n(t) - c_n t \in U_3$ and $\|c_n t\| = |c_n| \to \infty$ as $n \to \infty$. Then there exists $\lambda_n \in [0,1]$ such that

$$\lambda_n c_n + (1-\lambda_n)(QN)(c_n t) = 0.$$

$\{\lambda_n\}$ has a convergent subsequence, for simplicity we write $\lambda_n \to \lambda_0$. We show that $\lambda_0 \neq 1$. Indeed, otherwise we have

$$\lambda_n = -(1-\lambda_n)\frac{(QN)c_n t}{c_n},$$

and as previously we find that

$$(1-\lambda_n)\frac{\|(QN)c_n t\|}{|c_n|} \to 0 \quad (n \to \infty),$$

contradicting $\lambda_n \to 1$. Hence for n large enough, $1 - \lambda_n \neq 0$, and therefore

$$\frac{\lambda_n}{1-\lambda_n} c_n = Q\bigl(f(t, c_n t, c_n) + e(t)\bigr),$$

and so

$$\frac{\lambda_n}{1-\lambda_n} = \frac{2}{1-\eta}\int_0^1 \int_{\eta\tau}^{\tau} \frac{f(s, c_n s, c_n)}{c_n}\,\mathrm{d}s\,\mathrm{d}\tau + \frac{2}{c_n(1-\eta)}\int_0^1 \int_{\eta\tau}^{\tau} e(s)\,\mathrm{d}s\,\mathrm{d}\tau.$$

Since $|c_n| \to \infty$, we may assume that $|c_n| > \max\{N, R\}$. Then for large enough n we have

$$\left|\frac{f(s, c_n s, c_n)}{c_n}\right| \geqslant n - l - \frac{M}{|c_n|} \geqslant \frac{n-l}{2}.$$

By our assumption $c_n f(t, c_n t, c_n) \leqslant 0$ this yield $f(t, c_n t, c_n)/c_n \leqslant -(n-l)/2$. Hence, by Fatou's lemma,

$$
\begin{aligned}
&\limsup_{n\to\infty}\left\{\int_0^1\int_{\eta\tau}^{\tau}\frac{f(s, c_n s, c_n)}{c_n}\,\mathrm{d}s\,\mathrm{d}\tau+\frac{1}{c_n}\int_0^1\int_{\eta\tau}^{\tau}e(s)\,\mathrm{d}s\,\mathrm{d}\tau\right\}\\
&\quad\leqslant \limsup_{n\to\infty}\int_0^1\int_{\eta\tau}^{\tau}\frac{f(s, c_n s, c_n)}{c_n}\,\mathrm{d}s\,\mathrm{d}\tau\\
&\quad\leqslant \int_0^1\int_{\eta\tau}^{\tau}\limsup_{n\to\infty}\frac{f(s, c_n s, c_n)}{c_n}\,\mathrm{d}s\,\mathrm{d}\tau\\
&\quad\leqslant -\frac{(n-l)(1-\eta)}{4}.
\end{aligned}
$$

This is a contradiction with $\lambda_n/(1-\lambda_n) \geqslant 0$. Thus U_3 is bounded. □

PROOF OF THEOREM 3.21. Firstly, by Arzelá–Ascoli theorem, it can be shown that the linear operator $K : \mathrm{im}(L) \to D(L) \cap X_1$ in Lemma 3.22 is compact operator, so N is L-compact. Let Ω be a bounded open set containing $\overline{\bigcup_{i=1}^{3} U_i}$. Then by the above lemmas the conditions of Theorem 3.20 are satisfied and therefore $Lx + Nx = 0$ has at least one solution in $D(L) \cap \overline{\Omega}$ so that the BVP (3.26), (3.30) with $\alpha\eta = 1$ has at least one solution. □

THEOREM 3.26. *Assume that* $f : [0, 1] \times \mathbb{R}^2 \to \mathbb{R}$ *is continuous and has the decomposition*

$$f(t, x, p) = g(t, x, p) + h(t, x, p).$$

Assume that

(1) *There exists* $M_1 > 0$ *such that for* $x \in D(L) := \{x \in W^{2,1}(0, 1)\colon\ x(0) = 0,\ x(1) = \frac{1}{\eta}x(\eta)\}$, *if* $|x'(t)| > M_1$ *for all* $t \in [0, 1]$, *then*

$$\int_0^1\int_{\eta s}^{s}\bigl(f\bigl(t, x(t), x'(t)\bigr) + e(t)\bigr)\,\mathrm{d}t\,\mathrm{d}s \neq 0;$$

(2) *There exists* $M_2 > 0$, *such that for all* $v \in \mathbb{R}$ *with* $|v| > M_2$ *one has either*

$$v\bigl(f(t, vt, v) + e(t)\bigr) \geqslant 0 \quad \textit{for } t \in [0, 1],$$

or

$$v\bigl(f(t, vt, v) + e(t)\bigr) \leqslant 0 \quad \textit{for } t \in [0, 1];$$

(3) $pg(t, x, p) \leqslant 0$ *for all* $(t, x, p) \in [0, 1] \times [-M, M] \times \mathbb{R}$;

(4) $|h(t,x,p)| \leqslant a(t)|x| + b(t)|p| + u(t)|x|^r + v(t)|p|^k + c(t)$ *for* $(t,x,p) \in [0,1] \times [-M,M] \times \mathbb{R}$ *where* a, b, u, v *are in* $L^1[0,1]$ *and* $0 \leqslant r, k < 1$.

Then, the BVP (3.26), (3.30) *with* $\alpha\eta = 1$ *has at least one solution in* $C^1[0,1]$ *provided that*

$$\|a\|_1 + \|b\|_1 < \frac{1}{2}.$$

Theorem 3.21 is taken over from [16]. For the proof of Theorem 3.26 we refer to [15]. Other results are given by Gupta in [27,34]. In [16] is also proved the following uniqueness result.

THEOREM 3.27. *Suppose that the conditions* (1) *and* (2) *in Theorem* 3.21 *are replaced by the following conditions respectively*:

(1) *There exist functions* p, q *in* $L^1[0,1]$ *such that*

$$\bigl|f(t,u_1,v_1) - f(t,u_2,v_2)\bigr| \leqslant p(t)|u_1 - u_2| + q(t)|v_1 - v_2|$$

for $t \in [0,1]$ *and* $(u_1, v_1), (u_2, v_2) \in \mathbb{R}^2$.

(2) *There exists* $n > l \geqslant 0$ *such that*

$$\bigl|f(t,u_1,v_1) - f(t,u_2,v_2)\bigr| \geqslant -l|u_1 - u_2| + n|v_1 - v_2|$$

for $t \in [0,1]$, $(u_1, v_1), (u_2, v_2) \in \mathbb{R}^2$.

Then the BVP (3.26), (3.30) *with* $\alpha\eta = 1$ *has exactly one solution in* $C^1[0,1]$ *provided that*

$$2\bigl(\|p\|_1 + \|q\|_1\bigr) + \frac{l}{n} < 1.$$

3.7. *m-point boundary value problems reduced to four-point boundary value problems*

Let $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ be either a continuous or a Carathéodory's function and $e : [0,1] \to \mathbb{R}$ be a function in $L^1[0,1]$, $c_i, a_j \in \mathbb{R}$, with all of the c_i's, and all of a_j's, having the same sign, $\xi_i, \tau_j \in (0,1)$, $i = 1, 2, \dots, m-2$, $j = 1, 2, \dots, n-2$, $0 < \xi_1 < \xi_2 < \cdots < \xi_{m-2} < 1$, $0 < \tau_1 < \tau_2 < \cdots < \tau_{n-2} < 1$.

Consider the following second-order ordinary differential equation

$$x''(t) = f\bigl(t, x(t), x'(t)\bigr) + e(t), \quad t \in (0,1) \tag{3.38}$$

subject to one of the following boundary conditions

$$x(0) = \sum_{i=1}^{m-2} c_i x'(\xi_i), \qquad x(1) = \sum_{j=1}^{n-2} a_j x(\tau_j), \tag{3.39}$$

$$x(0)=\sum_{i=1}^{m-2} c_i x'(\xi_i), \qquad x'(1)=\sum_{j=1}^{n-2} a_i x'(\tau_j), \tag{3.40}$$

$$x(0)=\sum_{i=1}^{m-2} c_i x'(\xi_i), \qquad x(1)=\sum_{j=1}^{n-2} a_i x(\tau_j), \tag{3.41}$$

$$x(0)=\sum_{i=1}^{m-2} c_i x(\xi_i), \qquad x'(1)=\sum_{j=1}^{n-2} a_i x'(\tau_j). \tag{3.42}$$

It is well known that if a function $x \in C^1$ satisfies one of the boundary conditions (3.39)–(3.42) and c_i, a_j, $i = 1, 2, \dots, m-2$, $j = 1, 2, \dots, n-2$ are as above, then there exist $\zeta \in [\xi_1, \xi_{m-2}]$, $\eta \in [\tau_1, \tau_{n-2}]$ such that

$$x(0)=\gamma x'(\zeta), \qquad x(1)=\alpha x(\eta), \tag{3.43}$$

$$x(0)=\gamma x'(\zeta), \qquad x'(1)=\alpha x'(\eta), \tag{3.44}$$

$$x(0)=\gamma x(\zeta), \qquad x(1)=\alpha x(\eta), \tag{3.45}$$

$$x(0)=\gamma x(\zeta), \qquad x'(1)=\alpha x'(\eta), \tag{3.46}$$

respectively with $\gamma = \sum_{i=1}^{m-2} c_i$, $\alpha = \sum_{j=1}^{n-2} a_j$.

As in the case of m-point BVPs which are reduced to three-point BVPs, we shall prove an existence result the BVP (3.38)–(3.39), proving first an existence result for the BVP (3.38), (3.43), and using the a priori bounds obtained for this problem for the BVP (3.38)–(3.39). In the next lemma we obtain the needed a priori bounds, which are independent of ζ and η, for the BVP (3.38), (3.43).

LEMMA 3.28. *Let $\zeta, \eta \in (0, 1)$ is given and $x(t) \in W^{2,1}(0, 1)$ be such that $x(0) = \gamma x'(\zeta)$, $x(1) = \alpha x(\eta)$. Then*

$$\|x\|_\infty \leqslant A\|x'\|_\infty, \qquad \|x'\|_\infty \leqslant B\|x''\|_1$$

where

$$A = \begin{cases} 1, & \text{if } \alpha \leqslant 0, \\ L, & \text{if } \alpha > 0, \alpha \neq 1, \\ 1 + |\gamma|, & \text{if } \alpha = 1 \end{cases}$$

and

$$B = \begin{cases} 1, & \text{if } \alpha \leqslant 0, \gamma = 0, \\ \dfrac{1}{1-Q}, & \text{if } \alpha \leqslant 0, \gamma \neq 0, \\ \dfrac{1}{1-S}, & \text{if } \alpha > 0, \alpha \neq 1, \\ 1, & \text{if } \alpha = 1, \end{cases}$$

where for $\alpha > 0, \alpha \neq 1$,

$$\begin{cases} M = \min\left\{\alpha, \dfrac{1}{\alpha}\right\} < 1, \\ L = \min\left\{\dfrac{1}{1-M}, 1 + \dfrac{1-\eta}{|1-\alpha|}, 1 + \dfrac{|\alpha|(1-\eta)}{|1-\alpha|}, 1 + |\gamma|\right\}, \\ S = \min\left\{\dfrac{|1-\alpha|}{1-\eta}L, \dfrac{|1-\alpha|}{\alpha(1-\eta)}L, \dfrac{1}{|\gamma|}L\right\} \end{cases}$$

for $\alpha < 0$,

$$Q = \min\left\{\frac{1-\alpha}{1-\eta}, \frac{1-\alpha}{|\alpha|(1-\eta)}, \frac{1}{|\gamma|}\right\}$$

and for $\alpha = 0$, $Q = \frac{1}{|\gamma|}$ *provided* $Q < 1$ *and* $S < 1$.

PROOF. We consider the following cases:

Case 1: $\alpha \leqslant 0$. In this case $x(1)x(\eta) \leqslant 0$ and accordingly there exists a $\theta \in [\eta, 1]$ such that $x(\theta) = 0$. Hence it follows that $\|x\|_\infty \leqslant \|x'\|_\infty$. Also if $\gamma = 0$, we have from $x(0) = 0$ and $x(\theta) = 0$ that there exists a $z \in (0, \theta)$ such that $x'(z) = 0$. Accordingly, we get that $\|x'\|_\infty \leqslant \|x''\|_1$. Suppose, now, that $\alpha < 0$ and $\gamma \neq 0$. Next we see from Mean Value Theorem there exists $\omega \in (\eta, 1)$ such that

$$(\alpha - 1)x(\eta) = x(1) - x(\eta) = (1 - \eta)x'(\omega)$$

and hence

$$x(\eta) = \frac{1-\eta}{\alpha - 1}x'(\omega).$$

Also, since $x(1) = \alpha x(\eta)$ we get

$$x(1) = \frac{\alpha(1-\eta)}{\alpha - 1}x'(\omega).$$

From the relations

$$x'(t) = x'(\omega) + \int_\omega^t x''(s)\,\mathrm{d}s = \frac{\alpha - 1}{1 - \eta}x(\eta) + \int_\omega^t x''(s)\,\mathrm{d}s,$$
$$x'(t) = x'(\omega) + \int_\omega^t x''(s)\,\mathrm{d}s = \frac{\alpha - 1}{\alpha(1 - \eta)}x(1) + \int_\omega^t x''(s)\,\mathrm{d}s$$

and

$$x'(t) = x'(\zeta) + \int_0^t x''(s)\,\mathrm{d}s = \frac{1}{\gamma}x(0) + \int_0^t x''(s)\,\mathrm{d}s$$

then follows that

$$\|x'\|_\infty \leqslant \frac{1}{1-Q}\|x''\|_1,$$

where $Q=\min\{\frac{1-\alpha}{1-\eta}, \frac{1-\alpha}{|\alpha|(1-\eta)}, \frac{1}{|\gamma|}\}$ if $Q<1$. Finally, for $\alpha=0$, $\gamma\neq 0$ it is easy to see that $Q=\frac{1}{|\gamma|}$ since we require that $Q<1$ and $\frac{1}{1-\eta}>1$.

Case 2: $\alpha>0, \alpha\neq 1$. We first consider the relations

$$x(t)=x(1)+\int_1^t x'(s)\,\mathrm{d}s=\alpha x(\eta)+\int_1^t x'(s)\,\mathrm{d}s$$

and

$$x(t)=x(\eta)+\int_\eta^t x'(s)\,\mathrm{d}s=\frac{1}{\alpha}x(1)+\int_\eta^t x'(s)\,\mathrm{d}s.$$

Since, now, $M=\min\{\alpha, \frac{1}{\alpha}\}<1$, we get from the above relations that

$$\|x\|_\infty \leqslant \frac{1}{1-M}\|x'\|_\infty.$$

Next, we get the relations

$$x(t)=x(1)+\int_1^t x'(s)\,\mathrm{d}s=\frac{\alpha(1-\eta)}{\alpha-1}x'(\omega)+\int_1^t x'(s)\,\mathrm{d}s$$

and

$$x(t)=x(\eta)+\int_\eta^t x'(s)\,\mathrm{d}s=\frac{1-\eta}{\alpha-1}x'(\omega)+\int_\eta^t x'(s)\,\mathrm{d}s.$$

From these relations and

$$x(t)=x(0)+\int_0^t x'(s)\,\mathrm{d}s=\gamma x'(\zeta)+\int_0^t x'(s)\,\mathrm{d}s$$

it is immediate that

$$\|x\|_\infty \leqslant L\|x'\|_\infty,$$

where $L=\min\{\frac{1}{1-M}, 1+\frac{1-\eta}{|\alpha-1|}, 1+\frac{|\alpha|(1-\eta)}{|\alpha-1|}, 1+|\gamma|\}$. Further, we have

$$\|x'\|_\infty \leqslant \frac{1}{1-S}\|x''\|_1,$$

where $S = \min\{\frac{|1-\alpha|}{1-\eta}L, \frac{|1-\alpha|}{\alpha(1-\eta)}L, \frac{1}{|\gamma|}L\}$ if $S < 1$.

Case 3: $\alpha = 1$. Since $x(1) = x(\eta)$ there exists an $\omega \in (\eta, 1)$ with $x'(\omega) = 0$. It is then immediate that $\|x'\|_\infty \leqslant \|x''\|_1$. Also since $x(t) = x(0) + \int_0^t x'(s)\,\mathrm{d}s = \gamma x'(\zeta) + \int_0^t x'(s)\,\mathrm{d}s$, it is immediate that $\|x\|_\infty \leqslant (1 + |\gamma|)\|x'\|_\infty$.

This completes the proof of the lemma. □

THEOREM 3.29. *Let* $f : [0, 1] \times \mathbb{R}^2 \to \mathbb{R}$ *be a function satisfying Carathéodory's conditions. Assume that* (3.12) *holds. Also let* $\eta \in (0, 1)$ *be given and* $\alpha, \gamma \in \mathbb{R}$ *with* $1 + \gamma \neq \alpha(\gamma + \eta)$. *Moreover we assume that* $Q < 1$ *and* $S < 1$.

Then the BVP (3.38), (3.43) *has at least one solution in* $C^1[0, 1]$ *provided*

$$\begin{cases} \|p\|_1 + \|q\|_1 < 1, & \alpha \leqslant 0, \gamma = 0, \\ \|p\|_1 + \|q\|_1 < 1 - Q, & \alpha \leqslant 0, \gamma \neq 0, \\ L\|p\|_1 + \|q\|_1 < 1 - S, & \alpha > 0, \alpha \neq 1, \\ \big(1 + |\gamma|\big)\|p\|_1 + \|q\|_1 < 1, & \alpha = 1. \end{cases}$$

PROOF. Let X be the Banach space $C^1[0, 1]$ and Y denote the Banach space $L^1(0, 1)$ with their usual norms. We denote a linear mapping $L : D(L) \subset X \to Y$ by setting

$$D(L) = \{x \in W^{2,1}(0, 1)\colon\ x(0) = \gamma x'(\zeta),\ x(1) = \alpha x(\eta)\},$$

and for $x \in D(L)$,

$$Lx = x''.$$

We also define a nonlinear mapping $N : X \to Y$ by setting

$$(Nx)(t) = f\big(t, x(t), x'(t)\big), \quad t \in [0, 1].$$

We note that N is a bounded mapping from X into Y. Next, it is easy to see that the linear mapping $L : D(L) \subset X \to Y$, is one-to-one mapping. Next, the linear mapping $K : Y \to X$, defined for $y \in Y$ by

$$\begin{aligned}(Ky)(t) = {}& \int_0^t (t-s)y(s)\,\mathrm{d}s + \gamma \int_0^\zeta y(s)\,\mathrm{d}s \\ & + \frac{\gamma + t}{1 + \gamma - \alpha(\gamma + \eta)} \bigg[\alpha \int_0^\eta (\eta - s)y(s)\,\mathrm{d}s \\ & - \int_0^1 (1-s)y(s)\,\mathrm{d}s + \gamma(\alpha - 1)\int_0^\zeta y(s)\,\mathrm{d}s\bigg], \quad t \in [0, 1]\end{aligned}$$

is such that for $y \in Y$, $Ky \in D(L)$ and $LKy = y$; and for $u \in D(L)$, $KLu = u$. Furthermore, it follows easily using the Arzelá–Ascoli theorem that KN maps bounded subsets of X into a relatively compact subsets of X. Hence $KN : X \to X$ is a compact mapping.

We, next, note that $x \in C^1[0,1]$ is a solution of the BVP (3.38), (3.43) if and only if x is a solution to the operator equation

$$Lx = Nx + e.$$

Now, the operator equation $Lx = Nx + e$ is equivalent to the equation

$$x = KNx + Ke.$$

We apply the Leray–Schauder continuation theorem to obtain the existence of a solution for $x = KNx + Ke$ or equivalently to the BVP (3.38), (3.43).

To do this, it suffices to verify that the set of all possible solutions of the family of equations

$$x''(t) = \lambda f\bigl(t, x(t), x'(t)\bigr) + \lambda e(t), \quad t \in (0,1), \tag{3.47}$$

$$x(0) = \gamma x'(0), \qquad x(1) = \alpha x(\eta) \tag{3.48}$$

is, a priori, bounded in $C^1[0,1]$ by a constant independent of $\lambda \in [0,1]$.

(I) Assume that $\alpha \leqslant 0$, $\gamma = 0$. From Lemma 3.28 we have

$$\|x\|_\infty \leqslant \|x'\|_\infty \leqslant \|x''\|_1.$$

Let, now, $x(t)$ be a solution of (3.47)–(3.48) for some $\lambda \in [0,1]$, so that $x \in W^{2,1}(0,1)$ with $x(0) = \gamma x'(\zeta)$, $x(1) = \alpha x(\eta)$. We then get

$$\begin{aligned}\|x''\|_1 &= \lambda \bigl\| f\bigl(t, x(t), x'(t)\bigr) + e(t)\bigr\|_1 \\ &\leqslant \|p\|_1\|x\|_\infty + \|q\|_1\|x'\|_\infty + \|r\|_1 + \|e\|_1 \\ &\leqslant \bigl(\|p\|_1 + \|q\|_1\bigr)\|x''\|_1 + \|r\|_1 + \|e\|_1.\end{aligned}$$

It follows from our assumption that there is a constant c, independent of $\lambda \in [0,1]$, such that

$$\|x''\|_1 \leqslant c.$$

It is now immediate that the set of solutions of the family of equations (3.47)–(3.48) is, a priori, bounded in $C^1[0,1]$ by a constant independent of $\lambda \in [0,1]$.

(II) Assume that $\alpha \leqslant 0$, $\gamma \neq 0$. Then we have, by Lemma 3.28 that

$$\|x\|_\infty \leqslant \|x'\|_\infty, \qquad \|x'\|_\infty \leqslant \frac{1}{1-Q}\|x''\|_1.$$

We then get

$$\|x''\|_1 = \lambda \bigl\| f\bigl(t, x(t), x'(t)\bigr) + e(t)\bigr\|_1$$

$$\leqslant \|p\|_1\|x\|_\infty + \|q\|_1\|x'\|_\infty + \|r\|_1 + \|e\|_1$$
$$\leqslant \big[\|p\|_1 + \|q\|_1\big]\frac{1}{1-Q}\|x''\|_1 + \|r\|_1 + \|e\|_1.$$

We proceed as in case (I).

The process for the other cases is similar to the previous cases and we omit the details. This completes the proof of the theorem. □

We study now the multi-point BVP (3.38)–(3.39) using the a priori estimates that can be obtained for a four-point BVP (3.38), (3.43). This is because for every solution $x(t)$ of the BVP (3.38)–(3.39), there exist $\eta \in [\xi_1, \xi_{m-2}]$, $\zeta \in [\tau_1, \tau_{n-2}]$, depending on, $x(t)$, such that $x(t)$ is also a solution of the BVP (3.38), (3.43) with $\gamma = \sum_{i=1}^{m-2} c_i$ and $\alpha = \sum_{j=1}^{n-2} a_j$. The proof is quite similar to the proof of Theorem 3.29 and uses the a priori estimates obtained in the proof of Lemma 3.28 for the set of solutions of the family of equations (3.38), (3.43).

THEOREM 3.30. *Let* $f:[0,1]\times\mathbb{R}^2 \to \mathbb{R}$ *be a function satisfying Carathéodory's conditions. Assume that* (3.12) *holds. Let* $c_i, a_j \in \mathbb{R}$, *with all of the* c_i*'s* (*respectively,* a_j*'s*), *having the same sign,* $\xi_i, \tau_j \in (0,1)$, $i = 1,2,\dots,m-2$, $j = 1,2,\dots,n-2$, $0 < \xi_1 < \xi_2 < \cdots < \xi_{m-2} < 1$, $0 < \tau_1 < \tau_2 < \cdots < \tau_{n-2} < 1$ *be given. Suppose that*

$$1 + \left(\sum_{i=1}^{m-2} c_i\right)\left(1 - \sum_{j=1}^{n-2} a_j\right) - \sum_{j=1}^{n-2} a_j\tau_j \neq 0.$$

Let $\gamma = \sum_{i=1}^{m-2} c_i$ *and* $\alpha = \sum_{j=1}^{n-2} a_j$. *Moreover we assume that* $Q' < 1$, *and* $S' < 1$, *where* $M = \min\{\alpha, \frac{1}{\alpha}\} < 1$,

$$L' = \min\left\{\frac{1}{1-M}, 1 + \frac{1-\tau_1}{|1-\alpha|}, 1 + \frac{|\alpha|(1-\tau_1)}{|1-\alpha|}, 1 + |\gamma|\right\},$$
$$S' = \min\left\{\frac{|1-\alpha|}{1-\tau_{n-2}}L, \frac{|1-\alpha|}{\alpha(1-\tau_{n-2})}L, \frac{1}{|\gamma|}L\right\},$$
$$Q' = \min\left\{\frac{1-\alpha}{1-\tau_{n-2}}, \frac{1-\alpha}{|\alpha|(1-\tau_{n-2})}, \frac{1}{|\gamma|}\right\}.$$

Then the BVP (3.38)–(3.39) *has at least one solution in* $C^1[0,1]$ *provided*

$$\begin{cases} \|p\|_1 + \|q\|_1 < 1, & \alpha \leqslant 0, \gamma = 0, \\ \|p\|_1 + \|q\|_1 < 1 - Q', & \alpha \leqslant 0, \gamma \neq 0, \\ L'\|p\|_1 + \|q\|_1 < 1 - S', & \alpha > 0, \alpha \neq 1, \\ \big(1 + |\gamma|\big)\|p\|_1 + \|q\|_1 < 1, & \alpha = 1. \end{cases}$$

PROOF. Let X be the Banach space $C^1[0,1]$ and Y denote the Banach space $L^1(0,1)$ with their usual norms. We denote a linear mapping $L: D(L) \subset X \to Y$ by setting

$$D(L) = \left\{ x \in W^{2,1}(0,1)\colon\ x(0) = \sum_{i=1}^{m-2} c_i x'(\xi_i),\ x(1) = \sum_{j=1}^{n-2} a_j x(\tau_j) \right\},$$

and for $x \in D(L)$,

$$Lx = x''.$$

We also define a nonlinear mapping $N: X \to Y$ by setting

$$(Nx)(t) = f\bigl(t, x(t), x'(t)\bigr), \quad t \in [0,1].$$

We note that N, is a bounded mapping from X into Y. Next, it is easy to see that the linear mapping $L: D(L) \subset X \to Y$, is one-to-one mapping. Next, the linear mapping $K: Y \to X$, defined for $y \in Y$ by

$$(Ky)(t) = \int_0^t (t-s)y(s)\,\mathrm{d}s + ct + k, \quad t \in [0,1],$$

where c and k are given by

$$\begin{aligned}
&\left[1 + \left(\sum_{i=1}^{m-2} c_i\right)\left(1 - \sum_{j=1}^{n-2} a_j\right) - \sum_{j=1}^{n-2} a_j \tau_j\right] c \\
&\quad = \left(\sum_{j=1}^{n-2} a_j - 1\right)\left(\sum_{i=1}^{m-2} c_i \int_0^{\xi_i} y(s)\,\mathrm{d}s\right) \\
&\qquad + \sum_{j=1}^{n-2} a_j \int_0^{\tau_j} (\tau_j - s)y(s)\,\mathrm{d}s - \int_0^1 (1-s)y(s)\,\mathrm{d}s
\end{aligned}$$

and

$$\begin{aligned}
&\left[1 + \left(\sum_{i=1}^{m-2} c_i\right)\left(1 - \sum_{j=1}^{n-2} a_j\right) - \sum_{j=1}^{n-2} a_j \tau_j\right] k = \sum_{i=1}^{m-2} c_i \sum_{j=1}^{n-2} a_j \int_0^{\tau_j} (\tau_j - s)y(s)\,\mathrm{d}s \\
&\qquad - \sum_{i=1}^{m-2} c_i \int_0^1 (1-s)y(s)\,\mathrm{d}s + \left(1 - \sum_{j=1}^{n-2} a_j \tau_j\right) \sum_{i=1}^{m-2} c_i \int_0^{\xi_i} y(s)\,\mathrm{d}s,
\end{aligned}$$

is such that for $y \in Y$, $Ky \in D(L)$and $LKy = y$; and for $u \in D(L)$, $KLu = u$. Furthermore, it follows easily using the Arzelá–Ascoli theorem that KN maps bounded subsets of X into a relatively compact subset of X. Hence $KN: X \to X$ is a compact mapping.

We, next, note that $x \in C^1[0,1]$ is a solution of the BVP (3.38)–(3.39) if and only if x is a solution to the operator equation

$$Lx = Nx + e.$$

Now, the operator equation $Lx = Nx + e$ is equivalent to the equation

$$x = KNx + Ke.$$

We apply the Leray–Schauder continuation theorem to obtain the existence of a solution for $x = KNx + Ke$ or equivalently to the BVP (3.38)–(3.39).

To do this, it suffices to verify that the set of all possible solutions of the family of equations

$$x''(t) = \lambda f\big(t, x(t), x'(t)\big) + \lambda e(t), \quad t \in (0,1), \tag{3.49}$$

$$x(0) = \sum_{i=1}^{m-2} c_i x'(\xi_i), \qquad x(1) = \sum_{j=1}^{n-2} a_j x(\tau_j) \tag{3.50}$$

is, a priori, bounded in $C^1[0,1]$ by a constant independent of $\lambda \in [0,1]$.

Let, now, $x(t)$ be a solution of (3.49)–(3.50) for some $\lambda \in [0,1]$, so that $x \in W^{2,1}(0,1)$ with $x(0) = \sum_{i=1}^{m-2} c_i x'(\xi_i)$, $x(1) = \sum_{j=1}^{n-2} a_j x(\tau_j)$. Accordingly, there exist $\zeta \in [\xi_1, \xi_{m-2}]$ and $\eta \in [\tau_1, \tau_{n-2}]$ depending on $x(t)$, such that $x(t)$ is a solution of the four point BVP

$$x''(t) = \lambda f\big(t, x(t), x'(t)\big) + \lambda e(t), \quad t \in (0,1),$$

$$x(0) = \gamma x'(\zeta), \qquad x(1) = \alpha x(\eta).$$

It then follows, as in the proof of Theorem 3.30 that there is a constant c, independent of $\lambda \in [0,1]$, and $\eta \in [\xi_1, \xi_{m-2}]$, $\zeta \in [\tau_1, \tau_{n-2}]$ such that

$$\|x\|_\infty \leqslant c_1 \|x'\|_\infty \leqslant c_2 \|x''\|_1 \leqslant c,$$

where c_1, c_2 are constants independent of λ, η, ζ. Thus the set of solutions of the family of equations is, a priori, bounded in $C^1[0,1]$ by a constant, independent of $\lambda \in [0,1]$. This completes the proof of the theorem. □

Theorems 3.29 and 3.30 are taken from [35]. Similar results for the BVP (3.38), (3.40) are proved in [35] and for the BVPs (3.38), (3.41) and (3.38), (3.42) in [36].

3.8. *The general case*

Let $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ be either a continuous or a Carathéodory's function and $e(t) \in L^1[0,1]$. Let ξ_i, η_j, c_i, a_j, $i = 1, 2, \ldots, m-2$, $j = 1, 2, \ldots, n-2$, $0 < \xi_1 < \xi_2 < \cdots <$

$\xi_{m-2} < 1$, $0 < \eta_1 < \eta_2 < \cdots < \eta_{n-2} < 1$ be given. Consider the following second-order ordinary differential equation:

$$x''(t) = f\bigl(t, x(t), x'(t)\bigr) + e(t), \quad t \in (0,1), \tag{3.51}$$

subject to one of the following boundary value conditions:

$$x(0) = \sum_{i=1}^{m-2} c_i x(\xi_i), \qquad x(1) = \sum_{j=1}^{n-2} a_j x(\eta_j), \tag{3.52}$$

$$x(0) = \sum_{i=1}^{m-2} c_i x(\xi_i), \qquad x'(1) = \sum_{j=1}^{n-2} a_j x(\eta_j), \tag{3.53}$$

$$x'(0) = \sum_{i=1}^{m-2} c_i x(\xi_i), \qquad x(1) = \sum_{j=1}^{n-2} a_j x(\eta_j), \tag{3.54}$$

$$x(0) = \sum_{i=1}^{m-2} c_i x'(\xi_i), \qquad x(1) = \sum_{j=1}^{n-2} a_j x'(\eta_j), \tag{3.55}$$

$$x(0) = \sum_{i=1}^{m-2} c_i x'(\xi_i), \qquad x'(1) = \sum_{j=1}^{n-2} a_j x(\eta_j), \tag{3.56}$$

$$x(0) = \sum_{i=1}^{m-2} c_i x(\xi_i), \qquad x'(1) = \sum_{j=1}^{n-2} a_j x'(\eta_j), \tag{3.57}$$

$$x(0) = \sum_{i=1}^{m-2} c_i x'(\xi_i), \qquad x'(1) = \sum_{j=1}^{n-2} a_j x'(\eta_j). \tag{3.58}$$

When all the c_i's have the same sign and also all the a_j's have the same sign, it is known that existence of a solution for (3.51) with boundary conditions (3.52)–(3.58) can be obtained via existence subject to the respective four-point boundary conditions

$$x(0) = \gamma x(\xi), \qquad x(1) = \alpha x(\eta), \tag{3.59}$$

$$x(0) = \gamma x(\xi), \qquad x'(1) = \alpha x(\eta), \tag{3.60}$$

$$x'(0) = \gamma x'(\xi), \qquad x(1) = \alpha x(\eta), \tag{3.61}$$

$$x'(0) = \gamma x'(\xi), \qquad x'(1) = \alpha x'(\eta), \tag{3.62}$$

$$x(0) = \gamma x'(\xi), \qquad x(1) = \alpha x(\eta), \tag{3.63}$$

$$x(0) = \gamma x'(\xi), \qquad x'(1) = \alpha x'(\eta), \tag{3.64}$$

$$x(0) = \gamma x(\xi), \qquad x'(1) = \alpha x'(\eta), \tag{3.65}$$

where $\xi \in [\xi_1, \xi_{m-2}]$, $\eta \in [\eta_1, \eta_{n-2}]$, $\gamma = \sum_{i=1}^{m-2} c_i$, $\alpha = \sum_{j=1}^{n-2} a_j$.

When all the c_i's have the same sign and also all the a_j's have no same sign, the existence of a solution for (3.51) with boundary value conditions (3.52)–(3.58) can be obtained via existence subject to the following boundary conditions

$$x(0) = \gamma x(\xi), \qquad x(1) = \sum_{j=1}^{n-2} a_j x(\eta_j), \tag{3.66}$$

$$x(0) = \gamma x(\xi), \qquad x'(1) = \sum_{j=1}^{n-2} a_j x'(\eta_j), \tag{3.67}$$

$$x'(0) = \gamma x(\xi), \qquad x(1) = \sum_{j=1}^{n-2} a_j x(\eta_j), \tag{3.68}$$

$$x'(0) = \gamma x'(\xi), \qquad x'(1) = \sum_{j=1}^{n-2} a_j x'(\eta_j), \tag{3.69}$$

$$x(0) = \gamma x'(\xi), \qquad x(1) = \sum_{j=1}^{n-2} a_j x(\eta_j), \tag{3.70}$$

$$x(0) = \gamma x'(\xi), \qquad x'(1) = \sum_{j=1}^{n-2} a_j x'(\eta_j), \tag{3.71}$$

$$x(0) = \gamma x(\xi), \qquad x'(1) = \sum_{j=1}^{n-2} a_j x'(\eta_j), \tag{3.72}$$

where $\xi \in [\xi_1, \xi_{m-2}]$, $\gamma = \sum_{i=1}^{m-2} c_i$.

When all the c_i's do not have the same sign and also all the a_j's have the same sign, the existence of a solution for (3.51) with boundary value conditions (3.52)–(3.58) can be obtained via existence subject to the following boundary conditions

$$x(0) = \sum_{i=1}^{m-2} c_i x(\xi_i), \qquad x(1) = \alpha x(\eta), \tag{3.73}$$

$$x(0) = \sum_{i=1}^{m-2} c_i x(\xi_i), \qquad x'(1) = \alpha x'(\eta), \tag{3.74}$$

$$x'(0) = \sum_{i=1}^{m-2} c_i x'(\xi_i), \qquad x(1) = \alpha x(\eta), \tag{3.75}$$

$$x'(0)=\sum_{i=1}^{m-2}c_i x'(\xi_i),\qquad x'(1)=\alpha x'(\eta), \tag{3.76}$$

$$x(0)=\sum_{i=1}^{m-2}c_i x'(\xi_i),\qquad x(1)=\alpha x(\eta), \tag{3.77}$$

$$x(0)=\sum_{i=1}^{m-2}c_i x'(\xi_i),\qquad x'(1)=\alpha x'(\eta), \tag{3.78}$$

$$x(0)=\sum_{i=1}^{m-2}c_i x(\xi_i),\qquad x'(1)=\alpha x'(\eta), \tag{3.79}$$

where $\eta \in [\eta_1, \eta_{n-2}], \alpha = \sum_{j=1}^{n-2} a_j$.

For certain boundary condition case such that the linear operator $Lx = x''$, defined in a suitable Banach space, is invertible, is the so-called nonresonance case. Otherwise, the so-called resonance case.

Existence results were given for the BVPs:

- (3.51)–(3.52) in [29] for the nonresonance case and in [59] for resonance case,
- (3.51), (3.54) and (3.51), (3.55) in [59] for resonance case,
- (3.51), (3.57) in [29] for nonresonance case,
- (3.51), (3.58) in [28] for nonresonance case,
- (3.51), (3.59), (3.51), (3.61), (3.51), (3.62) and (3.51), (3.65) in [61] for resonance case,
- (3.51), (3.66), (3.51), (3.67), (3.51), (3.68) and (3.51), (3.69) in [60] for resonance case,
- (3.51), (3.73), (3.51), (3.74), (3.51), (3.75) and (3.51), (3.76) in [62] for resonance case.

For the BVP (3.38), (3.52) we report an existence result for the nonresonance case from [29].

THEOREM 3.31. *Let* $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ *be a function satisfying Carathéodory's conditions. Assume that* (3.12) *holds. Let* $c_i, a_j \in \mathbb{R}$, *with all of the* c_i*'s* (*respectively,* a_j*'s*), *not having the same sign,* $\xi_i, \eta_j \in (0,1)$, $i = 1,2,\ldots,m-2$, $j = 1,2,\ldots,n-2$, $0 < \xi_1 < \xi_2 < \cdots < \xi_{m-2} < 1$, $0 < \eta_1 < \eta_2 < \cdots < \eta_{n-2} < 1$ *be given. Suppose that*

$$\left(\sum_{i=1}^{m-2} c_i\xi_i\right)\left(1-\sum_{j=1}^{n-2}a_j\right) \neq \left(1-\sum_{j=1}^{n-2}c_i\right)\left(\sum_{j=1}^{n-2}a_j\eta_j - 1\right).$$

Then the BVP (3.51)–(3.52) *has at least one solution in* $C^1[0,1]$ *provided*

$$Q\big(M\|p\|_1 + \|q\|_1\big) < 1,$$

where

$$M=\min\left\{\frac{1}{\left|\sum_{i=1}^{m-2}c_i\right|}\left(\sum_{i=1}^{m-2}|c_i|\lambda_i+\frac{\sum_{i=1}^{m-2}|c_i\xi_i|}{\left|1-\sum_{i=1}^{m-2}c_i\right|}\right),\right.$$
$$\frac{1}{\left|\sum_{j=1}^{n-2}a_j\right|}\left(\sum_{j=1}^{n-2}|a_i|\mu_j+\frac{\sum_{j=1}^{n-2}|c_j(1-\eta_j)|}{\left|1-\sum_{j=1}^{n-2}a_j\right|}\right),$$
$$\left.1+\frac{\sum_{i=1}^{m-2}|c_i\xi_i|}{\left|1-\sum_{i=1}^{m-2}c_i\right|},1+\frac{\sum_{j=1}^{n-2}|c_j(1-\eta_j)|}{\left|1-\sum_{j=1}^{n-2}a_j\right|}\right\},$$

$\lambda_i=\max\{\xi_i,1-\xi_i\}$, $i=1,\dots,m-2$, $\mu_j=\max\{\eta_j,1-\eta_j\}$, $j=1,\dots,n-2$ *and*

$$Q=\min\left\{\frac{\sum_{i=1}^{m-2}|c_i\xi_i|}{\left|\sum_{i=1}^{m-2}c_i\xi_i\right|-M\left|1-\sum_{i=1}^{m-2}c_i\right|},\right.$$
$$\left.\frac{\sum_{j=1}^{n-2}|a_j(1-\eta_j)|}{\left|\sum_{i=j}^{n-2}a_j(1-\eta_j)\right|-M\left|\sum_{j=1}^{n-2}a_j-1\right|}\right\}.$$

For resonance case Liu in [59] gave existence results for the BVP (3.51)–(3.52) in the following cases:

(1) $\sum_{i=1}^{m-2}c_i\xi_i=0$, $\gamma=1$, $\alpha=1$, $\sum_{i=1}^{m-2}c_i\xi_i^2\neq 0$, $\sum_{j=1}^{n-2}a_j\eta_j^2\neq 1$.

(2) $\sum_{i=1}^{m-2}c_i\xi_i\neq 0$, $\gamma=1$, $\alpha=1$, $\sum_{i=1}^{n-2}a_j\eta_j=1$, $\sum_{j=1}^{n-2}a_j\eta_j^2\neq 1$.

(3) $\sum_{i=1}^{m-2}c_i\xi_i=0$, $\gamma\neq 1$, $\alpha=1$, $\sum_{i=1}^{n-2}a_j\eta_j=1$, $\sum_{j=1}^{n-2}a_j\eta_j^2\neq 1$.

(4) $\sum_{i=1}^{m-2}c_i\xi_i=0$, $\gamma=1$, $\alpha\neq 1$, $\sum_{i=1}^{n-2}a_j\eta_j=1$, $\sum_{i=1}^{m-2}c_i\xi_i^2\neq 0$.

We give an existence result at resonance, for the BVP (3.51)–(3.52) in the case (1), i.e. when

$$\sum_{i=1}^{m-2}c_i\xi_i=0,\qquad \gamma=1,\qquad \alpha=1,\qquad \sum_{i=1}^{m-2}c_i\xi_i^2\neq 0,\qquad \sum_{j=1}^{n-2}a_j\eta_j^2\neq 1.$$

THEOREM 3.32. *Let* $f:[0,1]\times\mathbb{R}^2\to\mathbb{R}$ *be a continuous function. Assume that either there exists* $m_1\in\{1,2,\dots,m-3\}$ *such that* $c_i<0$ $(1\leqslant i\leqslant m_1)$, $c_i>0$ $(m_1+1\leqslant i\leqslant m-2)$, *or there exists* $n_1\in\{1,2,\dots,n-3\}$ *such that* $a_j>0$ $(1\leqslant j\leqslant n_1)$, $a_j<0$ $(n_1+1\leqslant j\leqslant n-2)$, *furthermore*

(H1) *There exist functions* a,b,c,r *in* $L^1[0,1]$, *and constant* $\theta\in[0,1)$ *such that for all* $(x,y)\in\mathbb{R}^2$, $t\in[0,1]$ *either*

$$\left|f(t,x,y)\right|\leqslant a(t)|x|+b(t)|y|+c(t)|y|^\theta+r(t)$$

or else

$$\left|f(t,x,y)\right|\leqslant a(t)|x|+b(t)|y|+c(t)|x|^\theta+r(t).$$

(H2) *There exists constant* $M > 0$ *such that, for* $x \in \operatorname{dom}(L)$, *if* $|x(t)| > M$ *for all* $t \in [0, 1]$, *then*

$$\sum_{i=1}^{m-2} c_i \int_0^{\xi_i} \int_0^s \left[f\big(\tau, x(\tau), x'(\tau)\big) + e(\tau) \right] \mathrm{d}\tau \, \mathrm{d}s \neq 0.$$

(H3) *There exists constant* $M^* > 0$ *such that for any* $d \in \mathbb{R}$, *if* $|d| > M^*$, *then either*

$$d \sum_{i=1}^{m-2} c_i \int_0^{\xi_i} \int_0^s \left[f(\tau, d, 0) + e(\tau) \right] \mathrm{d}\tau \, \mathrm{d}s < 0$$

or else

$$d \sum_{i=1}^{m-2} c_i \int_0^{\xi_i} \int_0^s \left[f(\tau, d, 0) + e(\tau) \right] \mathrm{d}\tau \, \mathrm{d}s > 0.$$

Then, for every $e \in L^1[0, 1]$, *the BVP* (3.51)–(3.52) *in the case* (1), *i.e. when* $\sum_{i=1}^{m-2} c_i \xi_i = 0$, $\gamma = 1$, $\alpha = 1$, $\sum_{i=1}^{m-2} c_i \xi_i^2 \neq 0$, $\sum_{j=1}^{n-2} a_j \eta_j^2 \neq 1$, *has at least one solution in* $C^1[0, 1]$ *provided*

$$\|a\|_1 + \|b\|_1 < \frac{1}{\Delta_1 + 1}, \quad \textit{where } \Delta_1 = a + \frac{\sum_{i=j}^{n-2} |a_j|(1 - \eta_j)}{\left|1 - \sum_{j=1}^{n-2} a_j \eta_j\right|}.$$

3.9. *Positive solutions of some nonlocal boundary value problems*

There is much attention focused on question of positive solutions of BVPs for ordinary differential equations. Much of interest is due to the applicability of certain Krasnosel'skii fixed point theorems or the Leggett–Williams multiple fixed point theorem, or a synthesis of both to obtain positive solutions or multiple positive solutions which lie in a cone. Here we present some of the results on positive solutions of some nonlocal BVPs.

Consider the differential equation

$$x'' + a(t) f(x) = 0, \quad t \in (0, 1) \tag{3.80}$$

with the boundary conditions

$$x(0) = 0, \qquad \alpha x(\eta) = x(1), \tag{3.81}$$

where $0 < \eta < 1$. Our purpose here is to give some existence results for positive solutions to (3.80)–(3.81), assuming that $\alpha\eta < 1$ and f is either superlinear or sublinear.

Set

$$f_0 = \lim_{x\to 0+} \frac{f(x)}{x}, \qquad f_\infty = \lim_{x\to\infty} \frac{f(x)}{x}.$$

Then $f_0 = 0$ and $f_\infty = \infty$ correspond to the superlinear case, and $f_0 = \infty$ and $f_\infty = 0$ correspond to the sublinear case. By the *positive solution* of the BVP (3.80)–(3.81) we understand a function $x(t)$ which is positive on $0 < t < 1$ and satisfies the differential equation (3.80) and the boundary conditions (3.81).

The key tool in our approach is the following Krasnosel'skii's fixed point theorem in a cone [54] (see also [21]).

THEOREM 3.33. *Let E be a Banach space, and let $K \subset E$ be a cone. Assume Ω_1, Ω_2 are open bounded subsets of E with $0 \in \Omega_1$, $\overline{\Omega}_1 \subset \Omega_2$, and let*

$$A : K \cap \left(\overline{\Omega}_2 \setminus \Omega_1\right) \longrightarrow K$$

be a completely continuous operator such that

(i) $\|Au\| \leqslant \|u\|$, $u \in K \cap \partial\Omega_1$, *and* $\|Au\| \geqslant \|u\|$, $u \in K \cap \partial\Omega_2$; *or*

(ii) $\|Au\| \geqslant \|u\|$, $u \in K \cap \partial\Omega_1$, *and* $\|Au\| \leqslant \|u\|$, $u \in K \cap \partial\Omega_2$.

Then A has a fixed point in $K \cap (\overline{\Omega}_2 \setminus \Omega_1)$.

We give first some preliminary lemmas.

LEMMA 3.34. *Let $\alpha\eta \neq 1$ then for $y \in C[0,1]$, the problem*

$$x''(t) + y(t) = 0, \quad t \in (0,1), \tag{3.82}$$

$$x(0) = 0, \qquad \alpha x(\eta) = x(1) \tag{3.83}$$

has a unique solution

$$x(t) = -\int_0^t (t-s)y(s)\,\mathrm{d}s - \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta - s)y(s)\,\mathrm{d}s + \frac{t}{1-\alpha\eta}\int_0^1 (1-s)y(s)\,\mathrm{d}s.$$

LEMMA 3.35. *Let $0 < \alpha < \frac{1}{\eta}$. If $y \in C[0,1]$ and $y \geqslant 0$, for $t \in (0,1)$ then the unique solution x of the problem* (3.82)–(3.83) *satisfies $x(t) \geqslant 0$, $t \in [0,1]$.*

PROOF. From the fact that $x''(t) = -y(t) \leqslant 0$, we know that the graph of $x(t)$ is concave down on $(0,1)$. So, if $x(1) \geqslant 0$, then the concavity of x and the boundary condition $x(0) = 0$ imply that $x \geqslant 0$ for $t \in [0,1]$. If $x(1) < 0$, then we have that $x(\eta) < 0$ and $x(1) = \alpha x(\eta) > \frac{1}{\eta}x(\eta)$ which contradicts to the concavity of x. □

LEMMA 3.36. *Let* $\alpha\eta > 1$. *If* $y \in C[0,1]$ *and* $y \geqslant 0$, *for* $t \in (0,1)$ *then* (3.82)–(3.83) *has no positive solution.*

PROOF. Assume that (3.82)–(3.83) has a positive solution x. If $x(1) > 0$, then $x(\eta) > 0$ and $\frac{x(1)}{1} = \frac{\alpha x(\eta)}{1} > \frac{x(\eta)}{\eta}$ which contradicts to the concavity of x. If $x(1) = 0$ and $x(\tau) > 0$ for some $\tau \in (0,1)$, then $x(\eta) - x(1) = 0$, $\tau = \eta$. If $\tau \in (0,\eta)$, then $x(\tau) > x(\eta) = x(1)$, which contradicts to the concavity of x again. □

LEMMA 3.37. *Let* $0 < \alpha < \frac{1}{\eta}$. *If* $y \in C[0,1]$ *and* $y \geqslant 0$, *for* $t \in (0,1)$ *then the unique solution* x *of the problem* (3.82)–(3.83) *satisfies*

$$\inf_{t\in[\eta,1]} x(t) \geqslant \gamma \|x\|,$$

where $\gamma = \min\{\alpha\eta, \frac{\alpha(1-\eta)}{1-\alpha\eta}, \eta\}$.

PROOF. We divide the proof into two steps.

Step 1. We deal with the case $0 < \alpha < 1$. In this case, by Lemma 3.35, we know that $x(\eta) \geqslant x(1)$. Set

$$x(\bar{t}) = \|x\|.$$

If $\bar{t} \leqslant \eta < 1$, then $\min_{t\in[\eta,1]} x(t) = x(1)$ and

$$x(\bar{t}) \leqslant x(1) + \frac{x(1) - x(\eta)}{1-\eta}(0-1) = x(1)\frac{1-\alpha\eta}{\alpha(1-\eta)}.$$

Then

$$\min_{t\in[\eta,1]} x(t) \geqslant \frac{\alpha(1-\eta)}{1-\alpha\eta}.$$

If $\eta < \bar{t} < 1$, then $\min_{t\in[\eta,1]} x(t) = x(1)$. From the concavity of x, we know that $\frac{x(\eta)}{\eta} \geqslant \frac{x(\bar{t})}{\bar{t}}$ which combined with boundary condition $\alpha x(\eta) = x(1)$, gives $\frac{x(1)}{\alpha\eta} \geqslant \frac{x(\bar{t})}{\bar{t}} \geqslant x(\bar{t}) = \|x\|$. Therefore

$$\min_{t\in[\eta,1]} x(t) \geqslant \alpha\eta\|x\|.$$

Step 2. We deal with the case $1 < \alpha < \frac{1}{\eta}$. In this case, we have $x(\eta) \leqslant x(1)$. Set

$$x(\bar{t}) = \|x\|.$$

Then we can choose $\bar{t}$ such that $\eta \leqslant \bar{t} \leqslant 1$. This contradicts to the concavity of x. From $x(\eta) \leqslant x(1)$ and the concavity of x, we know that $\min_{t\in[\eta,1]} x(t) = x(\eta)$. Using the concavity of x and Lemma 3.35, we have $\frac{x(\eta)}{\eta} \geqslant \frac{x(\bar{t})}{\bar{t}}$ which implies that

$$\min_{t\in[\eta,1]} x(t) \geqslant \eta\|x\|.$$

This completes the proof. □

The main result here is the following:

THEOREM 3.38. *Assume that*

(A1) $f \in C([0,\infty),[0,\infty))$;

(A2) $a \in C([0,1],[0,\infty))$ *and there exists* $x_0 \in [\eta,1]$ *such that* $a(x_0) > 0$.

Then the BVP (3.80)–(3.81) *has at least one positive solution in the case*

(i) $f_0 = 0$ *and* $f_\infty = \infty$ (*superlinear*) *or*

(ii) $f_0 = \infty$ *and* $f_\infty = 0$ (*sublinear*).

PROOF. (i) Suppose that $f_0 = 0$ and $f_\infty = \infty$. We wish to show the existence of a positive solution of (3.80)–(3.81). Now (3.80)–(3.81) has a solution $x = x(t)$ if and only if x solves the operator equation

$$\begin{aligned} x(t) &= -\int_0^t (t-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s - \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\ &\quad + \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\ &:= Ax(t). \end{aligned}$$

Denote

$$K = \Bigl\{x\colon\ x \in C[0,1], x \geqslant 0, \min_{\eta\leqslant t\leqslant 1} x(t) \geqslant \gamma\|x\|\Bigr\}.$$

It is obvious that K is a cone in $C[0,1]$. Moreover, by Lemma 3.37, $AK \subset K$. It is also easy to check that $A : K \to K$ is completely continuous.

Now $f_0 = 0$, we may choose $H_1 > 0$ so that $f(x) \leqslant \epsilon x$, for $0 < x < H_1$, where $\epsilon > 0$ satisfies

$$\frac{\epsilon}{1-\alpha\eta}\int_0^1 (1-s)a(s)\,\mathrm{d}s \leqslant 1.$$

Thus, if $x \in K$ and $\|x\| = H_1$ we get

$$Ax(t) \leqslant \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s$$

$$\begin{aligned}
&\leqslant \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)\epsilon x(s)\,\mathrm{d}s\\
&\leqslant \frac{\epsilon}{1-\alpha\eta}\int_0^1 (1-s)a(s)\,\mathrm{d}s\|x\|\\
&\leqslant \frac{\epsilon}{1-\alpha\eta}\int_0^1 (1-s)a(s)\,\mathrm{d}sH_1.
\end{aligned}$$

Now if we let

$$\Omega_1 = \bigl\{x \in C[0,1]:\ \|x\| < H_1\bigr\},$$

then $\|Ax\| \leqslant \|x\|$, for $x \in K \cap \partial\Omega_1$.

Further since $f_\infty = \infty$, there exists $\widehat{H}_2 > 0$ such that $f(x) \geqslant \rho x$, for $x \geqslant \widehat{H}_2$, where ρ is chosen so that

$$\rho\frac{\eta\gamma}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)\,\mathrm{d}s > 1.$$

Let $H_2 = \max\{2H_1, \frac{\widehat{H}_2}{\gamma}\}$ and $\Omega_2 = \{x \in C[0,1]:\ \|x\| < H_2\}$, then $x \in K$ and $\|x\| = H_2$ implies

$$\min_{\eta\leqslant t\leqslant 1} x(t) \geqslant \gamma\|x\| \geqslant \widehat{H}_2,$$

and so

$$\begin{aligned}
Ax(\eta) &= -\int_0^\eta (\eta-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s - \frac{\alpha\eta}{1-\alpha\eta}\int_0^\eta (\eta-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s\\
&\quad + \frac{\eta}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s\\
&= -\frac{1}{1-\alpha\eta}\int_0^\eta (\eta-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s\\
&\quad + \frac{\eta}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s\\
&= \frac{\eta}{1-\alpha\eta}\int_\eta^1 a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s + \frac{1}{1-\alpha\eta}\int_0^\eta sa(s)f\bigl(x(s)\bigr)\,\mathrm{d}s\\
&\quad - \frac{\eta}{1-\alpha\eta}\int_0^1 sa(s)f\bigl(x(s)\bigr)\,\mathrm{d}s\\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_\eta^1 a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s
\end{aligned}$$

$$- \frac{\eta}{1-\alpha\eta}\int_{\eta}^{1} sa(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \quad (\text{by } \eta < 1)$$

$$= \frac{\eta}{1-\alpha\eta}\int_{\eta}^{1} (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s.$$

Hence, for $x \in K \cap \partial\Omega_2$,

$$\|Ax\| \geqslant \rho\frac{\eta\gamma}{1-\alpha\eta}\int_{\eta}^{1}(1-s)a(s)\,\mathrm{d}s\,\|x\| \geqslant \|x\|.$$

Therefore, by the first part of Theorem 3.33, it follows that A has a fixed point in $K \cap (\overline{\Omega_2} \setminus \Omega_1)$, such that $H_1 \leqslant \|x\| \leqslant H_2$. This completes the superlinear part of the theorem.

(ii) Suppose next that $f_0 = \infty$ and $f_\infty = 0$. We first choose $H_3 > 0$ such that $f(x) \geqslant Mx$ for $0 < x < H_3$, where

$$M\gamma\left(\frac{\eta}{1-\alpha\eta}\right)\int_{\eta}^{1}(1-s)a(s)\,\mathrm{d}s \geqslant 1.$$

Then

$$\begin{aligned}
Ax(\eta) &= -\int_0^{\eta}(\eta-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s - \frac{\alpha\eta}{1-\alpha\eta}\int_0^{\eta}(\eta-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\quad + \frac{\eta}{1-\alpha\eta}\int_0^{1}(1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_{\eta}^{1}(1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_{\eta}^{1}(1-s)a(s)Mx(s)\,\mathrm{d}s \\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_{\eta}^{1}(1-s)a(s)M\gamma\,\mathrm{d}s\,\|x\| \\
&\geqslant H_3.
\end{aligned}$$

Thus, we may let $\Omega_3 = \{x \in C[0,1]\colon\ \|x\| < H_3\}$ so that

$$\|Ax\| \geqslant \|x\|, \quad x \in K \cap \partial\Omega_3.$$

Now, since $f_\infty = 0$, there exist $\widehat{H}_4 > 0$ so that $f(x) \leqslant \lambda x$ for $x \geqslant \widehat{H}_4$, where $\lambda > 0$ satisfies

$$\frac{\lambda}{1-\alpha\eta}\int_0^{1}(1-s)a(s)\,\mathrm{d}s \leqslant 1.$$

We consider two cases:

Case 1. Suppose f is bounded, say $f(x) \leqslant N$ for all $x \in [0, \infty)$. In this case choose

$$H_4 = \max\left\{2H_3, \frac{N}{1-\alpha\eta}\int_0^1 (1-s)a(s)\,\mathrm{d}s\right\}$$

so that for $x \in K$ with $\|x\| = H_4$ we have

$$\begin{aligned}
Ax(t) &= -\int_0^t (t-s)a(s)f\big(x(s)\big)\,\mathrm{d}s - \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta-s)a(s)f\big(x(s)\big)\,\mathrm{d}s \\
&\quad + \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\big(x(s)\big)\,\mathrm{d}s \\
&\leqslant \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\big(x(s)\big)\,\mathrm{d}s \\
&\leqslant \frac{1}{1-\alpha\eta}\int_0^1 (1-s)a(s)N\,\mathrm{d}s \\
&\leqslant H_4
\end{aligned}$$

and therefore $\|Ax\| \leqslant \|x\|$.

Case 2. If f is unbounded, then we know from (A1) that there is H_4: $H_4 > \max\{2H_3, \frac{1}{\gamma}\widehat{H}_4\}$ such that

$$f(x) \leqslant f(H_4) \quad \text{for } 0 < x \leqslant H_4.$$

Then for $x \in K$ and $\|x\| = H_4$ we have

$$\begin{aligned}
Ax(t) &= -\int_0^t (t-s)a(s)f\big(x(s)\big)\,\mathrm{d}s - \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta-s)a(s)f\big(x(s)\big)\,\mathrm{d}s \\
&\quad + \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\big(x(s)\big)\,\mathrm{d}s \\
&\leqslant \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f(H_4)\,\mathrm{d}s \\
&\leqslant \frac{1}{1-\alpha\eta}\int_0^1 (1-s)a(s)\lambda H_4\,\mathrm{d}s \\
&\leqslant H_4.
\end{aligned}$$

Therefore, in either case we may put

$$\Omega_4 = \big\{x \in C[0,1]:\ \|x\| < H_4\big\},$$

and for $x \in K \cap \partial\Omega_4$ we may have $\|Ax\| \leqslant \|x\|$. By the second part of Theorem 3.33, if follows that the BVP (3.80)–(3.81) has a positive solution. The proof of the theorem is completed. □

The previous results are taken from [69]. Theorem 3.38 was proved either for sub- or superlinear case. Therefore the following questions are natural:

- *Whether or not we can obtain similar conclusion, if* $f_0 = f_\infty = 0$ *or* $f_0 = f_\infty = \infty$;
- *Whether or not we can obtain similar conclusion, if* $f_0, f_\infty \notin \{0, \infty\}$.

Motivated by the results in [69], Liu in [57] established some simple criteria for the existence of positive solutions of the BVP (3.80)–(3.81), which gives a positive answer to the questions stated above. The key tool in his approach is the following fixed point index theorem [21].

THEOREM 3.39. *Let E be Banach space and $K \subset E$ be a cone in E. Let $r > 0$, and define $\Omega_r = \{x \in K\colon \|x\| < r\}$. Assume $A : \overline{\Omega}_r \to K$ is a completely continuous operator such that $Ax \neq x$ for $x \in \partial\Omega_r$.*

(i) *If $\|Ax\| \leqslant \|x\|$ for $x \in \partial\Omega_r$, then $i(A, \Omega_r, K) = 1$.*
(ii) *If $\|Ax\| \geqslant \|x\|$ for $x \in \partial\Omega_r$, then $i(A, \Omega_r, K) = 0$.*

In what follows, for the sake of convenience, set

$$\Lambda_1 = (1 - \alpha\eta)\left(\int_0^1 (1 - s)a(s)\,\mathrm{d}s\right)^{-1},$$

$$\Lambda_2 = (1 - \alpha\eta)\left(\eta\gamma \int_\eta^1 (1 - s)a(s)\,\mathrm{d}s\right)^{-1}.$$

THEOREM 3.40. *Assume that the following assumptions are satisfied*:

(H1) $f_0 = f_\infty = \infty$.
(H2) *There exist constants $\rho_1 > 0$ and $M_1 \in (0, \Lambda_1)$ such that $f(u) \leqslant M_1\rho_1$, $u \in [0, \rho_1]$.*

Then the BVP (3.80)–(3.81) *has at least two positive solutions x_1 and x_2 such that*

$$0 < \|x_1\| < \rho_1 < \|x_2\|.$$

PROOF. At first, in view of $f_0 = \infty$, then for any $M_* \in (\Lambda_2, \infty)$, there exist $\rho_* \in (0, \rho_1)$ such that

$$f(x) \geqslant M_* x, \quad 0 \leqslant x \leqslant \rho_*.$$

Set $\Omega_{\rho_*} = \{x \in K\colon \|x\| < \rho_*\}$. Since $x \in \partial\Omega_{\rho_*} \subset K$, we have $\min_{\eta \leqslant t \leqslant 1} x(t) \geqslant \gamma\|x\|$. Thus, for any $x \in \partial\Omega_{\rho_*}$, we have

$$Ax(\eta) = -\int_0^\eta (\eta - s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s - \frac{\alpha\eta}{1 - \alpha\eta}\int_0^\eta (\eta - s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s$$

$$
\begin{aligned}
&\quad + \frac{\eta}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&= \frac{\eta}{1-\alpha\eta}\int_\eta^1 a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s + \frac{1-\eta}{1-\alpha\eta}\int_0^\eta sa(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\quad - \frac{\eta}{1-\alpha\eta}\int_\eta^1 sa(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)M_* x(s)\,\mathrm{d}s \\
&\geqslant \frac{\gamma\eta M_*}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)\|x\|\,\mathrm{d}s \\
&> \|x\|,
\end{aligned}
$$

which yields

$$
\|Ax\| > \|x\|, \quad \text{for } y \in \partial\Omega_{\rho_*}.
$$

Hence, Theorem 3.39 implies $i(A, \Omega_{\rho_*}, K) = 0$.

Next, since $f_\infty = \infty$, then for any $M^* \in (\Lambda_2, \infty)$, then there exist $\rho^* > \rho_1$ such that

$$
f(x) \geqslant M^* x, \quad 0 \leqslant x \leqslant \gamma\rho^*.
$$

Set $\Omega_{\rho^*} = \{x \in K\colon \|x\| < \rho^*\}$ for $x \in \partial\Omega_{\rho^*}$. Since $x \in K$, we have $\min_{\eta \leqslant t \leqslant 1} x(t) \geqslant \gamma\|x\| = \gamma\rho^*$. Thus, for any $x \in \partial\Omega_{\rho^*}$, we have

$$
\begin{aligned}
Ax(\eta) &= -\int_0^\eta (\eta-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s - \frac{\alpha\eta}{1-\alpha\eta}\int_0^\eta (\eta-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\quad + \frac{\eta}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)M^* x(s)\,\mathrm{d}s \\
&\geqslant \frac{\gamma\eta M^*}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)\|x\|\,\mathrm{d}s \\
&> \|x\|,
\end{aligned}
$$

which implies

$$\|Ax\| > \|x\|, \quad \text{for } y \in \partial\Omega_{\rho^*}.$$

Hence, Theorem 3.39 yields $i(A, \Omega_{\rho_*}, K) = 0$.

Finally, set $\Omega_{\rho_1} = \{x \in K\colon \|x\| < \rho_1\}$. For any $x \in \partial\Omega_{\rho_1}$ we obtain

$$\begin{aligned} Ax(t) &\leqslant \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\ &\leqslant \frac{M_1\rho_1}{1-\alpha\eta}\int_0^1 (1-s)a(s)\,\mathrm{d}s \\ &< \rho_1 = \|x\|, \end{aligned}$$

which yields

$$\|Ax\| < \|x\|, \quad \text{for } y \in \partial\Omega_{\rho_1}.$$

Thus, Theorem 3.39 again shows that $i(A, \Omega_{\rho_1}, K) = 1$.

Hence, since $\rho_* < \rho_1 < \rho^*$ it follows from the additivity of the fixed point index that

$$i\bigl(A, \Omega_{\rho_1} \setminus \overline{\Omega}_{\rho_*}, K\bigr) = 1, \qquad i\bigl(A, \Omega_{\rho^*} \setminus \overline{\Omega}_{\rho_1}, K\bigr) = -1.$$

Thus, A has a fixed point x_1 in $\Omega_{\rho_1} \setminus \overline{\Omega}_{\rho_*}$, and a fixed point x_2 in $\Omega_{\rho^*} \setminus \overline{\Omega}_{\rho_1}$. Both are positive solutions, $0 < \|x_1\| < \rho_1 < \|x_2\|$, and the proof of the theorem is complete. □

THEOREM 3.41. *Assume that the following assumptions are satisfied*:

(H3) $f_0 = f_\infty = 0$.

(H4) *There exist constants* $\rho_2 > 0$ *and* $M_2 \in (\Lambda_2, \infty)$ *such that* $f(u) \geqslant M_2\rho_2$, $u \in [\gamma\rho_2, \rho_2]$.

Then the BVP (3.80)–(3.81) *has at least two positive solutions* x_1 *and* x_2 *such that*

$$0 < \|x_1\| < \rho_2 < \|x_2\|.$$

PROOF. First, since $f_0 = 0$, for any $\varepsilon \in (0, \Lambda_1)$, there exists $\rho_* \in (0, \rho_2)$ such that

$$f(x) \leqslant \varepsilon x, \quad \text{for } x \in [0, \rho_*].$$

Setting $\Omega_{\rho_*} = \{x \in K\colon \|x\| < \rho_*\}$ for any $x \in \partial\Omega_{\rho_*}$, we get

$$\begin{aligned} Ax(t) &\leqslant \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s \\ &\leqslant \frac{\varepsilon\rho_*}{1-\alpha\eta}\int_0^1 (1-s)a(s)\,\mathrm{d}s < \rho_* = \|x\|, \end{aligned}$$

which yields

$$\|Ax\| < \|x\|, \quad \text{for } x \in \partial\Omega_{\rho_*}.$$

Thus Theorem 3.39 implies $i(A, \Omega_{\rho_*}, K) = 1$.

Second, in view $f_\infty = 0$, then for any $\varepsilon \in (0, \Lambda_1)$, there exists $\rho_0 > \rho_2$ such that

$$f(x) \leqslant \varepsilon x, \quad \text{for } x \in [\rho_0, \infty),$$

and we consider two cases.

Case (i). Suppose that f is unbounded; then from $f \in C([0,\infty),[0,\infty))$, we know that there is $\rho^* > \rho_0$ such that $f(x) \leqslant f(\rho^*)$, for $x \in [0, \rho^*]$. Since $\rho^* > \rho_0$, one has $f(x) \leqslant f(\rho^*) \leqslant \varepsilon\rho^*$, for $x \in [0, \rho^*]$. For $x \in K$, $\|x\| = \rho^*$ we obtain

$$\begin{aligned} Ax(t) &\leqslant \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\big(x(s)\big)\,\mathrm{d}s \\ &\leqslant \frac{\varepsilon\rho_*}{1-\alpha\eta}\int_0^1 (1-s)a(s)\,\mathrm{d}s < \rho_* = \|x\|. \end{aligned}$$

Case (ii). Suppose that f is bounded, say $f(x) \leqslant L$. Taking $\rho^* \geqslant \max\{L/\varepsilon, \rho_2\}$, for $x \in K$, $\|x\| = \rho^*$ one has

$$\begin{aligned} Ax(t) &\leqslant \frac{t}{1-\alpha\eta}\int_0^1 (1-s)a(s)f\big(x(s)\big)\,\mathrm{d}s \\ &\leqslant \frac{L}{1-\alpha\eta}\int_0^1 (1-s)a(s)\,\mathrm{d}s < \rho_* = \|x\|. \end{aligned}$$

Hence, in either case, we always may set $\Omega_{\rho^*} = \{x \in K\colon \|x\| < \rho^*\}$ such that

$$\|Ax\| < \|x\|, \quad \text{for } x \in \partial\Omega_{\rho_*}.$$

Thus Theorem 3.39 implies $i(A, \Omega_{\rho_*}, K) = 1$.

Finally set $\Omega_{\rho_2} = \{x \in K\colon \|x\| < \rho_2\}$, for $x \in \partial\Omega_{\rho_2}$, since $x \in K$, $\min_{\eta\leqslant t\leqslant 1} x(t) \geqslant \gamma\|x\| = \gamma\rho_2$, and hence we can get

$$\begin{aligned} Ax(\eta) &\geqslant \frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)f\big(x(s)\big)\,\mathrm{d}s \\ &\geqslant \frac{\gamma\eta M_2}{1-\alpha\eta}\rho_2\int_\eta^1 (1-s)a(s)\,\mathrm{d}s > \rho_2 = \|x\|, \end{aligned}$$

which yields

$$\|Ax\| > \|x\|, \quad \text{for } x \in \partial\Omega_{\rho_2}.$$

Thus Theorem 3.39 implies $i(A, \Omega_{\rho_2}, K) = 0$.

Hence, since $\rho_* < \rho_2 < \rho^*$ it follows from the additivity of the fixed point index that

$$i\big(A, \Omega_{\rho_2} \setminus \overline{\Omega}_{\rho_*}, K\big) = -1, \qquad i\big(A, \Omega_{\rho^*} \setminus \overline{\Omega}_{\rho_2}, K\big) = 1.$$

Thus, A has a fixed point x_1 in $\Omega_{\rho_2} \setminus \overline{\Omega}_{\rho_*}$, and a fixed point x_2 in $\Omega_{\rho^*} \setminus \overline{\Omega}_{\rho_2}$. Both are positive solutions, $0 < \|x_1\| < \rho_2 < \|x_2\|$, and the proof of the theorem is complete. □

For the case when $f_0, f_\infty \notin \{0, \infty\}$ we have the following

THEOREM 3.42. *Suppose that* (H2) *and* (H4) *hold and that* $\rho_1 \neq \rho_2$. *Then the BVP* (3.80)–(3.81) *has at least one positive solution* x *satisfying* $\rho_1 < \|x\| < \rho_2$ *or* $\rho_2 < \|x\| < \rho_1$.

Theorems 3.40–3.42 were taken from [57]. A method to study the existence of positive solutions of some three-point BVPs is to write the BVP as an equivalent Hammerstein integral equation of the form

$$x(t) = \int_0^1 k(t,s)a(s)f\big(x(s)\big)\,\mathrm{d}s \equiv Tx(t),$$

and seek fixed points of T in the cone of positive functions in the space $C[0,1]$. We use the fixed point index for compact maps, which is based on Leray–Schauder degree theory, and use a well-known nonzero fixed point theorem in order to prove that T has a positive fixed point.

We apply this method for the BVP (3.80)–(3.81) using the cone defined by

$$K = \big\{x \in C[0,1]\colon\ x \geqslant 0,\ \min\big\{x(t)\colon\ a \leqslant t \leqslant b\big\} \geqslant c\|x\|\big\},$$

and the following kernel (Green's function) in the Hammerstein integral operator

$$k(t,s) = \frac{1}{1-\alpha\eta}t(1-s) - \begin{cases} \dfrac{\alpha t}{1-\alpha\eta}(\eta - s), & s \leqslant \eta \\ 0, & s > \eta \end{cases} - \begin{cases} t-s, & s \leqslant t \\ 0, & s > t. \end{cases}$$

Let $a, b \in (0,1]$ and suppose that $\int_a^b a(t)\,\mathrm{d}t > 0$. We are able to find upper and lower bounds for $k(t,s)$ with s fixed, of the same type. We have the same freedom in choosing the numbers a and b. Let us find the upper and lower bounds for the BVP (3.80)–(3.81). We have to exhibit $\Phi(s)$, a subinterval $[a,b] \subset [0,1]$ and a constant $c \in (0,1]$ such that

$$\begin{aligned} &k(t,s) \leqslant \Phi(s), && \text{for every } t, s \in [0,1], \\ &k(t,s) \geqslant c\Phi(s), && \text{for every } s \in [0,1], t \in [a,b]. \end{aligned}$$

3.9.1. *Upper bounds* We shall show that we may take $\Phi(s) = \max\{1, \alpha\}\frac{s(1-s)}{1-\alpha\eta}$. (In this case $\alpha\eta < 1$ but it is possible to have $\alpha > 1$.)

Case 1: $s > \eta$. $t < s$ is simple so consider $t \geqslant s$. Then $k(t,s) = s + \frac{(\alpha\eta - s)t}{1-\alpha\eta}$. If also $s \geqslant \alpha\eta$ then it is a decreasing function of t so the maximum occurs when $t = s$ and

$$k(t,s) \leqslant \frac{s - s\alpha\eta + s(\alpha\eta - s)}{1-\alpha\eta} = \frac{s(1-s)}{1-\alpha\eta} \leqslant \Phi(s).$$

If $s < \alpha\eta$ (which can happen only if $\alpha > 1$) the maximum occurs at $t = 1$ and then

$$k(t,s) \leqslant \frac{\alpha\eta(1-s)}{1-\alpha\eta} < \alpha\frac{s(1-s)}{1-\alpha\eta} = \Phi(s).$$

Case 2: $s \leqslant \eta$. For $t \leqslant s$ we clearly have $k(t,s) \leqslant \frac{s(1-s)}{1-\alpha\eta} \leqslant \Phi(s)$. So consider $t \geqslant s$. Then $k(t,s) = s + \frac{st(\alpha-1)}{1-\alpha\eta}$. For $\alpha > 1$ the maximum occurs when $t = 1$ and then

$$k(t,s) \leqslant \frac{\alpha s(1-\eta)}{1-\alpha\eta} \leqslant \alpha\frac{s(1-s)}{1-\alpha\eta} = \Phi(s),$$

using $s \leqslant \eta$. For $\alpha \leqslant 1$ the maximum occurs when $t = s$ and then

$$k(t,s) = \frac{\alpha s(s-\eta) + s(1-s)}{1-\alpha\eta} \leqslant \frac{s(1-s)}{1-\alpha\eta} = \Phi(s).$$

3.9.2. *Lower bounds* We will show that we may take arbitrary $a > 0$ and $b \leqslant 1$.

Case 1: $s > \eta$. For $t < s$, $k(t,s) = \frac{t(1-s)}{1-\alpha\eta} \geqslant a\frac{s(1-s)}{1-\alpha\eta} \geqslant a\eta\Phi(s)$. For $t \geqslant s$, $k(t,s) = s - \frac{t(s-\alpha\eta)}{1-\alpha\eta}$. If $s > \alpha\eta$ then the minimum occurs when $t = 1$ so

$$k(t,s) \geqslant \frac{\alpha\eta(1-s)}{1-\alpha\eta} \geqslant \alpha\eta\frac{s(1-s)}{1-\alpha\eta} = \begin{cases} \alpha\eta\Phi(s), & \alpha \leqslant 1, \\ \eta\Phi(s), & \alpha > 1. \end{cases}$$

If $s \leqslant \alpha\eta$ (only possible if $\alpha > 1$) the minimum is at $t = s$ and

$$k(t,s) \geqslant \frac{s(1-s)}{1-\alpha\eta} \geqslant \eta\Phi(s).$$

Case 2: $s \leqslant \eta$. First suppose $a \leqslant t \leqslant s$. (This case cannot occur if $\alpha \geqslant \eta$.) Then $k(t,s) = \frac{t[1-\alpha\eta+(\alpha-1)s]}{1-\alpha\eta}$. For $\alpha < 1$ we have $[1-\alpha\eta+(\alpha-1)s] \geqslant 1-\alpha\eta+(\alpha-1)\eta = 1-\eta$ so $k(t,s)$ is increasing in t and has a minimum when $t = a$. Thus

$$k(t,s) \geqslant \frac{a}{1-\alpha\eta}[1-\eta] \geqslant 4a(1-\eta)\Phi(s),$$

since $s(1-s) \leqslant 1/4$. Note that $4a(1-\eta) < 1$ since $\alpha < \eta$.

For $\alpha > 1$, $k(t,s)$ is clearly increasing so we have

$$k(t,s) \geqslant \frac{a}{1-\alpha\eta}\big[1-\alpha\eta+(\alpha-1)s\big] \geqslant \frac{a}{1-\alpha\eta}[1-\alpha\eta] \geqslant \frac{4(1-\alpha\eta)a}{\alpha}\Phi(s).$$

Note that $\frac{4(1-\alpha\eta)a}{\alpha} = \frac{4a\alpha\eta(1-\alpha\eta)}{\alpha^2\eta} < \frac{1}{\alpha^2} < 1$ since $\alpha < \eta$.

Now suppose that $t \geqslant s$. Then $k(t,s) = s + \frac{t(\alpha-1)s}{1-\alpha\eta}$. For $\alpha \leqslant 1$ the minimum occurs when $t = 1$ and then

$$k(t,s) \geqslant \frac{\alpha(1-\eta)s}{1-\alpha\eta} \geqslant \alpha(1-\eta)\Phi(s).$$

For $\alpha > 1$ the minimum occurs either at $t = s$ or $t = a$ but in both cases

$$k(t,s) \geqslant s \geqslant \frac{1-\alpha\eta}{\alpha}\Phi(s) \geqslant \eta(1-\alpha\eta)\Phi(s).$$

The conclusion is that we may take

$$c = \begin{cases} \min\bigl\{a, \alpha\eta, 4a(1-\eta), \alpha(1-\eta)\bigr\}, & \alpha < 1, \\ \min\bigl\{a\eta, 4a(1-\alpha\eta), \eta(1-\alpha\eta)\bigr\}, & \alpha \geqslant 1. \end{cases}$$

Hence we have:

THEOREM 3.43. *Let $a, b \in (0,1]$ and suppose that $\int_a^b a(t)\,\mathrm{d}t > 0$. Let c as defined above and*

$$m = \left(\max_{0\leqslant t\leqslant 1} \int_0^1 k(t,s)a(s)\,\mathrm{d}s\right)^{-1}, \qquad M = \left(\min_{0\leqslant t\leqslant 1} \int_a^b k(t,s)a(s)\,\mathrm{d}s\right)^{-1}.$$

Then for $0 < \alpha\eta < 1$ the BVP (3.80)–(3.81) has at least one positive solution if either

$$\text{(h1)} \quad 0 \leqslant \limsup_{x\to 0} \frac{f(x)}{x} < m \quad \textit{and} \quad M < \liminf_{x\to\infty} \frac{f(x)}{x} \leqslant \infty,$$

or

$$\text{(h2)} \quad 0 \leqslant \limsup_{x\to \infty} \frac{f(x)}{x} < m \quad \textit{and} \quad M < \liminf_{x\to 0} \frac{f(x)}{x} \leqslant \infty,$$

and has two positive solutions if there is $\rho > 0$ such that either

$$\text{(E1)} \quad \begin{cases} 0 \leqslant \limsup\limits_{x\to 0} \dfrac{f(x)}{x} < m, \\ \min\left\{\dfrac{f(x)}{\rho}\colon x \in [c\rho, \rho]\right\} \geqslant cM,\ x \neq Tx \quad \textit{for } x \in \partial\Omega_\rho, \quad \textit{and} \\ 0 \leqslant \limsup\limits_{x\to \infty} \dfrac{f(x)}{x} < m, \end{cases}$$

or

$$\text{(E2)}\quad \begin{cases} M < \limsup\limits_{x\to 0} \dfrac{f(x)}{x} \leqslant \infty, \\ \max\left\{\dfrac{f(x)}{\rho}\colon\ x \in [0,\rho]\right\} \leqslant m,\ x \neq Tx \quad \textit{for } x \in \partial K_\rho, \quad \textit{and} \\ M < \limsup\limits_{x\to \infty} \dfrac{f(x)}{x} \leqslant \infty, \end{cases}$$

where

$$\Omega_\rho = \Bigl\{x \in K\colon\ c\|x\| \leqslant \min_{a\leqslant t\leqslant b} x(t) < c\rho\Bigr\}, \qquad K_r = \bigl\{x \in K\colon\ \|x\| < \rho\bigr\}$$

and K is a closed convex set in a Banach space.

Theorem 3.43 is taken from Webb [92]. Similar results are given by Infante [45] who proved existence of eigenvalues of some nonlocal BVP, including nonlocal boundary conditions other than (3.81), as

(a) $x'(0) = 0, \alpha x'(\eta) = x(1), 0 < \eta < 1,$
(b) $x(0) = 0, \alpha x'(\eta) = x(1), 0 < \eta < 1,$
(c) $x'(0) = 0, \alpha x(\eta) = x(1), 0 < \eta < 1.$

Infante and Webb [47] obtained more general results for existence of positive solutions of some m-point BVPs in the case when all the parameters occurring in the boundary conditions are not positive and the nonlinear term allow more general behaviour than being either sub- or superlinear. Two four-points BVPs are studied in details in [47].

Also interesting results were given for the following three-point BVP

$$x'' + a(t)f(x) = 0, \quad t \in (0,1), \tag{3.84}$$

$$x'(0) = 0, \qquad \alpha x(\eta) = x(1), \tag{3.85}$$

where $0 < \eta < 1$, $0 < \alpha < 1$ by Liu in [56]. By applying Krasnosel'skii's fixed point theorem he proved existence

- of single positive solution under $f_0 = 0, f_\infty = \infty$ or $f_0 = \infty, f_\infty = 0$,
- of two positive solution under $f_0 = f_\infty = \infty$ or $f_0 = f_\infty = 0$,
- of positive solutions under $f_0, f_\infty \notin \{0, \infty\}$.

In [56] some interesting examples were given to demonstrate the results.

A more general three-point BVP was studied by Ma and Wang. In [79] they studied the existence of positive solutions of the following BVP

$$x''(t) + a(t)x'(t) + b(t)x(t) + h(t)f(x) = 0, \quad t \in (0,1), \tag{3.86}$$

$$x(0) = 0, \qquad x(1) = \alpha x(\eta), \tag{3.87}$$

where $f \in C([0,1],[0,\infty))$, $h \in C([0,1],[0,\infty))$ and there exists $x_0 \in [0,1]$ such that $h(x_0) > 0$, and $a \in C[0,1]$, $b \in C([0,1],(-\infty,0))$. Under these assumptions they proved the existence of at least one positive solution of the BVP (3.86)–(3.87) either in superlinear or sublinear case, by applying the Krasnosel'skii's fixed point theorem in cones.

If in the BVP (3.80)–(3.81) the boundary condition (3.81) is nonhomogeneous, i.e.

$$x(0) = 0, \qquad x(1) - \alpha x(\eta) = b, \tag{3.88}$$

Ma in [78], by using Schauder's fixed point theorem, proved in the superlinear case that, if (A1) and (A2) of Theorem 3.38 hold, then there exists a positive number b^* such that (3.80), (3.88) has at least one positive solution for b: $0 < b < b^*$ and no solution for $b > b^*$. This result was improved and complemented by Zhang and Wang in [95], where they studied the BVP (3.80), (3.88) under the following conditions:

(H1) $\alpha \in (0, 1/\eta)$, $\eta \in (0, 1)$;

(H2) $a(t)$ is a nonnegative measurable function defined on $(0, 1)$ and

$$0 \leqslant \int_0^\eta s a(s)\,\mathrm{d}s < +\infty, \qquad 0 < \int_\eta^1 (1-s)a(s)\,\mathrm{d}s < +\infty;$$

(H3) $f : [0, \infty) \to [0, \infty)$ is continuous;

(H3*) $f : [0, \infty) \to [0, \infty)$ is locally Lipschitz continuous;

(H4) $\limsup_{u\to+\infty} \frac{f(u)}{u} < \delta$;

(H5) $\liminf_{u\to+\infty} \frac{f(u)}{u} > M$;

(H6) $\lim_{u\to 0+} \frac{f(u)}{u} < \delta$,

where δ, M are suitable defined constants, and proved that:

- the BVP (3.80), (3.88) has a positive solution for all $b > 0$ if (H1)–(H4) hold;
- there exists a positive number b^* such that (3.80), (3.88) has at least one positive solution for b: $0 < b < b^*$ and no solution for $b > b^*$ if (H1)–(H3), (H5) and (H6) hold;
- there exists a positive number b^* such that (3.80), (3.88) has at least two solution for b: $0 < b < b^*$, at least one for $b = b^*$, none for $b > b^*$ if (H1), (H2), (H3*), (H5) and (H6) hold.

We remark that (H2) allows $a(t)$ to be singular at $t = 0$ and/or $t = 1$, and (H5) and (H6) allow but not require the nonlinearity $f(x)$ to be superlinear at zero and infinity.

Now, we consider the following three-point nonlinear second-order BVP

$$x''(t) + \lambda a(t) f\bigl(x(t)\bigr) = 0, \quad t \in (0, 1), \tag{3.89}$$

$$x(0) = 0, \qquad \alpha x(\eta) = x(1), \tag{3.90}$$

where $0 < \eta < 1$. For this problem, an open interval of eigenvalues is determined, which in return, imply the existence of a positive solution of (3.89)–(3.90) by appealing to Krasnosel'skii's fixed point theorem.

For the sake of simplicity, we let

$$A = (1-\alpha\eta)^{-1} \int_0^1 (1-\eta)a(s)\,\mathrm{d}s, \qquad B = \eta(1-\alpha\eta)^{-1} \int_0^1 (1-\eta)a(s)\,\mathrm{d}s.$$

THEOREM 3.44. *Assume that*

(A1) $f \in C([0,\infty),[0,\infty))$;

(A2) $a \in C([0,1],[0,\infty))$ *and does not vanish identically on any subinterval.*

(A3) $\lim_{x\to 0+} \frac{f(x)}{x} = l$ *with* $0 < l < \infty$.

(A4) $\lim_{x\to \infty} \frac{f(x)}{x} = L$ *with* $0 < L < \infty$.

Then, for each λ *satisfying either*

(i) $\frac{1}{\gamma BL} < \lambda < \frac{1}{Al}$

or

(ii) $\frac{1}{\gamma Bl} < \lambda < \frac{1}{AL}$

the BVP (3.89)–(3.90) *has at least one positive solution.*

PROOF. Let λ given as in (i), and choose $\epsilon > 0$ such that

$$\frac{1}{\gamma B(L-\epsilon)} \leqslant \lambda \leqslant \frac{1}{A(l+\epsilon)}.$$

Consider also the cone K and the operator A defined in the proof of Theorem 3.38. By (A3), there exists $H_1 > 0$ such that $f(x) \leqslant (l+\epsilon)x$, for $0 < x \leqslant H_1$. So, choosing $x \in K$ with $\|x\| = H_1$, we have

$$\begin{aligned}
Ax(t) &= \lambda \frac{t}{1-\alpha\eta} \int_0^1 (1-s)a(s) f\bigl(x(s)\bigr)\,\mathrm{d}s \\
&\leqslant \lambda \frac{t}{1-\alpha\eta} \int_0^1 (1-s)a(s)(l+\epsilon)x(s)\,\mathrm{d}s \\
&\leqslant \lambda \frac{1}{1-\alpha\eta} \int_0^1 (1-s)a(s)(l+\epsilon)\|x\|\,\mathrm{d}s \\
&\leqslant \lambda \frac{1}{1-\alpha\eta} \int_0^1 (1-s)a(s)(l+\epsilon)H_1\,\mathrm{d}s \\
&\leqslant \lambda A(l+\epsilon)\|x\| \leqslant \|x\|.
\end{aligned}$$

Consequently, $\|Ax\| \leqslant \|x\|$. So, if we set

$$\Omega_1 = \bigl\{x \in C[0,1]\colon\ \|x\| < H_1\bigr\},$$

then $\|Ax\| \leqslant \|x\|$, for $x \in K \cap \partial\Omega_1$.

Next we construct the set Ω_2. Considering (A4) there exists $\overline{H}_2$ such that $f(x) \geqslant (L-\epsilon)x$, for $x \geqslant \overline{H}_2$. Let $H_2 = \max\{2H_1, \frac{\widehat{H}_2}{\gamma}\}$ and $\Omega_2 = \{x \in C[0,1]\colon\ \|x\| < H_2\}$, then $x \in K$ and $\|x\| = H_2$ implies

$$\min_{\eta\leqslant t\leqslant 1} x(t) \geqslant \gamma\|x\| \geqslant \widehat{H}_2,$$

and so

$$\begin{aligned}
Ax(\eta) &= \lambda\frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)f\bigl(x(s)\bigr)\,\mathrm{d}s\\
&\geqslant \lambda\frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)(L-\epsilon)x(s)\,\mathrm{d}s\\
&\geqslant \lambda\frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)(L-\epsilon)\gamma\|x\|\,\mathrm{d}s\\
&= \lambda\frac{\gamma\eta}{1-\alpha\eta}\int_\eta^1 (1-s)a(s)(L-\epsilon)H_2\,\mathrm{d}s\\
&\geqslant \lambda B\gamma(L-\epsilon)\|x\|\\
&= \|x\|.
\end{aligned}$$

Hence, for $x \in K \cap \partial\Omega_2$,

$$\|Ax\| \geqslant \|x\|.$$

Therefore, by the first part of Theorem 3.33, it follows that A has a fixed point in $K \cap (\overline{\Omega}_2 \setminus \Omega_1)$, completing the proof in case (i).

Consider now the case (ii). Let λ given as in (ii) and choose $\epsilon > 0$ such that

$$\frac{1}{\gamma B(l-\epsilon)} \leqslant \lambda \leqslant \frac{1}{A(L+\epsilon)}.$$

We omit the rest of the proof, since it is similar to that of Theorem 3.38. □

Theorem 3.44 is taken from Raffoul [89]. He proved also that:

- if (A1), (A2), (A4) hold and $\lim_{x\to 0+}\frac{f(x)}{x} = \infty$ then the BVP (3.89)–(3.90) has at least one positive solution for each λ satisfying $0 < \lambda < \frac{1}{AL}$;
- if (A1), (A2), (A3) hold and $\lim_{x\to\infty}\frac{f(x)}{x} = \infty$ then the BVP (3.89)–(3.90) has at least one positive solution for each λ satisfying $0 < \lambda < \frac{1}{Al}$.

Other existence and multiplicities of positive solutions for the BVP (3.89)–(3.90) were proved by Ma in [70] by using fixed point index theory and the method of upper and lower solutions. He proved in [70] the following result.

THEOREM 3.45. *Assume that*:

(A1) λ *is a positive parameter*; $\eta \in (0,1)$ *and* $\alpha\eta < 1$.

(A2) $a : [0,1] \to [0,\infty)$ *is continuous and does not vanish identically on any subset of positive measure*.

(A3) $\lim_{x\to\infty}\frac{f(x)}{x} = \infty$.

Then there exists a positive λ^* *such that the BVP* (3.89)–(3.90) *has at least two positive solutions for* $0 < \lambda < \lambda^*$, *at least one positive solution for* $\lambda = \lambda^*$, *and no positive solutions for* $\lambda > \lambda^*$.

Interesting results on the existence of multiple positive solutions for three-point BVP can be obtained by using the Leggett–Williams fixed point theorem. We will give such a result proving the existence of triple positive solutions of the following BVP

$$x'' + f(t, x) = 0, \quad t \in (0, 1), \tag{3.91}$$

$$x(0) = 0, \qquad x(1) = \alpha x(\eta). \tag{3.92}$$

In this approach, we do not need the assumption that f is sublinear or superlinear, which was required in Theorem 3.38.

For convenience of the reader, we present here the necessary definitions from cone theory in Banach spaces.

DEFINITION 3.46. Let E be a Banach space over $\mathbb{R}$. A nonempty, closed set $P \subset E$ is said to be a cone provided that

(a) $\alpha u + \beta v \in P$ for all $u, v \in P$ and all $\alpha, \beta \geqslant 0$ and
(b) $u, -u \in P$ implies $u = 0$.

If $P \subset E$ is a cone, we denote the order induced by P on E by $\leqslant$. For $u, v \in P$, we write $u \leqslant v$ if and only if $v - u \in P$.

DEFINITION 3.47. The map ψ is said to be a nonnegative continuous concave functional on P provided $\psi : P \to [0, \infty)$ is continuous and

$$\psi\big(\lambda x + (1 - \lambda)y\big) \geqslant \lambda\psi(x) + (1 - \lambda)\psi(y)$$

for all $x, y \in P$ and $\lambda \in [0, 1]$.

Let $0 < a < b$ be given and let ψ be a nonnegative continuous concave functional on the cone P. Define the convex sets P_r and $P(\psi, a, b)$ by

$$P_r = \big\{y \in P\colon \ \|y\| < r\big\}$$

and

$$P(\psi, a, b) = \big\{y \in P\colon \ a \leqslant \psi(y), \ \text{and } \|y\| \leqslant b\big\}.$$

Our consideration is based on the following fixed point theorem given by Leggett and Williams in 1979 [55] (see also Guo and Lakshmikantham [21]).

THEOREM 3.48 (Legget–Williams fixed-point theorem). *Let* $T : \overline{P}_c \to \overline{P}_c$ *be a completely continuous operator and let* ψ *be a nonnegative continuous concave functional* *on P such that $\psi(y) \leqslant \|y\|$ for $y \in \overline{P}_c$. Suppose that there exist $0 < a < b < d \leqslant c$ such that*

(C1) $\{y \in P(\psi, b, d)\colon \psi(y) > b\} \neq \emptyset$ *and* $\psi(Ty) > b$ *for all* $y \in P(\psi, b, d)$;

(C2) $\|Ty\| < a$ *for all* $\|y\| \leqslant a$, *and*

(C3) $\psi(Ty) > b$ *for* $y \in P(\psi, b, c)$ *with* $\|Ty\| > d$.

Then T has at least three fixed points y_1, y_2, and y_3 such that $\|y_1\| < a$, $b < \psi(y_2)$ and $\|y_3\| > a$ with $\psi(y_3) < b$.

THEOREM 3.49. *Assume that $\eta \in (0,1)$, $\alpha > 0$ and $\alpha\eta < 1$. Assume that*:

(A1) $f:[0,1] \times [0,\infty) \to [0,\infty)$ *is continuous and* $f(t,\cdot)$ *does not vanish identically on any subset of* $[0,1]$ *with positive measure.*

Suppose also that there exist constants $0 < a < b < b/\gamma \leqslant c$ such that

(D1) $f(t,x) < ma$, *for* $0 \leqslant t \leqslant 1$, $0 \leqslant x \leqslant a$,

(D2) $f(t,x) \geqslant b/\delta$, *for* $\eta \leqslant t \leqslant 1$, $b \leqslant x \leqslant b/\gamma$,

(D3) $f(t,x) \leqslant mc$, *for* $0 \leqslant t \leqslant 1$, $0 \leqslant x \leqslant c$,

where

$$m = \left(\frac{2 - \alpha\eta + \alpha\eta^2}{2(1-\alpha\eta)}\right)^{-1}, \qquad \delta = \min\left\{\frac{\alpha\eta(1-\eta)^2}{2(1-\alpha\eta)}, \frac{\eta(1-\eta)^2}{2(1-\alpha\eta)}\right\},$$

and γ is defined in Lemma 3.37. *Then the BVP* (3.91)–(3.92) *has at least three positive solutions x_1, x_2 and x_3 satisfying $\|x_1\| < a, b < \psi(x_2)$ and $\|x_3\| > a$ with $\psi(x_3) < b$.*

PROOF. Let $E = C[0,1]$ be endowed with the maximum norm, $\|x\| = \max_{0 \leqslant t \leqslant 1} |x(t)|$, and the ordering $x \leqslant y$ if $x(t) \leqslant y(t)$ for all $t \in [0,1]$. From the fact $x''(t) = -f(t,x) \leqslant 0$, we know that $x(t)$ is concave on $[0,1]$. So, define the cone $P \subset E$ by

$$P = \bigl\{x \in E\colon x \text{ is concave and nonnegative valued on } [0,1]\bigr\}.$$

Finally, let the nonnegative concave functional $\psi : P \to [0,\infty)$ be defined by

$$\psi(x) = \min_{\eta \leqslant t \leqslant 1} x(t), \quad x \in P.$$

We notice that for each $x \in P$, $\psi(x) \leqslant \|x\|$. Define an operator $T : P \to E$ by

$$Tx(t) = -\int_0^t (t-s) f\bigl(s, x(s)\bigr)\,\mathrm{d}s - \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta - s) f\bigl(s, x(s)\bigr)\,\mathrm{d}s + \frac{t}{1-\alpha\eta}\int_0^1 (1-s) f\bigl(s, x(s)\bigr)\,\mathrm{d}s.$$

We note that, if $x \in P$ then $Tx(t) \geqslant 0$, $0 \leqslant t \leqslant 1$. In view of $(Tx)''(t) = -f(t, x(t)) \leqslant 0$, $0 \leqslant t \leqslant 1$, we see that $Tx \in P$; that is to say that $T : P \to P$. Also T is completely continuous.

We now show that all the conditions of Theorem 3.48 are satisfied. Now if $x \in \overline{P}_c$, then $\|x\| \leqslant c$ and (D3) implies $f(t,x(t)) \leqslant mc$, $0 \leqslant t \leqslant 1$. Consequently,

$$\begin{aligned}
\|Tx\| &= \max_{0\leqslant t\leqslant 1} -\int_0^t (t-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s - \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s \\
&\quad + \frac{t}{1-\alpha\eta}\int_0^1 (1-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s \\
&\leqslant \max_{0\leqslant t\leqslant 1}\int_0^t (t-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s + \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s \\
&\quad + \frac{t}{1-\alpha\eta}\int_0^1 (1-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s \\
&\leqslant \max_{0\leqslant t\leqslant 1}\biggl(\int_0^t (t-s)\,\mathrm{d}s + \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta-s)\,\mathrm{d}s \\
&\quad + \frac{t}{1-\alpha\eta}\int_0^1 (1-s)\,\mathrm{d}s\biggr)\cdot mc \\
&= \frac{2-\alpha\eta+\alpha\eta^2}{2(1-\alpha\eta)}\cdot mc \\
&= c.
\end{aligned}$$

Hence, $T : \overline{P}_c \to \overline{P}_c$. In the same way, if $x \in \overline{P}_c$, then assumption (D1) yields $f(t,x(t)) < ma$, $0 \leqslant t \leqslant 1$. As in the argument above, we can obtain that $T : \overline{P}_a \to P_a$. Therefore condition (C2) of Theorem 3.48 is satisfied.

To check condition (C1) of Theorem 3.48, we choose $x(t) = b/\gamma$, $0 \leqslant t \leqslant 1$. It is easy to see that $x(t) = b/\gamma \in P(\psi, b, b/\gamma)$ and $\psi(x) = \psi(b/\gamma) > b$, and so $\{x \in P(\psi, b, b/\gamma)\colon\ \psi(x) > b\} \neq \emptyset$. Hence, if $x \in P(\psi, b, b/\gamma)$, then $b \leqslant x(t) \leqslant b/\gamma$, $\eta \leqslant t \leqslant 1$. From assumption (D2), we have $f(t,x(t)) \geqslant b/\delta$, $\eta \leqslant t \leqslant 1$, and by the definition of ψ and the cone P, we have to distinguish two cases, (i) $\psi(Tx) = Tx(\eta)$ and (ii) $\psi(Tx) = Tx(1)$.

In case (i) we have

$$\begin{aligned}
\psi(Tx) &= Tx(\eta) \\
&= -\int_0^\eta (\eta-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s - \frac{\alpha\eta}{1-\alpha\eta}\int_0^\eta (\eta-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s \\
&\quad + \frac{\eta}{1-\alpha\eta}\int_0^1 (1-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s \\
&\leqslant -\frac{1}{1-\alpha\eta}\int_0^\eta (\eta-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s + \frac{\eta}{1-\alpha\eta}\int_0^1 (1-s) f\bigl(s,x(s)\bigr)\,\mathrm{d}s \\
&= -\frac{\eta}{1-\alpha\eta}\int_0^\eta f\bigl(s,x(s)\bigr)\,\mathrm{d}s + \frac{1}{1-\alpha\eta}\int_0^\eta s f\bigl(s,x(s)\bigr)\,\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&\quad + \frac{\eta}{1-\alpha\eta}\int_0^1 f\big(s,x(s)\big)\,\mathrm{d}s - \frac{\eta}{1-\alpha\eta}\int_0^1 sf\big(s,x(s)\big)\,\mathrm{d}s \\
&= \frac{\eta}{1-\alpha\eta}\int_\eta^1 f\big(s,x(s)\big)\,\mathrm{d}s + \frac{1}{1-\alpha\eta}\int_0^\eta sf\big(s,x(s)\big)\,\mathrm{d}s \\
&\quad - \frac{\eta}{1-\alpha\eta}\int_0^1 sf\big(s,x(s)\big)\,\mathrm{d}s \\
&> \frac{\eta}{1-\alpha\eta}\int_\eta^1 f\big(s,x(s)\big)\,\mathrm{d}s - \frac{\eta}{1-\alpha\eta}\int_\eta^1 sf\big(s,x(s)\big)\,\mathrm{d}s \\
&\geqslant \frac{\eta}{1-\alpha\eta}\int_\eta^1 (1-s)\,\mathrm{d}s\cdot\frac{b}{\delta} \\
&= \frac{\eta(1-\eta)^2}{2(1-\alpha\eta)}\cdot\frac{b}{\delta} \\
&\geqslant b.
\end{aligned}
$$

In case (ii), we have

$$
\begin{aligned}
\psi(Tx) &= Tx(1) \\
&= -\int_0^1 (1-s)f\big(s,x(s)\big)\,\mathrm{d}s - \frac{\alpha}{1-\alpha\eta}\int_0^\eta (\eta-s)f\big(s,x(s)\big)\,\mathrm{d}s \\
&\quad + \frac{1}{1-\alpha\eta}\int_0^1 (1-s)f\big(s,x(s)\big)\,\mathrm{d}s \\
&= \frac{\alpha\eta}{1-\alpha\eta}\int_0^1 (1-s)f\big(s,x(s)\big)\,\mathrm{d}s - \frac{\alpha}{1-\alpha\eta}\int_0^\eta (\eta-s)f\big(s,x(s)\big)\,\mathrm{d}s \\
&= \frac{\alpha\eta}{1-\alpha\eta}\int_0^1 f\big(s,x(s)\big)\,\mathrm{d}s - \frac{\alpha\eta}{1-\alpha\eta}\int_0^1 sf\big(s,x(s)\big)\,\mathrm{d}s \\
&\quad - \frac{\alpha\eta}{1-\alpha\eta}\int_0^\eta f\big(s,x(s)\big)\,\mathrm{d}s + \frac{\alpha}{1-\alpha\eta}\int_0^\eta sf\big(s,x(s)\big)\,\mathrm{d}s \\
&= \frac{\alpha\eta}{1-\alpha\eta}\int_\eta^1 f\big(s,x(s)\big)\,\mathrm{d}s - \frac{\alpha\eta}{1-\alpha\eta}\int_0^1 sf\big(s,x(s)\big)\,\mathrm{d}s \\
&\quad + \frac{\alpha}{1-\alpha\eta}\int_0^\eta sf\big(s,x(s)\big)\,\mathrm{d}s \\
&> \frac{\alpha\eta}{1-\alpha\eta}\int_\eta^1 f\big(s,x(s)\big)\,\mathrm{d}s - \frac{\alpha\eta}{1-\alpha\eta}\int_\eta^1 sf\big(s,x(s)\big)\,\mathrm{d}s \\
&= \frac{\alpha\eta}{1-\alpha\eta}\int_\eta^1 (1-s)f\big(s,x(s)\big)\,\mathrm{d}s
\end{aligned}
$$

$$\geqslant \frac{\alpha\eta(1-\eta)^2}{2(1-\alpha\eta)} \cdot \frac{b}{\delta}$$
$$\geqslant \delta \cdot \frac{b}{\delta} = b,$$

i.e.

$$\psi(Tx) > b, \quad \forall x \in P(\psi, b, b/\gamma).$$

This shows that condition (C1) of Theorem 3.48 is satisfied. We finally show that (C3) of Theorem 3.48 also holds. Suppose that $x \in P(\psi, b, c)$ with $\|Tx\| > b/\gamma$. Then we have

$$\psi(Tx) = \min_{\eta \leqslant t \leqslant 1} Tx(t) \geqslant \gamma \cdot \|Tx\| > \gamma \cdot b/\gamma = b.$$

So, condition (C3) of Theorem 3.48 is satisfied. Therefore an application of Theorem 3.48 completes the proof. □

Theorem 3.49 was taken from [42]. For the BVP (3.91)–(3.92) Ma in [76] studied multiplicity results at resonance, i.e. when $\alpha\eta = 1$, by developing the methods of lower and upper solutions when the lower and upper solutions are well ordered as well as opposite ordered, by the connectivity properties of the solution set of parameterized families of compact vector fields.

Let us discuss some extensions of the above results.

Consider the following m-point BVP

$$x'' + a(t)f(x) = 0, \quad t \in (0,1), \tag{3.93}$$

$$x(0) = 0, \qquad x(1) = \sum_{i=1}^{m-2} \alpha_i x(\xi_i) \tag{3.94}$$

where $\alpha_i \geqslant 0$ for $i = 1, 2, \ldots, m-3$ and $\alpha_{m-2} > 0$, $0 < \xi_1 < \xi_2 < \cdots < \xi_{m-2} < 1$, $\sum_{i=1}^{m-2} \alpha_i \xi_i < 1$. By using the Krasnosel'skii's fixed point theorem in cones Ma in [72] proved the existence of positive solution of the BVP (3.93)–(3.94) under the assumptions that $a \in C([0,1],[0,\infty))$, and there exists $x_0 \in [\xi_{m-2}, 1]$ such that $a(x_0) > 0$ and $\sum_{i=1}^{m-2} \alpha_i \xi_i < 1$, either in superlinear or sublinear case. The steps of the proof are parallel to that of the proof of Theorem 3.38.

A natural generalization of the boundary condition (3.94) is

$$x(0) = 0, \qquad \int_a^b h(t)x(t)\,\mathrm{d}t = x(1), \tag{3.95}$$

where $[a,b] \subset (0,1)$, $h \in C([a,b],[0,\infty))$, $\int_a^b t h(t)\,\mathrm{d}t \neq 1$ and $b\int_a^b h(t)\,\mathrm{d}t < 1$. The existence of positive solutions for the BVP (3.80), (3.95) is studied by Ma [73] when f is either superlinear or sublinear.

Another result for the following m-point BVP

$$x'' + a(t)f(x) = 0, \quad t \in (0,1), \tag{3.96}$$

$$x'(0) = \sum_{i=1}^{m-2} \beta_i x'(\xi_i), \qquad x(1) = \sum_{i=1}^{m-2} \alpha_i x(\xi_i) \tag{3.97}$$

where $\alpha_i, \beta_i \geqslant 0$ and $\sum_{i=1}^{m-2} \alpha_i < 1$, $\sum_{i=1}^{m-2} \beta_i < 1$, was given by Ma and Castaneda in [71] also by using the Krasnosel'skii's fixed point theorem in cones and either in superlinear or sublinear case. These results were generalized and improved by Liu [58] by applying topological degree methods.

The result for the BVP (3.80), (3.88) was generalized by Guo et al. in [23], where they studied the following BVP (3.80), (3.98) where (3.98) stands for the boundary condition

$$x(0) = 0, \qquad x(1) - \sum_{i=1}^{m-2} \alpha_i x(\xi_i) = b, \tag{3.98}$$

where $b, \alpha_i > 0$ $(i = 1, 2, \dots, m-2)$. In this problem $a(t)$ is allowed to be singular at $t = 0, 1$.

The m-point BVP

$$x'' + f(t,x) = 0, \quad t \in (0,1), \tag{3.99}$$

$$x(0) = 0, \qquad x(1) = \sum_{i=1}^{m-2} \alpha_i x(\xi_i) \tag{3.100}$$

where $\alpha_i > 0$ for $i = 1, 2, \dots, m-2$, $0 < \xi_1 < \xi_2 < \cdots < \xi_{m-2} < 1$, $\sum_{i=1}^{m-2} \alpha_i \xi_i < 1$ and $f : [0,1] \times [0,\infty) \to [0,\infty)$ a continuous function, was studied by Liu et al. in [64] by using the Legget–Williams fixed point theorem and Green's functions.

Existence and multiplicity results for positive solutions for the m-point BVP

$$\begin{aligned} &\big(p(t)x'\big)' - q(t)x + f(t,x) = 0, \quad 0 < t < 1, \\ &ax(0) - bp(0)x'(0) = \sum_{i=1}^{m-2} \alpha_i x(\xi_i), \\ &cx(1) - dp(1)x'(1) = \sum_{i=1}^{m-2} \beta_i x(\xi_i), \end{aligned}$$

where $p, q \in C([0,1], (0,\infty))$, $a, b, c, d \in [0,\infty)$ are given by Ma in [77]. See also [75] for results for superlinear semipositone m-point BVPs.

Three-point BVP of higher-order ordinary differential equations was studied by Liu and Ge in [65] where they consider BVP consisting of the equation

$$x^{(n)} + \lambda a(t) f\bigl(x(t)\bigr) = 0, \quad t \in (0,1)$$

with one of the following boundary value conditions:

$$\begin{aligned} &x(0) = \alpha x(\eta), \qquad x(1) = \beta x(\eta),\\ &x^{(i)}(0) = 0, \quad i = 1,2,\dots,n-2, \end{aligned}$$

and

$$\begin{aligned} &x^{(n-2)}(0) = \alpha x^{(n-2)}(\eta), \qquad x^{(n-2)}(1) = \beta x^{(n-2)}(\eta),\\ &x^{(i)}(0) = 0, \quad i = 1,2,\dots,n-3, \end{aligned}$$

where $\eta \in (0,1)$, $\alpha \geqslant 0$, $\beta \geqslant 0$, and $a:(0,1) \to \mathbb{R}$ may change sign, $f(0) > 0$, and $\lambda > 0$ is a parameter.

For other recent results on positive solutions for nonlocal BVP the interested reader is referred to [13,18,46,48,51,64,67,68] and the references cited therein.

3.10. *Positive solutions of nonlocal boundary value problems with dependence on the first-order derivative*

All the results in the previous section were proved under the assumption that the first derivative x' is not involved explicitly in the nonlinear term. In this section we are concerned with the existence of positive solutions for the second-order three-point BVP

$$x''(t) + f\bigl(t, x(t), x'(t)\bigr), \quad 0 < t < 1, \tag{3.101}$$

$$x(0) = 0, \qquad x(1) = \alpha x(\eta), \tag{3.102}$$

where $f:[0,1] \times [0,\infty) \times \mathbb{R} \to [0,\infty)$ is continuous, $\alpha > 0$, $0 < \eta < 1$ and $1 - \alpha\eta > 0$.

To show the existence of positive solutions to BVP (3.101)–(3.102), we use an extension of Krasnosel'skii's fixed point theorem in cones, proved in [19]. To state this fixed point theorem some notations are necessary.

Let X be a Banach space and $K \subset X$ a cone. Suppose $\alpha, \beta : X \to \mathbb{R}^+$ are two convex functionals satisfying

$$\alpha(\lambda x) = |\lambda| \alpha(x), \quad \beta(\lambda x) = |\lambda| \beta(x), \quad \lambda \in \mathbb{R},$$

and

$$\begin{aligned} &\|x\| \leqslant M \max\bigl\{\alpha(x), \beta(x)\bigr\} && \text{for } x \in X \quad \text{and}\\ &\alpha(x) \leqslant \alpha(y) && \text{for } x, y \in K,\ x \leqslant y, \end{aligned}$$

where $M > 0$ is a constant.

THEOREM 3.50. *Let $r_2 > r_1 > 0$, $L > 0$ be constants and*

$$\Omega_i = \{x \in X\colon\ \alpha(x) < r_i,\ \beta(x) < L\}, \quad i = 1, 2,$$

two bounded open sets in X. Set

$$D_i = \{x \in X\colon\ \alpha(x) = r_i\}.$$

Assume $T : K \to K$ is a completely continuous operator satisfying

(a) $\alpha(Tu) < r_1$, $u \in D_1 \cap K$; $\alpha(Tu) > r_2$, $u \in D_2 \cap K$;
(b) $\beta(Tu) < L$, $u \in K$;
(c) *there is a $p \in (\Omega \cap K) \setminus \{0\}$ such that $\alpha(p) \neq 0$ and $\alpha(x + \lambda p) \geqslant \alpha(x)$ for all $x \in K$ and $\lambda \geqslant 0$.*

Then T has at least one fixed point in $(\Omega_2 \setminus \overline{\Omega}_1) \cap K$.

From Lemmas 3.34 and 3.37 we know that the unique solution of the BVP

$$x'' + y(t) = 0, \quad 0 < t < 1, \qquad x(0) = 0, \qquad x(1) = \alpha x(\eta)$$

is given by

$$\begin{aligned} x(t) = &-\int_0^t (t-s)y(s)\,\mathrm{d}s - \frac{\alpha t}{1-\alpha\eta}\int_0^\eta (\eta - s)y(s)\,\mathrm{d}s \\ &+ \frac{t}{1-\alpha\eta}\int_0^1 (1-s)y(s)\,\mathrm{d}s \end{aligned}$$

and satisfies

$$\min_{t\in[\eta,1]} x(t) \geqslant \gamma\|x\|,$$

where $\gamma = \min\{\alpha\eta, \frac{\alpha(1-\eta)}{1-\alpha\eta}, \eta\}$.

Moreover it is easy to see that, if $1 - \alpha\eta \neq 0$, the Green function for the BVP

$$-x'' = 0, \quad 0 < t < 1, \qquad x(0) = 0, \qquad x(1) = \alpha x(\eta)$$

is given by

$$G(t,s)=\begin{cases}\dfrac{s[(1-t)-\alpha(\eta-t)]}{1-\alpha\eta}, & s\leqslant t,\ s\leqslant\eta,\\ \dfrac{s(1-t)+\alpha\eta(t-s)}{1-\alpha\eta}, & \eta\leqslant s\leqslant t,\\ \dfrac{t[(1-s)-\alpha(\eta-s)]}{1-\alpha\eta}, & t\leqslant s\leqslant\eta,\\ \dfrac{t(1-s)}{1-\alpha\eta}, & t\leqslant s,\ s\geqslant\eta.\end{cases}$$

Let $X=C^1([0,1],\mathbb{R})$ with $\|x\|=\max_{0\leqslant t\leqslant 1}[x^2(t)+(x'(t))^2]^{1/2}$, and $K=\{x\in X\colon\ x(t)\geqslant 0,\ \ x$ is concave on $[0,1]\}$. Define functionals $\alpha(x)=\max_{0\leqslant t\leqslant 1}|x(t)|$ and $\beta(x)=\max_{0\leqslant t\leqslant 1}|x'(t)|$ for each $x\in X$. Then $\|x\|\leqslant\sqrt{2}\max\{\alpha(x),\beta(x)\}$ and

$$\alpha(\lambda x)=|\lambda|\alpha(x),\quad \beta(\lambda x)=|\lambda|\beta(x),\quad x\in X,\ \lambda\in\mathbb{R},$$
$$\alpha(x)\leqslant\alpha(y),\quad \text{for } x,y\in K,\ x\leqslant y.$$

We set

$$M=\max_{0\leqslant t\leqslant 1}\int_0^1 G(t,s)\,\mathrm{d}s,\qquad m=\max_{0\leqslant t\leqslant 1}\int_\eta^1 G(t,s)\,\mathrm{d}s,$$
$$Q=\frac{3-\alpha\eta+\alpha\eta^2}{2(1-\alpha\eta)}.$$

THEOREM 3.51. *Let* $f:[0,1]\times[0,\infty)\times\mathbb{R}\to[0,\infty)$ *be a continuous function,* $\alpha>0$, $0<\eta<1$ *and* $1-\alpha\eta>0$. *Suppose that there are* $L>b>\gamma b>c>0$ *such that* $f(t,u,v)$ *satisfies the growth conditions*:

(1) $f(t,u,v)<c/M$ *for* $(t,u,v)\in[0,1]\times[0,c]\times[-L,L]$;
(2) $f(t,u,v)\geqslant b/m$ *for* $(t,u,v)\in[0,1]\times[\gamma b,b]\times[-L,L]$;
(3) $f(t,u,v)<L/Q$ *for* $(t,u,v)\in[0,1]\times[0,b]\times[-L,L]$.

Then the BVP (3.101)–(3.102) *has at least one positive solution* $y(t)$ *satisfying*

$$c<\alpha(x)<b,\qquad \left|y'(t)\right|<L.$$

PROOF. Take

$$\Omega_1=\left\{x\in X\colon\ \left|x(t)\right|<c,\ \left|x'(t)\right|<L\right\},$$
$$\Omega_2=\left\{x\in X\colon\ \left|x(t)\right|<b,\ \left|x'(t)\right|<L\right\}$$

two bounded open sets in X, and

$$D_1=\left\{x\in X\colon\ \alpha(x)=c\right\},\qquad D_2=\left\{x\in X\colon\ \alpha(x)=b\right\}.$$

Let

$$f^*(t,u,v)=\begin{cases} f(t,u,v), & (t,u,v)\in[0,1]\times[0,b]\times(-\infty,\infty),\\ f(t,b,v), & (t,u,v)\in[0,1]\times(b,\infty)\times(-\infty,\infty),\end{cases}$$

and

$$f_1(t,u,v)=\begin{cases} f^*(t,u,v), & (t,u,v)\in[0,1]\times[0,\infty)\times[-L,L],\\ f^*(t,u,-L), & (t,u,v)\in[0,1]\times[0,\infty)\times(-\infty,-L],\\ f^*(t,u,L), & (t,u,v)\in[0,1]\times[0,\infty)\times[L,\infty).\end{cases}$$

Then $f_1\in C([0,1]\times[0,\infty)\times\mathbb{R},\mathbb{R}^+)$. Define

$$(Tx)(t)=\int_0^1 G(t,s)f_1\big(s,x(s),x'(s)\big)\,\mathrm{d}s.$$

Obviously, $T:K\to K$ is completely continuous, and there is a $p\in(\Omega_2\cap K)\setminus\{0\}$ such that $\alpha(x+\lambda p)\geqslant\alpha(x)$ for all $x\in K$ and $\lambda\geqslant 0$. For $x\in D_1\cap K$, $\alpha(x)=c$. From (1), we get

$$\begin{aligned}\alpha(Tx)&=\max_{t\in[0,1]}\left|\int_0^1 G(t,s)f_1\big(s,x(s),x'(s)\big)\,\mathrm{d}s\right|\\ &<\max_{t\in[0,1]}\int_0^1 G(t,s)\frac{c}{M}\,\mathrm{d}s=\frac{c}{M}\max_{t\in[0,1]}\int_0^1 G(t,s)\,\mathrm{d}s=c.\end{aligned}$$

Whereas for $x\in D_2\cap K$, $\alpha(x)=b$. We have $x(t)\geqslant\gamma\alpha(x)=\gamma b$ for $t\in[\eta,1]$. So, from (2), we get

$$\begin{aligned}\alpha(Tx)&=\max_{t\in[0,1]}\left|\int_0^1 G(t,s)f_1\big(s,x(s),x'(s)\big)\,\mathrm{d}s\right|\\ &>\max_{t\in[0,1]}\left|\int_\eta^1 G(t,s)f_1\big(s,x(s),x'(s)\big)\,\mathrm{d}s\right|\\ &>\max_{t\in[0,1]}\int_\eta^1 G(t,s)\frac{b}{m}\,\mathrm{d}s=\frac{b}{m}\max_{t\in[0,1]}\int_\eta^1 G(t,s)\,\mathrm{d}s=b.\end{aligned}$$

For $x\in K$, from (3), we get

$$\begin{aligned}\beta(Tx)=\max_{t\in[0,1]}\bigg|&-\int_0^t f_1\big(s,x(s),x'(s)\big)\,\mathrm{d}s\\ &+\frac{1}{1-\alpha\eta}\int_0^1(1-s)f_1\big(s,x(s),x'(s)\big)\,\mathrm{d}s\\ &-\frac{\alpha}{1-\alpha\eta}\int_0^\eta(\eta-s)f_1\big(s,x(s),x'(s)\big)\,\mathrm{d}s\bigg|\end{aligned}$$

$$< \left[1 + \frac{1}{1-\alpha\eta}\int_0^1 (1-s)\,\mathrm{d}s + \frac{\alpha}{1-\alpha\eta}\int_0^\eta (\eta - s)\,\mathrm{d}s\right]\frac{L}{Q}$$
$$= \frac{3-\alpha\eta+\alpha\eta^2}{2(1-\alpha\eta)}\frac{L}{Q} = L.$$

Theorem 3.50 implies there is $y \in (\Omega_2 \setminus \overline{\Omega}_1) \cap K$ such that $y = Ty$. So, y is a positive solution for BVP (3.101)–(3.102) satisfying

$$c < \alpha(x) < b, \qquad \left|y'(t)\right| < L.$$

Thus, the proof of the theorem is complete. □

Theorem 3.51 was proved in [19], where one more result was proved in the case when $0 < \alpha \leqslant 1$ and $\eta \in (0,1)$, under the assumptions (1), (2) and the following one

(4) $f(t,u,v) < L^2/(2b)$ for $(t,u,v) \in [0,1] \times [0,b] \times [-L,L]$.

Consider now a three-point BVP

$$\big(p(t)x'(t)\big)' = f\big(t, x(t), x'(t)\big), \quad t \in [0,1], \tag{3.103}$$

$$x'(0) = 0, \qquad x(1) = x(\eta) \tag{3.104}$$

where $f : [0,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous, $p : [0,1] \to \mathbb{R}$ is a positive continuous differentiable function, $\min_{t\in[0,1]} p(t) := p_1 > 0$ and $\eta \in (0,1)$. For this problem the linear operator $Lx(t) = (p(t)x'(t))'$ is not invertible, so the three-point BVP (3.103)–(3.104) is a resonance problem. By using a theorem of a fixed point index for A-proper semilinear operators Bai and Fang [6] proved the following theorem.

THEOREM 3.52. *Suppose*

(1) $f : [0,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *is continuous and there exist constants* $a > 0$, $b > 0$, *and* $0 < c < p_1^2/(2\|p\|_\infty)$ *such that*

$$\left|f(t,x,y)\right| \leqslant a + b|x| + c|y|, \quad \forall t \in [0,1],\ x, y \in \mathbb{R}.$$

(2) *There exists* $M > 0$ *such that*

$$f(t,x,0) > 0, \quad \forall t \in [0,1],\ x > M.$$

(3) $f(t,x,y) \geqslant 0$, $\forall t \in [0,1]$, $y \in \mathbb{R}$ *and there exist constants* $0 < \sigma < b$ *and* $\varepsilon > 0$ *such that*

$$f(t,x,y) \geqslant \sigma x, \quad \forall t \in [0,1],\ 0 \leqslant x \leqslant \varepsilon,\ y \in \mathbb{R}.$$

Then the BVP (3.103)–(3.104) *has at least one positive solution.*

A more general nonlocal BVP was studied by Palamides [87]. He consider the following BVP

$$x''(t) = f\bigl(t, x(t), x'(t)\bigr), \quad t \in (0,1),$$
$$\alpha x(0) - \beta x'(0) = 0, \qquad x(1) = \sum_{i=0}^{m-2} \alpha_i x(\xi_i)$$

with $\alpha \geqslant 0$, $\beta > 0$. Existence of positive solution is given, under superlinear and/or sublinear growth rate in f. A different method is employed based on analysis of the corresponding vector field on the (x, x')-face plane and Kneser's property of the solutions funnel.

Finally positive solutions for $2n$th order nonlocal BVP was studied by Guo et al. in [22] and [20] where they consider BVPs consisting of the equation

$$x^{(2n)}(t) = f\bigl(t, x(t), x''(t), \dots, x^{2(n-1)}(t)\bigr), \quad 0 \leqslant t \leqslant 1,$$

and one of the following boundary conditions:

$$x^{(2i)}(0) = 0, \quad x^{(2i)}(1) = \sum_{j=1}^{m-2} k_{ij} x^{(2i)}(\xi_j), \quad 0 \leqslant i \leqslant n-1,$$

and

$$x^{(2i)}(0) - \beta_i x^{(2i+1)}(0) = 0, \quad x^{(2i)}(1) = \sum_{j=1}^{m-2} k_{ij} x^{(2i)}(\xi_j), \quad 0 \leqslant i \leqslant n-1.$$

They use the Legget–Williams fixed point theorem and a fixed point theorem in double cones respectively.

For other recent results we refer to [49,50,52,53].

3.11. *Positive solutions of nonlocal problems for p-Laplacian*

The purpose of this section is to establish the existence of positive solutions to the following three-point boundary value problem for p-Laplacian

$$\bigl(g(u')\bigr)' + a(t)f(u) = 0, \quad \text{for } 0 < t < 1, \tag{3.105}$$
$$u(0) = 0, \quad \text{and} \quad u(\nu) = u(1), \tag{3.106}$$

where $g(v) = |v|^{p-2}v$, with $p > 1$, and $\nu \in (0,1)$.

We will prove the existence of at least three positive pseudo-symmetric solutions of the BVP (3.105)–(3.106) where we now define what we mean by a pseudo-symmetric function.

DEFINITION 3.53. For $\nu \in (0, 1)$ a function $u \in C[0, 1]$ is said to be pseudo-symmetric if u is symmetric over the interval $[\nu, 1]$. That is, for $t \in [\nu, 1]$ we have $u(t) = u(1 - (t - \nu))$.

In this setting we are able to verify that for all x in our cone $\sigma_u = \frac{1+\nu}{2}$.

In our approach we will use a new multiple fixed point theorem, now called the five functionals fixed point theorem, obtained by Avery [2], which generalized the Leggett–Williams fixed point theorem [55] in terms of functionals rather than norms. We state first the five functional fixed point theorem.

Let γ, β, θ be nonnegative, continuous, convex functionals on P and α, ψ be nonnegative, continuous, concave functionals on P. Then, for nonnegative real numbers h, a, b, d and c, we define the convex sets,

$$\begin{aligned}
P(\gamma, c) &= \{x \in P\colon\ \gamma(x) < c\},\\
P(\gamma, \alpha, a, c) &= \{x \in P\colon\ a \leqslant \alpha(x),\ \gamma(x) \leqslant c\},\\
Q(\gamma, \beta, d, c) &= \{x \in P\colon\ \beta(x) \leqslant d,\ \gamma(x) \leqslant c\},\\
P(\gamma, \theta, \alpha, a, b, c) &= \{x \in P\colon\ a \leqslant \alpha(x),\ \theta(x) \leqslant b,\ \gamma(x) \leqslant c\},
\end{aligned}$$

and

$$Q(\gamma, \beta, \psi, h, d, c) = \{x \in P\colon\ h \leqslant \psi(x),\ \beta(x) \leqslant d,\ \gamma(x) \leqslant c\}.$$

THEOREM 3.54. *Let P be a cone in a real Banach space E. Suppose there exist positive numbers c and M, nonnegative, continuous, concave functionals α and ψ on P, and nonnegative, continuous, convex functionals γ, β, and θ on P, with*

$$\alpha(x) \leqslant \beta(x) \quad \textit{and} \quad \|x\| \leqslant M\gamma(x)$$

for all $x \in \overline{P(\gamma, c)}$. Suppose

$$A : \overline{P(\gamma, c)} \to \overline{P(\gamma, c)}$$

is completely continuous and there exist nonnegative numbers h, a, k, b, with $0 < a < b$ such that:

(i) *$\{x \in P(\gamma, \theta, \alpha, b, k, c)\colon\ \alpha(x) > b\} \neq \emptyset$ and $\alpha(Ax) > b$ for $x \in P(\gamma, \theta, \alpha, b, k, c)$;*
(ii) *$\{x \in Q(\gamma, \beta, \psi, h, a, c)\colon\ \beta(x) < a\} \neq \emptyset$ and $\beta(Ax) < a$ for $x \in Q(\gamma, \beta, \psi, h, a, c)$;*
(iii) *$\alpha(Ax) > b$ for $x \in P(\gamma, \alpha, b, c)$ with $\theta(Ax) > k$;*
(iv) *$\beta(Ax) < a$ for $x \in Q(\gamma, \beta, a, c)$ with $\psi(Ax) < h$.*

Then A has at least three fixed points $x_1, x_2, x_3 \in \overline{P(\gamma, c)}$ such that,

$$\beta(x_1) < a, \qquad b < \alpha(x_2),$$

and

$$a < \beta(x_3) \quad \textit{with}\ \alpha(x_3) < b.$$

Assume $f:[0,\infty)\to[0,\infty)$ is continuous, and let $a:[0,1]\to[0,\infty)$ be continuous. Let $\nu\in(0,1)$, and let $E=C[0,1]$ be the Banach space with the sup-norm, $\|u\|=\sup\{|u(x)|:\ 0\leqslant x\leqslant 1\}$, and define the cone $P\subset E$ by

$$P=\bigl\{u\in E:\ u(0)=0,\ u \text{ is concave, and } u \text{ is symmetric on } [\nu,1]\bigr\}.$$

Hereafter, suppose $\nu\in(0,1)$, with $\int_0^\nu a(r)\,\mathrm{d}r>0$, and we will choose a $\delta\in(0,\nu)$, such that $\int_\delta^\nu a(t)\,\mathrm{d}t>0$, and the constants M^* and m^*, defined by

$$M^*=\int_0^\nu G\biggl(\int_s^\sigma a(r)\,\mathrm{d}r\biggr)\,\mathrm{d}s,$$

and

$$m^*=\int_0^\delta G\biggl(\int_s^\sigma a(r)\,\mathrm{d}r\biggr)\,\mathrm{d}s,$$

satisfy the inequality $\nu<\frac{m^*}{M^*}$. Trivially, we have $\delta<\frac{m^*}{M^*}$. We let

$$m=\int_0^\delta G\biggl(\int_\delta^\nu a(r)\,\mathrm{d}r\biggr)\,\mathrm{d}s=\delta G\biggl(\int_\delta^\nu a(r)\,\mathrm{d}r\biggr)$$

and

$$M=\int_0^\sigma G\biggl(\int_s^\sigma a(r)\,\mathrm{d}r\biggr)\,\mathrm{d}s,$$

as well as,

$$h_1=\int_\delta^\nu G\biggl(\int_s^\sigma a(r)\,\mathrm{d}r\biggr)\,\mathrm{d}s,$$

$$h_2=\int_0^\delta G\biggl(\int_s^\delta a(r)\,\mathrm{d}r\biggr)\,\mathrm{d}s,$$

and

$$h_3=\int_0^\delta G\biggl(\int_\delta^\sigma a(r)\,\mathrm{d}r\biggr)\,\mathrm{d}s.$$

Define the nonnegative, continuous, concave functionals α, ψ, and the nonnegative, continuous, convex functionals β, θ, γ on the cone P by:

$$\gamma(x)=\theta(x):=\max_{t\in[0,1]}x(t)=x(\sigma),$$

$$\beta(x):=\max_{t\in[\delta,\nu]}x(t)=x(\nu),$$

$$\alpha(x)=\psi(x):=\min_{t\in[\delta,\nu]}x(t)=x(\nu).$$

In our main result, we will make use of the following lemma. The lemma is easily proved using the concavity and the pseudo-symmetry of all $u\in P$.

LEMMA 3.55. *Let $x\in P$. Then*
(D1) $u(\delta)\geqslant\delta u(1)=\delta u(\nu)$, *and*
(D2) $\sigma u(\nu)\geqslant\nu u(\sigma)=\nu\|u\|$.

We are now ready to apply the five functionals fixed point theorem to an operator A to give sufficient conditions for the existence of at least three positive pseudo-symmetric solutions to (3.105)–(3.106).

THEOREM 3.56. *Assume that $\nu\in(0,1)$, $a:[0,1]\to[0,\infty)$ is a pseudo-symmetric continuous function, $\delta\in(0,\nu)$ such that*

$$\int_\delta^\nu a(t)\,\mathrm{d}t>0\quad \textit{with } \nu<\frac{m^*}{M^*},$$

and $f:[0,\infty)\to[0,\infty)$ is continuous. Let $0<a<b<c$, with $ch_2<aM$, and suppose that f satisfies the following conditions:
(i) $f(x)>g(\frac{b}{m})$ *for all* $b\leqslant x\leqslant\frac{b}{\delta}$,
(ii) $f(x)<g(\frac{Ma-ch_2}{M(h_1+h_3)})$ *for all* $a\delta\leqslant x\leqslant\frac{a\sigma}{\nu}$, *and*
(iii) $f(x)\leqslant g(\frac{c}{M})$ *for all* $0\leqslant x\leqslant c$.
Then the three-point boundary value problem (3.105)–(3.106) *has at least three positive pseudo-symmetric solutions u_1,u_2,u_3 such that*

$$\max_{t\in[\delta,\nu]}u_1(t)<a<\max_{t\in[\delta,\nu]}u_2(t)\quad\textit{and}\quad\min_{t\in[\delta,\nu]}u_2(t)<b<\min_{t\in[\delta,\nu]}u_3(t).$$

PROOF. Define the completely continuous operator A on P by

$$Au(t)=w(t)=\begin{cases}\displaystyle\int_0^t G\left(\int_s^\sigma a(r)f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s, & 0\leqslant t\leqslant\sigma,\\ \displaystyle w(\nu)+\int_t^1 G\left(\int_\sigma^s a(r)f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s, & \sigma\leqslant t\leqslant 1,\end{cases}$$

where

$$\sigma=\frac{\nu+1}{2}.$$

We first note that for $u\in P$ we have $Au(t)\geqslant 0$, $Au(0)=0$, and applying the Fundamental Theorem of Calculus we have that Au is concave. Furthermore, for $t\in[\nu,1]$

$$Au(t)=Au\big(1-(t-\nu)\big).$$

Consequently, $Au \in P$, that is, $A : P \to P$. Moreover, since

$$\big(g\big((Au)'\big)\big)'(t) = -a(t)f\big(u(t)\big) \leqslant 0 \quad \text{for all } 0 < t < 1,$$

we have that all fixed points of A are solutions of (3.105)–(3.106). Thus we set out to verify that the operator A satisfies the five functionals fixed point theorem which will prove the existence of three fixed points of A which satisfy the conclusion of the theorem.

If $u \in \overline{P(\gamma, c)}$, then $\gamma(u) = \max_{t\in[0,1]} u(t) \leqslant c$. Thus

$$\begin{aligned}
\gamma(Au) &= Au(\sigma) \\
&= \int_0^\sigma G\bigg(\int_s^\sigma a(r)f\big(u(r)\big)\,\mathrm{d}r\bigg)\,\mathrm{d}s \\
&\leqslant \int_0^\sigma G\bigg(\int_s^\sigma a(r)g\bigg(\frac{c}{M}\bigg)\,\mathrm{d}r\bigg)\,\mathrm{d}s \\
&= \bigg(\frac{c}{M}\bigg)\int_0^\sigma G\bigg(\int_s^\sigma a(r)\,\mathrm{d}r\bigg)\,\mathrm{d}s \\
&= c.
\end{aligned}$$

Hence,

$$A : \overline{P(\gamma, c)} \to \overline{P(\gamma, c)}.$$

Next, let $N = \frac{m^* + \delta M^*}{2}$. Thus $m^* > N > \delta M^*$, and if we define

$$u_P(t) = \begin{cases} \bigg(\dfrac{b}{N}\bigg)\displaystyle\int_0^t G\bigg(\int_s^\sigma a(r)\,\mathrm{d}r\bigg)\,\mathrm{d}s, & 0 \leqslant t \leqslant \sigma, \\ u_P(\nu) + \bigg(\dfrac{b}{N}\bigg)\displaystyle\int_t^1 G\bigg(\int_\sigma^s a(r)\,\mathrm{d}r\bigg)\,\mathrm{d}s, & \sigma \leqslant t \leqslant 1, \end{cases}$$

and

$$u_Q(t) = \begin{cases} \bigg(\dfrac{a\delta}{N}\bigg)\displaystyle\int_0^t G\bigg(\int_s^\sigma a(r)\,\mathrm{d}r\bigg)\,\mathrm{d}s, & 0 \leqslant t \leqslant \sigma, \\ u_Q(\nu) + \bigg(\dfrac{a\delta}{N}\bigg)\displaystyle\int_t^1 G\bigg(\int_\sigma^s a(r)\,\mathrm{d}r\bigg)\,\mathrm{d}s, & \sigma \leqslant t \leqslant 1, \end{cases}$$

then, clearly $u_P, u_Q \in P$. Furthermore,

$$\alpha(u_P) = u_P(\delta) = \bigg(\frac{b}{N}\bigg)\int_0^\delta G\bigg(\int_s^\sigma a(r)\,\mathrm{d}r\bigg)\,\mathrm{d}s = \frac{bm^*}{N} > b,$$

and

$$\theta(u_P) = u_P(\nu) = \bigg(\frac{b}{N}\bigg)\int_0^\nu G\bigg(\int_s^\sigma a(r)\,\mathrm{d}r\bigg)\,\mathrm{d}s = \frac{bM^*}{N} < \frac{b}{\delta},$$

as well as,

$$\psi(u_Q) = u_Q(\delta) = \left(\frac{a\delta}{N}\right)\int_0^\delta G\left(\int_s^\sigma a(r)\,\mathrm{d}r\right)\mathrm{d}s = \frac{a\delta m^*}{N} > a\delta,$$

and

$$\beta(u_Q) = u_Q(\nu) = \left(\frac{a\delta}{N}\right)\int_0^\nu G\left(\int_s^\sigma a(r)\,\mathrm{d}r\right)\mathrm{d}s = \frac{aM^*\delta}{N} < a.$$

Therefore,

$$u_P \in \left\{u \in P\left(\gamma, \theta, \alpha, b, \frac{b}{\delta}, c\right) \colon \alpha(u) > b\right\},$$

and

$$u_Q \in \left\{u \in Q(\gamma, \beta, \psi, a\delta, a, c) \colon \beta(u) < a\right\},$$

hence, these sets are nonempty.

If $u \in P(\gamma, \theta, \alpha, b, \frac{b}{\delta}, c)$, then $b \leqslant u(t) \leqslant \frac{b}{\delta}$, for all $t \in [\delta, \nu]$, and thus by condition (i) of this theorem,

$$\begin{aligned}
\alpha(Au) &= Au(\delta) \\
&= \int_0^\delta G\left(\int_s^\sigma a(r) f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s \\
&\geqslant \int_0^\delta G\left(\int_\delta^\nu a(r) f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s \\
&> \int_0^\delta G\left(\int_\delta^\nu a(r) g\left(\frac{b}{m}\right)\mathrm{d}r\right)\mathrm{d}s \\
&= \left(\frac{b}{m}\right)\int_0^\delta G\left(\int_\delta^\nu a(r)\,\mathrm{d}r\right)\mathrm{d}s \\
&= b.
\end{aligned}$$

Hence, condition (i) of the five functionals fixed point theorem is satisfied.

If $u \in P(\gamma, \alpha, b, c)$ with $\theta(Au) > \frac{b}{\delta}$, then by Lemma 3.55(D1), we have

$$\alpha(Au) = Au(\delta) \geqslant \delta Au(1) = \delta Au(\nu) = \delta\theta(Au) > b.$$

Thus, condition (iii) of the five functionals fixed point theorem is satisfied.

If $u \in Q(\gamma, \beta, \psi, a\delta, a, c)$, then $a\delta \leqslant u(t) \leqslant a$, for all $t \in [\delta, \nu]$, and thus by Lemma 3.55(D2), $a\delta \leqslant u(t) \leqslant \frac{a\sigma}{\nu}$, for all $t \in [\delta, \sigma]$. Thus by condition (ii) of this theorem,

$$\begin{aligned}
\beta(Au) &= Au(\nu) \\
&= \int_0^\nu G\left(\int_s^\sigma a(r) f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s \\
&= \int_0^\delta G\left(\int_s^\sigma a(r) f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s + \int_\delta^\nu G\left(\int_s^\sigma a(r) f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s \\
&\leqslant \int_0^\delta G\left(\int_s^\delta a(r) f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s + \int_0^\delta G\left(\int_\delta^\sigma a(r) f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s \\
&\quad + \int_\delta^\nu G\left(\int_s^\sigma a(r) f\big(u(r)\big)\,\mathrm{d}r\right)\mathrm{d}s \\
&< \left(\frac{c}{M}\right)\int_0^\delta G\left(\int_s^\delta a(r)\,\mathrm{d}r\right)\mathrm{d}s \\
&\quad + \left(\frac{Ma - ch_2}{M(h_1 + h_3)}\right)\int_0^\delta G\left(\int_\delta^\sigma a(r)\,\mathrm{d}r\right)\mathrm{d}s \\
&\quad + \left(\frac{Ma - ch_2}{M(h_1 + h_3)}\right)\int_\delta^\nu G\left(\int_s^\sigma a(r)\,\mathrm{d}r\right)\mathrm{d}s \\
&= \frac{ch_2}{M} + \frac{(Ma - ch_2)h_3}{M(h_1 + h_3)} + \frac{(Ma - ch_2)h_1}{M(h_1 + h_3)} \\
&= a.
\end{aligned}$$

Hence, condition (ii) of the five functionals fixed point theorem is satisfied.

If $u \in Q(\gamma, \beta, a, c)$, with $\psi(Au) < a\delta$, then by Lemma 3.55, we have

$$\beta(Au) = Au(\nu) \leqslant \frac{Au(\delta)}{\delta} = \frac{\psi(Au)}{\delta} < a.$$

Consequently, condition (iv) of the five functionals fixed point theorem is also satisfied. Therefore, the hypotheses of the five functionals fixed point theorem 3.54 are satisfied, and there exist at least three positive pseudo-symmetric solutions $u_1, u_2, u_3 \in \overline{P(\gamma, c)}$ for the three-point boundary value problem (3.105)–(3.106) such that,

$$\beta(u_1) < a < \beta(u_2) \quad \text{and} \quad \alpha(u_2) < b < \alpha(u_3). \qquad \square$$

Theorem 3.56 was taken from [3]. Existence of positive solutions for the BVP (3.105)–(3.106) was proved also in [91] and [41], relying on Krasnosel'skii's and Legget–Williams fixed point theorems respectively. BVP with p-Laplacian and more general boundary conditions was studied in [66], where the boundary conditions $u(0) - B_0(u'(\eta)) = 0, u'(1) = 0$

and $u'(0)=0, u(1)+B_1(u'(\eta))=0$ are used and in [4,5] where the m-point boundary conditions $u(0)=0, u(1)=\sum_{i=1}^{m-2} a_i u(\xi_i)$ and $u'(0)=\sum_{i=1}^{m-2} b_i u'(\xi_i)$, $u(1)=\sum_{i=1}^{m-2} a_i u(\xi_i)$ are used. Finally multi-point BVP for p-Laplacian at resonance was proved by Ni and Ge in [85] and Garcia-Huidobro et al. in [17].

References

[1] A. Aftabizadeh, C. Gupta and J.-M. Xu, *Existence and uniqueness theorems for three-point boundary value problems*, SIAM J. Math. Anal. **20** (1989), 716–726.

[2] R. Avery, *A generalization of the Legget–Willaimas fixed point theorem*, Math. Res. Hot-Line **2** (1998), 9–14.

[3] R. Avery and J. Henderson, *Existence of three positive pseudo-symmetric solutions for a one-dimensional p-Laplacian*, J. Math. Anal. Appl. **277** (2003), 395–404.

[4] C. Bai and J. Fang, *Existence of multiple positive solutions for nonlinear m-point boundary value problems*, J. Math. Anal. Appl. **281** (2003), 76–85.

[5] C. Bai and J. Fang, *Existence of multiple positive solutions for nonlinear m-point boundary-value problems*, Appl. Math. Comput. **140** (2003), 297–305.

[6] C. Bai and J. Fang, *Existence of positive solutions for three-point boundary value problems at resonance*, J. Math. Anal. Appl. **291** (2004), 538–549.

[7] A. Bitsadze, *On the theory of nonlocal boundary value problems*, Russian Acad. Sci. Dokl. Math. **30** (1984), 8–10.

[8] A. Bitsadze and A. Samarskii, *On some simple generalizations of linear elliptic boundary problems*, Russian Acad. Sci. Dokl. Math. **10** (1969), 398–400.

[9] A. Boucherif and S. Bouguima, *Nonlinear second order ordinary differential equations with nonlocal boundary conditions*, Comm. Appl. Nonlinear Anal. **5** (1998), 73–85.

[10] L. Byszewski, *Theorems about existence and uniqueness of solutions of a semilinear evolution nonlocal Cauchy problem*, J. Math. Anal. Appl. **162** (1991), 494–505.

[11] L. Byszewski, *Existence and uniqueness of a classical solution to a functional-differential abstract nonlocal Cauchy problem*, J. Appl. Math. Stochastic Anal. **12** (1999), 91–97.

[12] L. Byszewski and V. Lakshmikantham, *Theorem about the existence and uniqueness of a solution of a nonlocal abstract Cauchy problem in a Banach space*, Appl. Anal. **40** (1990), 11–19.

[13] B. Calvert and C. Gupta, *Multiple solutions for a super-linear three-point boundary value problem*, Nonlinear Anal. **50** (2002), 115–128.

[14] J. Dugundji and A. Granas, *Fixed Point Theory*, Vol. I, Monographie Matematyczne, PNW, Warsawa (1982).

[15] W. Feng and J.R.L. Webb, *Solvability of three-point boundary value problems at resonance*, Proceedings of the Second World Congress of Nonlinear Analysts, Part 6 (Athens, 1996), Nonlinear Anal. **30** (1997), 3227–3238.

[16] W. Feng and J.R.L. Webb, *Solvability of m-point boundary value problems with nonlinear growth*, J. Math. Anal. Appl. **212** (1997), 467–480.

[17] M. Garcia-Huidobro, C.P. Gupta and R. Manasevich, *An m-point boundary value problem of Neumann type for a p-Laplacian like operator*, Nonlinear Anal. **56** (2004), 1071–1089.

[18] J. Graef, C. Qian and B. Yang, *A three-point boundary value problem for nonlinear fourth order differential equations*, J. Math. Anal. Appl. **287** (2003), 217–233.

[19] Y. Guo and W. Ge, *Positive solutions for three-point boundary value problems with dependence on the first order derivative*, J. Math. Anal. Appl. **290** (2004), 291–301.

[20] Y. Guo, W. Ge and Y. Gao, *Twin positive solutions for higher order m-point boundary value problems with sign changing nonlinearities*, Appl. Math. Comput. **146** (2003), 299–311.

[21] D. Guo and V. Lakshmikantham, *Nonlinear Problems in Abstract Cones*, Springer-Verlag, New York (1985).

[22] Y. Guo, X. Liu and J. Qiu, *Three positive solutions for higher order m-point boundary value problems*, J. Math. Anal. Appl. **289** (2004), 545–553.

[23] Y. Guo, W. Shan and W. Ge, *Positive solutions for second-order m-point boundary value problems*, J. Comput. Appl. Math. **151** (2003), 415–424.

[24] C. Gupta, *On a third-order three-point boundary value problem at resonance*, Differential Integral Equations **2** (1989), 1–12.

[25] C. Gupta, *Solvability of a three-point nonlinear boundary value problem for second order ordinary differential equations*, J. Math. Anal. Appl. **168** (1992), 540–551.

[26] C. Gupta, *A note on a second order three-point boundary value problem,* J. Math. Anal. Appl. **186** (1994), 277–281.

[27] C. Gupta, *A second order m-point boundary value problem at resonance*, Nonlinear Anal. **24** (1995), 1483–1489.

[28] C. Gupta, *Solvability of a generalized multi-point boundary value problem of mixed type for second order ordinary differential equations*, Proceedings of Dynamic Systems and Applications, Vol. 2 (Atlanta, GA, 1995), 215–222, Dynamic Publishers, Atlanta, GA (1996).

[29] C. Gupta, *A generalized multi-point boundary value problem for second order ordinary differential equations*, Appl. Math. Comput. **89** (1998), 133–146.

[30] C. Gupta, *A new a priori estimate for multi-point boundary-value problems*, Proceedings of the 16th Conference on Applied Mathematics (Edmond, OK, 2001), Electron. J. Differ. Equ. Conf. **7** (2001), 47–59 (electronic).

[31] C. Gupta and V. Lakshmikantham, *Existence and uniqueness theorems for a third-order three-point boundary value problem*, Nonlinear Anal. **16** (1991), 949–957.

[32] C. Gupta, S.K. Ntouyas and P.Ch. Tsamatos, *Existence results for m-point boundary value problems,* Differential Equations Dynam. Systems **2** (1994), 289–298.

[33] C. Gupta, S.K. Ntouyas and P.Ch. Tsamatos, *On an m-point boundary-value problem for second-order ordinary differential equations,* Nonlinear Anal. **23** (1994), 1427–1436.

[34] C. Gupta, S.K. Ntouyas and P.Ch. Tsamatos, *Solvability of an m-point boundary value problem for second order ordinary differential equations,* J. Math. Anal. Appl. **189** (1995), 575–584.

[35] C. Gupta, S.K. Ntouyas and P.Ch. Tsamatos, *On the solvability of some multi-point boundary value problems*, Appl. Math. **41** (1996), 1–17.

[36] C. Gupta, S.K. Ntouyas and P.Ch. Tsamatos, *Existence results for multi-point boundary value problems for second order ordinary differential equations*, Bull. Greek Math. Soc. **43** (2000), 105–123.

[37] C. Gupta and S. Trofimchuk, *Solvability of a multi-point boundary value problem and related a priori estimates*, Geoffrey J. Butler Memorial Conference in Differential Equations and Mathematical Biology (Edmonton, AB, 1996), Canad. Appl. Math. Quart. **6** (1998), 45–60.

[38] C. Gupta and S. Trofimchuk, *Existence of a solution of a three-point boundary value problem and the spectral radius of a related linear operator*, Nonlinear Anal. **34** (1998), 489–507.

[39] C. Gupta and S. Trofimchuk, *A Wirtinger type inequality and a three-point boundary value problem*, Dynam. Systems Appl. **8** (1999), 127–132.

[40] C. Gupta and S. Trofimchuk, *An interesting example for a three-point boundary value problem*, Bull. Belg. Math. Soc. Simon Stevin **7** (2000), 291–302.

[41] X. He and W. Ge, *A remark on some three-point boundary value problems for the one-dimensional p-Laplacian*, Z. Angew. Math. Mech. **82** (2002), 728–731.

[42] X. He and W. Ge, *Triple solutions for second-order three-point boundary value problems,* J. Math. Anal. Appl. **268** (2002), 256–265.

[43] V. Il'in and E. Moiseev, *Nonlocal boundary value problem of the first kind for a Sturm–Liouville operator in its differential and finite difference aspects*, Differential Equations **23** (1987), 803–810.

[44] V. Il'in and E. Moiseev, *Nonlocal boundary value problem of the second kind for a Sturm–Liouville operator,* Differential Equations **23** (1987), 979–987.

[45] G. Infante, *Eigenvalues of some nonlocal boundary-value problems,* Proc. Edinburgh Math. Soc. (2) **46** (2003), 75–86.

[46] G. Infante and J.R.L. Webb, *Nonzero solutions of Hammerstein integral equations with discontinuous kernels*, J. Math. Anal. Appl. **272** (2002), 30–42.

[47] G. Infante and J.R.L. Webb, *Positive solutions of some nonlocal boundary value problems,* Abstr. Appl. Anal. **18** (2003), 1047–1060.

[48] G.L. Karakostas and P.Ch. Tsamatos, *Positive solutions for a nonlocal boundary-value problem with increasing response*, Electron. J. Differential Equations **73** (2000), 8 pp. (electronic).

[49] G.L. Karakostas and P.Ch. Tsamatos, *Positive solutions of a boundary-value problem for second order ordinary differential equations*, Electron. J. Differential Equations **49** (2000), 9 pp. (electronic).

[50] G.L. Karakostas and P.Ch. Tsamatos, *On a nonlocal boundary value problem at resonance*, J. Math. Anal. Appl. **259** (2001), 209–218.

[51] G.L. Karakostas and P.Ch. Tsamatos, *Multiple positive solutions for a nonlocal boundary-value problem with response function quiet at zero*, Electron. J. Differential Equations **13** (2001), 10 pp. (electronic).

[52] G.L. Karakostas and P.Ch. Tsamatos, *Sufficient conditions for the existence of nonnegative solutions of a nonlocal boundary value problem*, Appl. Math. Lett. **15** (2002), 401–407.

[53] G.L. Karakostas and P.Ch. Tsamatos, *Nonlocal boundary vector value problems for ordinary differential systems of higher order*, Nonlinear Anal. **51** (2002), 1421–1427.

[54] M. Krasnosel'skii, *Positive Solutions of Operator Equations*, Noordhoff, Groningen (1964).

[55] R. Legget and L. Williams, *Multiple positive fixed points of nonlinear operators on ordered Banach spaces*, Indiana Univ. Math. J. **28** (1979), 673–688.

[56] B. Liu, *Positive solutions of a nonlinear three-point boundary value problem*, Appl. Math. Comput. **132** (2002), 11–28.

[57] B. Liu, *Positive solutions of a nonlinear three-point boundary value problem*, Comput. Math. Appl. **44** (2002), 201–211.

[58] B. Liu, *Positive solutions of a nonlinear three-point boundary value problem*, Appl. Math. Comput. **132** (2002), 11–28.

[59] B. Liu, *Solvability of multi-point boundary value problem at resonance IV*, Appl. Math. Comput. **143** (2003), 275–299.

[60] B. Liu, *Solvability of multi-point boundary value problem at resonance II*, Appl. Math. Comput. **136** (2003), 353–377.

[61] B. Liu and J. Yu, *Solvability of multi-point boundary value problems at resonance I*, Indian J. Pure Appl. Math. **33** (2002), 475–494.

[62] B. Liu and J. Yu, *Solvability of multi-point boundary value problem at resonance III*, Appl. Math. Comput. **129** (2002), 119–143.

[63] X. Liu, *Nontrivial solutions of singular m-point boundary value problems*, J. Math. Anal. Appl. **284** (2003), 576–590.

[64] X. Liu, J. Qiu and Y. Guo, *Three positive solutions for second-order m-point boundary value problems*, Appl. Math. Comput. **156** (2004), 733–742.

[65] Y. Liu and W. Ge, *Positive solutions for $(n-1,1)$ three-point boundary value problems with coefficient that changes sign*, J. Math. Anal. Appl. **282** (2003), 816–825.

[66] Y. Liu and W. Ge, *Multiple positive solutions to a three-point boundary value problem with p-Laplacian*, J. Math. Anal. Appl. **277** (2003), 293–302.

[67] R. Ma, *Existence theorems for a second order m-point boundary value problem*, J. Math. Anal. Appl. **211** (1997), 545–555.

[68] R. Ma, *Existence theorems for a second order three-point boundary value problem*, J. Math. Anal. Appl. **212** (1997), 430–442.

[69] R. Ma, *Positive solutions of a nonlinear three-point boundary value problem*, Electron. J. Differential Equations **34** (1999), 8 pp. (electronic).

[70] R. Ma, *Multiplicity of positive solutions for second-order three-point boundary value problems*, Comput. Math. Appl. **40** (2000), 193–204.

[71] R. Ma, *Existence of solutions of a nonlinear m-point boundary value problems*, J. Math. Anal. Appl. **256** (2001), 556–567.

[72] R. Ma, *Positive solutions of a nonlinear m-point boundary value problem*, Comput. Math. Appl. **42** (2001), 755–765.

[73] R. Ma, *Positive solutions for second order functional differential equations*, Dynam. Systems Appl. **10** (2001), 215–223.

[74] R. Ma, *Existence and uniqueness of solutions to first-order three-point boundary value problems*, Appl. Math. Lett. **15** (2002), 211–216.

[75] R. Ma, *Existence of positive solutions for superlinear semipositone m-point boundary value problems*, Proc. Edinburgh Math. Soc. **46** (2003), 279–292.

[76] R. Ma, *Multiplicity results for a three-point boundary value problem at resonance*, Nonlinear Anal. **53** (2003), 777–789.

[77] R. Ma, *Multiple positive solutions for nonlinear m-point boundary value problems*, Appl. Math. Comput. **148** (2004), 249–262.

[78] R. Ma and N. Castaneda, *Positive solutions for second-order three-point boundary value problems*, Appl. Math. Lett. **14** (2001), 1–5.

[79] R. Ma and H. Wang, *Positive solutions of nonlinear three-point boundary-value problems*, J. Math. Anal. Appl. **279** (2003), 216–227.

[80] S. Marano, *A remark on a second order three-point boundary value problem*, J. Math. Anal. Appl. **183** (1994), 518–522.

[81] S. Marano, *A remark on a third-order three-point boundary value problem*, Bull. Austral. Math. Soc. **49** (1994), 1–5.

[82] J. Mawhin, *Topological Degree Methods in Nonlinear Boundary Value Problems*, NSF-CBMS Reg. Conf. Ser. Math., Vol. 40, Amer. Math. Soc., Providence, RI (1979).

[83] M. Moshinsky, *Sobre los problemas de condiciones a la frontiera en una dimension de caracteristicas discontinuas*, Bol. Soc. Mat. Mexicana **7** (1950), 1–25.

[84] K. Murty and S. Sivasundaram, *Existence and uniqueness of solutions to three-point boundary value problems associated with nonlinear first order systems of differential equations*, J. Math. Anal. Appl. **173** (1993), 158–164.

[85] X. Ni and W. Ge, *Multi-point boundary value problems for the p-Laplacian at resonance*, Electron. J. Differential Equations **112** (2003), 7 pp. (electronic).

[86] S.K. Ntouyas and P.Ch. Tsamatos, *Global existence for semilinear evolution equations with nonlocal conditions*, J. Math. Anal. Appl. **210** (1997), 679–687.

[87] P.K. Palamides, *Positive and monotone solutions of an m-point boundary value problem*, Electron. J. Differential Equations **18** (2002), 1–16 pp. (electronic).

[88] B. Przeradski and R. Stanczy, *Solvability of a multi-point boundary value problem at resonance*, J. Math. Anal. Appl. **264** (2001), 253–261.

[89] Y. Raffoul, *Positive solutions of three-point nonlinear second order boundary value problem*, Electron. J. Qual. Theory Differ. Equ. **15** (2002), 11 pp. (electronic).

[90] S. Timoshenko, *Theory of Elastic Stability*, McGraw-Hill, New York (1961).

[91] J. Wang and D. Zheng, *On the existence of positive solutions to a three-point boundary value problem for the one-dimensional p-Laplacian*, Z. Angew. Math. Mech. **77** (1997), 477–479.

[92] J.R. Webb, *Positive solutions of some three-point boundary value problems via fixed point index theory*, Nonlinear Anal. **47** (2001), 4319–4322,

[93] M. Zhang and Y. Han, *On the applications of Leray–Schauder continuation theorem to boundary value problems of semilinear differential equations*, Ann. Differential Equations **13** (1997), 189–207.

[94] Z. Zhang and J. Wang, *The upper and lower solution method for a class of singular nonlinear second order three-point boundary value problems*, J. Comput. Appl. Math. **147** (2002), 41–52.

[95] Z. Zhang and J. Wang, *Positive solutions to a second order three-point boundary value problem*, J. Math. Anal. Appl. **285** (2003), 237–249.

This page intentionally left blank

Author Index

Roman numbers refer to pages on which the author (or his/her work) is mentioned. Italic numbers refer to reference pages. Numbers between brackets are the reference numbers. No distinction is made between first and co-author(s).

Abbondandolo, A. 80, 106, 121, *142* [1]
Adams, R.A. 98, 126, 129, *142* [2]
Aftabizadeh, A. 489, *554* [1]
Agrachev, A.A. 3, 16, 18, 57, 67, *74* [1]; *74* [2]
Ahmad, S. 439, *459* [1]; *459* [2]
Albrecht, F. 297, *348* [1]
Alexandroff, A. 315, *348* [2]
Alikakos, N. 314, 321, 332, *348* [3]
Alvárez, C. 439, *459* [3]
Amann, H. 104, 106, 121, *142* [3]; *142* [4]; 275, 314, 315, 320, 332, 338, 340, 345, *348* [4]; *349* [5]; *349* [6]; *349* [7]; *349* [8]; *349* [9]; *349* [10]; *349* [11]; 398, *459* [4]
Ambrosetti, A. 80, *142* [5]
Amine, Z. 458, *459* [5]
Angenent, S. *349* [12]; *349* [27]
Arioli, G. 90, 129, 138, *142* [6]
Arnold, V.I. 5, 6, 8, 19, 27, *74* [3]
Aubin, J.P. 154, 188, 228, 229, 233, 234, *235* [1]; *235* [2]; *235* [3]; *235* [4]
Avery, R. 548, 553, *554* [2]; *554* [3]

Bahri, A. 91, *142* [7]; *142* [8]
Bai, C. 546, 554, *554* [4]; *554* [5]; *554* [6]
Barbu, V. 3, 11, 38, 47, 67, 70, *74* [4]; *74* [5]; *74* [6]; 230, *235* [5]
Bartsch, T. 88–90, 97, 107, 125, *142* [9]; *142* [10]; *142* [11]; *143* [12]; *143* [13]; *143* [14]; *143* [15]; *143* [16]
Batty, C. *349* [13]
Bebernes, W. 227, *235* [6]
Bellman, R. 3, *74* [7]; 223, *235* [7]; *235* [8]; *235* [9]
Benaïm, M. 296, 302, *349* [14]; *349* [15]; *349* [16]; *349* [17]
Benci, V. 80, 88, 97, 106, 121, *143* [17]; *143* [18]; *143* [19]
Berestycki, H. 91, 115, 121, *142* [7]; *142* [8]; *143* [20]
Berliocchi, H. 190, *235* [10]
Berman, A. 282, 284, 286–289, *349* [18]; *349* [19]
Birkhoff, G. 315, *349* [20]
Bitsadze, A. 470, *554* [7]; *554* [8]
Bliss, G.A. 4, *74* [8]
Bolotin, S.V. 139, *143* [21]
Boltyanskii, V.G. 29, *75* [30]; 223, *235* [11]; *235* [12]; *237* [83]
Bony, J.M. 153, *235* [13]
Bothe, D. 233, 234, *235* [14]; *235* [15]
Boucherif, A. 482, 484, *554* [9]
Bouguima, S. 482, 484, *554* [9]
Bouligand, H. 159, 160, *235* [16]
Bressan, A. 233, *235* [17]
Brezis, H. 11, 13, *74* [9]; 153, 154, 156, 157, *235* [18]; *235* [19]
Browder, F. 156, 157, *235* [19]
Brunovsky, P. 296, 314, *349* [21]; *349* [22]
Byszewski, L. 465, 469, *554* [10]; *554* [11]; *554* [12]

Calvert, B. 542, *554* [13]
Carathéodory, C. 223, *235* [29]
Cârjă, O. 154, 158, 176, 178, 190, 203, 210, 213, 225, 227, 228, 231, 233, 234, *235* [20]; *235* [21]; *235* [22]; *235* [23]; *235* [24]; *235* [25]; *235* [26]; *235* [27]; *235* [28]
Castaneda, N. 533, *557* [78]
Cellina, A. 154, 188, 228, 229, 233, 234, *235* [3]
Cerami, G. 96, *143* [22]
Cesari, L. 223, *236* [30]
Chang, K.C. 81, 82, 106, 110, 121, 124, *143* [23]
Chen, M. 314, *349* [24]
Chen, X. 314, *349* [24]
Chen, X.-Y. 314, *349* [23]

Chiş-Şter, I. 232, *236* [31]
Cholewa, J. 332, 345, *349* [25]
Chow, S.N. 331, *349* [26]
Clarke, F.H. 3, 14, 15, 32, 50, 73, *74* [10]; *74* [11]; *74* [12]; *74* [13]; *74* [14]; 163, 175, 223, 226, *236* [32]; *236* [33]; *236* [34]; *236* [35]
Clements, Ph. *349* [27]
Coddington, E.A. 16, *74* [15]; *349* [28]
Cohen, D.S. 314, *352* [95]
Colombo, F. 340, *349* [29]
Conley, C. 124, *143* [24]; 293, *349* [30]; *349* [31]
Coppel, W. *349* [32]
Cornet, B. 166, *236* [36]
Cosner, C. 345, *349* [33]
Coti Zelati, V. 80, 126, 135, 136, 138–140, *142* [5]; *143* [25]; *143* [26]
Courant, R. 7, 8, *74* [16]; 223, *236* [37]; 314, *349* [34]
Crandall, M.G. 68, 70, *74* [17]; *75* [18]; *75* [19]; *75* [20]; 151, 152, 154, 233, *236* [38]; *236* [39]; 371, 404, 407, *459* [6]
Cushing, J.M. 392, 439, *459* [7]

Da Prato, G. 70, *74* [5]
Dafermos, C.M. 327, *350* [35]
Dancer, E.N. 316, 318, 320–322, 332, *350* [36]; *350* [37]; *350* [38]; 392, 397, 458, *459* [8]; *459* [9]; *459* [10]
Dancer, N. 107, 109, *143* [27]
de Leenheer, P. 302, *350* [41]
de Mottoni, P. 314, 329–331, *350* [42]; 438, 454, *460* [29]
de Pagter, B. *349* [27]
DeAngelis, D.L. 242, *350* [39]
Degiovanni, M. 121, *143* [28]
Deimling, K. *350* [40]
Delgado, M. 438, *459* [11]
Ding, Y.H. 97, 129, 132, 135, 136, 138, 140, *143* [12]; *143* [13]; *143* [29]; *143* [30]; *143* [31]
Dlotko, T. 332, 345, *349* [25]
Dugundji, J. 482, 483, *554* [14]
Dunford, N. 201, 234, *236* [40]

Eden, A. *350* [43]
Eilbeck, J.C. 363, 384, 439, 452, *459* [12]; *459* [13]
Ekeland, I. 15, 50, *74* [13]; *75* [21]; 80, 115, 118, 126, 135, 139, 140, *143* [25]; *143* [32]; *143* [33]; *143* [34]; *143* [35]
Elsner, L. 286, *350* [44]
Evans, L.C. 68, 70, *74* [17]

Fadell, E.R. 88, 109, 121, *143* [36]
Fang, J. 546, 554, *554* [4]; *554* [5]; *554* [6]
Federer, H. 161, *236* [41]
Feferman, S. 156, *236* [42]
Felmer, P. 118, 124, *143* [37]; *143* [38]; *144* [39]
Feng, W. 497, 505, *554* [15]; *554* [16]
Fife, P. *350* [49]; *350* [50]
Filippov, A.F. 213, *236* [43]
Fleckinger-Pellé, J. 327, *350* [45]; *350* [46]
Fleming, W.H. 3, 73, *75* [22]
Foias, C. *350* [43]
Fomin, S.V. 4, 8, *75* [24]
Fonda, A. 124, *144* [40]
Fournier, G. 92, 94, 124, *144* [42]
Frankowska, H. 190, 228, 234, *235* [4]; *236* [44]; *236* [45]
Freedman, H.I. 320, 313, *356* [235]; *357* [236]
Friedman, A. *350* [47]; *350* [48]
Frobenius, G. 315, *350* [51]
Fryszkowski, A. 233, *236* [46]
Fukaya, K. 125, *144* [41]
Fukuhara, M. 219, 220, *236* [47]
Furter, J.E. 384, *459* [13]

Gamkrelidze, R.V. 16, 29, *74* [1]; *75* [23]; *75* [30]; 223, *237* [83]
Gao, Y. 547, *554* [20]
Garcia-Huidobro, M. 554, *554* [17]
Gatzke, H. 297, *348* [1]
Gautier, S. 228, *236* [48]
Ge, W. 540–542, 546, 547, 553, 554, *554* [19]; *554* [20]; *555* [23]; *555* [41]; *555* [42]; *556* [65]; *556* [66]; *557* [85]
Gelfand, I.M. 4, 8, *75* [24]
Ginzburg, V.L. 110, 113, *144* [43]; *144* [44]; *144* [45]
Girardi, M. 115, 136, 138, *143* [30]; *144* [46]
Girsanov, I.V. 161, *236* [49]
Glas, D. 254, *350* [55]
Gonzales, R. 223, *236* [50]
Goursat, E. 223, *236* [51]
Graef, J. 542, *554* [18]
Granas, A. 482, 483, *554* [14]
Greiner, G. *350* [52]
Grossberg, S. 297, *350* [53]; *350* [54]
Guckenheimer, J. 80, 138, 142, *144* [47]
Guo, D. 519, 525, 536, *554* [21]
Guo, Y.X. 121, *144* [48]; 541, 542, 546, 547, *554* [19]; *554* [20]; *554* [22]; *555* [23]; *556* [64]
Gupta, C.P. 481, 489, 494, 495, 497, 505, 513, 516, 542, 554, *554* [1]; *554* [13]; *554* [17]; *555* [24]; *555* [25]; *555* [26]; *555* [27]; *555* [28]; *555* [29]; *555* [30]; *555* [31]; *555* [32]; *555* [33]; *555* [34]; *555* [35]; *555* [36]; *555* [37]; *555* [38]; *555* [39]; *555* [40]

Gürel, B.Z. 113, *144* [44]

Haddad, A. 297, *348* [1]
Haddad, G. 227, *236* [52]
Hadeler, K. 254, *350* [55]
Hájek, O. 223, *236* [53]
Hale, J.K. 301–303, 305, 307, 309, 313, 314, 329, 331, 332, 344, *349* [24]; *349* [26]; *350* [56]; *350* [57]; *350* [58]; *350* [59]; *350* [60]; *350* [61]; 439, *459* [14]
Hall, G.R. 80, *145* [75]
Han, Y. 482, *557* [93]
Hartman, P. 151, 152, 220, *236* [54]; *236* [55]
He, X. 540, 553, *555* [41]; *555* [42]
Heijmans, H. *349* [27]
Henderson, J. 553, *554* [3]
Henry, D. 332–334, *350* [62]
Herman, M. 113, *144* [49]
Hess, P. 314, 316, 318, 320, 321, 329, 332, *348* [3]; *350* [38]; *350* [63]; *350* [64]; *350* [65]; 439, *459* [15]
Hilbert, D. 7, 8, *74* [16]; 223, *236* [37]; 314, *349* [34]
Hirsch, M.W. 242, 247, 249, 250, 254–256, 258, 264, 265, 269, 271, 272, 275, 281, 291, 292, 296, 297, 299, 300, 302, 317, 319, *349* [16]; *349* [17]; *351* [66]; *351* [67]; *351* [68]; *351* [69]; *351* [70]; *351* [71]; *351* [72]; *351* [73]; *351* [74]; *351* [75]; *351* [76]; *351* [77]; *351* [78]; *351* [79]; *351* [80]
Hobson, E.W. 175, *236* [56]
Hofbauer, J. 301, 302, *351* [81]; *351* [82]
Hofer, H. 80, 113, 115, 125, *144* [50]; *144* [51]; *144* [52]
Holmes, P. 80, 138, 142, *144* [47]
Hopf, H. 315, *348* [2]
Howard, P. 156, *236* [57]
Hsu, S.-B. 302, 320, 332, *351* [83]; *351* [84]

Il'in, V. 470, *555* [43]; *555* [44]
Infante, G. 532, 542, *555* [45]; *555* [46]; *555* [47]
Ishii, H. 70, *75* [18]; *75* [19]; 223, *236* [58]
Isidori, A. 3, *75* [25]
Izydorek, M. 121, 124, *144* [53]; *144* [54]

Jeanjean, L. 133, *144* [55]
Jentzsch, R. 315, *351* [85]
Jiang, J. 265, 296, 314, 319, 332, *351* [86]; *351* [87]; *351* [88]; *351* [89]; *351* [90]; *352* [120]; *352* [121]; *356* [229]; *356* [230]; *356* [231]
Jurdjevic, V. 3, 67, *75* [26]

Kamke, E. 223, *237* [59]; 241, 285, *351* [91]
Karakostas, G.L. 542, 547, *556* [48]; *556* [49]; *556* [50]; *556* [51]; *556* [52]; *556* [53]
Kato, T. 274, 275, *351* [92]
Keller, H.B. 314, *351* [93]; *352* [94]; *352* [95]
Kenmochi, N. 181, *237* [60]
Kerman, E. 110, *144* [45]
Kerscher, W. *352* [96]
Kneser, H. 219, 221, *237* [61]
Kobayashi, J. 213, *237* [62]
Kolmogorov, A. 297, *352* [97]
Krasnosel'skii, M.A. 88, *144* [56]; 314, 323, 324, *352* [98]; *352* [99]; *352* [100]; *352* [101]; 363, 411, *459* [16]; 519, *556* [54]
Krause, U. 315, 323, 324, 327, *352* [102]; *352* [103]
Krein, M.G. 274, 315, *352* [104]
Krisztin, T. 303, 313, *352* [105]; *352* [106]; *352* [107]; *352* [108]
Kryszewski, W. 95, 106, 121, 136, *144* [57]; *144* [58]
Kucia, A. 190, *237* [63]
Kunisch, K. 305, *352* [109]
Kunze, H. 289, *352* [110]; *352* [111]; *352* [112]
Kuznetsov, Y.A. 328, *352* [113]

Ladyzhenskaya, L. 314, *352* [100]; *352* [114]
Lakshmikantham, V. 465, 489, 519, 525, 536, *554* [12]; *554* [21]; *555* [31]
LaSalle, J.P. 316, *352* [116]
Lasry, J.M. 115, 121, *143* [20]; *143* [35]; 190, *235* [10]
Lassoued, L. 115, *143* [34]
Lazer, A.C. 332, *350* [64]; 439, *459* [2]; *459* [3]; *459* [15]
Ledyaev, Yu.S. 175, 226, 228, *236* [33]; *236* [34]; *237* [64]
Legget, R. 536, 548, *556* [55]
Lemmens, B. 315, *352* [117]
Lemmert, R. 284, *352* [118]
Leung, A.W. *352* [119]
Levinson, N. 16, *74* [15]; *349* [28]
Li, G. 96, *144* [59]
Li, M. 301, *352* [115]
Li, S.J. 83, 86, 91, 106, 118, 121, *144* [60]; *144* [61]; *144* [62]
Li, W. 301, *357* [237]
Li, Y. 54, *75* [36]
Li, Y.Q. 92, *144* [63]
Liang, X. 314, 332, *352* [120]; *352* [121]
Lieb, E.H. 129, *144* [64]
Lions, J.-L. 3, *75* [27]
Lions, P.-L. 68, 70, *74* [17]; *75* [18]; *75* [19]; *75* [20]; 95, 129, *144* [65]; 233, *236* [39]

Liu, B. 516, 517, 525, 529, 532, 541, *556* [56]; *556* [57]; *556* [58]; *556* [59]; *556* [60]; *556* [61]; *556* [62]
Liu, C.G. 115, *144* [67]
Liu, J.Q. 83, 91, 106, 121, 124, *144* [60]; *144* [61]; *144* [66]
Liu, X. 541, 542, 547, *554* [22]; *556* [63]; *556* [64]
Liu, Y. 542, 553, *556* [65]; *556* [66]
Loewy, R. *353* [122]
Long, Y. 80, 106, 115, 118, 121, *144* [67]; *145* [68]; *145* [69]; *145* [70]; *145* [71]
López-Gómez, J. 363, 384, 390, 392, 394–397, 404, 411, 438, 439, 452, 458, *459* [10]; *459* [11]; *459* [12]; *459* [13]; *459* [17]; *459* [18]; *459* [19]; *459* [20]; *459* [21]; *459* [22]; *459* [23]; *459* [24]; *459* [25]; *459* [26]
Loss, M. 129, *144* [64]
Lui, R. 315, *353* [123]
Lunardi, A. 332, 334, 339–342, 345, *353* [124]
Lupo, D. 92, 94, 124, *144* [42]
Lupulescu, V. 154, *237* [65]
Lyapunov, A.M. 106, *145* [72]

Ma, R. 475, 525, 532, 533, 535, 540–542, *556* [67]; *556* [68]; *556* [69]; *556* [70]; *556* [71]; *556* [72]; *556* [73]; *556* [74]; *557* [75]; *557* [76]; *557* [77]; *557* [78]; *557* [79]
Manasevich, R. 554, *554* [17]
Mancini, G. 115, 121, *143* [20]
Marano, S. 481, 489, *557* [80]; *557* [81]
Marsden, J.E. 6, 19, *75* [28]
Martin, R.H. 305, 332, 335–338, 343, *353* [125]; *353* [126]; *353* [127]; *353* [128]; *353* [129]; *353* [130]
Martin, R.H., Jr. 153, 154, *237* [66]
Matano, H. 247, 314, 320, 321, 332, *348* [3]; *349* [23]; *353* [131]; *353* [132]; *353* [133]; *353* [134]; *353* [135]
Matzeu, M. 115, *144* [46]
Mawhin, J. 80, 89, 124, *144* [40]; *145* [73]; 363, *459* [27]; *460* [28]; 474, 476, 477, 486, 490, 491, 498, *557* [82]
May, R. 315, *353* [136]
McDuff, D. 80, *145* [74]
McShane, E.J. 293, *353* [137]
Meyer, K.R. 80, *145* [75]
Michael, E. 212, 229, *237* [67]; *237* [68]
Mierczyński, J. 265, 280, 296, 327, 328, *353* [138]; *353* [139]; *353* [140]; *353* [141]; *353* [142]; *353* [143]
Miller, R.K. 320, *357* [236]
Mincheva, M. 345, *353* [144]; *353* [145]
Mischaikow, K. *353* [146]
Mishchenko, E.F. 29, *75* [30]; 223, *237* [83]
Moiseev, E. 470, *555* [43]; *555* [44]
Molina-Meyer, M. 363, 390, *459* [21]; *459* [22]; *459* [23]; *459* [24]
Monteiro Marques, M.D.P. 190, 227, 228, 231, *235* [21]; *235* [22]; *235* [23]
Mora, X. 332, 340, *353* [147]
Mordukhovich, B.S. 153, *237* [69]
Moreau, J. 13, *75* [29]
Moser, J. 106, 107, *145* [76]
Moshinsky, M. 470, *557* [83]
Muldowney, J. 301, *352* [115]
Müller, M. 241, 285, *353* [148]
Murty, K. 475, *557* [84]

Nagel, R. *352* [96]
Nagumo, M. 151, 166, *237* [70]
Necula, M. 154, 163, 178, 231, 233, *235* [24]; *237* [65]; *237* [71]; *237* [72]; *237* [73]; *237* [74]
Neumann, M. 282, 284, 286–289, *349* [18]
Ni, X. 554, *557* [85]
Nicolaenko, B. *350* [43]
Ntouyas, S.K. 469, 481, 494, 497, 505, 513, *555* [32]; *555* [33]; *555* [34]; *555* [35]; *555* [36]; *557* [86]
Nussbaum, R.D. 315, 327, *352* [102]; *352* [117]; *354* [149]; *354* [150]; *354* [151]

Ohta, Y. 305, *354* [152]
Olian Fannio, L. 121, *143* [28]
Ono, K. 125, *144* [41]
Ortega, R. 302, *354* [153]; 439, 447, 458, *459* [5]; *459* [10]; *459* [25]; *460* [30]
Oster, G. 315, *353* [136]
Oxtoby, J. 257, *354* [154]

Palamides, P.K. 547, *557* [87]
Palmer, K.J. 138, 140, *145* [77]
Pao, C.V. 341, 343, *354* [155]
Pardo, R. 458, *459* [26]
Pavel, N.H. 230–232, 234, *235* [5]; *237* [75]; *237* [76]; *237* [77]; *237* [78]; *237* [79]
Pazy, A. 333, *354* [156]
Peano, G. 149, *237* [80]
Peixoto, M. 246, *354* [157]
Perron, O. 223, *237* [81]; 315, *354* [158]
Pituk, M. 313, *354* [159]
Plaskacz, S. 190, 227, 228, *236* [45]; *237* [82]
Plemmons, R. 287–289, *349* [19]
Poláčik, P. 280, 281, 314, 320, 329, 332, 343, 345, *349* [22]; *350* [65]; *354* [160]; *354* [161]; *354* [162]; *354* [163]; *354* [164]; *354* [165]
Pontryagin, L.S. 29, *75* [30]; 223, *237* [83]
Post, W.M. 242, *350* [39]

Precupanu, Th. 11, 47, *74* [6]
Protter, M.H. 346, *354* [166]
Przeradski, B. *557* [88]

Qian, C. 542, *554* [18]
Qiu, J. 541, 542, 547, *554* [22]; *556* [64]
Quincampoix, M. 161, *237* [84]

Rabinowitz, P.H. 79, 80, 82, 87, 88, 99, 100, 109, 111, 118, 121, 136, 138, 139, *143* [19]; *143* [21]; *143* [26]; *143* [36]; *145* [78]; *145* [79]; *145* [80]; *145* [81]; 363, 371, 392, 395–397, 404, 407, *459* [6]; *460* [31]
Raffoul, Y. 535, *557* [89]
Ramos, M. 92, 94, 124, *144* [42]
Ranft, P. 323, 324, 327, *352* [103]
Ratiu, T.S. 6, 19, *75* [28]
Raugel, G. *354* [167]
Redheffer, R. 153, 154, *237* [85]
Rescigno, A. 297, *354* [168]
Richardson, I. 297, *354* [168]
Rishel, R.W. 3, 73, *75* [22]
Robinson, C. *349* [13]; *354* [169]
Robinson, J. *354* [170]
Rockafellar, R.T. 11, 14, 47, *75* [31]; *75* [32]
Roxin, E. 174, *237* [86]
Rubin, J.E. 156, *236* [57]
Rudin, W. 306, *354* [171]
Ruelle, D. *354* [172]
Ruf, B. 115, 121, *143* [20]
Rutman, M.A. 274, 315, *352* [104]; *354* [173]
Rybicki, S. 107, 109, *143* [27]
Rzeżuchowski, T. 190, 228, *236* [45]

Sachkov, Y.L. 3, 16, 18, 57, 67, *74* [2]
Sacker, R. 332, *354* [174]
Salamon, D. 80, *145* [74]
Samarskii, A. 470, *554* [8]
Sánchez, L. 302, *354* [153]
Sandholm, W. 302, *351* [81]
Sandstede, B. 314, *349* [22]
Saperstone, S.H. 246, *354* [175]
Sattinger, D. 314, *354* [176]
Schappacher, W. 305, *352* [109]
Schiaffino, A. 314, 329–331, *350* [42]; 438, 454, *460* [29]
Schneider, H. 282, 284, 286, 301, *353* [122]; *354* [177]
Schuur, I.D. 227, *235* [6]
Schwartz, J.T. 201, 234, *236* [40]
Scorza Dragoni, G. 190, *237* [87]; *237* [88]
Secchi, S. 135, *145* [82]
Selgrade, J. 297, 315, *354* [178]; *355* [179]; *355* [180]; *355* [181]
Sell, G. 332, *354* [174]
Séré, E. 126, 135, 139, 140, 142, *143* [25]; *145* [83]
Severi, F. 159, 222, *237* [89]; *238* [90]
Shan, W. 541, *555* [23]
Shi, S.Z. 231, 232, *238* [91]
Siegel, D. 289, 345, *352* [110]; *352* [111]; *352* [112]; *353* [145]
Sivasundaram, S. 475, *557* [84]
Sleeman, B.D. 332, *357* [246]
Slemrod, M. 327, *350* [35]
Smale, S. 294, 297, *351* [80]; *355* [182]
Smith, H.L. 242, 247, 249, 250, 255, 266, 267, 269, 271, 272, 281, 293, 296, 297, 299, 301, 302, 304, 305, 311–315, 318, 320, 327–329, 332, 337, 343, 345, 346, *350* [41]; *351* [83]; *353* [129]; *353* [130]; *353* [146]; *355* [183]; *355* [184]; *355* [185]; *355* [186]; *355* [187]; *355* [188]; *355* [189]; *355* [190]; *355* [191]; *355* [192]; *355* [193]; *355* [194]; *355* [195]; *355* [196]; *355* [197]; *355* [198]; *355* [199]; *355* [200]; *355* [201]; *355* [202]; *355* [203]; *357* [248]; 439, *460* [32]; *460* [33]
Smith, R.A. 302, *355* [204]
Smoller, J. 345, *355* [205]
So, J.W.-H. 301, *351* [82]
Somolinos, A.S. 314, 329, 332, *350* [60]; 439, *459* [14]
Sontag, E.D. 3, *75* [33]
Spanier, E.H. 83, *145* [84]
Staicu, V. 233, *235* [17]
Stanczy, R. *557* [88]
Stern, R.J. 175, 226, *236* [33]; *236* [34]; 282, 284–289, *349* [18]; *355* [206]
Stewart, H. *355* [207]
Struwe, M. 87, 113, *145* [85]; *145* [86]
Stuart, C.A. 126, 131, 135, *145* [82]; *145* [87]
Suárez, A. 438, *459* [11]
Szlenk, W. 142, *145* [88]
Szulkin, A. 86, 90, 92, 95, 96, 106, 109, 115, 118, 121, 124, 129, 135, 136, 138, *142* [6]; *144* [57]; *144* [58]; *144* [59]; *144* [62]; *145* [89]; *145* [90]; *145* [91]; *145* [92]; *145* [93]; *145* [94]

Takáč, P. 274, 281, 315, 316, 318, 321, 322, 327, 328, 332, *350* [45]; *350* [46]; *356* [208]; *356* [209]; *356* [210]; *356* [211]; *356* [212]; *356* [213]; *356* [214]; *356* [215]; *356* [216]
Takahashi, T. 181, *237* [60]
Tallos, P. 228, *238* [92]
Tanaka, K. 135, *145* [95]

Tang, M. *350* [50]
Temam, R. *350* [43]
Tereščák, I. 329, *354* [164]; *354* [165]; *356* [217]
Thieme, H.R. 247, 250, 255, 266, 267, 269, 271, 272, 281, 304, 313, 315, 320, 332, *353* [146]; *355* [197]; *355* [198]; *355* [199]; *355* [200]; *355* [201]; *356* [218]
Thompson, A.C. 324, *356* [219]
Tikhomirov, V.M. 4, *75* [34]
Timoshenko, S. 470, *557* [90]
Tineo, A. 296, *356* [220]; 439, 447, 458, *459* [25]; *460* [30]; *460* [34]; *460* [35]
Travis, C.C. 242, *350* [39]
Treiman, J.S. 166, *238* [93]
Trofimchuk, S. 494, 495, *555* [37]; *555* [38]; *555* [39]; *555* [40]
Tsamatos, P.Ch. 469, 481, 494, 497, 505, 513, 542, 547, *555* [32]; *555* [33]; *555* [34]; *555* [35]; *555* [36]; *556* [48]; *556* [49]; *556* [50]; *556* [51]; *556* [52]; *556* [53]; *557* [86]

Uhl, R. 282, *356* [221]
Ursescu, C. 151, 152, 154, 158, 161, 162, 166, 175, 176, 210, 213, 222, 225, 227, 228, 234, *235* [25]; *235* [26]; *237* [77]; *238* [94]; *238* [95]; *238* [96]; *238* [97]; *238* [98]; *238* [99]; *238* [100]; *238* [101]
Uryson, P.S. 314, *356* [222]

van den Driessche, P. 296, 301, *356* [223]
van Duijn, C. *349* [27]
Verduyn Lunel, S.M. 302, 303, 305, 307, 309, 315, *350* [61]; *352* [117]
Vespri, V. 340, *349* [29]
Vidyasagar, M. 282, 284, 286, 301, *354* [177]
Vinter, R.B. 73, *74* [14]; 223, *236* [35]
Viterbo, C. 113, 115, *145* [96]; *146* [97]
Volkmann, P. 282, 284, *352* [118]; *356* [224]
Vrabie, I.I. 21, *75* [35]; 151, 152, 154, 171, 173, 174, 177, 178, 228, 230–234, *235* [24]; *235* [27]; *235* [28]; *237* [74]; *237* [78]; *237* [79]; *238* [102]; *238* [103]; *238* [104]; *238* [105]; *238* [106]
Vulikh, B. 338, *356* [225]

Walcher, S. 282, *356* [226]
Walter, W. 282, 289, 343, *356* [227]; *356* [228]
Walther, H.-O. 303, 313, *352* [105]
Waltman, P. 249, 302, 320, 332, 343, *351* [83]; *351* [84]; *355* [202]; *355* [203]
Wang, H.Z. 54, *75* [36]; 532, *557* [79]
Wang, J. 533, 553, *557* [91]; *557* [94]; *557* [95]
Wang, Y. 296, 314, 332, *356* [229]; *356* [230]; *356* [231]
Wang, Z.-Q. 125, *143* [14]; *143* [15]; *144* [39]
Wax, N. 297, *348* [1]
Ważewski, T. 227, *238* [107]
Webb, J.R.L. 497, 505, 532, 542, *554* [15]; *554* [16]; *555* [46]; *555* [47]; *557* [92]
Weinberger, H.F. 315, 346, *354* [166]; *356* [232]
Weinstein, A. 106, 107, 111, *146* [98]; *146* [99]
Wiggins, S. 138, 140, *146* [100]
Wilder, R. 322, *356* [233]
Willem, M. 80–82, 89, 90, 92, 94, 99, 124, 127, 129, 132, 135, 140, *143* [16]; *143* [31]; *144* [42]; *145* [73]; *146* [101]; *146* [102]
Williams, L. 536, 548, *556* [55]
Wolenski, P.R. 226, *236* [34]
Wolkowicz, H. 285, *355* [206]
Wu, J. 303, 313, 320, *352* [105]; *352* [106]; *352* [107]; *352* [108]; *356* [234]; *356* [235]; *357* [236]
Wysocki, K. 115, *144* [50]

Xiao, D. 301, *357* [237]
Xu, J.-M. 489, *554* [1]
Xu, X. 118, 135, 138, *145* [70]; *146* [103]

Yang, B. 542, *554* [18]
Yorke, J.A. 151–154, 217, 220, *238* [108]; *238* [109]
Yoshizawa, T. 216, *238* [110]
Yosida, S. 47, *75* [37]
Yu, J. 516, *556* [61]; *556* [62]
Yu, S. 265, 319, *351* [90]

Zabczyk, J. 3, 67, *75* [38]
Zabreiko, P.P. 314, *352* [101]
Zanolin, F. 332, *357* [238]; 439, *460* [36]
Zaremba, S.C. 227, *238* [111]
Zeeman, E.C. 296, 301, *357* [239]; *357* [240]; *357* [241]; *357* [242]
Zeeman, M.L. 296, 301, *356* [223]; *357* [240]; *357* [241]; *357* [242]; *357* [243]
Zehnder, E. 80, 104, 106, 113, 115, 121, 124, 125, *142* [3]; *142* [4]; *143* [24]; *144* [50]; *144* [51]; *144* [52]
Zeidler, E. 274, *357* [244]
Zhang, M. 482, *557* [93]
Zhang, Z. 533, *557* [94]; *557* [95]
Zhao, X.-Q. 314, 332, *357* [245]; *357* [246]
Zheng, D. 553, *557* [91]
Zhu, C. 115, *144* [67]; *145* [71]
Zhu, H.-R. 302, *357* [247]; *357* [248]
Ziehe, M. 315, *355* [181]
Zou, W. 121, 135, *145* [93]; *145* [94]

Subject Index

τ-topology, 95

accessible point, 244
adapted space, 335
admissible pair, 9
Arnold conjecture, 125
asymptotically order stable, 262
attainable set, 55
attract, 245
attractor, 245
axiom
– of choice, 156
– of countable choice, 156
– of dependent choice, 156

Banach lattice, 338
basin, 258
Bernoulli shift, 140
Bochner integral, 21
Bolza problem, 9, 30, 40, 56
Brezis–Browder ordering principle, 157

C_0-semigroup, 229
Carathéodory function, 471
Cartan formula, 24
Cerami sequence, 96
chain recurrent, 293
chronological exponential
– left, 22
– right, 19
classical solution, 468
closed loop system, 62
closed order interval, 244
cocycle identities, 333
coexistence state, 361, 401
compact map, 335
comparison property
– Carathéodory, 198
– of a function, 180
– of a multifunction, 212
competing species, 379, 438
competitive map, 329
completely continuous map, 266
component of coexistence states, 404
component of positive solutions, 394
concave functional, 536
conditionally completely continuous map, 266
cone, 536
– conjugate, 163
– contingent, 160
– in the sense of Bony, 162
– normal, 11, 15
– proximal normal, 161
– tangent, 15
conjugate mapping, 141
conjugate point, 8, 67
Conley–Zehnder index, 106
contingent derivative, 227
contraction, 325
control, 9
– bang-bang, 9, 47
– closed loop, 62
– constraints, 9
– feedback, 62
– open loop, 62
– optimal, 9, 55, 62
convergent, 246, 317
cooperative map, 329
cost functional, 9
critical orbit, 88
crossing number, 393
curve of change of stability, 376–378
curve of neutral stability, 376, 377

deformation lemma, 81
derivation, 18
derivative
– contingent, 227
– Dini right lower, 175
– right directional
– – of the norm, 162
diameter, 258
directional derivative, 11, 14
domain, 149
doubly accessible point, 257
dual cone, 281

eigenvalue, 393
Ekeland variational principle, 15
entire orbit, 320
equation
– Bellman, 63
– dynamic programming, 63
– eikonal, 72
– Euler–Lagrange, 5
– generalized Bellman, 224
– Hamilton, 6, 27
– Hamilton–Jacobi, 7, 63
– semilinear differential, 230
equilibrium, 245, 316
equivariant mapping, 87
escape time, 334
eventually strongly monotone, 247
eventually strongly monotone map, 316
extremal, 5

falling interval, 249, 317
five functional fixed point theorem, 548
fixed point, 316
– index, 393
– index theorem, 525
– set, 87
Floquet multipliers, 375, 377, 415
flow, 17, 245
– Hamiltonian, 27
function
– Carathéodory comparison, 197
– comparison, 154
– conjugate, 12
– convex, 10
– effective domain, 10
– epigraph, 10
– integrable, 21
– locally Lipschitz, 14
– lower semicontinuous, 10
– of Carathéodory type, 190
– positively sublinear, 173
– proper, 10
– regulated, 229
– Severi differentiable, 222
– strongly measurable, 21
– synthesis, 63
– weakly measurable, 21
functional ordering, 245
fundamental matrix, 414

generalized gradient, 14
genus, 86
geometrically distinct homoclinic solutions, 128
geometrically distinct periodic solutions, 100
global attractor, 246
global process, 333
group representation, 87

Hamiltonian, 6, 27
Hamiltonian lift, 28
high intensity competition, 384
homoclinic solutions, 126
– geometrically distinct, 128
hyperbolic point solution, 126
hysteresis point, 431

index, 88
infimum, 244
infinitesimal generator, 229
invariant, 245
– forward, 152
– local, 152
– locally right, 152
– mapping, 87
– right, 152
– set, 87
invariantly connected set, 316
isometric group representation, 87

k-bump solution, 138
K-competitive, 289
K-cooperative, 285
K-irreducible matrix, 288
K-positive matrix, 284
Krasnoselskii’s fixed point theorem, 519
Krein–Rutman operator, 274

Lagrange problem, 9, 56
Lebesgue point, 34
Legendre transform, 6, 12
Legget–Williams fixed point theorem, 536
lemma of Cârjă–Ursescu, 177, 210
limit genus, 92
limit index, 92
Liouville formula, 414
Lipschitz system, 292
local linking, 83
local semiflow, 251
locally closed, 335
locally closed set, 86
locally Hölder map, 333
locally Lipschitz map, 333
locally monotone map, 290
locally strongly monotone map, 290
logistic equation, 364
Lotka–Volterra periodic system, 362
low intensity competition, 384
lower boundary, 244

$\mathcal{M}$-maximal element, 157
m-point BVP, 470
mapping
– positively sublinear, 208
Maslov index, 106
maximal element, 244
Mayer problem, 9, 56
mild solution, 334, 468
minimal element, 244
minimal set, 269
monotone map, 316
monotone system, 282
multi-point BVP, 470
multibump solution, 138
multifunction, 164
– l.s.c. at a point, 200
– l.s.c. on a set, 200
– u.s.c. at a point, 200
– u.s.c. on a set, 200

necessary conditions, 5, 29
– Jacobi, 8
– Legendre, 6
– Weierstrass–Erdmann, 6
negative semiorbit, 289
non-cooperative periodic system, 456
nonexpansive map, 325
nonlocal BVP, 470
nonlocal condition, 463
nonlocal initial value problem, 465
nonresonance, 489
nonwandering set, 319
normality constant, 259
normally ordered space, 259

omega limit set, 246
open order interval, 244
operator pencil, 393
optimal control problem, 9
optimal pair, 9
optimal state, 9
orbit, 87, 245, 333
– of a set, 266
order bounded, 244
order compact, 322
order neighborhood, 261
order norm, 261
order related, 244
order stable, 262
order topology, 261
ordered Banach space, 245
ordered space, 243
ordered subspace, 244

p-convex, 245
Palais–Smale condition, 81
Palais–Smale sequence, 81
period, 245
periodic, 245, 316
periodic solutions
– geometrically distinct, 100
Poincaré map, 377, 414, 417, 441
point dissipative, 329
Poisson bracket, 27
polyhedral, 281
positive operator, 316
positive semiorbit, 289
positive solution, 519
positively invariant, 245, 333
positivity properties, 364
predator–prey model, 386, 455
preorder, 176
– closed, 176
principal characteristic exponent, 428, 449
principal eigenvalue, 417
problem
– abnormal, 30
– normal, 30
– of local invariance, 152
process, 333
projection
– subordinated to $\mathbb{V}$, 158
proposition
– Roxin, 174
proximal neighborhood, 158
(PS)-attractor, 96
$(\mathrm{PS})_c$-condition, 81
$(\mathrm{PS})^*$-condition, 91
$(\mathrm{PS})^*$-sequence, 91
pseudo-gradient vector field, 82
pseudo-symmetric function, 547

quasi-cooperative system, 440
quasiconvergent, 246, 317
quasiderivative map, 273
quasimonotone condition, 282
quasimonotone condition (QMD), 304

reachable set, 55
reaction–diffusion systems, 339
recurrence, 293
regularization, 13
relation
– preorder, 156
relative Morse index, 105, 106
representation
– diffeomorphisms, 18

– points, 18
– tangent vectors, 18
– vector fields, 18
reproducing cone, 322
resolvent operator, 366
resonance, 489
rising interval, 249, 317

sectorial, 333
semi-trivial positive state, 361
semiconjugate mapping, 141
semiflow, 245
semiorbit, 251
set
– reachable, 224
– target, 224
Severi differential
– at x, 222
– at x in the direction u, 222
singular value of an operator pencil, 393
solution
– ε-approximate, 168, 206
– Carathéodory, 190
– contingent, 227, 233
– flow, 245
– global, 172
– mild, or C^0, 229
– noncontinuable, 172, 207
– of a differential inclusion, 202
– of inequality, 224
– process, 335
– right funnel, 219
– semigroup, 229
– to a differential equation, 149
spectral radius, 274
stable fixed point, 320
stable from above, 258
stable from below, 258
standard cone, 245
state of the system, 9
strictly monotone, 246
strictly monotone map, 316
strong approximation, 265
strong maximum principle, 393
strong ordering, 244
strongly accessible, 265
strongly increasing, 370
strongly monotone, 246
strongly negative, 366
strongly order-preserving, 246
strongly order-preserving (SOP) map, 316
strongly positive, 366
strongly positive operator, 316
strongly sublinear map, 324
sub-critical turning point, 431
subdifferential, 11
subgradient, 11
sublinear map, 323
subset
– Carathéodory invariant, 197
– Carathéodory locally invariant, 197
– closed relative to $\mathbb{D}$, 151
– flow invariant, 153
– invariant, 178, 210
– locally closed, 151, 158
– locally invariant, 178, 210
– right viable with respect to f, 149
– right viable with respect to F, 202
– viable, 166
subsolution, 367, 368
super-critical turning point, 431
superlinear indefinite model, 388
supersolution, 367, 368
supremum, 244
symbiotic species, 387, 411
synthesis problem, 63
system
– strongly decreasing, 226
– weakly decreasing, 226

t-cross-section of $F_{\tau,\xi}$, 219
theorem
– Bony, 153
– Brezis–Browder, 157
– Hukuhara, 220
– Kneser, 221
– Nagumo, 151, 166
– Pavel, 231
– Peano, 149
– Redheffer, 154
– Scorza Dragoni, 190
– Shi, 232
– Vitali, 234
– Yorke, 217, 220
– Yoshizawa, 216
topological degree, 393
topologically equivalent, 292
totally competitive, 294
totally ordered arc, 255
trajectory, 245, 333
transversality conditions, 7, 30, 61
trivial state, 361

uniform global attractor, 321
unilateral bifurcation theorem, 397
unordered, 244
unstable from above, 268

unstable from below, 269
upper boundary, 244

variation of constants formula, 24
variation of parameters method, 24
vector
– metric normal to $\mathbb{K}$ at ξ, 153, 161
– tangent in the sense of
– – Bony, 162
– – Bouligand–Severi, 159
– – Clarke, 163
– – Federer, 161
vector field
– complete, 17
– Hamiltonian, 27
– nonautonomous, 16
vector ordering, 245
very strongly order preserving (VSOP) process, 336
viability, 149

Yosida approximation, 13

This page intentionally left blank

HANDBOOK
OF DIFFERENTIAL EQUATIONS

ORDINARY DIFFERENTIAL EQUATIONS

VOLUME III

Contents

Preface v
List of Contributors vii
Contents of Volume 1 xi
Contents of Volume 2 xiii

1. Topological principles for ordinary differential equations 1
J. Andres
2. Heteroclinic orbits for some classes of second and fourth order differential equations 103
D. Bonheure and L. Sanchez
3. A qualitative analysis, via lower and upper solutions, of first order periodic evolutionary equations with lack of uniqueness 203
C. De Coster, F. Obersnel and P. Omari
4. Bifurcation theory of limit cycles of planar systems 341
M. Han
5. Functional differential equations with state-dependent delays: Theory and applications 435
F. Hartung, T. Krisztin, H.-O. Walther and J. Wu
6. Global solution branches and exact multiplicity of solutions for two point boundary value problems 547
P. Korman
7. Singularities and Laplacians in boundary value problems for nonlinear ordinary differential equations 607
I. Rachůnková, S. Staněk and M. Tvrdý

Author index 725
Subject index 735

Contents of Volume 1

Preface v
List of Contributors vii

1. A survey of recent results for initial and boundary value problems singular in the dependent variable 1
R.P. Agarwal and D. O'Regan
2. The lower and upper solutions method for boundary value problems 69
C. De Coster and P. Habets
3. Half-linear differential equations 161
O. Došlý
4. Radial solutions of quasilinear elliptic differential equations 359
J. Jacobsen and K. Schmitt
5. Integrability of polynomial differential systems 437
J. Llibre
6. Global results for the forced pendulum equation 533
J. Mawhin
7. Ważewski method and Conley index 591
R. Srzednicki

Author index 685
Subject index 693

Contents of Volume 2

Preface v
List of Contributors vii
Contents of Volume 1 xi

1. Optimal control of ordinary differential equations 1
   V. Barbu and C. Lefter
2. Hamiltonian systems: periodic and homoclinic solutions by variational methods 77
   T. Bartsch and A. Szulkin
3. Differential equations on closed sets 147
   O. Cârjă and I.I. Vrabie
4. Monotone dynamical systems 239
   M.W. Hirsch and H. Smith
5. Planar periodic systems of population dynamics 359
   J. López-Gómez
6. Nonlocal initial and boundary value problems: a survey 461
   S.K. Ntouyas

Author index 559
Subject index 565

CHAPTER 1

Topological Principles for Ordinary Differential Equations

Jan Andres*

Department of Mathematical Analysis, Faculty of Science, Palacký University, Tomkova 40,
779 00 Olomouc-Hejčín, Czech Republic
E-mail: andres@inf.upol.cz

Contents

1. Introduction 3
2. Preliminaries 8
    2.1. Elements of ANR-spaces 8
    2.2. Elements of multivalued maps 10
    2.3. Some further preliminaries 13
3. Applied fixed point principles 15
    3.1. Lefschetz fixed point theorems 15
    3.2. Nielsen fixed point theorems 18
    3.3. Fixed point index theorems 23
4. General methods for solvability of boundary value problems 29
    4.1. Continuation principles to boundary value problems 29
    4.2. Topological structure of solution sets 45
    4.3. Poincaré’s operator approach 59
5. Existence results 61
    5.1. Existence of bounded solutions 61
    5.2. Solvability of boundary value problems with linear conditions 70
    5.3. Existence of periodic and anti-periodic solutions 73
6. Multiplicity results 76
    6.1. Several solutions of initial value problems 76
    6.2. Several periodic and bounded solutions 79
    6.3. Several anti-periodic solutions 93
7. Remarks and comments 95
    7.1. Remarks and comments to general methods 95
    7.2. Remarks and comments to existence results 96
    7.3. Remarks and comments to multiplicity results 97
Acknowledgements 97
References 97

---

[*] Supported by the Council of Czech Government (MSM 6198959214).

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 3
Edited by A. Cañada, P. Drábek and A. Fonda
© 2006 Elsevier B.V. All rights reserved

1. Introduction

The classical courses of ordinary differential equations (ODEs) start either with the Peano existence theorem (see, e.g., [54]) or with the Picard–Lindelöf existence and uniqueness theorem (see, e.g., [71]), both related to the Cauchy (initial value) problems

$$\begin{cases} \dot{x} = f(t, x), \\ x(0) = x_0, \end{cases} \tag{1.1}$$

where $f \in C([0, \tau] \times \mathbb{R}^n, \mathbb{R}^n)$, and

$$\bigl|f(t, x) - f(t, y)\bigr| \leqslant L|x - y|, \quad \text{for all } t \in [0, \tau] \text{ and } x, y \in \mathbb{R}^n, \tag{1.2}$$

in the latter case.

In fact, if f satisfies the Lipschitz condition (1.2), then "*uniqueness implies existence*" even for boundary value problems with linear conditions that are "close" to $x(0) = x_0$, as observed in [53]. Moreover, *uniqueness implies* in general (i.e. not necessarily, under (1.2)) *continuous dependence of solutions on initial values* (see, e.g., [54, Theorem 4.1 in Chapter 4.2]), and subsequently the *Poincaré translation operator* $T_\tau : \mathbb{R}^n \to \mathbb{R}^n$, at the time $\tau > 0$, along the trajectories of $\dot{x} = f(t, x)$, defined as follows:

$$T_\tau(x_0) := \bigl\{x(\tau) \mid x(.) \text{ is a solution of } (1.1)\bigr\}, \tag{1.3}$$

is a homeomorphism (cf. [54, Theorem 4.4 in Chapter 4.2]).

Hence, besides the existence, uniqueness is also a very important problem. W. Orlicz [92] showed in 1932 that the set of continuous functions $f : U \to \mathbb{R}^n$, where U is an open subset relative to $[0, \tau] \times \mathbb{R}^n$, for which problem (1.1) with $(0, x_0) \in U$ is not uniquely solvable, is meager, i.e. a set of the first Baire category. In other words, the generic continuous Cauchy problems (1.1) are solvable in a unique way. Therefore, no wonder that the first example of nonuniqueness was constructed only in 1925 by M.A. Lavrentev (cf. [71] and, for more information, see, e.g., [1]). The same is certainly also true for Carathéodory ODEs, because the notion of a *classical* (C^1-) *solution* can be just replaced by the *Carathéodory solution*, i.e. absolutely continuous functions satisfying (1.1), almost everywhere (a.e.). The change is related to the application of the Lebesgue integral, instead of the Riemann integral.

On the other hand, H. Kneser [80] proved in 1923 that the sets of solutions to continuous Cauchy problems (1.1) are, at every time, continua (i.e. compact and connected). This result was later improved by M. Hukuhara [75] who proved that the solution set itself is a continuum in $C([0, \tau], \mathbb{R}^n)$. N. Aronszajn [41] specified in 1942 that these continua are R_δ-sets (see Definition 2.3 below), and as a subsequence, multivalued operators T_τ in (1.3) become *admissible* in the sense of L. Górniewicz (see Definition 2.5 below).

Obviously if, for $f(t, x) \equiv f(t + \tau, x)$, operator T_τ admits a *fixed point*, say $\hat{x} \in \mathbb{R}^n$, i.e. $\hat{x} \in T_\tau(\hat{x})$, then $\hat{x}$ determines a τ-periodic solution of $\dot{x} = f(t, x)$, and vice versa. This is one of stimulations why to study the fixed point theory for multivalued mappings in order to obtain periodic solutions of nonuniquely solvable ODEs. Since the regularity of (multivalued) Poincaré's operator T_τ is the same (see Theorem 4.17 below) for differential inclusions $\dot{x} \in F(t,x)$, where F is an upper Carathéodory mapping with nonempty, convex and compact values (see Definition 2.10 below), it is reasonable to study directly such differential inclusions with this respect. Moreover, initial value problems for differential inclusions are, unlike ODEs, typically nonuniquely solvable (cf. [42]) by which Poincaré's operators are multivalued.

In this context, an interesting phenomenon occurs with respect to the Sharkovskii cycle coexistence theorem [95]. This theorem is based on a new ordering of the positive integers, namely

$$\begin{aligned} &3 \rhd 5 \rhd 7 \rhd \cdots \rhd 2\cdot 3 \rhd 2\cdot 5 \rhd 2\cdot 7 \rhd \cdots \rhd 2^2\cdot 3 \rhd 2^2\cdot 5 \rhd 2^2\cdot 7 \rhd \cdots \\ &\quad \rhd 2^n\cdot 3 \rhd 2^n\cdot 5 \rhd 2^n\cdot 7 \rhd \cdots \rhd 2^{n+1}\cdot 3 \rhd 2^{n+1}\cdot 5 \rhd 2^{n+1}\cdot 7 \rhd \cdots \\ &\quad \rhd 2^{n+1} \rhd 2^n \rhd \cdots \rhd 2^2 \rhd 2 \rhd 1, \end{aligned}$$

saying that *if a continuous function* $g : \mathbb{R} \to \mathbb{R}$ *has a point of period* m *with* $m \rhd k$ (in the above Sharkovskii ordering), *then it has also a point of period* k.

By a period, we mean the least period, i.e. a point $a \in \mathbb{R}$ is a *periodic point of period* m if $g^m(a) = a$ and $g^j(a) \neq a$, for $0 < j < m$.

Now, consider the scalar ODE

$$\dot{x} = f(t,x), \quad f(t,x) \equiv f(t+\tau, x), \tag{1.4}$$

where $f : [0,\tau] \times \mathbb{R} \to \mathbb{R}$ is a continuous function.

Since

$$T_\tau^m = \underbrace{T_\tau \circ \cdots \circ T_\tau}_{m \text{ times}} = T_{m\tau}$$

holds for the Poincaré translation operator T_τ along the trajectories of Eq. (1.4), defined in (1.3), there is (in the case of uniqueness) an apparent one-to-one correspondence between m-periodic points of T_τ and (subharmonic) $m\tau$-periodic solutions of (1.4). Nevertheless, the analogy of classical Sharkovskii's theorem does not hold for subharmonics of (1.4). In fact, we only obtain an empty statement, because every bounded solution of (1.4) is, under the uniqueness assumption, either τ-periodic or asymptotically τ-periodic (see, e.g., [94, pp. 120–122]).

This handicap is due to the assumed uniqueness condition. On the other hand, in the lack of uniqueness, the multivalued operator T_τ in (1.3) is admissible (see Theorem 4.17 below) which in $\mathbb{R}$ means (cf. Definition 2.5 below) that T_τ is upper semicontinuous (cf. Definition 2.4 below) and the sets of values consist either of single points or of compact intervals. In a series of our papers [16,29,36], we developed a version of the Sharkovskii cycle coexistence theorem which applies to (1.4) as follows:

THEOREM 1.1. *If Eq.* (1.4) *has an* $m\tau$*-periodic solution, then it also admits a* $k\tau$*-periodic solution, for every* $k \lhd m$, *with at most two exceptions, where* $k \lhd m$ *means that* k *is less*

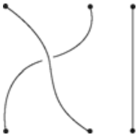


Fig. 1. Braid σ.

than m in the above Sharkovskii ordering of positive integers. In particular, if $m \neq 2^k$, *for all* $k \in \mathbb{N}$, *then infinitely many* (*subharmonic*) *periodic solutions of* (1.4) *coexist.*

REMARK 1.1. As pointed out, Theorem 1.1 holds only in the lack of uniqueness; otherwise, it is empty. On the other hand, the right-hand side of the given (multivalued) ODE can be a (multivalued upper) Carathéodory mapping with nonempty, convex and compact values (see Definition 2.10 below).

REMARK 1.2. Although, e.g., a 3τ-periodic solution of (1.4) implies, for every $k \in \mathbb{N}$, with a possible exception for $k = 2$ or $k = 4, 6$, the existence of a $k\tau$-periodic solution of (1.4), it is very difficult to prove that such a solution exists. Observe that a 3τ-periodic solution of (1.4) implies the existence of at least two more 3τ-periodic solutions of (1.4).

The Sharkovskii phenomenon is essentially one-dimensional. On the other hand, it follows from T. Matsuoka's results in [87–89] that three (harmonic) τ-periodic solutions of the planar (i.e. in $\mathbb{R}^2$) system (1.4) imply "generically" the coexistence of infinitely many (subharmonic) $k\tau$-periodic solutions of (1.4), $k \in \mathbb{N}$. "Genericity" is this time understood in terms of the Artin braid group theory, i.e. with the exception of certain simplest braids, representing the three given harmonics.

The following theorem was presented in [8], on the basis of T. Matsuoka's results in papers [87–89].

THEOREM 1.2. *Assume that a uniqueness condition is satisfied for planar system* (1.4). *Let three* (*harmonic*) τ*-periodic solutions of* (1.4) *exist whose graphs are not conjugated to the braid* σ^m *in* B_3/Z, *for any integer* $m \in \mathbb{N}$, *where* σ *is shown in Fig.* 1, B_3/Z *denotes the factor group of the Artin braid group* B_3 *and* Z *is its center* (*for definitions*, *see*, *e.g.*, [22, Chapter III.9]). *Then there exist infinitely many* (*subharmonic*) $k\tau$*-periodic solutions of* (1.4), $k \in \mathbb{N}$.

REMARK 1.3. In the absence of uniqueness, there occur serious obstructions, but Theorem 1.2 still seems to hold in many situations; for more details see [8].

REMARK 1.4. The application of the Nielsen theory considered in Section 3.2 below might determine the desired three harmonic solutions of (1.4). More precisely, it is more realistic to detect two harmonics by means of the related Nielsen number (see again Section 3.2 below), and the third one by means of the related fixed point index (see Section 3.3 below).

For $n > 2$, statements like Theorem 1.1 or Theorem 1.2 appear only rarely. Nevertheless, if $f = (f_1, f_2, \dots, f_n)$ has a special triangular structure, i.e.

$$f_i(x) = f_i(x_1, \dots, x_n) = f_i(x_1, \dots, x_i), \quad i = 1, \dots, n, \tag{1.5}$$

then Theorem 1.1 can be extended to hold in $\mathbb{R}^n$ (see [35]).

THEOREM 1.3. *Under assumption* (1.5), *the conclusion of Theorem* 1.1 *remains valid in* $\mathbb{R}^n$.

REMARK 1.5. Similarly to Theorem 1.1, Theorem 1.3 holds only in the lack of uniqueness. Without the special triangular structure (1.5), there is practically no chance to obtain an analogy to Theorem 1.1, for $n \geqslant 2$.

There is also another motivation for the investigation of *multivalued ODEs*, i.e. differential inclusions, because of the strict connection with

(i) optimal control problems for ODEs,
(ii) Filippov solutions of discontinuous ODEs,
(iii) implicit ODEs, etc.

ad (i): Consider a *control problem* for

$$\dot{x} = f(t, x, u), \quad u \in U, \tag{1.6}$$

where $f : [0, \tau] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ and $u \in U$ are control parameters such that $u(t) \in \mathbb{R}^n$, for all $t \in [0, \tau]$. In order to solve a control problem for (1.6), we can define a multivalued map $F(t, x) := \{f(t, x, u)\}_{u \in U}$. The solutions of (1.6) are those of

$$\dot{x} \in F(t, x), \tag{1.7}$$

and the same is true for a given control problem. For more details, see, e.g., [27,79].

ad (ii): If function f is discontinuous in x, then Carathéodory theory cannot be applied for solving, e.g., (1.1). Making, however, the *Filippov regularization* of f, namely

$$F(t, x) := \bigcap_{\delta > 0} \bigcap_{\substack{r \subset [0,\tau] \times \mathbb{R}^n \\ \mu(r) = 0}} \overline{\operatorname{conv}}\, f\big(O_\delta\big((t, x) \setminus r\big)\big), \tag{1.8}$$

where $\mu(r)$ denotes the Lebesgue measure of the set $r \subset \mathbb{R}^n$ and

$$O_\delta(y) := \big\{z \in [0, \tau] \times \mathbb{R}^n \mid |y - z| < \delta\big\},$$

multivalued F is well known (see [60]) to be again upper Carathéodory with nonempty, convex and compact values (cf. Definition 2.10 below), provided only f is measurable and satisfies $|f(t, x)| \leqslant \alpha + \beta|x|$, for all $(t, x) \in [0, \tau] \times \mathbb{R}^n$, with some nonnegative constants α, β. Thus, by a *Filippov solution* of $\dot{x} = f(t, x)$, it is so understood a Carathéodory solution of (1.7), where F is defined in (1.8). As an example from physics, dry friction problems (see, e.g., [84,91]) can be solved in this way.

ad (iii): Let us consider the *implicit differential equation*

$$\dot{x} = f(t, x, \dot{x}), \tag{1.9}$$

where $f : [0, \tau] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is a compact (continuous) map and the solutions are understood in the sense of Carathéodory. We can associate with (1.9) the following two differential inclusions:

$$\dot{x} \in F_1(t, x) \tag{1.10}$$

and

$$\dot{x} \in F_2(t, x), \tag{1.11}$$

where $F_1(t, x) := \operatorname{Fix}(f(t, x, \cdot))$, i.e. the (nonempty, see [22, p. 560]) fixed point set of $f(t, x, \cdot)$ w.r.t. the last variable, and $F_2 \subset F_1$ is a (multivalued) lower semicontinuous (see Definition 2.4 below) selection of F_1. The sufficient condition for the existence of such a selection F_2 reads (see, e.g., [22, Chapter III.11, pp. 558–559]):

$$\dim \operatorname{Fix}\bigl(f(t, x, \cdot)\bigr) = 0, \quad \text{for all } (t, x) \in [0, \tau] \times \mathbb{R}^n, \tag{1.12}$$

where dim denotes the topological (covering) dimension.

Denoting by $S(f)$, $S(F_1)$, $S(F_2)$ the sets of all solutions of initial value problems to (1.9), (1.10), (1.11), respectively, one can prove (see [22, p. 560]) that, under (1.12), $S(f) = S(F_1) \subset S(F_2) \neq \emptyset$. For more details, see [19] (cf. [22, Chapter III.11]).

Although there are several monographs devoted to multivalued ODEs (see, e.g., [22,42, 45,58,61,74,79,91,96,97]), topological principles were presented mainly for single-valued ODEs (besides [22,45,58] and [61] for differential inclusions, see, e.g., [62,64,65,82,83, 90]). Hence our main object will be topological principles for (multivalued) ODEs; whence the title. We will consider without special distinguishing differential equations as well as inclusions; both in Euclidean and Banach spaces. All solutions of problems under our consideration (even in Banach spaces) will be understood at least in the sense of Carathéodory. Thus, in view of the indicated relationship with problems (i)–(iii), many obtained results can be also employed for solving optimal control problems, problems for systems with variable structure, implicit boundary value problems, etc.

The reader exclusively interested in single-valued ODEs can simply read “continuous”, instead of “upper semicontinuous” or “lower semicontinuous”, and replace the inclusion symbol $\in$ by the equality $=$, in the given differential inclusions. If, in the single-valued case, the situation simplifies dramatically or if the obtained results can be significantly improved, then the appropriate remarks are still supplied.

We wished to prepare an as much as possible self-contained text. Nevertheless, the reader should be at least familiar with the elements of nonlinear analysis, in particular of fixed point theory, in order to understand the degree arguments, or so. Otherwise, we recommend the monographs [69] (in the single-valued case) and [22] (in the multivalued case).

Furthermore, one is also expected to know several classical results and notions from the standard courses of ODEs, functional analysis and the theory of integration like the Gronwall inequality, the Arzelà–Ascoli lemma, the Mazur Theorem, the Bochner integral, etc.

We will study mainly existence and multiplicity of bounded, periodic and anti-periodic solutions of (multivalued) ODEs. Since our approach consists in the application of the fixed point principles, these solutions will be either determined by, (e.g., τ-periodic solutions $x(t)$ by the initial values $x(0)$ via (1.3)) or directly identified (e.g., solutions of initial value problems (1.1)) with fixed points of the associated (Cauchy, Hammerstein, etc.) operators.

Although the usage of the relative degree (i.e. the fixed point index) arguments is rather traditional in this framework, it might not be so when the maps, representing, e.g., problems on noncompact intervals, operate in nonnormable Fréchet spaces. This is due to the unpleasant locally convex topology possessing bounded subsets with an empty interior. We had therefore to develop with my colleagues our own fixed point index theory. The application of the Nielsen theory, for obtaining multiplicity criteria, is very delicate and quite rare, and the related problem is named after Jean Leray who posed it in 1950, at the first International Congress of Mathematics held after World War II in Cambridge, Mass. We had also to develop a new multivalued Nielsen theory suitable for applications in this field. Before presenting general methods for solvability of boundary value problems in Section 4, we therefore make a sketch of the applied fixed point principles in Section 3. Hence besides Section 4, the main results are contained in Section 5 (Existence results) and Section 6 (Multiplicity results). The reference sources to our results and their comparison with those of other authors are finally commented in Section 7 (Remarks and comments).

2. Preliminaries

2.1. *Elements of ANR-spaces*

In the entire text, all topological spaces will be metric and, in particular, all topological vector spaces will be at least Fréchet. Let us recall that by a *Fréchet space*, we understand a complete (metrizable) locally convex space. Its topology can be generated by a countable family of seminorms. If it is normable, then it becomes *Banach*.

DEFINITION 2.1. A (metrizable) space X is an *absolute neighbourhood retract* (ANR) if, for each (metrizable) Y and every closed $A \subset Y$, each continuous mapping $f : A \to X$ is extendable over some neighbourhood of A.

PROPOSITION 2.1.

(i) *If X is an ANR, then any open subset of X is an ANR and any neighbourhood retract of X is an ANR.*

(ii) *X is an ANR if and only if it is a neighbourhood retract of every* (*metrizable*) *space in which it is embedded as a closed subset.*

(iii) *X is an ANR if and only if it is a neighbourhood retract of some normed linear space, i.e. if and only if it is a retract of some open subset of a normed space.*

(iv) *If X is a retract of an open subset of a convex set in a Fréchet space, then it is an ANR.*
(v) *If X_1, X_2 are closed ANRs such that $X_1 \cap X_2$ is an ANR, then $X_1 \cup X_2$ is an ANR.*
(vi) *Any finite union of closed convex sets in a Fréchet space is an ANR.*
(vii) *If each $x \in X$ admits a neighbourhood that is an ANR, then X is an ANR.*

DEFINITION 2.2. A (metrizable) space X is an *absolute retract* (AR) if, for each (metrizable) Y and every closed $A \subset Y$, each continuous mapping $f : A \to X$ is extendable over Y.

PROPOSITION 2.2.
(i) *X is an AR if and only if it is a contractible (i.e. homotopically equivalent to a one point space) ANR.*
(ii) *X is an AR if and only if it is a retract of every (metrizable) space in which it is embedded as a closed subset.*
(iii) *If X is an AR and A is a retract of X, then A is an AR.*
(iv) *If X is homeomorphic to Y and X is an AR, then so is Y.*
(v) *X is an AR if and only if it is a retract of some normed space.*
(vi) *If X is a retract of a convex subset of a Fréchet space, then it is an AR.*
(vii) *If X_1, X_2 are closed ARs such that $X_1 \cap X_2$ is an AR, then $X_1 \cup X_2$ is an AR.*

Furthermore, it is well known that every ANR X is *locally contractible* (i.e. for each $x \in X$ and a neighbourhood U of x, there exists a neighbourhood V of x that is contractible in U) and, as follows from Proposition 2.2(i) that every AR X is *contractible* (i.e. if $\mathrm{id}_X : X \to X$ is homotopic to a constant map).

DEFINITION 2.3. X is called an *R_δ-set* if, there exists a decreasing sequence $\{X_n\}$ of compact, contractible sets X_n such that $X = \bigcap\{X_n \mid n = 1, 2, \ldots\}$.

Although contractible spaces need not be ARs, X is an R_δ-set if and only if it is an intersection of a decreasing sequence of compacts ARs. Moreover, every R_δ-set is *acyclic* w.r.t. any continuous theory of homology (e.g., the Čech homology), i.e. homologically equivalent to a one point space, and so it is in particular nonempty, compact and connected.

The following hierarchies hold for metric spaces:

$$\begin{array}{ccccc} & & \text{contractible} & \subset & \text{acyclic} \\ & & \cup & & \\ \text{convex} & \subset & \text{AR} & \subset & \text{ANR,} \end{array}$$

compact + convex $\subset$ compact AR $\subset$ compact + contractible $\subset R_\delta \subset$ compact + acyclic, and all the above inclusions are proper.

For more details, see [47] (cf. also [22,67,69]).

2.2. *Elements of multivalued maps*

In what follows, by a multivalued map $\varphi : X \multimap Y$, i.e. $\varphi : X \to 2^Y \backslash \{0\}$, we mean the one with at least nonempty, closed values.

DEFINITION 2.4. A map $\varphi : X \multimap Y$ is said to be *upper semicontinuous* (u.s.c.) if, for every open $U \subset Y$, the set $\{x \in X \mid \varphi(x) \subset U\}$ is open in X. It is said to be *lower semicontinuous* (l.s.c.) if, for every open $U \subset Y$, the set $\{x \in X \mid \varphi(x) \cap U \neq \emptyset\}$ is open in X. If it is both u.s.c. and l.s.c., then it is called *continuous*.

Obviously, in the single-valued case, if $f : X \to Y$ is u.s.c. or l.s.c., then it is continuous. Moreover, the compact-valued map $\varphi : X \multimap Y$ is continuous if and only if it is *Hausdorff-continuous*, i.e. continuous w.r.t. the metric d in X and the Hausdorff-metric d_H in $\{B \subset Y \mid B$ is nonempty and bounded$\}$, where $d_H(A, B) := \inf\{\varepsilon > 0 \mid A \subset O_\varepsilon(B)$ and $B \subset O_\varepsilon(A)\}$ and $O_\varepsilon(B) := \{x \in X \mid \exists y \in B\colon\ d(x, y) < \varepsilon\}$. Every u.s.c. map $\varphi : X \multimap Y$ has a closed graph Γ_φ, but not vice versa. Nevertheless, if the graph Γ_φ of a compact map $\varphi : X \multimap Y$ is closed, then φ is u.s.c.

The important role will be played by the following class of admissible maps in the sense of L. Górniewicz.

DEFINITION 2.5. Assume that we have a diagram $X \overset{p}{\Longleftarrow} \Gamma \xrightarrow{q} Y$ (Γ is a metric space), where $p : \Gamma \Rightarrow X$ is a continuous *Vietoris map*, namely

(i) p is onto, i.e. $p(\Gamma) = X$,
(ii) p is proper, i.e. $p^{-1}(K)$ is compact, for every compact $K \subset X$,
(iii) $p^{-1}(x)$ is acyclic, for every $x \in X$, where acyclicity is understood in the sense of the Čech homology functor with compact carriers and coefficients in the field $\mathbb{Q}$ of rationals,

and $q : \Gamma \to Y$ is a continuous map. The map $\varphi : X \multimap Y$ is called *admissible* if it is induced by $\varphi(x) = q(p^{-1}(x))$, for every $x \in X$. We, therefore, identify the admissible map φ with the pair (p, q) called an *admissible* (*selected*) *pair*.

DEFINITION 2.6. Let $X \overset{p_0}{\Longleftarrow} \Gamma_0 \xrightarrow{q_0} Y$ and $X \overset{p_1}{\Longleftarrow} \Gamma_1 \xrightarrow{p_1} Y$ be two admissible maps, i.e. $\varphi_0 = q_0 \circ p_0^{-1}$ and $\varphi_1 = q_1 \circ p_1^{-1}$. We say that φ_0 is *admissibly homotopic* to φ_1 (written $\varphi_0 \sim \varphi_1$ or $(p_0, q_0) \sim (p_1, q_1)$) if there exists an admissible map $X \times [0, 1] \overset{p}{\Longleftarrow} \Gamma_0 \xrightarrow{q} Y$ such that the following diagram is commutative:

$$\begin{tikzcd}
X \arrow[d, "k_i"'] & \Gamma_i \arrow[l, Leftarrow, "p_i"'] \arrow[r, "q_i"] \arrow[d, "f_i"'] & Y \\
X \times [0,1] & \Gamma \arrow[l, Leftarrow, "p"] \arrow[ur, dashed, "q"'] &
\end{tikzcd}$$

for $k_i(x) = (x, i)$, $i = 0, 1$, and $f_i : \Gamma_i \to \Gamma$ is a homeomorphism onto $p^{-1}(X \times i)$, $i = 0, 1$, i.e. $k_0 p_0 = p f_0$, $q_0 = q f_0$, $k_1 p_1 = p f_1$ and $q_1 = q f_1$.

Thus, admissible maps are always u.s.c. with nonempty, compact and connected values. Moreover, their class is closed w.r.t. finite compositions, i.e. a finite composition of admissible maps is also admissible. In fact, a map is admissible if and only if it is a finite composition of *acyclic maps* with compact values, i.e. u.s.c. maps with acyclic and compact values.

The class of admissible maps so contains u.s.c. maps with convex and compact values, u.s.c. maps with contractible and compact values, R_δ*-maps* (i.e. u.s.c. maps with R_δ-values), acyclic maps with compact values and their compositions.

The class of compact admissible maps $\varphi : X \multimap Y$, i.e. $\overline{\varphi(X)}$ is compact, will be denoted by $\mathbb{K}(X, Y)$, or simply by $\mathbb{K}(X)$, provided φ is a self-map (an endomorphism). If the admissible homotopy in Definition 2.6 is still compact, then we say that $\varphi_0 \in \mathbb{K}(X, Y)$ and $\varphi_1 \in \mathbb{K}(X, Y)$ are *compactly admissibly homotopic*.

Another important class of admissible maps are condensing admissible maps denoted by $\mathbb{C}(X, Y)$. For this, we need to recall the notion of a measure of noncompactness (MNC).

Let E be a Fréchet space endowed with a countable family of seminorms $\|.\|_s$, $s \in S$ (S is the index set), generating the locally convex topology. Denoting by $\mathcal{B} = \mathcal{B}(E)$ the set of nonempty, bounded subsets of E, we can give

DEFINITION 2.7. The family of functions $\alpha = \{\alpha_s\}_{s\in S} : \mathcal{B} \to [0, \infty)^S$, where $\alpha_s(B) := \inf\{\delta > 0 \mid B \in \mathcal{B}$ admits a finite covering by the sets of $\operatorname{diam}_s \leqslant \delta\}$, $s \in S$, for $B \in \mathcal{B}$, is called the *Kuratowski measure of noncompactness* and the family of functions $\gamma = \{\gamma_s\}_{s\in S} : \mathcal{B} \to [0, \infty)^S$, where $\gamma_s(B) := \inf\{\delta > 0 \mid B \in \mathcal{B}$ has a finite ε_s-net$\}$, $s \in S$, for $B \in \mathcal{B}$, is called the *Hausdorff measure of noncompactness*.

These MNC are related as follows:

$$\gamma(B) \leqslant \alpha(B) \leqslant 2\gamma(B), \quad \text{i.e.} \quad \gamma_s(B) \leqslant \alpha_s(B) \leqslant 2\gamma_s(B), \quad \text{for each } s \in S.$$

Moreover, they satisfy the following properties:

PROPOSITION 2.3. *Assume that* $B, B_1, B_2 \in \mathcal{B}$. *Then we have* (*component-wise*):

(μ_1) (regularity) $\mu(B) = 0 \Leftrightarrow \overline{B}$ *is compact*,

(μ_2) (nonsingularity) $\{b\} \in \mathcal{B} \Rightarrow \{b\} \cup B \in \mathcal{B}$ *and* $\mu(\{b\} \cup B) = \mu(B)$,

(μ_3) (monotonicity) $B_1 \subset B_2 \Rightarrow \mu(B_1) \leqslant \mu(B_2)$,

(μ_4) (closed convex hull) $\mu(\overline{\operatorname{conv}} B) = \mu(B)$,

(μ_5) (closure) $\mu(\overline{B}) = \mu(B)$,

(μ_6) (Kuratowski condition) *decreasing sequence of closed sets* $B_n \in \mathcal{B}$ *with*

$$\lim_{n\to\infty} \mu(B_n) = 0 \quad \Longrightarrow \quad \bigcap\{B_n \mid n = 1, 2, \ldots\} \neq \emptyset,$$

(μ_7) (semiadditivity) $\mu(B_1 + B_2) \leqslant \mu(B_1) + \mu(B_2)$,

(μ_8) (union) $\mu(B_1 \cup B_2) = \max\{\mu(B_1), \mu(B_2)\}$,

(μ_9) (intersection) $\mu(B_1 \cap B_2) = \min\{\mu(B_1), \mu(B_2)\}$,

(μ_{10}) (seminorm) $\mu(\lambda B) = |\lambda|\mu(B)$, *for every* $\lambda \in \mathbb{R}$, *and* $\mu(B_1 \cup B_2) \leqslant \mu(B_1) + \mu(B_2)$, *where μ denotes either α or γ.*

DEFINITION 2.8. A bounded mapping $\varphi : E \supset U \multimap E$, i.e. $\varphi(B) \in \mathcal{B}$, for $\mathcal{B} \ni B \subset U$, is said to be *$\mu$-condensing* (shortly, *condensing*) if $\mu(\varphi(B)) < \mu(B)$, whenever $\mathcal{B} \ni B \subset U$ and $\mu(B) > 0$, or equivalently, if $\mu(\varphi(B)) \geqslant \mu(B)$ implies $\mu(B) = 0$, whenever $\mathcal{B} \ni B \subset U$, where $\mu = \{\mu_s\}_{s \in S} : \mathcal{B} \to [0, \infty)^S$ is a family of functions satisfying at least conditions (μ_1)–(μ_5). Analogously, a bounded mapping $\varphi : E \supset U \multimap E$ is said to be a *k-set contraction* w.r.t. $\mu = \{\mu_s\}_{s \in S} : \mathcal{B} \to [0, \infty)^S$ satisfying at least conditions (μ_1)–(μ_5) (shortly, a *k-contraction* or a *set-contraction*) if $\mu(\varphi(B)) \leqslant k\mu(B)$, for some $k \in [0, 1)$, whenever $\mathcal{B} \ni B \subset U$.

Obviously, any set-contraction is condensing and both α-condensing and γ-condensing maps are μ-condensing. Furthermore, compact maps or contractions with compact values (in vector spaces, also their sum) are well known to be (α, γ)-set-contractions, and so (α, γ)-condensing.

Besides semicontinuous maps, measurable and semi-Carathéodory maps will be also of importance. Hence, assume that Y is a separable metric space and $(\Omega, \mathcal{U}, \nu)$ is a *measurable space*, i.e. a set Ω equipped with σ-algebra $\mathcal{U}$ of subsets and a countably additive measure ν on $\mathcal{U}$. A typical example is when Ω is a bounded domain in $\mathbb{R}^n$, equipped with the Lebesgue measure.

DEFINITION 2.9. A map $\varphi : \Omega \multimap Y$ is called *strongly measurable* if there exists a sequence of step multivalued maps $\varphi_n : \Omega \multimap Y$ such that $d_H(\varphi_n(\omega), \varphi(\omega)) \to 0$, for a.a. $\omega \in \Omega$, as $n \to \infty$. In the single-valued case, one can simply replace multivalued step maps by single-valued step maps and $d_H(\varphi_n(\omega), \varphi(\omega))$ by $\|\varphi_n(\omega) - \varphi(\omega)\|$.

A map $\varphi : \Omega \multimap Y$ is called *measurable* if $\{\omega \in \Omega \mid \varphi(\omega) \subset V\} \in \mathcal{U}$, for each open $V \subset Y$.

A map $\varphi : \Omega \multimap Y$ is called *weakly measurable* if $\{\omega \in \Omega \mid \varphi(\omega) \subset V\} \in \mathcal{U}$, for each closed $V \subset Y$.

Obviously, if φ is strongly measurable, then it is measurable and if φ is measurable, then it is also weakly measurable. If φ has compact values, then the notions of measurability and weak measurability coincide. In separable Banach spaces Y, the notions of strong measurability and measurability coincide for multivalued maps with compact values as well as for single-valued maps (see [78, Theorem 1.3.1 on pp. 45–49]). If Y is a not necessarily separable Banach space, then a strongly measurable map $\varphi : \Omega \multimap Y$ with compact values has a single-valued strongly measurable selection (see, e.g., [58, Proposition 3.4(b) on pp. 25–26]). Furthermore, if Y is a separable complete space, then every measurable $\varphi : \Omega \multimap Y$ has, according to the Kuratowski–Ryll-Nardzewski theorem (see, e.g., [22, Theorem 3.49 in Chapter I.3]), a single-valued measurable selection.

Now, let $\Omega = [0, a]$ be equipped with the Lebesgue measure and X, Y be Banach.

DEFINITION 2.10. A map $\varphi : [0, a] \times X \multimap Y$ with nonempty, compact and convex values is called *u-Carathéodory* (resp. *l-Carathéodory*, resp. *Carathéodory*) if it satisfies

(i) $t \multimap \varphi(t, x)$ is strongly measurable, for every $x \in X$,

(ii) $x \multimap \varphi(t, x)$ is u.s.c. (resp. l.s.c., resp. continuous), for almost all $t \in [0, a]$,

(iii) $\|y\|_Y \leqslant r(t)(1+\|x\|_X)$, for every $(t,x) \in [0,a] \times X$, $y \in \varphi(t,x)$, where $r : [0,a] \to [0,\infty)$ is an integrable function.

For $X = \mathbb{R}^m$ and $Y = \mathbb{R}^n$, one can state

PROPOSITION 2.4.

(i) *Carathéodory maps are product-measurable* (*i.e. measurable as the whole* $(t,x) \multimap \varphi(t,x)$), *and*

(ii) *they possess a single-valued Carathéodory selection.*

It need not be so for u-Carathéodory or l-Carathéodory maps. Nevertheless, for u-Carathéodory maps, we have at least (again $X = \mathbb{R}^m$ and $Y = \mathbb{R}^n$).

PROPOSITION 2.5. *u-Carathéodory maps* (*in the sense of Definition* 2.10) *are weakly superpositionally measurable*, *i.e. the composition* $\varphi(t,q(t))$ *admits*, *for every* $q \in C([0,a],\mathbb{R}^m)$, *a single-valued measurable selection. If they are still product-measurable*, *then they are also superpositionally measurable*, *i.e. the composition* $\varphi(t,q(t))$ *is measurable*, *for every* $q \in C([0,a],\mathbb{R}^m)$.

REMARK 2.1. If X, Y are separable Banach spaces and $\varphi : X \multimap Y$ is a Carathéodory mapping, then φ is also superpositionally measurable, i.e. $\varphi(t,q(t))$ is measurable, for every $q \in C([0,a],X)$ (see [78, Theorem 1.3.4 on p. 56]). Under the same assumptions, Proposition 2.4 can be appropriately generalized (see [73, Proposition 7.9 on p. 229 and Proposition 7.23 on pp. 234–235]).

If $\varphi : X \multimap Y$ is only u-Carathéodory and X, Y are (not necessarily separable) Banach spaces, then φ is weakly superpositionally measurable, i.e. $\varphi(t,q(t))$ admits a single-valued measurable selection, for every $q \in C([0,a],X)$ (see, e.g., [58, Proposition 3.5 on pp. 26–27] or [78, Theorem 1.3.5 on pp. 57–58]).

For more details, see [22,40,58,67,73,78].

2.3. *Some further preliminaries*

Assume we have again a diagram (see Definition 2.5)

$$\begin{tikzcd} X & \Gamma \arrow[l, Rightarrow, "p"'] \arrow[r, "q"] & Y \end{tikzcd}$$

where $p : \Gamma \Longrightarrow X$ is a Vietoris map and $q : \Gamma \longrightarrow Y$ is continuous.

Taking $\varphi(x) = q(p^{-1}(x))$, for every $x \in X$, and denoting as

$$\mathrm{Fix}(p,q) = \mathrm{Fix}(\varphi) := \bigl\{x \in X \mid x \in \varphi(x)\bigr\},$$
$$C(p,q) := \bigl\{z \in \Gamma \mid p(z) = q(z)\bigr\}$$

the sets of *fixed points* and *coincidence points* of the admissible pair (p, q), it is clear that $p(C(p, q)) = \mathrm{Fix}(p, q)$, and so

$$\mathrm{Fix}(p, q) \neq \emptyset \quad \Longleftrightarrow \quad C(p, q) \neq \emptyset.$$

The following Aronszajn–Browder–Gupta-type result (see [21, Theorem 3.15]; cf. [22, Theorem 1.4]) is very important in order to say something about the topological structure of $\mathrm{Fix}(\varphi)$.

PROPOSITION 2.6. *Let X be a metric space, E a Fréchet space, $\{U_k\}$ a base of open convex symmetric neighbourhoods of the origin in E, and let $\varphi : X \multimap E$ be a u.s.c proper map with compact values. Assume that there is a convex symmetric subset C of E and a sequence of compact, convex-valued u.s.c. proper maps $\varphi_k : X \multimap E$ such that*

(i) $\varphi_k(x) \subset \varphi(O_{1/k}(x)) + U_k$, *for every $x \in X$, where*

$$O_{1/k}(x) = \left\{y \in X \mid d(x, y) < \tfrac{1}{k}\right\},$$

(ii) *for every $k \geqslant 1$, there is a convex, symmetric set $V_k \subset \overline{U}_k \cap C$ such that V_k is closed in E and $0 \in \varphi(x)$ implies $\varphi_k(x) \cap V_k \neq \emptyset$,*

(iii) *for every $k \geqslant 1$ and every $u \in V_k$, the inclusion $u \in \varphi_k(x)$ has an acyclic set of solutions.*

Then the set $S = \{x \in X \mid \varphi(x) \cap \{0\} \neq \emptyset\}$ is compact and acyclic.

Now, let us assume that E is a Fréchet space, C is a convex subset of E, U is an open subset of C, $\mu : \mathcal{B} \to [0, \infty)^S$ is a measure of noncompactness satisfying at least conditions (μ_1)–(μ_5) in Proposition 2.3 (see Definitions 2.7 and 2.8).

If $\varphi \in \mathbb{C}(U, C)$, then $\mathrm{Fix}(\varphi)$ can be proved relatively compact. We can say more about $\mathrm{Fix}(\varphi)$.

DEFINITION 2.11. Let $(p, q) \in \mathbb{C}(U, C)$. A nonempty, compact, convex set $S \subset C$ is called a *fundamental set* if:

(i) $q(p^{-1}(U \cap S)) \subset S$,

(ii) if $x \in \overline{\mathrm{conv}}(\varphi(x) \cup S)$, then $x \in S$.

For a homotopy $\chi \in \mathbb{C}(U \times [0, 1], C)$, $S \subset C$ is called fundamental if it is fundamental to $\chi(., \lambda)$, for each $\lambda \in [0, 1]$.

PROPOSITION 2.7. *Assume $(p, q) \in \mathbb{C}(U, C)$.*

(i) *If S is a fundamental set for (p, q), then* $\mathrm{Fix}(p, q) \subset S$.

(ii) *Intersection of fundamental sets, for (p, q), is also fundamental, for (p, q).*

(iii) *The family of all fundamental sets for (p, q) is nonempty.*

(iv) *If S is a fundamental set for $\chi \in \mathbb{C}(U \times [0, 1], C)$ and $P \subset S$, then the set $\overline{\mathrm{conv}}(\chi((U \cap S) \times [0, 1]) \cup P)$ is also fundamental.*

For more details, see [22,67], and the references therein.

3. Applied fixed point principles

3.1. *Lefschetz fixed point theorems*

We start with the Lefschetz theory, because it is a base for our further investigation. More precisely, the generalized Lefschetz number can be used for the definition of essential classes in the Nielsen theory as well as the possible normalization property of the fixed point index. We restrict ourselves only to the presentation of necessary facts.

Consider a multivalued map $\varphi : X \multimap X$ and assume that

(i) X is a (metric) ANR-space, e.g., a retract of an open subset of a convex set in a Fréchet space,

(ii) φ is a compact (i.e. $\overline{\varphi(X)}$ is compact) composition of an R_δ-map $p^{-1} : X \multimap \Gamma$ and a continuous (single-valued) map $q : \Gamma \to X$, namely $\varphi = q \circ p^{-1}$, where Γ is a metric space.

Then an integer $\Lambda(\varphi) = \Lambda(p, q)$, called the *generalized Lefschetz number* for $\varphi \in \mathbb{K}(X)$, is well-defined (see, e.g., [12; 22, Chapter I.6; 67]) and $\Lambda(\varphi) \neq 0$ implies that

$$\mathrm{Fix}(\varphi) := \{x \in X \mid x \in \varphi(x)\} \neq \emptyset.$$

Moreover, Λ is a homotopy invariant, namely if φ is compactly homotopic (in the same class of maps) with $\widetilde{\varphi} : X \multimap X$, then $\Lambda(\varphi) = \Lambda(\widetilde{\varphi})$.

In order to define the generalized Lefschetz number, one should be familiar with the elements of algebraic topology, in particular, of homology theory. Therefore, we only briefly sketch this definition without proofs. For more details, we recommend [51,68] (in the single-valued case) and [12,22,67] (in the multivalued case).

At first, we recall some algebraic preliminaries. In what follows, all vector spaces are taken over $\mathbb{Q}$. Let $f : E \to E$ be an endomorphism of a finite-dimensional vector space E. If $v_1, \dots, v_n$ is a basis for E, then we can write

$$f(v_i) = \sum_{j=1}^{n} a_{ij} v_j, \quad \text{for all } i = 1, \dots, n.$$

The matrix $[a_{ij}]$ is called the matrix of f (with respect to the basis $v_1, \dots, v_n$). Let $A = [a_{ij}]$ be an $(n \times n)$-matrix; then the trace of A is defined as $\sum_{i=1}^{n} a_{ii}$. If $f : E \to E$ is an endomorphism of a finite-dimensional vector space E, then the trace of f, written $\mathrm{tr}(f)$, is the trace of the matrix of f with respect to some basis for E. If E is a trivial vector space then, by definition, $\mathrm{tr}(f) = 0$. It is a standard result that the definition of the trace of an endomorphism is independent of the choice of the basis for E.

Hence, let $E = \{E_q\}$ be a graded vector space of a finite type.

If $f = \{f_q\}$ is an endomorphism of degree zero of such a graded vector space, then the (*ordinary*) *Lefschetz number* $\lambda(f)$ of f is defined by

$$\lambda(f) = \sum_q (-1)^q \,\mathrm{tr}(f_q).$$

Let $f : E \to E$ be an endomorphism of an arbitrary vector space E. Denote by $f^n : E \to E$ the nth iterate of f and observe that the kernels

$$\ker f \subset \ker f^2 \subset \cdots \subset \ker f^n \subset \cdots$$

form an increasing sequence of subspaces of E. Let us now put

$$N(f) = \bigcup_n \ker f^n \quad \text{and} \quad \widetilde{E} = E/N(f).$$

Clearly, f maps $N(f)$ into itself and, therefore, induces the endomorphism $\widetilde{f} : \widetilde{E} \to \widetilde{E}$ on the factor space $\widetilde{E} = E/N(f)$.

Let $f : E \to E$ be an endomorphism of a vector space E. Assume that $\dim \widetilde{E} < \infty$. In this case, we define the generalized trace $\mathrm{Tr}(f)$ of f by putting $\mathrm{Tr}(f) = \mathrm{tr}(\widetilde{f})$.

LEMMA 3.1. *Let $f : E \to E$ be an endomorphism. If* $\dim E < \infty$, *then* $\mathrm{Tr}(f) = \mathrm{tr}(f)$.

For the proof, see [22].

Let $f = \{f_q\}$ be an endomorphism of degree zero of a graded vector space $E = \{E_q\}$. We say that f is a *Leray endomorphism* if the graded vector space $\widetilde{E} = \{\widetilde{E}_q\}$ is of finite type. For such an f, we define the (*generalized*) *Lefschetz number* $\Lambda(f)$ of f by putting

$$\Lambda(f) = \sum_q (-1)^q \,\mathrm{Tr}(f_q).$$

It is immediate from Lemma 3.1 that

LEMMA 3.2. *Let $f : E \to E$ be an endomorphism of degree zero, i.e., $f = \{f_q\}$ and $f_q : E_q \to E_q$ is a linear map. If E is a graded vector space of finite type, then* $\Lambda(f) = \lambda(f)$.

Now, the *Lefschetz number* will be defined for admissible compact mappings. For our needs in the sequel, it is enough to consider only the compact compositions of R_δ-maps and continuous single-valued maps as above (by which Lefschetz sets simplify into Lefschetz numbers). Let $\varphi : E \multimap E$ be an admissible compact map and $(p, q) \subset \varphi$ be a selected pair of φ. Then the induced homomorphism $q_* \circ p_*^{-1} : H_*(E) \to H_*(E)$ is an endomorphism of the graded vector space $H_*(E)$ into itself. So, we can define the Lefschetz number $\Lambda(p, q)$ of the pair (p, q) by putting $\Lambda(p, q) = \Lambda(q_* \circ p_*^{-1})$, provided the Lefschetz number $\Lambda(q_* \circ p_*^{-1})$ is well-defined.

It allows us to define the Lefschetz set Λ of φ as follows:

$$\Lambda(\varphi) = \big\{\Lambda(p, q) \mid (p, q) \subset \varphi\big\}.$$

In what follows, we say that the Lefschetz set $\Lambda(\varphi)$ of φ is *well-defined* if, for every $(p, q) \subset \varphi$, the Lefschetz number $\Lambda(p, q)$ of (p, q) is defined.

Moreover, from the homotopy property of Λ, we get:

LEMMA 3.3.

(i) *If* $\varphi, \psi : E \multimap E$ *are compactly homotopic* ($\varphi \sim \psi$), *then*: $\Lambda(\varphi) \cap \Lambda(\psi) \neq \emptyset$.

(ii) *If* $\varphi : E \multimap E$ *is admissible and* E *is acyclic*, *then the Lefschetz set* $\Lambda(\varphi)$ *is well-defined and* $\Lambda(\varphi) = \{1\}$.

It is useful to formulate

THEOREM 3.1 (Coincidence theorem). *Let* U *be an open subset of a finite dimensional normed space* E. *Consider the following diagram*:

$$U \xLeftarrow{p} \Gamma \xrightarrow{q} U$$

in which q *is a compact map. Then the Lefschetz number* $\Lambda(p, q)$ *of the pair* (p, q), *given by the formula*

$$\Lambda(p, q) = \Lambda\big(q_* \circ p_*^{-1}\big),$$

is well-defined, *and* $\Lambda(p, q) \neq 0$ *implies that* $p(y) = q(y)$, *for some* $y \in \Gamma$.

Theorem 3.1 can be reformulated in terms of multivalued mappings as follows.

Let $U \subset E$ be the same as in Theorem 3.1 and let $\varphi : U \multimap U$ be a compact, admissible map, i.e., $\varphi \in \mathbb{K}(U)$. We let $\Lambda(\varphi) = \{\Lambda(p, q) \mid (p, q) \subset \varphi\}$, where $\Lambda(p, q) = \Lambda(q_* \circ p_*^{-1})$. Then we have:

THEOREM 3.2.

(i) *The set* $\Lambda(\varphi)$ *is well-defined*, *i.e. for every* $(p, q) \subset \varphi$, *the generalized Lefschetz number* $\Lambda(p, q)$ *of the pair* (p, q) *is well-defined*, *and*

(ii) $\Lambda(\varphi) \neq \{0\}$ *implies that the set* $\mathrm{Fix}(\varphi) := \{x \in U \mid x \in \varphi(x)\}$ *is nonempty.*

Theorem 3.2 can be generalized, by means of the Schauder-like approximation technique (for more details, see [22]), for compact admissible maps $\varphi \in \mathbb{K}$ on ANR-spaces, e.g., on retracts of open subsets of convex sets in Fréchet spaces, as follows.

THEOREM 3.3 (The Lefschetz fixed point theorem). *Let* X *be an ANR-space*, *e.g.*, *a retract of an open subset* U *of a convex set in a Fréchet space. Assume*, *furthermore*, *that* $\varphi \in \mathbb{K}(X)$. *Then*:

(i) *the Lefschetz set* $\Lambda(\varphi)$ *of* φ *is well-defined*,

(ii) *if* $\Lambda(\varphi) \neq \{0\}$, *then* $\mathrm{Fix}(\varphi) \neq \emptyset$.

REMARK 3.1. If admissible map $\varphi \in \mathbb{K}(X)$ is a composition of an R_δ-map and a continuous single-valued map, then $\Lambda(\varphi)$ is an integer. If, in particular, X is an AR-space, then $\Lambda(\varphi) = 1$.

REMARK 3.2. The definition of a generalized Lefschetz number for condensing maps is far from to be obvious, and so it can not be used as a normalization property for the related fixed points index. Roughly speaking, it requires to assume additionally the existence of a compact attractor or to impose some additional restrictions on the set X like to be a special neighbourhood retract of a Fréchet space (cf. [68]).

3.2. *Nielsen fixed point theorems*

The standard Nielsen theory allows us to obtain the lower estimate of the number of fixed points. More precisely, if $f : X \to X$ is a compact (continuous) map on a (metric) ANR-space X, then a nonnegative integer $N(f)$, called the *Nielsen number* of f, is defined such that

- $N(f) \leqslant \#\mathrm{Fix}(f) := \mathrm{card}\{\hat{x} \in X \mid f(\hat{x}) = \hat{x}\}$,
- $N(f) = N(\widetilde{f})$, for any compact $\widetilde{f} : X \to X$ which is *compactly homotopic* to f, i.e. if there is a compact map $h : X \times [0,1] \to X$ such that $h_0 = f$, $h_1 = \widetilde{f}$, where $h_t(x) := h(x,t)$, for $t \in [0,1]$.

Given a compact $f : X \to X$ on $X \in$ ANR, we say that $x, y \in \mathrm{Fix}(f)$ are *Nielsen related* if there exists a path $u : [0,1] \to X$ such that $u(0) = x$, $u(1) = y$, and u, $f(u)$ are homotopic keeping the endpoints fixed. Since the Nielsen relation is an equivalence, $\mathrm{Fix}(f)$ splits into fixed point classes. Since the classes are open and f is compact, we have a finite number of fixed point classes.

If, for a Nielsen class $\mathcal{N} \subset \mathrm{Fix}(f)$, we have $\mathrm{ind}(\mathcal{N}, f) \neq 0$, i.e. if the associated fixed point index is nontrivial, then $\mathcal{N}$ is called *essential*. The *Nielsen number* $N(f)$ is then defined to be the number of essential Nielsen classes. For more details, see, e.g., [77].

To compute $N(f)$ can be a difficult task. In the multivalued case, the situation is even more delicate, because the above definition can not be directly generalized. Thus, we only indicate this subtle definition again. Nevertheless, in the single-valued case, these definitions are equivalent.

Consider a multivalued map $\varphi : X \multimap X$ and assume that

(i) X is a connected ANR-space, e.g., a connected retract of an open subset of a convex set in a Fréchet space,
(ii) X has a finitely generated abelian fundamental group,
(iii) φ is a compact (i.e. $\overline{\varphi(X)}$ is compact) composition of an R_δ-map $p^{-1} : X \multimap \Gamma$ and a continuous (single-valued) map $q : \Gamma \to X$, namely $\varphi = q \circ p^{-1}$, where Γ is a metric space.

Then a nonnegative integer $N(\varphi) = N(p,q)$,[1] called the *Nielsen number* for $\varphi \in \mathbb{K}$, exists (see [24] and [22, Chapter I.10] or [12]) such that $N(\varphi) \leqslant \#C(\varphi)$, where

$$\#C(\varphi) = \#C(p,q) := \mathrm{card}\big\{z \in \Gamma \mid p(z) = q(z)\big\}$$

and $N(\varphi_0) = N(\varphi_1)$, for compactly homotopic maps $\varphi_0 \sim \varphi_1$.

REMARK 3.3. Condition (ii) is satisfied, provided X is the torus $\mathbb{T}^n$ ($\pi_1(\mathbb{T}^n) = \mathbb{Z}^n$) and it can be avoided if X is compact and $q = \mathrm{id}$ is the identity (cf. [5]).

---

[1] We should write more correctly $N_H(\varphi) = N_H(p,q)$, because it is in fact (mod H)-Nielsen number, as can be seen below. For the sake of simplicity, we omit the index H in the following sections.

REMARK 3.4 (Important). We have a counter-example in [24] (cf. [22, Example 10.1 in Chapter I.10]) that, under the above assumptions (i)–(iii), the Nielsen number $N(\varphi)$ is rather the topological invariant for the number of essential classes of coincidences than of fixed points. On the other hand, for a compact X and $q = \mathrm{id}$, $N(\varphi)$ gives even without (ii) a lower estimate of the number of fixed points of φ (see [5]), i.e. $N(\varphi) \leqslant \#\mathrm{Fix}(\varphi)$, where $\mathrm{Fix}(\varphi) := \mathrm{card}\{x \in X \mid x \in \varphi(x)\}$. We have conjectured in [38] that if $\varphi = q \circ p^{-1}$ assumes only simply connected values, then also $N(\varphi) \leqslant \#\mathrm{Fix}(\varphi)$.

The following sketch demonstrates how subtle is the definition of the Nielsen number for multivalued maps. Let

$$X \xLeftarrow{p_0} \Gamma \xrightarrow{q_0} Y \quad \text{and} \quad X \xLeftarrow{p_1} \Gamma \xrightarrow{q_1} Y$$

be two admissible maps.

If $(p_0, q_0) \sim (p_1, q_1)$, i.e. if (p_0, q_0) is admissibly homotopic to (p_1, q_1) (see Definition 2.6), and $h: Y \to Z$ is a continuous map, then we write $(p_0, hq_0) \sim (p, hq)$. We say that a multivalued map $X \xLeftarrow{p} \Gamma \xrightarrow{q} Y$ *represents a single-valued map* $\rho: X \to Y$ if $q = p\rho$. Now, we assume that $X = Y$ and we are going to estimate the cardinality of the coincidence set

$$C(p,q) := \bigl\{z \in \Gamma \mid p(z) = q(z)\bigr\}.$$

We begin by defining a Nielsen-type relation on $C(p,q)$. This definition requires the following conditions on $X \xLeftarrow{p} \Gamma \xrightarrow{q} Y$:

(i′) X, Y are connected, locally contractible metric spaces (observe that then they admit universal coverings),

(iii′) $p: \Gamma \Longrightarrow X$ is a Vietoris map,

(iii″) for any $x \in X$, the restriction $q_1 = q|_{p^{-1}(x)}: p^{-1}(x) \to Y$ admits a lift $\widetilde{q}_1$ to the universal covering space $(p_Y: \widetilde{Y} \to Y)$:

$$\begin{tikzcd}
 & \widetilde{Y} \arrow[d, Rightarrow, "p_Y"] \\
p^{-1}(x) \arrow[ur, dashed, "\widetilde{q}_1"] \arrow[r, "q_1"'] & Y
\end{tikzcd}$$

Let us note that the following implications hold: (i) $\Rightarrow$ (i′), (iii) $\Rightarrow$ (iii′), (iii″).

Consider a single-valued map $\rho: X \to Y$ between two spaces admitting universal coverings $p_X: \widetilde{X} \Rightarrow X$ and $p_Y: \widetilde{X} \Rightarrow Y$. Let $\theta_X = \{\alpha: \widetilde{X} \to \widetilde{X} \mid p_X\alpha = p_X\}$ be the group of natural transformations of the covering p_X. Then the map ρ admits a lift $\widetilde{\rho}: \widetilde{X} \to \widetilde{Y}$. We can define a homomorphism $\widetilde{\rho}_!: \theta_X \to \theta_Y$ by the equality

$$\widetilde{q}(\alpha \cdot \widetilde{x}) = \widetilde{q}_!(\alpha)\widetilde{q}(\widetilde{x}) \quad \bigl(\alpha \in \theta_X, \widetilde{x} \in \widetilde{X}\bigr).$$

It is well known that there is an isomorphism between the fundamental group $\pi_1(X)$ and θ_X which may be described as follows. We fix points $x_0 \in X$, $\widetilde{x} \in \widetilde{X}$ and a loop $\omega : I \to X$ based at x_0. Let $\widetilde{\omega}$ denote the unique lift of ω starting from $\widetilde{x}_0$. We subordinate to $[\omega] \in \pi_1(X, x_0)$ the unique transformation from θ_X sending $\widetilde{\omega}(0)$ to $\widetilde{\omega}(1)$. Then the homomorphism $\widetilde{\rho}_! : \theta_X \to \theta_Y$ corresponds to the induced homomorphism between the fundamental groups $\rho_\# : \pi_1(X, x_0) \to \pi_1(Y, \rho(x_0))$.

It can be shown that, under the assumptions (i$'$), (iii$'$), (iii$''$), a multivalued map (p, q) admits a lift to a multivalued map between the universal coverings. These lifts will split the coincidence set $C(p, q)$ into Nielsen classes. Besides that the pair (p, q) induces a homomorphism $\theta_X \to \theta_Y$ giving the Reidemeister set.

We start with the following lemma.

LEMMA 3.4. *Suppose we are given Y, a locally contractible metric space, Γ a metric space, $\Gamma_0 \subset \Gamma$ a compact subspace, $q : \Gamma \to Y$, $\widetilde{q}_0 : \Gamma_0 \to Y$ continuous maps for which the diagram*

$$\begin{tikzcd}
\Gamma_0 \arrow[r, "\widetilde{q}_0"] \arrow[d, "i"'] & \widetilde{Y} \arrow[d, "p_Y"] \\
\Gamma \arrow[r, "q"'] & Y
\end{tikzcd}$$

commutes (here, $p_Y : \widetilde{Y} \to Y$ denotes the universal covering). In other words, $\widetilde{q}_0$ is a partial lift of q. Then $\widetilde{q}_0$ admits an extension to a lift onto an open neighbourhood of Γ_0 in Γ.

Consider again a multivalued map $X \xleftarrow{p} \Gamma \xrightarrow{q} Y$ satisfying (i$'$). Define (a pullback)

$$\widetilde{\Gamma} = \big\{(\widetilde{x}, z) \in \widetilde{X} \times \Gamma \mid p_X(\widetilde{x}) = p(z)\big\}.$$

Now, we can apply Lemma 3.4 to the multivalued map $\widetilde{X} \overset{\widetilde{p}}{\Longleftarrow} \widetilde{\Gamma} \xrightarrow{qp_\Gamma} Y$, and so we get a lift $\widetilde{q} : \widetilde{\Gamma} \to \widetilde{Y}$ such that the diagram

$$\begin{tikzcd}
\widetilde{X} \arrow[d, "p_X"'] & \widetilde{\Gamma} \arrow[l, Rightarrow, "\widetilde{p}"'] \arrow[r, "\widetilde{q}"] \arrow[d, "p_\Gamma"'] & \widetilde{Y} \arrow[d, "p_Y"] \\
X & \Gamma \arrow[l, Rightarrow, "p"] \arrow[r, "q"'] & Y
\end{tikzcd}$$

is commutative, where $\widetilde{p}(\widetilde{x}, z) = \widetilde{x}$ and $p_\Gamma(\widetilde{x}, z) = z$. Let us note that the lift $\widetilde{p}$ is given by the above formula, but $\widetilde{q}$ is not precised. We fix such a $\widetilde{q}$.

Observe that $p:\Gamma \Longrightarrow X$ and the lift $\widetilde{p}$ induce a homomorphism $\widetilde{p}^!:\theta_X \to \theta_\Gamma$ by the formula $\widetilde{p}^!(\alpha)(\widetilde{x},z)=(\alpha\widetilde{x},z)$. It is easy to check that the homomorphism $\widetilde{p}^!$ is an isomorphism (any natural transformation of $\widetilde{\Gamma}$ is of the form $\alpha\cdot(\widetilde{x},z)=(\alpha\widetilde{x},z)$) and that $\widetilde{p}^!$ is inverse to $\widetilde{p}_!$. Recall that the lift $\widetilde{q}$ defines a homomorphism $\widetilde{q}_!:\theta_\Gamma \to \theta_Y$ by the equality $\widetilde{q}(\lambda)=\widetilde{q}_!(\lambda)\widetilde{q}$.

In the sequel, we will consider the composition $\widetilde{q}_!\widetilde{p}^!:\theta_X \to \theta_Y$.

LEMMA 3.5. *Let a multivalued map* (p,q) *satisfying* (i′) *represent a single-valued map* ρ, *i.e.* $q=\rho p$. *Let* $\widetilde{\rho}$ *be the lift of* ρ *which satisfies* $\widetilde{q}=\widetilde{\rho}\widetilde{p}$. *Then* $\widetilde{\rho}_!\widetilde{p}^!=\rho_!$.

Now, we are in a position to define the Nielsen classes. Consider a multivalued self-map $X \overset{p}{\Longleftarrow} \Gamma \xrightarrow{q} X$ satisfying (i′). By the above consideration, we have a commutative diagram

$$\begin{tikzcd}
\widetilde{\Gamma} \arrow[r, "\widetilde{p},\widetilde{q}"] \arrow[d, "p_\Gamma"'] & \widetilde{X} \arrow[d, "p_X"] \\
\Gamma \arrow[r, "p,q"] & X
\end{tikzcd}$$

Following the single-valued case (see, e.g., [77]), we can prove (see [24] and cf. [22]) the following lemma.

LEMMA 3.6.
- $C(p,q)=\bigcup_{\alpha\in\theta_X} p_\Gamma C(\widetilde{p},\alpha\widetilde{q})$,
- *if* $p_\Gamma C(\widetilde{p},\alpha\widetilde{q})\cap p_\Gamma C(\widetilde{p},\beta\widetilde{q})$ *is not empty, then there exists a* $\gamma\in\theta_X$ *such that* $\beta=\gamma\circ\alpha\circ(\widetilde{q}_!\widetilde{p}^!\gamma)^{-1}$,
- *the sets* $p_\Gamma C(\widetilde{p},\alpha\widetilde{q})$ *are either disjoint or equal.*

Define an action of θ_X on itself by the formula $\gamma\circ\alpha=\gamma\alpha(\widetilde{q}_!p^!\gamma)$. The quotient set will be called the *set of Reidemeister classes* and will be denoted by $R(p,q)$. The above lemma defines an injection:

$$\text{set of Nielsen classes} \to R(p,q),$$

given by $A \to [\alpha]\in R(p,q)$, where $\alpha\in\theta_X$ satisfies $A=p_\Gamma(C(\widetilde{p},\alpha\widetilde{q}))$.

One can prove that our definition does not depend on $\widetilde{q}$.

Let us recall that the homomorphism $\widetilde{q}_!:\theta_\Gamma \to \theta_Y$ is defined by the relation $\widetilde{q}\alpha=\widetilde{q}_!(\alpha)\widetilde{q}$, for $\alpha\in\theta_\Gamma$. If $\widetilde{q}'=\gamma\widetilde{q}$ is another lift of q ($\widetilde{\gamma}\in\theta_\Gamma$), then the induced homomorphism $\widetilde{q}'_!:\theta_\Gamma \to \theta_Y$ is defined by the relation $\widetilde{q}'\alpha=\widetilde{q}'_!(\alpha)\widetilde{q}'$.

One can also show that the Reidemeister sets obtained by different lifts of q are canonically isomorphic. That is why we write $R(p,q)$ omitting tildes.

PROPOSITION 3.1. *If* $X\times[0,1] \xleftarrow{p} \Gamma \xrightarrow{q} Y$ *is a homotopy satisfying* (i′), (iii′), (iii″), *then the homomorphism* $\widetilde{q}_{t!}\widetilde{p}_t^!:\theta_X \to \theta_Y$ *does not depend on* $t\in[0,1]$, *where the lifts used* *in the definitions of these homomorphisms are restrictions of some fixed lifts* p, q *of the given homotopy.*

REMARK 3.5. If (p,q) represents a single-valued map $\rho : X \to Y$ ($q = \rho p$), then $\widetilde{q}_! \widetilde{p}^!$ equals $\widetilde{\rho}_!$ (here the chosen lifts satisfy $\widetilde{q} = \widetilde{\rho}\widetilde{p}$).

Let us point out that the above theory can be modified into the relative case (i.e. modulo a normal subgroup $H \subset \theta_X$). The index H will denote the relative modification.

Assuming $X = Y$, we can give

LEMMA 3.7.
- (i) $C(p,q) = \bigcup_{\alpha \in \theta_{XH}} p_{\Gamma_H} C(\widetilde{p}_H, \alpha\widetilde{q}_H)$,
- (ii) *if* $p_{\Gamma_H} C(\widetilde{p}_H, \alpha\widetilde{q}_H) \cap p_{\Gamma_H} C(\widetilde{p}_H, \beta\widetilde{q}_H)$ *is not empty, then there exists a* $\gamma \in \theta_{XH}$ *such that* $\beta = \gamma \circ \alpha \circ (\widetilde{q}_{H!} \widetilde{p}^!_H \gamma)^{-1}$,
- (iii) *the sets* $p_{\Gamma_H} C(\widetilde{p}_H, \alpha\widetilde{q}_H)$ *are either disjoint or equal.*

Hence, we get the splitting of $C(p,q)$ into the H-Nielsen classes and the natural injection from the set of H-Nielsen classes into the set of Reidemeister classes modulo H, namely, $R_H(p,q)$.

Now, we would like to exhibit the classes which do not disappear under any compact (admissible) homotopy. For this, we need however (besides (i′), (iii′), (iii″)) the following two assumptions on the pair $X \overset{p}{\Longleftarrow} \Gamma \xrightarrow{q} Y$:
- (i″) Let X be a connected ANR-space, e.g., a connected retract of an open subset of (a convex set in) a Fréchet space, p is a Vietoris map and $\mathrm{cl}(q(\Gamma)) \subset X$ is compact, i.e. q is a compact map.
- (ii′) There exists a normal subgroup $H \subset \theta_X$ of a finite index satisfying $\widetilde{q}_! \widetilde{p}^!(H) \subset H$.

Let us note that the following implications hold: (i), (ii) ⇒ (ii′), (i), (iii) ⇒ (i″).

DEFINITION 3.1. We call a pair (p,q) *N-admissible* if it satisfies (i′), (i″), (ii′), (iii′), (iii″) (⇐ (i)–(iii)).

Let us recall that, under the assumption (ii′), the Lefschetz number $\Lambda(p,q) \in \mathbb{Q}$ is defined (see Section 3.1). This is a homotopy invariant (with respect to the homotopies satisfying (ii′)) and $\Lambda(p,q) \neq 0$ implies $C(p,q) \neq \emptyset$ (cf. Section 3.1).

Let $A = p_{\Gamma_H} C(\widetilde{p}, \alpha\widetilde{q})$ be a Nielsen class of an N-admissible pair (p,q). We say that (the N-Nielsen class) A is *essential* if $\Lambda(\widetilde{p}, \alpha\widetilde{q}) \neq 0$. This definition it correct, i.e. does not depend on the choice of α.

DEFINITION 3.2. Let (p,q) be an N-admissible multivalued map (for a subgroup $H \subset \theta_X$). We define the *Nielsen number modulo H* as the number of essential classes in θ_{X_H}. We denote this number by $N_H(p,q)$.

The following theorem is an easy consequence of the homotopy invariance of the Lefschetz number.

THEOREM 3.4. *$N_H(p,q)$ is a homotopy invariant (with respect to N-admissible homotopies*

$$X \times [0,1] \xLeftarrow{p} \Gamma \xrightarrow{q} X).$$

Moreover, (p,q) has at least $N_H(p,q)$ coincidences.

The following theorem shows that the above definition is consistent with the classical Nielsen number for single-valued maps.

THEOREM 3.5. *If an N-admissible map (p,q) is N-admissibly homotopic to a pair (p',q'), representing a single-valued map p (i.e. $q' = \rho p'$), then (p,q) has at least $N_H(\rho)$ coincidences (here H denotes also the subgroup of $\pi_1 X$ corresponding to the given $H \subset \theta_X$ in* (ii$'$)).

Although in the general case the theory requires special assumptions on the considered pair (p,q), in the case of multivalued self-maps on a torus it is enough to assume that this pair satisfies only (i$''$), i.e. it is admissible. This is due to the fact that any pair satisfying (i$''$) is homotopic to a pair representing a single-valued map.

THEOREM 3.6. *Any multivalued self-map (p,q) on the torus satisfying* (i$''$) *is admissibly homotopic to a pair representing a single-valued map.*

THEOREM 3.7. *Let $\mathbb{T}^n \xLeftarrow{p} \Gamma \xrightarrow{q} \mathbb{T}^n$ be such that p is a Vietoris map. Let $\rho : \mathbb{T}^n \to \mathbb{T}^n$ be a single-valued map representing a multivalued map homotopic to (p,q) (according to Theorem* 3.6, *such a map always exists). Then (p,q) has at least $N(\rho)$ coincidences.*

REMARK 3.6. Let us also recall that, on the torus $\mathbb{T}^n$, $N(\rho) = |\Lambda(\rho)| = |\det(I - A)|$, where A is an integer $(n \times n)$-matrix representing the induced homotopy homomorphism $\rho_\# : \pi_1 \mathbb{T}^n \to \pi_1 \mathbb{T}^n$. Moreover, if $\det(I - A) \neq 0$, then

$$\operatorname{card}\big(\pi_1(\mathbb{T}^n) / \operatorname{Im}(\rho_\#)\big) = \big|\det(I - A)\big|.$$

In particular, for $\rho = \mathrm{id}$, we have

$$N(\mathrm{id}) = \big|\Lambda(\mathrm{id})\big| = \big|\chi(\mathbb{T}^n)\big| = |\det O| = 0,$$

while for $\rho = -\mathrm{id}$, we have $N(-\mathrm{id}) = |\Lambda(-\mathrm{id})| = |\det 2I| = 2^n$.

For more details, see [22,12].

3.3. *Fixed point index theorems*

Consider a multivalued map $\varphi : X \multimap X$ and assume, similarly as in Section 3.1, that

(i) X is ANR-space, e.g., a retract of an open subset of a convex set in a Fréchet space,

(ii) φ is a compact composition of an R_δ-map $\Phi : X \multimap Y$ and a continuous single-valued map $f : Y \to X$, namely $\varphi = f \circ \Phi$, where Y is an ANR-space.

Let $D \subset X$ be an open subset of X with no fixed points of φ on its boundary ∂D. Then and integer $\operatorname{ind}(\varphi, X, D)$, called the *fixed point index* over X w.r.t. D exists such that the following proposition holds (see, e.g., [12,22,44]).

PROPOSITION 3.2. *Let $\varphi : X \multimap X$ be a map satisfying* (i), (ii). *Then* $\operatorname{ind}(\varphi, X, D) \in \mathbb{Z}$ *is well-defined satisfying the following properties*:

- (Existence) *If* $\operatorname{ind}(\varphi, X, D) \neq 0$, *then* $\operatorname{Fix}(\varphi) \neq \emptyset$.
- (Localization) *If $D_1 \subset D$ are open subsets of X such that* $\operatorname{Fix}(\varphi) \subset D_1 \subset D$, *then*

$$\operatorname{ind}(\varphi, X, D) = \operatorname{ind}(\varphi, X, D_1).$$

- (Additivity) *If D_j, $j = 1, \dots, n$, are open disjoint subsets of D and all fixed points of $\varphi|_D$ are located in $\bigcup_{j=1}^n D_j$, then* $\operatorname{ind}(\varphi, X, D_j)$, $j = 1, \dots, n$, *are well-defined satisfying*

$$\operatorname{ind}(\varphi, X, D) = \sum_{j=1}^{n} \operatorname{ind}(\varphi, X, D_j).$$

- (Homotopy) *If there is a compact homotopy $\chi : X \times [0, 1] \multimap X$ (in the same class of maps under consideration) with $\chi(\cdot, 0) = \varphi$, $\chi(\cdot, 1) = \psi$, and ∂D is fixed point free w.r.t. χ, then*

$$\operatorname{ind}(\varphi, X, D) = \operatorname{ind}(\psi, X, D).$$

- (Multiplicity) *If $\psi : \widetilde{X} \multimap \widetilde{X}$ satisfies* (i), (ii) *and an open $\widetilde{D} \subset \widetilde{X}$ is fixed point free w.r.t. ψ, then*

$$\operatorname{ind}\bigl(\varphi \times \psi, X \times \widetilde{X}, D \times \widetilde{D}\bigr) = \operatorname{ind}(\varphi, X, D) \cdot \operatorname{ind}\bigl(\psi, \widetilde{X}, \widetilde{D}\bigr).$$

- (Contraction) *If $X' \subset X$ are ANR-spaces such that $\varphi(X) \subset X'$ and $\varphi|_{X'}$ satisfies* (ii) *with* $\operatorname{Fix}(\varphi|_{X'}) \cap \partial(D \cap X') = \emptyset$, *then*

$$\operatorname{ind}(\varphi, X, D) = \operatorname{ind}\bigl(\varphi|_{X'}, X', D \cap X'\bigr).$$

- (Normalization) *If $X = D$, then*

$$\operatorname{ind}(\varphi, X, D) = \operatorname{ind}(\varphi, X, X) = \Lambda(\varphi).$$

Because of possible applications, it is very useful to formulate sufficiently general continuation principles.

For compact admissible maps from open subsets of a neighbourhood retract of a Fréchet space E into E, the fixed point index was just indicated.

Now, we will apply Proposition 3.2 to formulating the appropriate continuation principles. We restrict ourselves only to a particular class of admissible maps, namely to R_δ-maps $\Phi : D \multimap E$ (written here $\Phi \in J(D, E)$), i.e. u.s.c. maps with R_δ-values.

We often need to study fixed points for maps defined on sufficiently fine sets (possibly with an empty interior), but with values out of them. Making use of the previous results, we are in position to make the following construction.

Assume that X is a retract of a Fréchet space E and D is an open subset of X. Let $\Phi \in J(D, E)$ be locally compact, $\mathrm{Fix}(\Phi)$ be compact and let the following condition hold:

$$\forall x \in \mathrm{Fix}(\Phi)\ \exists U_x \ni x,\ U_x \text{ is open in } D \text{ such that } \Phi(U_x) \subset X. \tag{A}$$

The class of locally compact J-maps from D to E with the compact fixed point set and satisfying (A) will be denoted by the symbol $J_A(D, E)$. We say that $\Phi, \Psi \in J_A(D, E)$ are *homotopic in* $J_A(D, E)$ if there exists a homotopy $H \in J(D \times [0, 1], E)$ such that $H(\cdot, 0) = \Phi$, $H(\cdot, 1) = \Psi$, for every $x \in D$, there is an open neighbourhood V_x of x in D such that $H|_{V_x \times [0,1]}$ is compact, and

$$\forall x \in D\ \forall t \in [0, 1]$$
$$\big[x \in H(x, t) \Rightarrow \exists U_x \ni x,\ U_x \text{ is open in } D,\ H\big(U_x \times [0, 1]\big) \subset X\big]. \tag{A$_H$}$$

Note that the condition (A_H) is equivalent to the following one:

- If $\{x_j\}_{j \geqslant 1} \subset D$ converges to $x \in H(x, t)$, for some $t \in [0, 1]$, then $H(\{x_j\} \times [0, 1]) \subset X$, for j sufficiently large.

Let $\Phi \in J_A(D, E)$. Then $\mathrm{Fix}(\Phi) \subset \bigcup\{U_x \mid x \in \mathrm{Fix}(\Phi)\} \cap V =: D' \subset D$ and $\Phi(D') \subset X$, where V is a neighbourhood of the set $\mathrm{Fix}(\Phi)$ such that $\Phi|_V$ is compact (by the compactness of $\mathrm{Fix}(\Phi)$ and local compactness of Φ) and U_x is a neighbourhood of x as in (A). Define

$$\mathrm{Ind}_A(\Phi, X, D) = \mathrm{ind}\big(\Phi|_{D'}, X, D'\big),$$

where $\mathrm{ind}(\Phi|_{D'}, X, D')$ is defined as in Proposition 3.2. This definition is independent of the choice of D'.

In the following theorem, we give some properties of Ind_A which will be used in the proof of the continuation Theorem 3.8. The simple proof is omitted.

PROPOSITION 3.3.

(i) (Existence) *If* $\mathrm{Ind}_A(\Phi, X, D) \neq 0$, *then* $\mathrm{Fix}(\Phi) \neq \emptyset$.

(ii) (Localization) *If* $D_1 \subset D$ *are open subsets of a retract* X *of a space* E, $\Phi \in J_A(D, E)$ *is compact, and* $\mathrm{Fix}(\Phi)$ *is a compact subset of* D_1, *then*

$$\mathrm{Ind}_A(\Phi, X, D) = \mathrm{Ind}_A(\Phi, X, D_1).$$

(iii) (Homotopy) *If* H *is a homotopy in* $J_A(D, E)$, *then*

$$\mathrm{Ind}_A\big(H(\cdot, 0), X, D\big) = \mathrm{Ind}_A\big(H(\cdot, 1), X, D\big).$$

(iv) (Normalization) *If* $\Phi \in J(X)$ *is a compact map, then* $\mathrm{Ind}_A(\Phi, X, X) = 1$.

THEOREM 3.8 (Continuation principle). *Let* X *be a retract of a Fréchet space* E, D *be an open subset of* X *and* H *be a homotopy in* $J_A(D, E)$ *such that*

(i) $H(\cdot, 0)(D) \subset X$,

(ii) *there exists* $H' \in J(X)$ *such that* $H'|_D = H(\cdot, 0)$, H' *is compact and* $\mathrm{Fix}(H') \cap (X \setminus D) = \emptyset$.

Then there exists $x \in D$ *such that* $x \in H(x, 1)$.

PROOF. Applying the localization property (ii), we obtain

$$\mathrm{Ind}_A\bigl(H(\cdot, 0), X, D\bigr) = \mathrm{Ind}_A\bigl(H(\cdot, 0), X, X\bigr).$$

By the normalization property (iv), $\mathrm{Ind}_A(H(\cdot, 0), X, X) = 1$. Thus, by the homotopy property (iii), $\mathrm{Ind}_A(H(\cdot, 0), X, D) = \mathrm{Ind}_A(H(\cdot, 1), X, D) = 1$, which implies by (i) that $H(\cdot, 1)$ has a fixed point. □

COROLLARY 3.1. *Let* X *be a retract of a Fréchet space* E *and* H *be a homotopy in* $J_A(X, E)$ *such that* $H(x, 0) \subset X$, *for every* $x \in X$, *and* $H(\cdot, 0)$ *is compact. Then* $H(\cdot, 1)$ *has a fixed point.*

COROLLARY 3.2. *Let* X *be a retract of a Fréchet space* E, D *be an open subset of* X *and* H *be a homotopy in* $J_A(D, E)$. *Assume that* $H(x, 0) = x_0$, *for every* $x \in D$. *Then there exists* $x \in D$ *such that* $x \in H(x, 1)$.

COROLLARY 3.3. *Let* X *be a retract of a Fréchet space* E *and* $\Phi \in J(X)$ *be compact. Then* Φ *has a fixed point.*

REMARK 3.7. If $E = X$ is a Banach space, then it follows from Proposition 3.2 that the "pushing" condition (A_H), related to $J_A(D, E)$, can be reduced to $\mathrm{Fix}(\varphi) \cap \partial D \neq \emptyset$.

Some applications motivate us to consider weaker than (A_H) condition on H. Unfortunately, we cannot use the fixed point index technique described above. The proof of the following theorem is based on a Schauder-type approximation technique (for more details, see [19,22]).

THEOREM 3.9 (Continuation principle). *Let* X *be a closed, convex subset of a Fréchet space* E *and let* $H \in J(X \times [0, 1], E)$ *be compact. Assume that*

(i) $H(x, 0) \subset X$, *for every* $x \in X$,

(ii) *for any* $(x, t) \in \partial X \times [0, 1)$ *with* $x \in H(x, t)$, *there exist open neighbourhoods* U_x *of* x *in* X *and* I_t *of* t *in* $[0, 1)$ *such that* $H((U_x \cap \partial X) \times I_t) \subset X$.

Then there exists a fixed point of $H(\cdot, 1)$.

REMARK 3.8. Note that the convexity of X in Theorem 3.9 is essential only in the infinite-dimensional case. For the proof, we have namely to intersect X with a finite-dimensional subspace L.

Now, we would like to consider condensing maps. Hence, let this time $\varphi : X \multimap X$ be a multivalued map such that

(I) X is a closed, convex subset of a Fréchet space E,

(II) φ is a condensing composition of an R_δ-map $\Phi : X \multimap Y$ and a continuous single-valued map $f : Y \to X$, namely $\varphi = f \circ \Phi$, where Y is an ANR-space.

Assume that $\mathrm{Fix}(\varphi) \cap \partial D \neq \emptyset$, for some open subset $D \subset X$. Since φ is condensing, it has a nonempty compact fundamental set T (see Definition 2.11 and Proposition 2.7(iii)). Let $\mathrm{ind}(\varphi, X, D) = 0$, whenever $\mathrm{Fix}(\varphi) = \emptyset$. Since T is an AR-space, we may choose a retraction $r : X \to T$ in order to define the fixed point index, for the composition

$$\widetilde{\varphi} : \overline{D} \cap T \overset{\Phi}{\multimap} Y \overset{r \circ f}{\multimap} T,$$

by putting

$$\mathrm{ind}(\varphi, X, D) := \mathrm{ind}\big(\widetilde{\varphi}, \widetilde{X}, D \cap T\big),$$

where ind on the right-hand side is defined as in Proposition 3.2. This correct definition is independent of the chosen fundamental set, and so the index has all the appropriate properties as in Proposition 3.2, but (see Remark 3.2 and cf. [37]) the normalization property. Instead of it, a weak normalization property can be formulated as follows:

- (Weak normalization) If f in $\varphi = f \circ \Phi$ is a constant map, i.e. $f(y) = a \notin \partial D$, for each $y \in Y$, then

$$\mathrm{ind}(\varphi, X, D) = \begin{cases} 1, & \text{for } a \in D, \\ 0, & \text{for } a \notin D. \end{cases}$$

As already pointed out above, in the applications of the fixed point theory, we often need to consider maps with values in a Fréchet space and not in a closed convex set. We will also extend our theory to this case.

Again, let E be a Fréchet space and X be a closed and convex subset of E. Let $U \subset X$ be open and consider the map $\varphi \in J_A(U, E)$, where the symbol $J_A(U, E)$ is again reserved for J-maps from U to E satisfying condition (A). The notion of homotopy in J_A will be understood analogously. Thus, $\mathrm{Fix}(\varphi)$ is compact and φ has a compact fundamental set T (see Section 2.3). Set $\mathrm{Ind}_A(\varphi, X, U) = 0$, whenever $\mathrm{Fix}(\varphi) = \emptyset$. Otherwise, let $x_1, \dots, x_n \in \mathrm{Fix}(\varphi)$ such that $\mathrm{Fix}(\varphi) \subset \bigcup_{i=1}^n U_{x_i} =: V$, where U_{x_i} are neighbourhoods of x_i such that $\overline{U}_{x_i} \subset U$ and satisfy condition (A). Then $\varphi|_V : V \multimap X$ is a J-map with compact fundamental set T and satisfies $\mathrm{Fix}(\varphi) \cap \partial V = \emptyset$. Thus, we can define

$$\mathrm{Ind}_A(\varphi, X, U) := \mathrm{ind}(\varphi|_{\overline{V}}, X, V).$$

The independence of this definition of the chosen set V follows from the additivity property. Furthermore, if $\varphi : \overline{U} \multimap X$ has a compact fundamental set and $\mathrm{Fix}(\varphi) \cap \partial U = \emptyset$, then $\mathrm{Ind}_A(\varphi, X, U)$ is defined and

$$\mathrm{Ind}_A(\varphi, X, U) = \mathrm{ind}(\varphi, X, U).$$

The following proposition easily follows from the above argumentation.

PROPOSITION 3.4.

(i) (Existence) *If* $\mathrm{Ind}_A(\varphi, X, U) \neq 0$, *then* $\mathrm{Fix}(\varphi) \neq \emptyset$.

(ii) (Additivity) *Let* $\mathrm{Fix}(\varphi) \subset U_1 \cup U_2$, *where* U_1, U_2 *are open disjoint subsets of* U. *Then*

$$\mathrm{Ind}_A(\varphi, X, U) = \mathrm{Ind}_A(\varphi|_{U_1}, X, U_1) + \mathrm{Ind}_A(\varphi|_{U_2}, X, U_2).$$

(iii) (Homotopy) *Let* $\psi : U \multimap E$ *be homotopic in* J_A *to the map* φ. *Assume that the homotopy* $\chi : U \times [0,1] \multimap E$ *has a compact fundamental set and the set*

$$\Sigma := \big\{(x,t) \in U \times [0,1] \mid x \in \chi(x,t)\big\}$$

is compact. Then

$$\mathrm{Ind}_A(\varphi, X, U) = \mathrm{Ind}_A(\psi, X, U).$$

(iv) (Weak normalization) *Assume that* $\varphi : U \to E$ *is a constant map* $\varphi(x) = a \in E$, *for all* $x \in U$. *Then*

$$\mathrm{Ind}_A(\varphi, X, U) = \begin{cases} 1, & \text{for } a \in U, \\ 0, & \text{for } a \notin U. \end{cases}$$

Using Proposition 3.4, we can easily formulate a continuation principle which is convenient for various applications.

THEOREM 3.10 (Continuation principle). *Let* X *be a closed, convex subset of a Fréchet space* E, *let* $U \subset X$ *be open and let* $\chi : U \times [0,1] \multimap E$ *be a homotopy in* J_A *such that* Σ (*see* (iii) *above*) *is compact. Let* χ *be condensing and assume that there is a condensing* $\varphi \in J(X)$ *such that* $\varphi|_U = \chi(\cdot, 0)$ *and* $\mathrm{Fix}(\varphi) \cap (X \setminus U) = \emptyset$. *Then* $\chi(\cdot, 1)$ *has a fixed point.*

PROOF. The proof follows, in view of the existence property (i) in Proposition 3.4, from the following equations:

$$\mathrm{Ind}_A\big(\chi(\cdot,1), X, U\big) = \mathrm{Ind}_A\big(\chi(\cdot,0), X, U\big),$$

by the homotopy property (iii),

$$\mathrm{Ind}_A\big(\chi(\cdot,0), X, U\big) = \mathrm{Ind}_A(\varphi|_U, X, U) = \mathrm{Ind}_A(\varphi, X, X),$$

by the additivity property (ii). Finally, we see that

$$\mathrm{Ind}_A(\varphi, X, X) = \mathrm{ind}(\varphi, X, X) = 1. \qquad \square$$

COROLLARY 3.4. *Let* $\chi : X \times [0,1] \multimap E$ *be a condensing homotopy in* J_A *such that* $\chi(x,0) \subset X$, *for every* $x \in X$. *Then* $\chi(\cdot,1)$ *has a fixed point.*

REMARK 3.9. If $E = X$ is a Banach space, then it follows from the properties of the above fixed point index $\operatorname{ind}(\chi, X, D)$ that the "pushing" condition (A_H), related to $J_A(D, E)$, can be reduced to $\operatorname{Fix}(\chi) \cap \partial D \neq \emptyset$. If, in particular, $U \subset E$ is an open convex subset such that $\chi(\cdot,0) = \varphi|_U : U \multimap U$, then the same is true even without requiring $\operatorname{Fix}(\varphi) \cap (X \setminus U) = \emptyset$.

4. General methods for solvability of boundary value problems

4.1. *Continuation principles to boundary value problems*

In this part, fixed point principles in Section 3 will be applied to differential equations and inclusions.

At first, we are interested in the existence problems for ordinary differential equations and inclusions in Euclidean spaces on not necessarily compact intervals. Let us start with some definitions.

Let J be an interval in $\mathbb{R}$. We say that a map $x : J \to \mathbb{R}^n$ is *locally absolutely continuous* if x is absolutely continuous on every compact subset of J. The set of all locally absolutely continuous maps from J to $\mathbb{R}^n$ will be denoted by $\mathrm{AC}_{\mathrm{loc}}(J, \mathbb{R}^n)$.

Consider the inclusion

$$\dot{x} \in F(t,x), \tag{4.1}$$

where F is a set-valued *u-Carathéodory map*, i.e. it has i.a. the following properties:

- the set of values of F is nonempty, compact and convex, for all $(t,x) \in J \times \mathbb{R}^n$,
- the map $F(t,\cdot)$ is u.s.c., for almost all $t \in J$,
- the map $F(\cdot,x)$ is measurable, for all $x \in \mathbb{R}^n$.

By a *solution* of the inclusion (4.1), we mean a locally absolutely continuous function x such that (4.1) holds, for almost all $t \in J$.

We recall two known results which are needed in the sequel.

PROPOSITION 4.1 (Cf. [42, Theorem 0.3.4] and Lemma 4.4 below). *Assume that the sequence of absolutely continuous functions* $x_k : K \to \mathbb{R}^n$ (*K is a compact interval*) *satisfies the following conditions*:

- *the set* $\{x_k(t) \mid k \in \mathbb{N}\}$ *is bounded, for every* $t \in K$,
- *there is an integrable function* (*in the sense of Lebesgue*) $\alpha : K \to \mathbb{R}$ *such that*

$$\left|\dot{x}_k(t)\right| \leqslant \alpha(t), \quad \text{for a.a. } t \in K \text{ and for all } k \in \mathbb{N}.$$

Then there exists a subsequence (*denoted just the same*) $\{x_k\}$ *convergent to an absolutely continuous function* $x : K \to \mathbb{R}^n$ *in the following sense*:

(i) $\{x_k\}$ *uniformly converges to* x, *and*
(ii) $\{\dot{x}_k\}$ *weakly converges in* $L^1(K, \mathbb{R}^n)$ *to* $\dot{x}$.

The second one is the well-known (see, e.g., [22, Theorem 1.33 in Chapter I.1]) Mazur theorem.

The following result is crucial.

PROPOSITION 4.2. *Let* $G : J \times \mathbb{R}^n \times \mathbb{R}^m \multimap \mathbb{R}^n$ *be a u-Carathéodory map and let* S *be a nonempty subset of* $\mathrm{AC}_{\mathrm{loc}}(J, \mathbb{R}^n)$. *Assume that*

(i) *there exists a subset* Q *of* $C(J, \mathbb{R}^n)$ *such that, for any* $q \in Q$, *the set* $T(q)$ *of all solutions of the boundary value problem*

$$\begin{cases} \dot{x}(t) \in G\big(t, x(t), q(t)\big), & \text{for a.a. } t \in J, \\ x \in S \end{cases} \tag{4.2}$$

is nonempty,

(ii) $T(Q)$ *is bounded in* $C(J, \mathbb{R}^n)$,

(iii) *there exists a locally integrable function* $\alpha : J \to \mathbb{R}$ *such that*

$$\big|G\big(t, x(t), q(t)\big)\big| = \sup\big\{|y| \mid y \in G\big(t, x(t), q(t)\big)\big\} \leqslant \alpha(t), \quad \text{a.e. in } J,$$

for any pair $(q, x) \in \Gamma_T$, *where* Γ_T *denotes the graph of* T.

Then $T(Q)$ *is a relatively compact subset of* $C(J, \mathbb{R}^n)$. *Moreover, under the assumptions* (i)–(iii), *the multivalued operator* $T : Q \multimap S$ *is u.s.c. with compact values if and only if the following condition is satisfied*:

(iv) *given a sequence* $\{(q_k, x_k)\} \subset \Gamma_T$, *if* $\{(q_k, x_k)\}$ *converges to* (q, x) *with* $q \in Q$, *then* $x \in S$.

PROOF. For the relative compactness of $T(Q)$, it is sufficient to show that all elements of $T(Q)$ are equicontinuous.

By (iii), for every $x \in T(Q)$, we have $|\dot{x}(t)| \leqslant \alpha(t)$, for a.a. $t \in J$, and

$$\big|x(t_1) - x(t_2)\big| \leqslant \left|\int_{t_1}^{t_2} \alpha(s)\, \mathrm{d}s\right|.$$

This implies an equicontinuity of all $x \in T(Q)$.

We show that the set Γ_T is closed (cf. Section 2.1).

Let $\Gamma_T \supset \{(q_k, x_k)\} \to (q, x)$. Let K be an arbitrary compact interval such that α is integrable on K. By conditions (ii) and (iii), the sequence $\{x_k\}$ satisfies the assumptions of Proposition 4.1.

Thus, there exists a subsequence (denoted just the same) $\{x_k\}$, uniformly convergent to x on K (because the limit is unique) and such that $\{\dot{x}_k\}$ weakly converges to $\dot{x}$ in L^1. Therefore, $\dot{x}$ belongs to the weak closure of the set $\mathrm{conv}\{\dot{x}_m \mid m \geqslant k\}$, for every $k \geqslant 1$. By the mentioned Mazur theorem, $\dot{x}$ also belongs to the strong closure of this set. Hence, for every $k \geqslant 1$, there is $z_k \in \mathrm{conv}\{\dot{x}_m \mid m \geqslant k\}$ such that $\|z_k - \dot{x}\|_{L^1} \leqslant 1/k$. This implies that there exists a subsequence $z_{k_l} \to \dot{x}$ a.e. in K.

Let $s \in K$ be such that

$$G(s, \cdot, \cdot) \text{ is u.s.c.}, \quad \lim_{l \to \infty} z_{k_l}(s) = \dot{x}(s), \quad \dot{x}_k(s) \in G\big(s, x_k(s), q_k(s)\big).$$

Let $\varepsilon > 0$. There is $\delta > 0$ such that $G(s, z, p) \subset N_\varepsilon(G(s, x(s), q(s)))$, whenever $|x(s) - z| < \delta$ and $|q(s) - p| < \delta$. But we know that there exists $N \geqslant 1$ such that $|x(s) - x_m(s)| < \delta$ and $|q(s) - q_m(s)| < \delta$, for every $m \geqslant N$. Hence,

$$\dot{x}_k(s) \in G\big(s, x_k(s), q_k(s)\big) \subset N_\varepsilon\big(G\big(s, x(s), q(s)\big)\big).$$

By the convexity of $G(s, x(s), q(s))$, for $k_l \geqslant N$, we have

$$z_{k_l}(s) \in N_\varepsilon\big(G\big(s, x(s), q(s)\big)\big).$$

Thus, $\dot{x}(s) \in \overline{N_\varepsilon(G(s, x(s), q(s)))}$, for every $\varepsilon > 0$. This implies

$$\dot{x}(s) \in G\big(s, x(s), q(s)\big).$$

Since K was arbitrary, $\dot{x}(t) \in G(t, x(t), q(t))$, a.e. in J. □

We can now state one of the main results of this subsection.

THEOREM 4.1. *Consider the boundary value problem*

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in J, \\ x \in S, \end{cases} \tag{4.3}$$

where J is a given real interval, $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ is a u-Carathéodory map and S is a subset of $\mathrm{AC}_{\mathrm{loc}}(J, \mathbb{R}^n)$.

Let $G : J \times \mathbb{R}^n \times \mathbb{R}^n \times [0, 1] \multimap \mathbb{R}^n$ be a u-Carathéodory map (*cf. Definition* 2.10) *such that*

$$G(t, c, c, 1) \subset F(t, c), \quad \text{for all } (t, c) \in J \times \mathbb{R}^n.$$

Assume that

(i) *there exist a retract Q of $C(J, \mathbb{R}^n)$ and a closed bounded subset S_1 of S such that the associated problem*

$$\begin{cases} \dot{x}(t) \in G\big(t, x(t), q(t), \lambda\big), & \text{for a.a. } t \in J, \\ x \in S_1 \end{cases} \tag{4.4}$$

is solvable with an R_δ-set of solutions, for each $(q, \lambda) \in Q \times [0, 1]$,

(ii) *there exists a locally integrable function $\alpha : J \to \mathbb{R}$ such that*

$$\big|G\big(t, x(t), q(t), \lambda\big)\big| \leqslant \alpha(t), \quad \text{a.e. in } J,$$

for any $(q, \lambda, x) \in \Gamma_T$, where T denotes the set-valued map which assigns to any $(q, \lambda) \in Q \times [0, 1]$ the set of solutions of (4.4),

(iii) $T(Q \times \{0\}) \subset Q$,

(iv) *if* $Q \ni q_j \to q \in Q$, $q \in T(q,\lambda)$, *then there exists* $j_0 \in \mathbb{N}$ *such that, for every* $j \geqslant j_0$, $\theta \in [0,1]$ *and* $x \in T(q_j,\theta)$, *we have* $x \in Q$.

Then problem (4.3) *has a solution.*

PROOF. Consider the set

$$Q' = \{ y \in C(J,\mathbb{R}^{n+1}) \mid y(t) = (q(t),\lambda),\ q \in Q,\ \lambda \in [0,1] \}.$$

By Proposition 4.2, we obtain that the set-valued map $T : Q \times [0,1] \multimap S_1$ is u.s.c., and so it belongs to the class $J(Q \times [0,1], C(J,\mathbb{R}^n))$. Moreover, it has a relatively compact image. Assumption (iv) implies that T is a homotopy in $J_A(Q, C(J,\mathbb{R}^n))$. Corollary 3.1 in Section 3 now gives the existence of a fixed point of $T(\cdot,1)$. However, by the hypothesis, it is a solution of (4.3). ☐

Note that the conditions (iii) and (iv) in the above theorem hold if $S_1 \subset Q$.

COROLLARY 4.1. *Consider the boundary value problem* (4.3).
Let $G : J \times \mathbb{R}^n \times \mathbb{R}^n \multimap \mathbb{R}^n$ *be a u-Carathéodory map such that*

$$G(t,c,c) \subset F(t,c), \quad \textit{for all } (t,c) \in J \times \mathbb{R}^n.$$

Assume that

(i) *there exists a retract* Q *of* $C(J,\mathbb{R}^n)$ *such that the associated problem*

$$\begin{cases} \dot{x}(t) \in G(t,x(t),q(t)), & \textit{for a.a. } t \in J, \\ x \in S \cap Q \end{cases} \tag{4.5}$$

has an R_δ*-set of solutions, for each* $q \in Q$,

(ii) *there exists a locally integrable function* $\alpha : J \to \mathbb{R}$ *such that*

$$|G(t,x(t),q(t))| \leqslant \alpha(t), \quad \textit{a.e. in } J,$$

for any $(q,x) \in \Gamma_T$,

(iii) $T(Q)$ *is bounded in* $C(J,\mathbb{R}^n)$ *and* $\overline{T(Q)} \subset S$.

Then problem (4.3) *has a solution.*

Making use of the special case of Theorem 3.3 and modifying appropriately the proof of Theorem 4.1, we can easily obtain the following:

COROLLARY 4.2. *Consider problem* (4.3) *and assume that all the assumptions of Corollary* 4.1 *hold with the convex closed set* Q *and nonempty acyclic sets of solutions* (4.5). *Then the problem* (4.3) *has a solution.*

Let us note that in applications solution sets are, in fact, R_δ-sets.

If, in particular, $J = [a,b]$ (i.e. compact), then Theorem 4.1 can be easily reformulated (see Remark 3.7) as follows.

COROLLARY 4.3. *Consider the boundary value problem* (4.3), *where* $J = [a, b]$ *is a compact interval,* $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ *is a u-Carathéodory map and* $S \subset \mathrm{AC}(J, \mathbb{R}^n)$.

Let $G : J \times \mathbb{R}^n \times \mathbb{R}^n \times [0, 1] \multimap \mathbb{R}^n$ *be a u-Carathéodory map such that* $G(t, c, c, 1) \subset F(t, c)$, *for all* $(t, c) \in J \times \mathbb{R}^n$. *Assume that*

(i) *there exist a* (*bounded*) *retract* Q *of* $C(J, \mathbb{R}^n)$ *such that* $Q \setminus \partial Q$ *is nonempty* (*open*) *and a closed bounded subset* S_1 *of* S *such that the associated problem* (4.4) *is solvable with an* R_δ*-set of solutions, for each* $(q, \lambda) \in Q \times [0, 1]$, *and conditions* (ii) *and* (iii) *in Theorem* 4.1 *hold true,*

(ii) *the solution map* T (*defined in condition* (ii) *of Theorem* 4.1) *has no fixed points on the boundary* ∂Q *of* Q, *for every* $(q, \lambda) \in Q \times [0, 1]$.

Then problem (4.3) *has a solution.*

REMARK 4.1. In the (single-valued) case of Carathéodory ODEs, we can only assume in Theorem 4.1(i), Corollary 4.1(i), Corollary 4.2 and Corollary 4.3(i) that the related linearized problems are uniquely solvable.

Since $C^{(n-1)}(J)$ can be considered as a subspace of $C(J, \mathbb{R}^n)$, we can also apply the previous results to nth-order scalar differential equations and inclusions. To solve an existence problem, one should check suitable a priori bounds for all the derivatives up to the order $n - 1$. Our technique simplifies a work. Let us describe it below.

We need the following lemma [52, Lemma 2.1] related to the Banach space $H^{n,1}(I)$:[2]

LEMMA 4.1. *Let* I *be a compact real interval and let* $a_0, a_1, \dots, a_{n-1} : I \times \mathbb{R}^n \to \mathbb{R}$ *be u-Carathéodory functions. Given any* $q \in C^{(n-1)}(I)$, *consider the following linear* n*th-order differential operator* $L_q : H^{n,1}(I) \to L^1(I)$:

$$L_q(x)(t) = x^{(n)}(t) + \sum_{i=0}^{n-1} a_i\bigl(t, q(t), \dots, q^{(n-1)}(t)\bigr)x^{(i)}(t).$$

Assume there exists a subset Q *of* $C^{(n-1)}(I)$ *and an* L^1*-function* $\beta : I \to \mathbb{R}$ *such that, for any* $q \in Q$ *and any* $i = 0, 1, \dots, n - 1$, *we have*

$$\bigl|a_i\bigl(t, q(t), \dots, q^{(n-1)}(t)\bigr)\bigr| \leqslant \beta(t), \quad \text{a.e. in } I.$$

Then the following two norms are equivalent in $H^{n,1}(I)$:

$$\|x\| = \sum_{i=0}^{n-1} \sup_{t\in I}\bigl|x^{(i)}(t)\bigr| + \int_I \bigl|x^{(n)}(t)\bigr|\,\mathrm{d}t,$$

$$\|x\|_Q = \sup_{t\in I}\bigl|x(t)\bigr| + \sup_{q\in Q} \int_I \bigl|L_q(x)(t)\bigr|\,\mathrm{d}t.$$

---

[2] By $H^{n,1}(I)$, we denote the Banach space of all $C^{(n-1)}$-functions $x : I \to \mathbb{R}$, where I is a compact interval, with absolutely continuous $(n - 1)$th derivatives.

COROLLARY 4.4. *Consider the scalar problem*

$$\begin{cases} x^{(n)}(t) \in \displaystyle\sum_{i=0}^{n-1} a_i\big(t, x(t), \ldots, x^{(n-1)}(t)\big)x^{(i)}(t) \\ \qquad\qquad + F\big(t, x(t), \ldots, x^{(n-1)}(t)\big), \quad \text{for a.a. } t \in J, \\ x \in S, \end{cases} \tag{4.6}$$

where $J \subset \mathbb{R}$, $S \subset C(J)$ and a_i, F are u-Carathéodory maps on $J \times \mathbb{R}^n$.

Suppose that there exists a u-Carathéodory map $G : J \times \mathbb{R}^n \times \mathbb{R}^n \times [0,1] \multimap \mathbb{R}^n$ such that, for every $c \in \mathbb{R}^n$ and $\lambda \in [0,1]$, $G(t, c, c, 1) \subset F(t, c)$, a.e. in J. Then problem (4.6) *has a solution, provided the following conditions are satisfied*:

(i) *there is a retract Q of the space $C^{(n-1)}(J)$ such that, for every $(q, \lambda) \in Q \times [0,1]$, the following problem,*

$$\begin{cases} x^{(n)}(t) \in \displaystyle\sum_{i=0}^{n-1} a_i\big(t, q(t), \ldots, q^{(n-1)}(t)\big)x^{(i)}(t) \\ \qquad\qquad + G\big(t, x(t), \ldots, x^{(n-1)}(t), q(t), \ldots, q^{(n-1)}(t), \lambda\big), \quad \text{for a.a. } t \in J, \\ x \in S \cap Q, \end{cases} \tag{4.7}$$

has an R_δ-set of solutions,

(ii) *there is a locally integrable function $\alpha : J \to \mathbb{R}$ such that, for every $i = 0, \ldots, n-1$:*

$$\big|a_i\big(t, q(t), \ldots, q^{(n-1)}(t)\big)\big| \leqslant \alpha(t), \quad \text{a.e. in } J,$$

and

$$\big|G\big(t, x(t), \ldots, x^{(n-1)}(t), q(t), \ldots, q^{(n-1)}(t), \lambda\big)\big| \leqslant \alpha(t), \quad \text{for a.e. } t \in J,$$

for each $(q, \lambda, x) \in Q \times [0,1] \times C^{(n-1)}(J)$ satisfying (4.7),

(iii) *$T(Q \times \{0\}) \subset Q$, where T denotes the set-valued map which assigns to any $(q, \lambda) \in Q \times [0,1]$ the set of solutions of* (4.7),

(iv) *the set $T(Q \times [0,1])$ is bounded in $C(J)$ and its $C^{(n-1)}$-closure is contained in S (in particular, this holds if $S \cap C^{(n-1)}(J)$ is closed in $C^{(n-1)}(J)$),*

(v) *if $\{q_j\} \subset Q$ converges to $q \in Q$, $q \in T(q, \lambda)$ in $C^{(n-1)}(J)$, then there exists $j_0 \in \mathbb{N}$ such that, for every $j \geqslant j_0$, $\theta \in [0,1]$ and $x \in T(q_j, \theta)$, we have $x \in Q$.*

PROOF. We construct a new problem in the following way:

Define $\widetilde{F} : J \times \mathbb{R}^n \multimap \mathbb{R}^n$,

$$\widetilde{F}\big(t, x(t), \ldots, x^{(n-1)}(t)\big) = F\big(t, x(t), \ldots, x^{(n-1)}(t)\big) + \sum_{i=0}^{n-1} a_i\big(t, x(t), \ldots, x^{(n-1)}(t)\big)x^{(i)}(t).$$

Denote $\overline{x}(t) = (x(t), \dots, x^{(n-1)}(t)) \in \mathbb{R}^n$ and define $F' : J \times \mathbb{R}^n \multimap \mathbb{R}^n$,

$$F'\big(t, \overline{x}(t)\big) = \big\{\big(\dot{x}(t), \dots, x^{(n-1)}(t), y\big) \mid y \in \widetilde{F}\big(t, x(t), \dots, x^{(n-1)}(t)\big)\big\}.$$

So, we have a problem

$$\begin{cases} \dot{\overline{x}}(t) \in F'\big(t, \overline{x}(t)\big), & \text{for a.a. } t \in J, \\ \overline{x} \in \overline{S}, \end{cases} \tag{4.8}$$

where $\overline{S}$ is an image of $S \cap C^{(n-1)}(J)$ via the inclusion $i : C^{(n-1)}(J) \to C(J, \mathbb{R}^n)$.

Analogously, we find the associated problem

$$\begin{cases} \dot{\overline{x}}(t) \in G'\big(t, \overline{x}(t), \overline{q}(t), \lambda\big), & \text{for a.a. } t \in J, \\ \overline{x} \in \overline{S} \cap \overline{Q}. \end{cases} \tag{4.9}$$

Notice that

(1) $G'(t, \overline{x}(t), \overline{q}(t), 1) \subset F'(t, \overline{x}(t))$,
(2) the set $\overline{Q} = i(Q)$ is a retract of $C(J, \mathbb{R}^n)$,
(3) $\overline{S} \subset \mathrm{AC}_{\mathrm{loc}}(J, \mathbb{R}^n)$,
(4) for every $(q, \lambda) \in Q \times [0, 1]$, the sets of solutions of the problems (4.7) and (4.9) are the same,
(5) $\overline{T}(\overline{Q} \times [0, 1]) \subset \overline{S}$, where $\overline{T}$ is a suitable map corresponding to T

and

$$\begin{aligned} \big|G'\big(t, \overline{x}(t), \overline{q}(t), \lambda\big)\big| &\leqslant \big|G\big(t, x(t), \dots, x^{(n-1)}, q(t), \dots, q^{(n-1)}(t), \lambda\big)\big| \\ &\quad + \sum_{i=0}^{n-1} \big|a_i\big(t, q(t), \dots, q^{(n-1)}(t)\big)\big| \big|x^{(i)}(t)\big| \\ &\leqslant \alpha(t) + \alpha(t) \sum_{i=0}^{n-1} \big|x^{(i)}(t)\big|. \end{aligned}$$

Since $T(Q \times [0, 1])$ is bounded in $C(J)$, there exists a positive continuous function $m : J \to \mathbb{R}$ such that $|x(t)| \leqslant m(t)$, for all $t \in J$ and any $x \in T(Q \times [0, 1])$. We will show that $T(Q \times [0, 1])$ is also bounded in $C^{(n-1)}(J)$. It is sufficient to prove that, for any compact subinterval I in J, there is a constant $M > 0$ such that

$$p_I(x) = \sum_{i=0}^{n-1} \sup \big|x^{(i)}(t)\big| \leqslant M, \quad \text{for all } x \in T\big(Q \times [0, 1]\big).$$

Let $I \subset J$ be an arbitrary compact interval. Using the notation in Lemma 4.1, we see that $p_I(x) \leqslant \|x\|$ and, by the equivalence of norms,

$$\|x\| \leqslant c\|x\|_Q \leqslant c\left(\max_{t \in I} m(f) + \int_I \alpha(t)\, \mathrm{d}t\right) \leqslant M.$$

We conclude that $T(Q \times [0,1])$ is bounded in $C^{(n-1)}(J)$ which implies that $\overline{T}(\overline{Q} \times [0,1])$ is bounded in $C(J, \mathbb{R}^n)$. Moreover, there exists a continuous function $\varphi : J \to \mathbb{R}$ such that

$$\left|G'\left(t, \overline{x}(t), \overline{q}(t), \lambda\right)\right| \leqslant \alpha(t)\left(1 + \varphi(t)\right).$$

Obviously, the right-hand side of the above inequality is a locally integrable function.

Finally, an easy computation shows that the condition (iv) in Theorem 4.1 holds for $\overline{Q}$ and $\overline{T}$. By Theorem 4.1, there exists a solution of (4.8) as well as the one of (4.6). □

The same argument as in Corollary 4.1 shows how to modify Corollary 4.4 for the following scalar problem, namely

$$\begin{cases} x^{(n)}(t) \in \displaystyle\sum_{i=0}^{n-1} a_i\left(t, x(t), \dots, x^{(n-1)}(t)\right)x^{(i)}(t) \\ \qquad\qquad + F\left(t, x(t), \dots, x^{(n-1)}(t)\right), \quad \text{for a.a. } t \in J, \\ x \in S, \end{cases} \tag{4.10}$$

where $J \subset \mathbb{R}$, $S \subset C(J)$ and a_i, F are u-Carathéodory maps on $J \times \mathbb{R}^n$, by means of the following linearized problem

$$\begin{cases} x^{(n)}(t) \in \displaystyle\sum_{i=0}^{n-1} a_i\left(t, q(t), \dots, q^{(n-1)}(t)\right)x^{(i)}(t) \\ \qquad\qquad + G\left(t, x(t), \dots, x^{(n-1)}(t), q(t), \dots, q^{(n-1)}(t)\right), \quad \text{for a.a. } t \in J, \\ x \in S \cap Q, \end{cases} \tag{4.11}$$

where Q is a retract of the space $C^{(n-1)}(J)$.

Theorem 3.9 in Section 3.3 gives similar consequences as those of Theorem 3.8 in Section 3.3. Unfortunately, the weakness of the assumption on solutions causes that we have to assume the convexity of the set Q. In spite of it, the results given below are important because of the applications.

THEOREM 4.2. *Consider the boundary value problem* (4.3), *where J is a given real interval, $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ is a u-Carathéodory map and S is a subset of* $\mathrm{AC}_{\mathrm{loc}}(J, \mathbb{R}^n)$.

Let $G : J \times \mathbb{R}^n \times \mathbb{R}^n \times [0,1] \multimap \mathbb{R}^n$ be as in Theorem 4.1. *Assume that the assumptions* (i)–(iii) *of Theorem* 4.1 *hold, with the convexity of the set Q, and*

(iv) *if $\partial Q \times [0,1] \supset \{(q_j, \lambda_j)\}$ converges to $(q, \lambda) \in \partial Q \times [0,1]$, $q \in T(q, \lambda)$, then there exists $j_0 \in \mathbb{N}$ such that, for every $j \geqslant j_0$, and $x_j \in T(q_j, \lambda_j)$, we have $x_j \in Q$.*

Then problem (4.3) *has a solution.*

The proof can be obtained immediately by using our continuation principle presented in Theorem 3.9 in Section 3.3.

REMARK 4.2. In the (single-valued) case of Carathéodory ODEs, we can again only assume in Corollary 4.4(i) and Theorem 4.2 that the linearized problems are uniquely solvable. If the associated problem (4.4) for G is so uniquely solvable, for every $(q, \lambda) \in Q \times [0, 1]$, then, by continuity of T, we can reformulate the above condi tion (iv) as follows:

(iv′) if $\{(x_j, \lambda_j)\}$ is a sequence in $S_1 \times [0, 1]$, with $\lambda_j \to \lambda \in [0, 1)$ and x_j is converging to a solution $x \in Q$ of (4.4) (corresponding to (x, λ)), then x_j belongs to Q, for j sufficiently large.

Now, we are interested in the existence of several solutions of problem (4.3). For this, the Nielsen theory developed in Section 3.2 will be applied. It will be convenient to use the following definition.

DEFINITION 4.1. We say that the mapping $T : Q \multimap U$ is *retractible onto* Q, where U is an open subset of $C(J, \mathbb{R}^n)$ containing Q, if there is a (continuous) retraction $r : U \to Q$ and $p \in U \setminus Q$ with $r(p) = q$ implies that $p \notin T(q)$.

Its advantage consists in the fact that, for a retractible mapping $T : Q \multimap U$ onto Q with a retraction r in the sense of Definition 4.1, its composition with r, $r|_{T(Q)} \circ T : Q \multimap Q$, has a fixed point $\widehat{q} \in Q$ if and only if $\widehat{q}$ is a fixed point of T.

The following principal statement characterizes the matter.

THEOREM 4.3. *Let* $G : J \times \mathbb{R}^n \times \mathbb{R}^n \multimap \mathbb{R}^n$ *be u-Carathéodory map* (*cf. Definition* 2.10) *and assume that*

(i) *there exists a closed, connected subset* Q *of* $C(J, \mathbb{R}^n)$ *with a finitely generated abelian fundamental group such that, for any* $q \in Q$, *the set* $T(q)$ *of all solutions of the linearized problem* (4.2) *is* R_δ,

(ii) $T(Q)$ *is bounded in* $C(J, \mathbb{R}^n)$ *and* $\overline{T(Q)} \subset S$,

(iii) *there exists a locally integrable function* $\alpha : J \to \mathbb{R}$ *such that*

$$\bigl|G\bigl(t, x(t), q(t)\bigr)\bigr| := \sup\bigl\{|y| \mid y \in G\bigl(t, x(t), q(t)\bigr)\bigr\} \leqslant \alpha(t), \quad \textit{a.e. in } J,$$

for any pair $(q, x) \in \Gamma_T$, *where* Γ_T *denotes the graph of* T.

Assume, furthermore, that

(iv) *the operator* $T : Q \multimap U$, *related to* (4.2), *is retractible onto* Q *with a retraction* r *in the sense of Definition* 4.1.

At last, let

$$G(t, c, c) \subset F(t, c) \tag{4.12}$$

for a.a. $t \in J$ *and any* $c \in \mathbb{R}^n$. *Then the original problem* (4.3) *admits at least* $N(r|_{T(Q)} \circ T)$ *solutions belonging to* Q, *where* N *stands for the Nielsen number defined in Definition* 3.2 *in Section* 3.2.

PROOF. By the hypothesis, Q is a connected (metric) ANR-space with a finitely generated abelian fundamental group and $T(q)$ is an R_δ-mapping. Since T is also, according to

Proposition 4.2, u.s.c. and such that $\overline{T(Q)}$ is compact, $r \circ T$ is compact and admissible. This follows from the commutativity of the following diagram:

$$\begin{tikzcd}
Q \arrow[r, "T"] & U \arrow[r, "r"] & Q \\
& \Gamma_T \arrow[ul, Rightarrow, "p_T"'] \arrow[u, "q_T"'] \arrow[ur, "r\circ q_T"'] &
\end{tikzcd}$$

where (p_T, q_T) is a pair of natural projections of the graph Γ_T and p_T is Vietoris. Therefore, according to Theorem 3.4 in Section 3.2, $(p_T, r|_{T(Q)} \circ q_T)$ admits at least $N(r|_{T(Q)} \circ T(\cdot))$ coincidence points. Because of Definition 4.1, they represent the solutions of problem (4.2) and, in view of (4.12), they also satisfy the original problem (4.3). $\square$

REMARK 4.3. In the (single-valued) case of Carathéodory ODEs, we can only assume in Theorem 4.3(i) that the linearized problem (4.2) is uniquely solvable. Moreover, the requirement that the fundamental group $\pi(Q)$ of Q to be finitely generated and abelian can be then omitted (see Section 3.2).

Furthermore, we will consider boundary value problems on arbitrary (possibly infinite) intervals for differential inclusions in Banach spaces. We start with some definitions.

Let E be a Banach space with the norm $\|\cdot\|$. Denote by $C(J, E)$ the space of all continuous functions $x : J \to E$ with the locally convex topology generated by the uniform convergence on compact subintervals of J (possibly, the whole $\mathbb{R}$). This topology is completely metrizable, and thus $C(J, E)$ is a Fréchet space.

Recall that a mapping $x : J \to E$ is locally absolutely continuous if x is absolutely continuous on every compact subinterval of J. Unfortunately, in general, on each interval $[a, b] \subset J$, there need not exist $\dot{x}(t)$ (in the sense of Fréchet), for almost all (a.a.) $t \in [a, b]$ with $\dot{x} \in L^1([a, b], E)$ (the set of all Bochner integrable functions $[a, b] \to E$) and so need not be

$$x(t) = x_0 + \int_a^t \dot{x}(s)\,\mathrm{d}s.$$

It is so if E satisfies the Radon–Nikodym property, in particular, if E is reflexive. Moreover, we have the following result (cf. [48]).

LEMMA 4.2. *Suppose $x : [a, b] \to E$ is absolutely continuous, $\dot{x}$ exists a.e., and*

$$\big\|\dot{x}(t)\big\| \leqslant y(t), \quad \textit{a.e., for some } y \in L^1\big([a, b], \mathbb{R}\big).$$

Then $\dot{x} \in L^1([a, b], E)$ and

$$\int_\tau^t \dot{x}(s)\,\mathrm{d}s = x(t) - x(\tau) \quad \big(t, \tau \in [a, b]\big). \tag{4.13}$$

The set of all locally absolutely continuous functions from J to E, satisfying all the above properties, will be denoted by $\mathrm{AC}_{\mathrm{loc}}(J, E)$.

Consider now the differential inclusion

$$\dot{x} \in F(t, x), \tag{4.14}$$

where $F : J \times E \multimap E$ is a u-Carathéodory map, i.e.

(C1) $F(t, x)$ is nonempty, compact and convex, for every $(t, x) \in J \times E$,

(C2) $F(t, \cdot)$ is u.s.c., for a.a. $t \in J$,

(C3) $F(\cdot, x)$ is strongly measurable (cf. Definition 2.9), on every compact interval $[a, b]$, for each $x \in E$.

By a *solution* of this differential inclusion we mean again a map $x \in \mathrm{AC}_{\mathrm{loc}}(J, E)$ satisfying (4.14), for a.a. $t \in J$.

To a u-Carathéodory map F, we associate the Nemytskiĭ (or superposition) operator $N_F : C(J, E) \multimap L^1_{\mathrm{loc}}(J, E)$ given by

$$N_F(x) := \left\{ f \in L^1_{\mathrm{loc}}(J, E) \mid f(t) \in F\big(t, x(t)\big), \text{ a.e. on } J \right\},$$

for each $x \in C(J, E)$.

In the sequel, we will need the following lemma (see [100, p. 88] and cf. Remark 2.1).

LEMMA 4.3. *Let $[a, b]$ be a compact interval. Let $F : [a, b] \times E \multimap E$ be a u-Carathéodory mapping and assume in addition that, for every nonempty bounded set $\Omega \subset E$, there exists $\nu = \nu(\Omega) \in L^1([a, b])$ such that*

$$\big\| F(t, x) \big\| := \sup\big\{ \|z\| \mid z \in F(t, x) \big\} \leqslant \nu(t),$$

for a.e. $t \in [a, b]$ and every $x \in \Omega$. Then the Nemytskiĭ operator

$$N_F : C\big([a, b], E\big) \multimap L^1\big([a, b], E\big)$$

has nonempty, convex values. Moreover, given sequences $\{x_n\} \subset C([a, b], E)$ and $\{f_n\} \subset L^1([a, b], E)$, $f_n \in N_F(x_n)$, $n \geqslant 1$, such that $x_n \to x$ in $C([a, b], E)$ and $f_n \to f$ weakly in $L^1([a, b], E)$, then $f \in N_F(x)$.

The following lemma extends Proposition 4.1 to infinite-dimensional spaces (see again [42, Theorem 0.3.4]).

LEMMA 4.4. *Assume that a sequence $\{x_k \mid [a, b] \to E\}$ of AC-maps satisfies the following conditions*:

(i) *$\{x_k(t)\}$ is relatively compact, for each $t \in [a, b]$*,

(ii) *there exists $\alpha \in L^1([a, b])$ such that $\|\dot{x}_k(t)\| \leqslant \alpha(t)$, for a.a. $t \in [a, b]$.*

Then there exists a subsequence (*again denoted by $\{x_k\}$*) *that converges to an absolutely continuous map $x : [a, b] \to E$ in the following sense*:

(iii) *$x_k \to x$ in $C([a, b], E)$*,

(iv) $\dot{x}_k \to \dot{x}$ *weakly in* $L^1([a,b], E)$.

PROPOSITION 4.3. *Let* $G : J \times E \times E \multimap E$ *be a u-Carathéodory map and let* S *be a nonempty subset of* $\mathrm{AC}_{\mathrm{loc}}(J, E)$. *Assume that*:

(i) *there exists a closed* $Q \subset C(J, E)$ *such that, for any* $q \in Q$, *the boundary value problem*

$$\begin{cases} \dot{x}(t) \in G\big(t, x(t), q(t)\big), & \text{for a.a. } t \in J, \\ x \in S, \end{cases}$$

has a solution. Denote by $T : Q \multimap S$ *the solution mapping.*

(ii) *There exist* $\alpha, \beta, \gamma \in L^1_{\mathrm{loc}}(J)$ *such that*

$$\big\|G(t, x, y)\big\| \leqslant \alpha(t) + \beta(t)\|x\| + \gamma(t)\|y\|,$$

for a.a. $t \in J$ *and every* $(x, y) \in E^2$.

(iii) *If* $\{(q_n, x_n)\}$ *is a sequence in the graph of* T *and* $(q_n, x_n) \to (q, x)$, *then* $x \in S$.

Then $T : Q \multimap S$ *has a closed graph* (S *is endowed with the topology of* $C(J, E)$).

PROOF. Let $\{(q_n, x_n)\}$ be an arbitrary sequence in the graph of T, i.e. $x_n \in T(q_n)$, for every $n \in \mathbb{N}$, and assume that $(q_n, x_n) \to (q_0, x_0)$. Thus, we see that

$$\dot{x}_n(t) \in G\big(t, x_n(t), q_n(t)\big), \quad \text{for a.a. } t \in J,$$

and $x_n \in S$. Then $q_0 \in Q$ and, by assumption (iii), $x_0 \in S$.

Now, let $[a, b]$ be an interval in J. Using assumption (ii), we see that the sequence $\{x_n\}$ satisfies the assumptions of Lemma 4.4. Thus, $\{x_n\}$ converges uniformly on $[a, b]$ to x_0 (because this limit is unique) and $\{\dot{x}_n\}$ converges to $\dot{x}_0$, weakly in $L^1([a, b], E)$. Using Lemma 4.3, it follows that $\dot{x}_0(t) \in G(t, x_0(t), q_0(t))$, for a.a. $t \in [a, b]$. Since $[a, b]$ was arbitrary, we see that indeed $\dot{x}_0(t) \in G(t, x_0(t), q_0(t))$, for a.a. $t \in J$ and $x_0 \in T(q_0)$. □

As another of the main results of this subsection, we can formulate the following continuation principle.

THEOREM 4.4. *Consider the boundary value problem* (*e.g., in a reflexive Banach space* E; *cf. Lemma* 4.2):

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in J, \\ x \in S, \end{cases} \tag{4.15}$$

where $F : J \times E \multimap E$ *is a u-Carathéodory map and* S *is a subset of* $\mathrm{AC}_{\mathrm{loc}}(J, E)$. *Let* $G : J \times E \times E \times [0, 1] \multimap E$ *be a u-Carathéodory map* (*cf. Definition* 2.10) *such that*

$$G(t, c, c, 1) \subset F(t, c), \quad \text{for all } (t, c) \in J \times E. \tag{4.16}$$

Assume that:

(i) *there exists a closed, convex* $Q \subset C(J, E)$ *and a closed subset* S_1 *of* S *such that the problem*

$$\begin{cases} \dot{x}(t) \in G\big(t, x(t), q(t), \lambda\big), & \text{for a.a. } t \in J, \\ x \in S_1 \end{cases}$$

is solvable with an R_δ*-set* $T(q, \lambda)$, *for each* $(q, \lambda) \subset Q \times [0, 1]$.

(ii) *There exist* $\alpha, \beta, \gamma \in L^1_{\mathrm{loc}}(I)$ *such that*

$$\big\|G(t, x, y, \lambda)\big\| \leqslant \alpha(t) + \beta(t)\|x\| + \gamma(t)\|y\|,$$

for a.a. $t \in J$, *every* $(x, y) \in E^2$ *and every* $\lambda \in [0, 1]$.

(iii) T *is quasi-compact, i.e.* T *maps compact subsets onto compact subsets, and there exists a measure of noncompactness* μ *in the sense of Definitions* 2.7 *and* 2.8 *in Section* 2.2 *such that, for each* $\Omega \subset Q$, *if*

$$\mu\big(T\big(\Omega \times [0, 1]\big)\big) \geqslant \mu(\Omega),$$

then Ω *is relatively compact.*

(iv) $T(Q \times \{0\}) \subset Q$.

(v) *For each* $\lambda_0 \in [0, 1]$ *and* $q \in T(q_0, \lambda_0)$, *if* $q_n \to q_0$ *in* Q, *then there is* $n_0 \in \mathbb{N}$ *such that, for each* $n \geqslant n_0$, $\lambda \in [0, 1]$ *and* $x \in T(q_n, \lambda)$, *we have* $x \in Q$.

Then problem (4.15) *has a solution.*

PROOF. Using Proposition 4.3, we see that the map $T : Q \times [0, 1] \multimap S_1$ has a closed graph. Since T is also quasi-compact (assumption (iii)), we can easily derive that T is indeed an u.s.c. set-valued map (see, e.g., [78, Theorem 1.1.12]). From assumption (i), we get therefore that $T \in J(Q \times [0, 1], C(\mathbb{R}, E))$ and assumption (iii) implies that T is also μ-condensing. By (v), we finally see that T is a homotopy in J_A, and thus Corollary 3.4 in Section 3.3 implies the existence of a fixed point of $T(\cdot, 1)$. However, by the inclusion (4.16), it is a solution of (4.15). □

REMARK 4.4. As we can see, Theorem 4.4 extends Theorem 4.1 into the infinite-dimensional setting, when replacing $\mathbb{R}^n$ by a real Banach space. On the other hand, this is possible with some loss, namely Q is only convex and the solution operator T is assumed to be quasi-compact, additionally. Because of those restrictions, we are unfortunately unable to establish a full infinite dimensional analogy of Theorem 4.1.

If, in particular, $J = [a, b]$ (i.e. compact), then Theorem 4.4 can be simplified, in view of Remark 3.9, similarly as Corollary 4.3 w.r.t. Theorem 4.1, as follows.

COROLLARY 4.5. *Consider the problem*

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in [a, b], \\ x \in S, \end{cases} \tag{4.17}$$

where $F : [a, b] \times E \multimap E$ *is a u-Carathéodory map and* S *is a subset of absolutely continuous functions* $x : [a, b] \to E$, *all in a reflexive Banach space* E. *Let* $G : [a, b] \times E \times E \times [0, 1] \multimap E$ *be a u-Carathéodory map such that*

$$G(t, c, c, 1) \subset F(t, c), \quad \textit{for all } (t, c) \in [a, b] \times E. \tag{4.18}$$

Furthermore, assume that

(i) *there exists a convex, bounded subset* $Q \subset C([a, b], E)$ *such that* $Q \setminus \partial Q$ *is nonempty and a closed subset* S_1 *of* S *such that the problem*

$$\begin{cases} \dot{x}(t) \in G\big(t, x(t), q(t), \lambda\big), & \textit{for a.a. } t \in [a, b], \\ x \in S_1 \end{cases}$$

is solvable with R_δ*-set* $T(q, \lambda)$, *for each* $(q, \lambda) \in Q \times [0, 1]$.

(ii) *There exists* $\alpha, \beta, \gamma \in Ł^1([a, b])$ *such that*

$$\big\|G(t, x, y, \lambda)\big\| \leqslant \alpha(t) + \beta(t)\|x\| + \gamma(t)\|y\|,$$

for a.a. $t \in [a, b]$, *every* $(x, y) \in E^2$ *and every* $\lambda \in [0, 1]$.

(iii) T *is quasi-compact, i.e.* T *maps compact subsets onto compact subsets, and there exists a measure of noncompactness* μ (*see Definitions* 2.7 *and* 2.8) *such that, for each* $\Omega \subset Q$, *if*

$$\mu\big(T\big(\Omega \times [0, 1]\big)\big) \geqslant \mu(\Omega),$$

then Ω *is relatively compact.*

(iv) $T(Q \times \{0\}) \subset Q$.

(v) *The map* T *has no fixed points on the boundary* ∂Q *of* Q, *for every* $(q, \lambda) \in Q \times [0, 1]$.

Then problem (4.17) *has a solution.*

PROOF. We can proceed quite analogously as in the proof of the foregoing Theorem 4.4. The only difference consists of modifying Corollary 3.4 in the sense of Remark 3.9, both in Section 3.3. □

REMARK 4.5. In the (single-valued) case of Carathéodory ODEs, we can again only assume in Theorem 4.4(i) and Corollary 4.5(i) that the linearized problems are uniquely solvable.

Sometimes it is convenient to consider the asymptotic problems sequentially. For this purpose, it can be useful to employ

PROPOSITION 4.4. *Let* $J_1 \subset J_2 \subset \cdots$ *be compact intervals such that* $J = \bigcup_{m=1}^{\infty} J_m$ *and* $t_0 \in J_1$. *Let* $F : J \times E \multimap E$ *be a u-Carathéodory mapping with nonempty, compact and convex values. Assume, furthermore, that*

(i) *There are* $\alpha, \beta \in L^1_{\mathrm{loc}}(J, \mathbb{R})$ *with*

$$\|F(t,x)\| \leqslant \alpha(t) + \beta(t)\|x\|.$$

(ii) *There is some Carathéodory mapping* $g : J \times [0,\infty) \to [0,\infty)$ (*called a Kamke function*) *such that the only nonnegative measurable solution of*

$$x(t) \leqslant \left| \int_{t_0}^{t} g\big(s, x(s)\big)\, \mathrm{d}s \right|$$

is 0 (*a.e.*), *and such that, for a.a.* $t \in J$, $\gamma(F(\{t\} \times C)) \leqslant g(t, \gamma(C))$, *for countable, bounded subsets* $C \subset E$, *where* γ *denotes the Hausdorff measure of noncompactness.*

(iii) E *has the so called retraction property in the sense of [99], e.g.,* E *is separable or reflexive.*

If $x_m \in \mathrm{AC}(J_m, E)$ *satisfies*

$$\dot{x}_m(t) \in F\big(t, x_m(t)\big), \quad \textit{for a.a. } t \in J_m,\ m \in \mathbb{N},$$

and $\{x_m(t_0) \mid m \in \mathbb{N}\}$ *is a relatively compact set, then there is a solution* $x \in \mathrm{AC}_{\mathrm{loc}}(J, E)$ *of the inclusion* $\dot{x}(t) \in F(t, x(t))$, *for a.a.* $t \in J$, *such that, for some subsequence,*

$$x_{m_k} \to x, \quad \textit{uniformly on each } J_m,$$

and

$$\dot{x}_{m_k} \to \dot{x}, \quad \textit{weakly in } L^1(J, E).$$

If still

(iv) $\sup\{\|x_m(t)\| \mid m \in \mathbb{N},\ t \in J_m\} < \infty$

and the values of x_m, $m \in \mathbb{N}$, *are located in a closed subdomain* $\mathcal{D}$ *of* E, *then there exists an entirely bounded solution* x *on* J *with* $x(t) \in \mathcal{D}$, *for all* $t \in \mathbb{R}$.

PROOF. By (i) and the well-known Gronwall inequality (see, e.g., [72]), we get the a priori estimates

$$\|x_m(t)\| \leqslant \widetilde{x}(t) \quad \text{and} \quad \|\dot{x}_m(t)\| \leqslant \widetilde{x}(t),$$

for some $\widetilde{x} \in L^1_{\mathrm{loc}}(J, \mathbb{R})$.

We claim that $\{x_m(t) \mid m \geqslant m_t\}$ is a relatively compact set, for a.a. $t \in J$, where $m_t = \min\{m \mid t \in J_m\}$. To show it, put $h(t) := \gamma(\{x_m(t) \mid m \geqslant m_t\})$. Then h is measurable (for more details, see Proposition 11.12 in [99]). Moreover, by means of Proposition 11.12 in [99] and (ii), we obtain

$$h(t) = \gamma\big(\{x_m(t) - x_m(t_0) \mid m \geqslant m_t\}\big) = \gamma\left(\left\{ \int_{t_0}^{t} \dot{x}_m(s)\, \mathrm{d}s \;\middle|\; m \geqslant m_t \right\}\right)$$

$$\leqslant \left| \int_{t_0}^{t} \gamma\big(\{\dot{x}_m(s) \mid m \geqslant m_t\}\big)\,\mathrm{d}s \right| \leqslant \left| \int_{t_0}^{t} g\big(s, h(s)\big)\,\mathrm{d}s \right|.$$

Applying (ii) again, we arrive at $h(t) = 0$, for a.a. $t \in J$, as claimed.

Since $F(t, \cdot)$ maps compact sets into compact sets, $\{\dot{x}_m(t) \mid m \geqslant m_t\}$ becomes relatively compact as well, for a.a. $t \in J$. An application of the standard diagonalization argument implies, jointly with Lemma 4.4, the existence of a subsequence such that $x_{m_k} \to x$, uniformly on each J_m, and $\dot{x}_{m_k} \to \dot{x}$, weakly in $L^1(J, E)$, where $x \in \mathrm{AC}_{\mathrm{loc}}(J, E)$.

It follows from Lemma 4.3 that $\dot{x}(t) \in F(t, x(t))$. Since the remaining part of the assertion is implied by the foregoing one (just proved) and (iv), the proof is completed. □

REMARK 4.6. Let E be a Banach space and assume that $F : \mathbb{R} \times E \multimap E$ is a u-Carathéodory mapping (cf. Definition 2.10) such that

$$\mu\big(F(t, B)\big) \leqslant k(t)\mu(B), \quad \text{for bounded subsets } B \subset E,\ t \in \mathbb{R},$$

where $k \in L^1_{\mathrm{loc}}(\mathbb{R})$ and μ denotes either the Kuratowski MNC α or the Hausdorff MNC γ. Then it is well known (see, e.g., [58, Theorem 9.2 and Remark 9.5.4 in Chapter 4.9.3]) that the initial value problem

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in [-m, m],\ m \in \mathbb{N}, \\ x(0) = x_0, \end{cases}$$

admits a solution $x_m \in \mathrm{AC}([-m, m])$, for each $m \in \mathbb{N}$, i.e. $x \in \mathrm{AC}_{\mathrm{loc}}(\mathbb{R}, E)$. If, in particular, the values of x_m, $m \in \mathbb{N}$, are located in a given bounded, closed subdomain $\mathcal{D}$ of E, then there exists an entirely bounded solution $x \in \mathrm{AC}_{\mathrm{loc}}(\mathbb{R}, E)$ on $\mathbb{R}$, $x(0) = x_0$, with values in $\mathcal{D}$, provided E is separable or reflexive. It is namely enough to apply Proposition 4.4, for $t_0 = 0$ and $g(t, x) := 2k(t)x$. Such special g is a Kamke function by means of the Gronwall inequality.

If in particular, $E = \mathbb{R}^n$, then (ii) and (iii) hold automatically. Hence, Proposition 4.4 can be then simplified as follows.

PROPOSITION 4.5. *Let $F : \mathbb{R} \times \mathbb{R}^n \multimap \mathbb{R}^n$ be a u-Carathéodory mapping with nonempty, compact and convex values, satisfying* (i) *in Proposition* 4.4, *for $J = (-\infty, \infty)$. Then, for every $x_0 \in \mathbb{R}^n$, there exists a solution $x \in \mathrm{AC}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ of the Cauchy problem*

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big), & \textit{for a.a. } t \in (-\infty, \infty), \\ x(0) = x_0. \end{cases}$$

Let $\{x_m(t)\}$ be a sequence of absolutely continuous functions such that

(i) *For every $m \in \mathbb{N}$, $x_m \in \mathrm{AC}([-m, m], \mathbb{R}^n)$ is a solution of*

$$\dot{x}(t) \in F\big(t, x(t)\big), \quad \textit{for a.a. } t \in [-m, m],$$

(ii) $\sup\{|x_m(t)| \mid m \in \mathbb{N},\ t \in [-m,m]\} := M < \infty$ *and* $x_m(t) \in \mathcal{D} \subset \mathbb{R}^n$, *for every* $t \in [-m,m]$.

Then there exists an entirely bounded solution $x \in \mathrm{AC}_{\mathrm{loc}}(\mathbb{R},\mathbb{R}^n)$ *of the inclusion*

$$\dot{x}(t) \in F\bigl(t, x(t)\bigr), \quad \text{for a.a. } t \in (-\infty,\infty),$$

such that

$$\sup_{t\in\mathbb{R}} \bigl|x(t)\bigr| \leqslant M \ (<\infty) \quad \text{and} \quad x(t) \in \mathcal{D}, \quad \text{for all } t \in \mathbb{R}.$$

4.2. *Topological structure of solution sets*

In this part, various methods for investigating the topological structure of solution sets, required in statements of the foregoing subsection, will be presented. Both initial and boundary value problems will be considered.

The classical result, due to F.S. De Blasi and J. Myjak in [57], deals with Cauchy problems for the u-Carathéodory differential inclusions in Euclidean spaces:

$$\begin{cases} \dot{x} \in F(t,x), \\ x(0) = x_0, \end{cases} \tag{4.19}$$

where $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ is a u-Carathéodory mapping, i.e. a multivalued mapping, satisfying conditions from the beginning of Section 4.1 (cf. Definition 2.10), and such that

$$\bigl|F(t,x)\bigr| \leqslant \alpha + \beta|x|, \quad \text{for all } t \in J,\ x \in \mathbb{R}^n,$$

where α, β are nonnegative constants.

THEOREM 4.5. *Problem* (4.19), *where J is a compact interval*, *has under the above assumptions an R_δ-set of solutions.*

We omit the proof of this theorem, because below we will prove its generalized version (cf. Theorem 4.9).

We recall that a multivalued mapping $F : J \times \mathbb{R} \multimap \mathbb{R}^n$ is said to be *integrably bounded* (resp. *locally integrably bounded*) if there exists an integrable (resp. locally integrable) function $\mu : J \to [0,\infty)$ such that $|y| \leqslant \mu(t)$, for every $x \in \mathbb{R}^n$, $t \in J$ and $y \in F(t,x)$. We say that F has *at most a linear growth* (resp. *a local linear growth*) if there exist integrable (resp. locally integrable) functions $\mu, \nu : J \to [0,\infty)$ such that

$$|y| \leqslant \mu(t)|x| + \nu(t),$$

for every $x \in \mathbb{R}^n$, $t \in J$ and $y \in F(t,x)$.

It is obvious that F has at most a linear growth if there exists an integrable function $\mu : J \to [0,\infty)$ such that $|y| \leqslant \mu(t)(|x|+1)$, for every $x \in \mathbb{R}^n$, $t \in J$ and $y \in F(t,x)$.

Let us also recall that a single-valued map $f : J \times \mathbb{R}^n \to \mathbb{R}^n$ is said to be *measurable-locally Lipschitz* (mLL) if, for every $x \in \mathbb{R}^n$, there exists a neighbourhood V_x of x in $\mathbb{R}^n$ and an integrable function $L_x : J \to [0, \infty)$ such that

$$\left|f(t, x_1) - f(t, x_2)\right| \leqslant L_x(t)|x_1 - x_2|, \quad \text{for every } t \in J \text{ and } x_1, x_2 \in V_x,$$

where $f(\cdot, x)$ is measurable, for every $x \in \mathbb{R}^n$.

Now, for the considerations below, fix J as the halfline $[0, \infty)$ and assume that $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ is again a multivalued u-Carathéodory map. Consider the Cauchy problem (4.19). By $S(F, 0, x_0)$, we denote the set of solutions of (4.19). For the characterization of the topological structure of $S(F, 0, x_0)$, it will be useful to recall the following well-known uniqueness criterion (see, e.g., [60, Theorem 1.1.2]).

THEOREM 4.6. *If f is a single-valued, integrably bounded, measurable-locally Lipschitz map, then the set $S(f, 0, x_0)$ is a singleton, for every $x_0 \in \mathbb{R}^n$.*

The following result will be employed as well.

THEOREM 4.7. *If F is locally integrably bounded, mLL-selectionable* (*i.e. if there exists a measurable-locally Lipschitz single-valued selection*), *then $S(F, 0, x_0)$ is contractible, for every $x_0 \in \mathbb{R}^n$.*

PROOF. Let $f \subset F$ be measurable-locally Lipschitz. By Theorem 4.6, the following Cauchy problem

$$\begin{cases} \dot{x} = f(t, x), \\ x(t_0) = x_0, \end{cases} \tag{4.20}$$

has exactly one solution, for every $t_0 \in J$ and $x_0 \in \mathbb{R}^n$. For the proof, it is sufficient to define a homotopy $h : S(F, 0, x_0) \times [0, 1] \to S(F, 0, x_0)$ such that

$$h(x, s) = \begin{cases} x, & \text{for } s = 1 \text{ and } x \in S(F, 0, x_0), \\ \widetilde{x}, & \text{for } s = 0, \end{cases}$$

where $\widetilde{x} = S(f, 0, x_0)$ is exactly one solution of the problem (4.20).

Define $\gamma : [0, 1) \to [0, \infty)$, $\gamma(s) = \tan(\pi s/2)$ and put

$$h(x, s)(t) = \begin{cases} x(t), & \text{for } 0 \leqslant t \leqslant \gamma(s),\ s < 1, \\ S\big(f, \gamma(s), x\big(\gamma(s)\big)\big)(t), & \text{for } \gamma(s) \leqslant t < \infty,\ s < 1, \\ x(t), & \text{for } 0 \leqslant t < \infty,\ s = 1. \end{cases}$$

Then h is a continuous homotopy, contracting $S(F, 0, x_0)$ to the point $S(f, 0, x_0)$. □

Analogously, we can get the following result.

THEOREM 4.8. *If F is locally integrably bounded, Ca-selectionable (i.e. if there exists a Carathéodory single-valued selection), or in particular c-selectionable (i.e. if there exists a continuous single-valued selection), then $S(F, 0, x_0)$ is R_δ-contractible, for every $x_0 \in \mathbb{R}^n$.*

Observe that, if $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ is an intersection of the decreasing sequence of $F_k : J \times \mathbb{R}^n \multimap \mathbb{R}^n$, $F(t, x) = \bigcap_{k=1}^{\infty} F_k(t, x)$ and $F_{k+1}(t, x) \subset F_k(t, x)$, for almost all $t \in J$ and for all $x \in \mathbb{R}^n$, then

$$S(F, 0, x_0) = \bigcap_{k=1}^{\infty} S(F_k, 0, x_0). \tag{4.21}$$

From Theorems 4.7 and 4.8, we obtain

PROPOSITION 4.6. *Let $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ be a multivalued map with nonempty, closed values.*

(i) *If F is σ-mLL-selectionable (i.e. it is an intersection of a decreasing sequence of mLL-selectionable mappings), then the set $S(F, 0, x_0)$ is an intersection of a decreasing sequence of contractible sets,*

(ii) *if F is σ-Ca-selectionable, i.e., it is an intersection of a decreasing sequence of Ca-selectionable mappings, then the set $S(F, 0, x_0)$ is an intersection of a decreasing sequence of R_δ-contractible sets.*

Before formulating the following important theorem, recall that, for two metric spaces X, Y and the interval J, the multivalued map $F : J \times X \multimap Y$ is *almost upper semi-continuous* (a.u.s.c.), if for every $\varepsilon > 0$ there exists a measurable set $A_\varepsilon \subset J$ such that $m(J \setminus A_\varepsilon) < \varepsilon$ and the restriction $F|_{A_\varepsilon \times X}$ is u.s.c., where m stands for the Lebesgue measure.

It is clear that every a.u.s.c. map is u-Carathéodory. In general, the reverse is not true. The following Scorza–Dragoni type result describing possible regularizations of Carathéodory maps (see, e.g., [76]) will be employed.

PROPOSITION 4.7. *Let X be a separable metric space and J be an interval. Suppose that $F : J \times X \multimap \mathbb{R}^n$ is a nonempty, compact, convex valued u-Carathéodory map. Then there exists an a.u.s.c. map $\psi : J \times X \multimap \mathbb{R}^n$ with nonempty compact convex values and such that:*

(i) *$\psi(t, x) \subset F(t, x)$, for every $(t, x) \in J \times X$,*

(ii) *if $\Delta \subset J$ is measurable, $u : \Delta \to \mathbb{R}^n$ and $v : \Delta \to X$ are measurable maps and $u(t) \in F(t, v(t))$, for almost all $t \in \Delta$, then $u(t) \in \psi(t, v(t))$, for almost all $t \in \Delta$.*

The proof of the following statement can be found in [67].

PROPOSITION 4.8. *Let E, E_1 be two separable Banach spaces, J be an interval and $F : J \times E \multimap E$ be an a.u.s.c. map with compact convex values. Then F is σ-Ca-selectionable (i.e. it is an intersection of a decreasing sequence of Ca-selectionable mappings $F_k : J \times E \multimap E_1$). The maps $F_k : J \times E \multimap E_1$ are a.u.s.c., and we have $F_k(t, e) \subset$*

$\overline{\mathrm{conv}}(\bigcup_{x\in E} F(t,x))$, for all $(t,e) \in J \times E$. Moreover, if F is integrably bounded, then F is σ-mLL-selectionable, i.e., it is an intersection of a decreasing sequence of mLL-selectionable mappings.

Now, we are ready to give

THEOREM 4.9. *If $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ is a u-Carathéodory map with compact convex values having at most the linear growth, then $S(F,0,x_0)$ is an R_δ-set, for every $x_0 \in \mathbb{R}^n$.*

PROOF. By the hypothesis, there exists an integrable function $\mu : J \to [0,\infty)$ such that $\sup\{|y| \mid y \in F(t,x)\} \leqslant \mu(t)(|x|+1)$, for every $(t,x) \in J \times \mathbb{R}^n$. By means of the well-known Gronwall inequality (see [71]), we obtain that $|x(t)| \leqslant (|x_0| + \gamma)\exp(\gamma) = M$, where $x \in S(F,0,x_0)$ and $\gamma = \int_0^\infty \mu(s)\,\mathrm{d}s$.

Take $r > M$ and define $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ as follows:

$$\widetilde{F}(t,x) = \begin{cases} F(t,x), & \text{if } |x| \leqslant r, \\ F\left(t, r\dfrac{x}{|x|}\right), & \text{if } |x| > r. \end{cases}$$

One can see that $\widetilde{F}$ is an integrably bounded u-Carathéodory map and

$$S(\widetilde{F},0,x_0) = S(F,0,x_0).$$

By Proposition 4.7, there exists an a.u.s.c. map $G : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ with nonempty, convex, compact values such that $S(G,0,x_0) = S(\widetilde{F},0,x_0)$. Applying Proposition 4.8 to the map G, we obtain the sequence of maps G_k. As in Proposition 4.6, we see that $S(G,0,x_0)$ is an intersection of the decreasing sequence $S(G_k,0,x_0)$ of contractible sets. By the well-known Arzelà–Ascoli lemma and Theorem 4.6, we obtain that, for every $k \in \mathbb{N}$, the set $S(G_k,0,x_0)$ is compact and nonempty, which completes the proof. □

Using the above results and the unified approach to the u.s.c. and l.s.c. case due to A. Bressan (cf. [49,50]), we can obtain the following result.

PROPOSITION 4.9. *Let $G : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ be a l.s.c. bounded map with nonempty closed values. Then there exists a u.s.c. map $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ with compact convex values such that, for any $x_0 \in \mathbb{R}^n$, the set $S(G,0,x_0)$ contains an R_δ-set $S(F,0,x_0)$ as a subset.*

REMARK 4.7. In [22], topological structure of solution sets is also treated, provided F is not necessarily convex-valued. However, the absence of convexity seems to be a big handicap, because to prove the connectedness and compactness of the related solution set can be a difficult task (see, e.g., [22, Example 2.18 on p. 258]).

REMARK 4.8. It follows from the result in [98] (cf. also [55] or [78, Corollary 5.3.1], where mild solutions were considered for semilinear differential inclusions) that, in a real separable Banach space E, the solution sets to initial value problems are R_δ, provided conditions in Remark 4.6 hold. In general Banach spaces E, it is at least so when $F(t,.) : E \multimap E$ is completely continuous (cf. [58, Corollary 9.1(b) on pp. 118–119]).

We can also say something about the covering (topological) dimension of solution sets to the Cauchy problem (4.19).

Let Ω be an open set in $\mathbb{R}^{n+1}$ such that $[t_0, t_0+h] \times B(x_0, r) \subset \Omega$, where B denotes the closed ball centered at x_0 and with the radius r. Assume that $F : \Omega \multimap \mathbb{R}^n$ satisfies the following conditions:

(C1) the set of values of F is nonempty, compact and convex, for all $(t,x) \in \Omega$,
(C2) $F(t,\cdot) : B(x_0,r) \multimap \mathbb{R}^n$ is continuous, for a.a. $t \in [t_0, t_0+h]$,
(C3) $F(\cdot,x) : [t_0, t_0+h] \to \mathbb{R}^n$ is measurable, for all $x \in B(x_0,r)$,
(C4) there exist Lebesgue-integrable nonnegative functions $\alpha, \beta : [t_0, t_0+h] \to [0,\infty)$ such that, for any $x \in B(x_0,r)$, $|F(t,x)| \leqslant \alpha(t) + \beta(t)|x|$, for a.a. $t \in [t_0, t_0+h]$, where $|F(t,x)| \leqslant \sup\{|y| \mid y \in F(t,x)\}$.

Denote by $S([t_0, t_0+d], x_0)$ the set of solutions $x \in \mathrm{AC}([t_0, t_0+d], \mathbb{R}^n)$ of (4.19) on the interval $[t_0, t_0+d]$, $0 < d \leqslant h$.

The following two theorems are due to B.D. Gel'man [66] (cf. [22, Theorems 2.60 and 2.61 in Chapter III.2]).

THEOREM 4.10. *Let the assumptions* (C1)–(C4) *be satisfied. Assume that the set*

$$A = \bigl\{t \in [t_0, t_0+h] \mid \dim\bigl(F(t,x)\bigr) \geqslant 1,\ \textit{for any}\ x \in B(x_0,r)\bigr\}$$

is measurable and

$$\lim_{h\to 0} \frac{\mu(A \cap [t_0, t_0+h])}{h} > 0,$$

where $\dim(\cdot)$ *denotes the covering dimension and* $\mu(\cdot)$ *stands for the Lebesgue measure. Then there exists a number* d_0 *such that, for any* $0 < d \leqslant d_0$, *we have* $S = S([t_0, t_0+d], x_0) \neq \{\emptyset\}$ *and* $\dim(S) = \infty$.

THEOREM 4.11. *Let the assumptions of Theorem* 4.10 *be satisfied jointly with*
(C2′) $F(t,\cdot) : B(x_0,r) \multimap \mathbb{R}^n$ *is Lipschitz-continuous.*
Then there exists a number d_0 *such that, for any* $0 < d \leqslant d_0$, *any* $\varepsilon > 0$ *and any solution* $x \in S([t_0, t_0+d], x_0)$ $(\neq \{\emptyset\})$, *we have* $\dim(S_{x,\varepsilon}) = \infty$, *where* $S_{x,\varepsilon} = \{y \in S \mid \|x-y\| \leqslant \varepsilon\}$.

Now, we shall study the reverse Cauchy problem when, instead of the origin, the value of solutions is prescribed at infinity, namely

$$\begin{cases} \dot{x}(t) \in F\bigl(t, x(t)\bigr), & \text{for a.a. } t \in [0,\infty), \\ \lim\limits_{t\to\infty} x(t) = x_\infty \in \mathbb{R}^n, \end{cases} \tag{4.22}$$

where $F : [0,\infty) \times \mathbb{R}^n \multimap \mathbb{R}^n$ is a u-Carathéodory map, i.e.

(i) values of F are nonempty, compact and convex, for all $(t, x) \in [0, \infty) \times \mathbb{R}^n$,
(ii) $F(t, \cdot)$ is upper semicontinuous, for a.a. $t \in [0, \infty)$,
(iii) $F(\cdot, x)$ is measurable, for all $x \in \mathbb{R}^n$.

We will prove acyclicity of the solution set of problem (4.22).

Recalling that any contractible set is acyclic, we can give

THEOREM 4.12. *Consider the target* (*terminal*) *problem* (4.22), *where* $F : [0, \infty) \times \mathbb{R}^n \multimap \mathbb{R}^n$ *is a u-Carathéodory map and* $x_\infty \in \mathbb{R}^n$ *is arbitrary. Assume that there exists a globally integrable function* $\nu : [0, \infty) \to [0, \infty)$, *where* $\int_0^\infty \nu(t)\,\mathrm{d}t = E < 1$, *such that*

$$d_H\big(F(t, x), F(t, y)\big) \leqslant \nu(t)|x - y|, \quad \textit{for all } t \in [0, \infty) \textit{ and } x, y \in \mathbb{R}^n. \tag{4.23}$$

Moreover, assume that $d_H(F(\cdot, 0), 0)$ *can be absolutely estimated by some globally integrable function. If* E *is a sufficiently small constant, then the set of solutions to problem* (4.22) *is compact and acyclic, for every* $x_\infty \in \mathbb{R}^n$.

PROOF. Observe that condition (4.23) implies the existence a globally integrable function $\alpha : [0, \infty) \to [0, \infty)$ and a positive constant B such that

$$\big|F(t, x)\big| \leqslant \alpha(t)\big(B + |x|\big), \quad \text{for every } x \in \mathbb{R}^n \text{ and a.a. } t \in [0, \infty), \tag{4.24}$$

where $|F(t, x)| = \sup\{|y| \mid y \in F(t, x)\}$. Thus, problem (4.22) can be equivalently replaced by the problem

$$\begin{cases} \dot{x}(t) \in G\big(t, x(t)\big), & \text{for a.a. } t \in [0, \infty), \\ \lim\limits_{t \to \infty} x(t) = x_\infty \in \mathbb{R}^n, \end{cases} \tag{4.25}$$

where G is a suitable Carathéodory map which can be estimated by a sufficiently large positive constant M, i.e.

$$\big|G(t, x)\big| \leqslant M, \quad \text{for every } x \in \mathbb{R}^n \text{ and a.a. } t \in [0, \infty),$$

and which satisfies condition (4.23) as well. In other words, the solution set $\mathcal{S}$ for problem (4.22) is the same as for problem (4.25), where

$$\mathcal{S} = \big\{x \in C\big([0, \infty), \mathbb{R}^n\big) \mid \dot{x}(t) \in F\big(t, x(t)\big), \\ \text{for a.a. } t \in [0, \infty) \text{ and } x(\infty) = x_\infty\big\}.$$

For the structure of $\mathcal{S}$, we will modify the approach from above. Observe that, under the above assumptions, F as well G are well known to be product-measurable (see Proposition 2.4), and subsequently having a Carathéodory selection $g \subset G$ which is Lipschitzian with a not necessarily same, but again sufficiently small constant (see, e.g., [73, pp. 101–103]). By the sufficiency we mean that, besides others,

$$\big|g(t, x) - g(t, y)\big| \leqslant \gamma(t)|x - y|$$

holds, for all $x, y \in \mathbb{R}^n$ and a.a. $t \in [0, \infty)$, with a Lebesgue integrable function $\gamma : [0, \infty) \to [0, \infty)$ such that $\int_0^\infty \gamma(t)\,\mathrm{d}t < 1$.

Considering the single-valued problem ($g \subset G$)

$$\begin{cases} \dot{x}(t) = g\bigl(t, x(t)\bigr), & \text{for a.a. } t \in [0, \infty), \\ \lim\limits_{t\to\infty} x(t) = x_\infty, \end{cases} \tag{4.26}$$

we can easily prove the existence of a unique solution $\overline{x}(t)$ of problem (4.26). The uniqueness can be verified in a standard manner by the contradiction, when assuming the existence of another solution $\overline{y}(t)$ of that problem, because so we would arrive at the false inequality

$$\begin{aligned} \sup_{t\in[0,\infty)} \bigl|\overline{x}(t) - \overline{y}(t)\bigr| &= \sup_{t\in[0,\infty)} \left| \int_\infty^t g\bigl(s, \overline{x}(s)\bigr)\,\mathrm{d}s - \int_\infty^t g\bigl(s, \overline{y}(s)\bigr)\,\mathrm{d}s \right| \\ &\leqslant \int_0^\infty \bigl|g\bigl(s, \overline{x}(s)\bigr) - g\bigl(s, \overline{y}(s)\bigr)\bigr|\,\mathrm{d}t \\ &\leqslant \int_0^\infty \gamma(t) \sup_{t\in[0,\infty)} \bigl|\overline{x}(t) - \overline{y}(t)\bigr|\,\mathrm{d}t \\ &\leqslant \sup_{t\in[0,\infty)} \bigl|\overline{x}(t) - \overline{y}(t)\bigr| \int_0^\infty \gamma(t)\,\mathrm{d}t \\ &< \sup_{t\in[0,\infty)} \bigl|\overline{x}(t) - \overline{y}(t)\bigr|. \end{aligned}$$

Hence, according to the definition of contractibility in Section 2.1, it is sufficient to show that the solution set $\mathcal{S}$ of problem (4.25) is homotopic to a unique solution $\overline{x}(t)$ of problem (4.26), which is at the same time a solution of problem (4.25) as well. The desired homotopy reads ($\lambda \in [0, 1]$)

$$h(x, \lambda)(t) = \begin{cases} x(t), & \text{for } t \geqslant 1/\lambda - \lambda,\ \lambda \neq 0, \\ \overline{z}(t), & \text{for } 0 < t \leqslant 1/\lambda - \lambda,\ \lambda \neq 0, \\ \overline{x}(t), & \text{for } \lambda = 0, \end{cases}$$

where $\overline{z}$ is a unique solution to the reverse Cauchy problem

$$\begin{cases} \dot{x}(t) = g\bigl(t, z(t)\bigr), & \text{for a.a. } t \in [0, 1/\lambda - \lambda], \\ z(1/\lambda - \lambda) = x(1/\lambda - \lambda), \end{cases}$$

for each $\lambda \in [0, 1]$. Then h is a continuous homotopy such that $h(x, 0) = \overline{x}$, $h(x, 1) = x$, as required, and subsequently, the set $\mathcal{S}$ is acyclic. Using the convexity assumption on values of F, we can prove by the standard manner [22, Mazur's Theorem 1.33 in Chapter I.1] that $\mathcal{S}$ is closed in $C([0, \infty), \mathbb{R}^n)$. By Arzelà–Ascoli's lemma, this set is compact, and the proof is complete. □

Now, we shall consider the boundary value problem

$$\begin{cases} \dot{x}(t) + A(t)x(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in [0, T], \\ Lx = r, \end{cases} \tag{4.27}$$

where

(i) $A : [0, T] \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is a measurable linear operator such that $|A(t)| \leqslant \gamma(t)$, for all $t \in [0, T]$ and some integrable function $\gamma : [0, T] \to [0, \infty)$,

(ii) the associated homogeneous problem

$$\begin{cases} \dot{x}(t) + A(t)x(t) = 0, & \text{for a.a. } t \in [0, T], \\ Lx = 0 \end{cases}$$

has only the trivial solution,

(iii) $F : [0, T] \times \mathbb{R}^n \multimap \mathbb{R}^n$ has nonempty, compact, convex values,

(iv) $F(\cdot, x)$ is measurable, for every $x \in \mathbb{R}^n$,

(v) there is a constant $M \geqslant 0$ such that

$$d_H\big(F(t, x), F(t, y)\big) \leqslant M|x - y|, \quad \text{for all } x, y \in \mathbb{R}^n \text{ and a.a. } t \in [0, T],$$

where d_H stands for the Hausdorff metric,

(vi) there are two nonnegative Lebesgue-integrable functions $\delta_1, \delta_2 : [0, T] \to [0, \infty)$ such that,

$$\big|F(t, x)\big| \leqslant \delta_1(t) + \delta_2(t)|x|, \quad \text{for a.a. } t \in [0, T] \text{ and all } x \in \mathbb{R}^n,$$

where $|F(t, x)| = \sup\{|y| \mid y \in F(t, x)\}$.

In [43], the authors have proved the functional generalization of following theorem.

THEOREM 4.13. *Under the assumptions* (i)–(vi), *a certain "critical" value* λ *exists such that if* $M < \lambda$, *then the set of solutions of* (4.27) *is a* (*nonempty*) *compact AR-space*. *Moreover, if the Lebesgue measure of the set* $\{t \mid \dim F(t, x) < 1,$ *for some* $x \in \mathbb{R}\}$ *is still zero, then the set of solutions of* (4.27) *is an infinite dimensional compact AR-space*, *where* $\dim X$ *denotes the covering* (*topological*) *dimension of a space* X.

REMARK 4.9. Observe that for $A \equiv 0$ and $Lx = x(0)$, the related Cauchy problem can have, under the assumptions of Theorem 4.13, infinitely many linearly independent solutions on the whole interval $[0, T]$.

REMARK 4.10. The first assertion of Theorem 4.13 can be still improved (see [22, Theorem 3.13 in Chapter III.3]), namely that, for the problem

$$\begin{cases} \dot{x}(t) + A(t)x(t) \in \alpha F\big(t, x(t)\big), & \text{for a.a. } t \in [0, T], \\ Lx = \theta, \end{cases} \tag{4.28}$$

where $\alpha \leqslant \lambda$, when λ is the critical value in Theorem 4.13, the set of solutions of (4.28) is nonempty, compact and acyclic.

REMARK 4.11. In the case of ODEs, the solution set in Theorem 4.13 consists, unlike in the critical case for $\alpha = \lambda$ in Remark 4.10, of a unique solution. The same is true for Theorem 4.12.

In view of Remark 4.11, a nontrivial structure of solution sets to single-valued boundary problems can be seen as a delicate problem. The following result in this field in [46] is rather rare.

THEOREM 4.14. *Consider the Floquet problem*

$$\begin{cases} \dot{x}(t) = f\big(t, x(t)\big), & \text{for a.a. } t \in [a,b], \\ x(a) + \lambda x(b) = \xi & \big(\lambda > 0, \xi \in \mathbb{R}^n\big), \end{cases} \tag{4.29}$$

where $f : [a,b] \times \mathbb{R}^n \to \mathbb{R}^n$ *is a bounded Carathéodory function. Assume, furthermore, that* f *satisfies*

$$\big|f(t,x) - f(t,y)\big| \leqslant p(t)|x-y|, \quad \text{for a.a. } t \in [a,b] \text{ and } x, y \in \mathbb{R}^n, \tag{4.30}$$

where $p : [a,b] \to [0,\infty)$ *is a Lebesgue-integrable function such that*

$$\int_a^b p(t)\,\mathrm{d}t \leqslant \sqrt{\pi^2 + \ln^2 \lambda}. \tag{4.31}$$

Then the set of solution to (4.29) *is an* R_δ*-set.*

REMARK 4.12. As pointed out in [46], if the sharp inequalities take place in (4.31), then problem (4.29) has a unique solution. On the other hand, for equalities (4.31), problem (4.29) can possess more solutions, respectively.

Unlike in the above theorems, the following problems can be regarded as those with "limiting" boundary conditions. In [53], the following result has been proved (as Theorem 3.1) for the boundary value problem

$$\begin{cases} \dot{x} = f(t,x), \\ Lx = r, \end{cases} \tag{4.32}$$

where $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ is a continuous function, $L : C^1(I, \mathbb{R}^n) \to \mathbb{R}^n$ is a linear operator and $I = [a,b]$ is a compact interval.

PROPOSITION 4.10. *Let* $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ *be a fixed continuous function such that, for every* $t_0 \in I$ *and* $x_0 \in \mathbb{R}^n$, *there exists a unique* (*smooth*) *solution* $x(t)$ *of the equation* $\dot{x} = f(t,x)$, *satisfying* $x(t_0) = x_0$.

Let $\mathcal{U}$ *be an open* (*in the norm topology*) *subset of the Banach space of all continuous linear operators* $L : C_1^0(I, \mathbb{R}^n) \to \mathbb{R}^n$, *where* C_1^0 *denotes the set* $C^1(\mathbb{R}^n) \subset C^0(I, \mathbb{R}^n)$, *topologized by the induced topology of* $C^0(I, \mathbb{R}^n)$ $(:= C(I, \mathbb{R}))$.

If, for every $L \in \mathcal{U}$ *and* $r \in \mathbb{R}^n$, *the boundary value problem* (4.32) *has at most one solution, then, for every* $L \in \mathcal{U}$ *and* $r \in \mathbb{R}^n$, *problem* (4.32) *has exactly one solution.*

Our aim is to prove, by means of Proposition 2.6 in Section 2.3, the following theorem.

THEOREM 4.15. *Let* $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ *be a fixed continuous function and* $p : I \to \mathbb{R}^n$ *be a continuous function such that, for every* $t_0 \in I$, $x_0 \in \mathbb{R}^n$ *and* $p \in C(I, \mathbb{R}^n)$ *with* $\|p\| \leqslant 1$, *there exists a unique solution* $x(t)$ *of*

$$\dot{x} = f(t, x) + p(t), \tag{4.33}$$

satisfying $x(t_0) = x_0$.

Let $\mathcal{U}$ *be an open* (*in the norm topology*) *subset of the Banach space of all continuous linear operators* $L : C_1^0(I, \mathbb{R}^n) \to \mathbb{R}^n$, *where* C_1^0 *has the same meaning as in Proposition* 4.10.

Assume that, for every $L \in \mathcal{U}$, $r \in \mathbb{R}^n$ *and* $p \in C^1(\mathbb{R}^n)$ *with* $\|p\| \leqslant 1$, *the boundary value problem*

$$\begin{cases} \dot{x} = f(t, x) + p(t), \\ Lx = r, \end{cases} \tag{4.34}$$

has at most one solution and that, for every $L \in \overline{\mathcal{U}}$, *all solutions of problem* (4.32) *are uniformly* (*i.e. independently of* $L \in \overline{\mathcal{U}}$) *a priori bounded, where* $\overline{\mathcal{U}}$ *denotes the closure of* $\mathcal{U}$ *in the* C_0^1*-topology.*

Then, for every $L \in \partial\mathcal{U}$ *and* $r \in \mathbb{R}^n$, *where* $\partial\mathcal{U}$ *denotes the boundary of* $\mathcal{U}$ *in the* C_1^0*-topology, problem* (4.32) *has an* R_δ*-set of solutions.*

PROOF. Since all assumptions of Proposition 4.10 are satisfied, problem (4.34) is solvable, for every $L \in \mathcal{U}$, $r \in \mathbb{R}^n$ and $p \in C(I, \mathbb{R}^n)$ with $\|p\| \leqslant 1$.

Furthermore, since $\overline{\mathcal{U}}$ is a closed subset of the Banach space of all continuous linear operators, each element $\widetilde{L} \in \partial\mathcal{U}$ can be regarded as a uniform limit of a suitable sequence $\{L_k\}$ such that $\widetilde{L} = \lim_{k\to\infty} L_k$, where $L_k \in \mathcal{U}$ $(= \operatorname{int}\mathcal{U})$, for every $k \in \mathbb{N}$.

Fix such an $\widetilde{L} \in \partial\mathcal{U}$ and consider the compact operators Φ_k, $\Phi : \mathcal{B} \to C_1^0(I, \mathbb{R}^n)$:

$$\Phi_k(x)(t) = x(a) + L_k x - r + \int_a^t f\bigl(s, x(s)\bigr)\,\mathrm{d}s$$

and

$$\Phi(x)(t) = x(0) + \widetilde{L}x - r + \int_a^t f\bigl(s, x(s)\bigr)\,\mathrm{d}s,$$

where $\mathcal{B} \subset C^0(I, \mathbb{R}^n)$ is a suitable closed ball centered at the origin, which is implied by the assumption of a uniform a priori boundedness of solutions. The compactness of operators follows directly by means of the well-known Arzelà–Ascoli lemma.

One can readily check the one-to-one correspondence between the fixed points of Φ and the solutions of problem (4.32) as well as those of Φ_k and the solutions of the equation $\dot{x} = f(t, x(t))$, satisfying

$$L_k x = r.$$

Thus, one can associate to Φ_k and Φ the proper maps $\varphi_k = \mathrm{id} - \Phi_k$ and $\varphi = \mathrm{id} - \Phi$, respectively, where id denotes the identity, namely

$$\varphi_k(x)(t) = x(t) - x(a) - L_k x + r - \int_a^t f\big(s, x(s)\big)\,\mathrm{d}s$$

and

$$\varphi(x)(t) = x(t) - x(a) - \widetilde{L}x + r - \int_a^t f\big(s, x(s)\big)\,\mathrm{d}s.$$

So, the nonempty kernel $\varphi^{-1}(0)$ of φ corresponds to the fixed points of Φ, and subsequently to solutions of (4.32), i.e.

$$x \in \Phi(x) \quad \Longleftrightarrow \quad 0 \in x - \Phi(x) = (\mathrm{id} - \Phi)(x) = \varphi(x).$$

We can assume without any loss of generality that, for a sufficiently large $k \in \mathbb{N}$, we have

$$\big|\varphi_k(x)(t) - \varphi(x)(t)\big| = \big|L_k x - \widetilde{L}x\big| = \big|\big(L_k - \widetilde{L}\big)x\big| \leqslant \frac{1}{k}, \tag{4.35}$$

because, otherwise, we can obviously select a subsequence with this property.

Since $\|\varphi_k(x)(t) - \varphi(x)(t)\| \leqslant 1/k$ holds, for every $x \in \mathcal{B}$, condition (i) of Proposition 2.6 in Section 2.3 is satisfied.

In order to prove (ii) in Proposition 2.6, it is sufficient to verify the following inequalities

$$\big|\varphi_k(x)(t)\big| \leqslant \frac{1}{k} \quad \text{and} \quad \big|\big(\varphi_k(x)\big)(t)\big| \leqslant \frac{1}{k}, \quad k \in \mathbb{N},$$

for every x with $\varphi(x) = 0$.

However, since $(\varphi_k(x))(t) = \dot{x}(t) - f(t, x(t)) = 0$, $k \in \mathbb{N}$, and the first inequality follows from (4.35), we are done.

In order to verify (iii) in Proposition 2.6, we should realize that, for any $u \in V_k = \{u \in C^1(I, \mathbb{R}^n) : \|u\|_{C^0} \leqslant 1/k$ and $\|\dot{u}\|_{C^0} \leqslant 1/k\}$, for some $k \in \mathbb{N}$, $x(t)$ is a solution of the equation $u(t) = \varphi_k(x)(t)$, i.e.

$$u(t) = x(t) - x(a) - L_k x + r - \int_a^t f\big(s, x(s)\big)\,\mathrm{d}s,$$

if and only if it satisfies

$$\begin{cases} \dot{x} = \dot{u}(t) + f(t, x), \\ L_k x = r - u(a). \end{cases} \tag{4.36}$$

By the hypothesis, problem (4.36) has a unique solution, for every $L_k \in \mathcal{U}$, $r \in \mathbb{R}^n$ and $u \in C^1(I, \mathbb{R}^n)$ with $\|u\| \leqslant 1$, as required. Therefore, applying Proposition 2.6 in Section 2.3, the set $\{\varphi(0)\}$ is R_δ. In other words, the solution set of the original problem (4.32) is R_δ as well. □

REMARK 4.13. One can observe that the sole existence can be easily proved by means of the well-known Schauder fixed point theorem.

EXAMPLE 4.1. According to Example 2 in [85], problem

$$\begin{cases} \dot{x}_i = f_i(t, x_1, x_2) + p_i(t), \quad i = 1, 2, \\ ax_1(0) + x_2(0) = r_1, \quad bx_1(1) + x_2(1) = r_2, \end{cases} \tag{4.37}$$

is uniquely solvable, for every $a^2 < 1$, $b^2 > 1$, $r_i \in \mathbb{R}$ $(i = 1, 2)$ and $p = (p_1, p_2) \in C([0, 1], \mathbb{R})$, provided $f_i \in C^1([0, 1], \mathbb{R})$, $i = 1, 2$, and

$$\frac{\partial f_1}{\partial x_1} u_1^2 + \frac{\partial f_1}{\partial x_2} u_1 u_2 - \frac{\partial f_2}{\partial x_1} u_1 u_2 - \frac{\partial f_2}{\partial x_2} u_2^2 \geqslant 0 \quad (i = 1, 2),$$

for each triple $(t, x_1, x_2) \in [0, 1] \times \mathbb{R}^2$ and each double $(u_1, u_2) \in \mathbb{R}^2$.

Therefore, according to Theorem 4.15 (more precisely, according to its modified version, where the set of all continuous linear operators can be restricted (see [85]) to the set of all real $(n \times n)$-matrices), problem

$$\begin{cases} \dot{x}_i = f_i(t, x_1, x_2), \quad i = 1, 2, \\ ax_1(0) + x_2(0) = r_1, \quad bx_1(1) + x_2(1) = r_2, \end{cases} \tag{4.38}$$

has an R_δ-set of solutions, for certain $(a, b) \in \mathbb{R}^2$ in a closed subset of $\mathbb{R}^2$ with $a^2 = 1$, $b^2 \geqslant 1$ or $a^2 \leqslant 1$, $b^2 = 1$ (r_1, r_2 can be arbitrary), whenever all solutions of problem (4.38) are uniformly a priori bounded, for such $a^2 \leqslant 1$, $b^2 \geqslant 1$.

This can be achieved for $a \neq b$, i.e. particularly with the exception of $a = b = 1$ or $a = b = -1$, and $b^2 \leqslant b_*^2$, for some $b_* > 1$, when, e.g.,

$$\left| f(t, x) \right| \leqslant \alpha |x| + \beta, \quad \text{for all } (t, x) \in [0, 1] \times \mathbb{R}^2, \tag{4.39}$$

where α, β are suitable nonnegative constants (α must be sufficiently small as below) and $x = (x_1, x_2)$, $f = (f_i, f_2)$.

Indeed. Since the linear homogeneous problem

$$\begin{cases} \dot{x}_i = 0, \quad i = 1, 2, \\ ax_1(0) + x_2(0) = 0, \quad bx_1(1) + x_2(1) = 0, \end{cases}$$

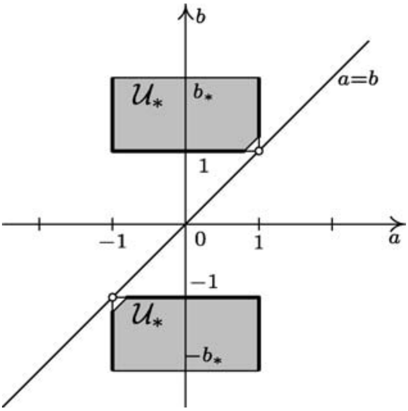


Fig. 2.

has for $a \neq b$ obviously only a trivial solution, every solution $x(t) = (x_1(t), x_2(t))$ of (4.38) takes the form (see, e.g., [22, Lemma 5.136 in Chapter III.5])

$$x_i(t) = \int_0^1 G_i(t, s, a, b) f_i\bigl(s, x_1(s), x_2(s)\bigr)\,\mathrm{d}s + \widetilde{x}_i, \quad i = 1, 2,$$

where $G = (G_1, G_2)$ is the related Green function of the linearized problem (4.38), namely

$$\begin{cases} \dot{x}_i = f_i(t, x_1, x_2), \quad i = 1, 2, \\ ax_1(0) + x_2(0) = 0, \quad bx_1(1) + x_2(1) = 0, \end{cases}$$

i.e.

$$G(t, s, a, b) = \begin{cases} \dfrac{1}{b-a}\begin{pmatrix} b & 1 \\ -ab & -a \end{pmatrix}, & \text{for } 0 \leqslant t \leqslant s \leqslant 1, \\ \dfrac{1}{b-a}\begin{pmatrix} a & 1 \\ -ab & -b \end{pmatrix}, & \text{for } 0 \leqslant s \leqslant t \leqslant 1, \end{cases}$$

and $\widetilde{x} = (\widetilde{x}_1, \widetilde{x}_2)$ is a unique solution of the problem

$$\begin{cases} \dot{x}_i = 0, \quad i = 1, 2, \\ ax_1(0) + x_2(0) = r_1, \quad bx_1(1) + x_2(1) = r_2, \end{cases}$$

i.e. $\widetilde{x}_1 = (r_2 - r_1)/(b - a)$, $\widetilde{x}_2 = r_1 - a(r_2 - r_1)/(b - a)$.

Let us fix (a, b) at the boundary $\partial\mathcal{U} = \{(a, b) \in \mathbb{R}^2 \mid a^2 = 1,\ b^2 \geqslant 1 \text{ or } a^2 \leqslant 1,\ b^2 = 1\}$ with $a \neq b$, for which we intend to get the result and cut off appropriately the corners with $a = b$, jointly with those (a, b) with $b^2 > b_*^2$ for some $b_* > 1$, as in Fig. 2. The bold curve in Fig. 2 so indicates the part of the boundary of our interest.

Denoting $g = \max_{(a,b)\in\overline{\mathcal{U}}_*} g_{a,b}$, $g_{a,b} = \max_{t,s\in[0,1]} |G(t,s,a,b)|$, where $\overline{\mathcal{U}}_*$ is indicated in Fig. 2 by the shaded region (observe that since $\overline{\mathcal{U}}_*$ is compact, g certainly exists), we obtain by means of (4.39) that

$$\|x(t)\| \leqslant g\big(\alpha\|x(t)\| + \beta\big) + \max_{(a,b)\in\overline{\mathcal{U}}_*} |\widetilde{x}|,$$

i.e.

$$\|x(t)\| \leqslant \frac{\beta g + \max_{(a,b)\in\overline{\mathcal{U}}_*} |\widetilde{x}|}{1 - \alpha g},$$

whenever $\alpha < g^{-1}$, as claimed. This completes the example.

REMARK 4.14. One can easily check that, for fixed values of $(a,b) \in \partial\mathcal{U}$ with $a \neq b$, the condition $\alpha < g^{-1}$ can take, e.g., the form

$$\alpha < \frac{|b-a|}{|b| + \max(1, |ab|)}.$$

A continuous function $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ can be approximated with an arbitrary accuracy by locally Lipschitzian (in the second variable) functions (see, e.g., [22, Theorem 3.37 in Chapter I.3] or [71]), say $(f + \varepsilon_k)$, $k \in \mathbb{N}$, such that $\lim_{k\to\infty} \|\varepsilon_k\| = 0$. Therefore, applying at first, for fixed $k \in \mathbb{N}$, Theorem 4.15 to the system

$$\dot{x} = f(t,x) + \varepsilon_k(t,x), \tag{4.40}$$

we can still avoid (for more details, see [22, Chapter III.3]) the uniqueness assumption in Theorem 4.15 as follows.

THEOREM 4.16. *Let $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ be a fixed continuous function and $\varepsilon_k : I \times \mathbb{R}^n \to \mathbb{R}^n$, $k \in \mathbb{N}$, be continuous functions with $\|\varepsilon_k\| \leqslant \varepsilon$ (ε – a sufficiently small constant) such that $(f + \varepsilon_k)(t,\cdot) : \mathbb{R}^n \to \mathbb{R}^n$, $k \in \mathbb{N}$, are locally Lipschitzian, for every $t \in I$, and $p : I \to \mathbb{R}^n$ be a continuous function with $\|p\| \leqslant 1$.*

Let $\mathcal{U}$ be an open (in the norm topology) subset of the Banach space of all continuous linear operators $L : C_1^0(I, \mathbb{R}^n) \to \mathbb{R}^n$.

Assume that, for every $L \in \mathcal{U}$, $r \in \mathbb{R}^n$, $k \in \mathbb{N}$ and $p \in C(I, \mathbb{R}^n)$ with $\|p\| \leqslant 1$, the boundary value problem

$$\begin{cases} \dot{x} = f(t,x) + \varepsilon_k(t,x) + p(t), \\ Lx = r \end{cases}$$

has at most one solution and that, for every $L \in \overline{\mathcal{U}}$ and $k \in \mathbb{N}$, all solutions of the problem

$$\begin{cases} \dot{x} = f(t,x) + \varepsilon_k(t,x), \\ Lx = r \end{cases}$$

are uniformly (*i.e. independently of* $L \in \overline{\mathcal{U}}$) *a priori bounded, where* $\overline{\mathcal{U}}$ *denotes the closure of* $\mathcal{U}$ *in the* C_1^0*-topology.*

Then, for every $L \in \partial\mathcal{U}$ *and* $r \in \mathbb{R}^n$, *where* $\partial\mathcal{U}$ *denotes the boundary of* $\mathcal{U}$ *in the* C_1^0*-topology, the problem* (4.32), *i.e.*

$$\begin{cases} \dot{x} = f(t,x), \\ Lx = r \end{cases}$$

has a (*nonempty*) *compact acyclic set of solutions.*

4.3. *Poincaré's operator approach*

By the Poincaré operators, we mean the *translation operators along the trajectories* of the associated differential systems. The translation operator is sometimes also called as *Poincaré–Andronov* or *Levinson* or, simply, *T-operator*.

In the classical theory, these operators are defined to be single-valued, when assuming among other things, the uniqueness of the initial value problems. At the absence of uniqueness, one usually approximates the right-hand sides of the given systems by the locally Lipschitzian ones (implying already uniqueness), and then applies the *standard limiting argument* (for more details, see, e.g., [71,81]).

On the other hand, set-valued analysis allows us to handle directly with multivalued Poincaré operators which become, under suitable natural restrictions imposed on the right-hand sides of given differential systems, admissible in the sense of Definition 2.5 in Section 2.2.

Hence, consider the u-Carathéodory system

$$\dot{x} \in F(t,x), \quad x \in \mathbb{R}^n, \tag{4.41}$$

where $F : [0,\tau] \times \mathbb{R}^n \multimap \mathbb{R}^n$ satisfies all conditions in Definition 2.10.

By a *solution* $x(t)$ of (4.41), we mean an absolutely continuous function $x(t) \in \mathrm{AC}([0,\tau],\mathbb{R}^n)$ satisfying (4.41), for a.a. $t \in [0,\tau]$, i.e. the one in the sense of Carathéodory, such solutions of (4.41) exist on $[0,\tau]$.

Hence, if $x(t,x_0) := x(t,0,x_0)$ is a solution of (4.41) with $x(0,x_0) = x_0 \in \mathbb{R}^n$, then the translation operator $T_\tau : \mathbb{R}^n \multimap \mathbb{R}^n$ at the time $\tau > 0$ along the trajectories of (4.41) is defined as follows:

$$T_\tau(x_0) := \{x(\tau,x_0) \mid x(\cdot,x_0) \text{ is a solution of (4.41) with } x(0,x_0) = x_0\}. \tag{4.42}$$

More precisely, T_τ can be considered as the composition of two maps, namely $T_\tau = \psi \circ \varphi$,

$$\mathbb{R}^n \overset{\varphi}{\multimap} \mathrm{AC}([0,\tau],\mathbb{R}^n) \xrightarrow{\psi} \mathbb{R}^n,$$

where $\varphi(x_0): x_0 \multimap \{x(t, x_0) \mid x(t, x_0)$ is a solution of (4.41) with $x(0, x_0) = x_0\}$ is well known to be an R_δ-mapping (see Theorem 4.5) and $\psi(y): y \to y(\tau)$ is obviously a continuous (single-valued) evaluation mapping.

In other words, we have the following commutative diagram:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, multimap, "\varphi"] \arrow[dr, multimap, dotted, "T_\tau"'] & \mathrm{AC}\big([0, \tau], \mathbb{R}^n\big) \arrow[d, "\psi"] \\
& \mathbb{R}^n
\end{tikzcd}$$

The following characterization of T_τ has been proved on various levels of abstraction in several papers (see, e.g., [22, Theorem 4.3 in Chapter III.4] and the references therein).

THEOREM 4.17. *T_τ defined by* (4.42) *is admissible and admissibly homotopic* (*see Definitions* 2.5 *and* 2.6) *to identity. More precisely, T_τ is a composition of an R_δ-mapping and a continuous* (*single-valued*) *evaluation mapping.*

PROOF. According to Theorem 4.5, the mapping φ has an R_δ-set of values. We will show that it is u.s.c. by proving the closedness of the graph Γ_φ of φ (cf. Section 2.2).

Let $(x_n, y_n) \in \Gamma_\varphi$, i.e. $y_n \in \varphi(x_n)$, and $(x_n, y_n) \to (x, y)$ as $n \to \infty$. Since the functions y_n are absolutely continuous on $[0, \tau]$, the application of the well-known Gronwall inequality (see, e.g., [71]) leads to the estimates (cf. (iii) in Definition 2.10)

$$\|y_n\| \leqslant M := \sup_{n \in \mathbb{N}} \big(|x_n| + \gamma\tau\big) \exp(\gamma\tau) \quad \text{and} \quad \|\dot{y}_n\| \leqslant \gamma(1 + M),$$

where $\gamma = \max\{\alpha, \beta\}$. It follows that $\{y_n\}$ are equibounded.

Proposition 4.1 guarantees the existence of a sequence $\{y_n\}$ such that $y_n \to y$, uniformly, and $\dot{y}_n \to \dot{y}$, weakly in $L^1([0, \tau], \mathbb{R}^n)$. According to Mazur's theorem (see, e.g., [22, Theorem 1.33 in Chapter I.1]), $\dot{y}$ belongs to the strong closure $\dot{y} \in \overline{\mathrm{conv}}\{\dot{y}_n \mid n \geqslant l\}$, for all $l \geqslant 1$. Thus, there also exists a subsequence $\{z_l\}$ such that $z_l \to \dot{y}$, in the L^1-topology, where $z_l \in \overline{\mathrm{conv}}\{\dot{y}_n \mid n \geqslant l\}$. Moreover, there exists a subsequence (for the simplicity, denoted again by $\{z_l\}$) satisfying $z_l \to \dot{y}$, a.e. on $[0, \tau]$.

Let $I \subset [0, \tau]$ be a set of a full measure on $[0, \tau]$, i.e. $\mu(I) = \tau$, where μ denotes the Lebesgue measure, such that $z_l \to \dot{y}$ as $l \to \infty$, for all $t \in I$. It follows from the definition of z_l that $z_l(t) \in \sum_i \lambda_i F(t, y_{n_i}(t))$, where $\sum_i \lambda_i = 1$.

Since $F(t, \cdot)$ is u.s.c., for a.a. $t \in [0, \tau]$, and $y_{n_i}(t)$ is sufficiently close to $x(t)$ as well as $z_l(t)$ to $\dot{x}(t)$, we obtain

$$\dot{x}(t) \in \sum_i \lambda_i F\big(t, x(t)\big) + \varepsilon B$$

for an arbitrary $\varepsilon > 0$, where B is an open unit ball. This already means that $\dot{x}(t) \in F(t, x(t))$, and subsequently the graph Γ_φ of φ is closed. Since the arbitrary closed set

$\{(x, y), (x_1, y_1), \ldots, (x_n, y_n), \ldots\}$ is, according to the well-known Arzelà–Ascoli lemma, compact, φ is u.s.c.

For the remaining part of the proof, it is sufficient to consider the admissible homotopy $T_{\lambda\tau}$, $\lambda \in [0, 1]$. □

REMARK 4.15. Since a composition of admissible maps is admissible as well (see Section 2.2), T_τ can be still composed with further admissible maps ϕ such that $\phi \circ T_\tau$ becomes an (admissible) self-map on a compact ENR-space (i.e. homeomorphic to ANR in $\mathbb{R}^n$), for computation of the well-defined (cf. Section 3.1) generalized Lefschetz number:

$$\Lambda(\phi \circ T_\tau) = \Lambda(\phi).$$

T_τ considered on ENRs can be even composed, e.g., with suitable homeomorphisms $\mathcal{H}$ (again considered on ENRs), namely $\mathcal{H} \circ T_\tau$, for computation of the well-defined (cf. Section 3.3) fixed point index:

$$\operatorname{ind}(\mathcal{H} \circ T_\tau) = \operatorname{ind} \mathcal{H},$$

provided the fixed point set of $\mathcal{H} \circ T_{\lambda\tau}$ is compact, for $\lambda \in [0, 1]$.

REMARK 4.16. In [10] (cf. also [22, Chapter III.4]), translation operators are also studied, e.g., for systems with constraints, systems in Banach spaces, for directionally semicontinuous systems, etc. In particular, in real separable Banach spaces, one can check that, under the conditions in Remark 4.6 (cf. also Remark 4.8 and [58, Corollary 9.1 in Chapter 9.4]), the related translation operator T_τ is like in Theorem 4.17. In order T_τ to be also condensing, one should however impose some further restrictions. Since these restrictions are rather technical (cf. [78, Theorem 6.3.1] or [22, Theorem 4.16 in Chapter III.4]), and so this Poincaré's translation operator will not be more employed, we omit them here.

5. Existence results

5.1. *Existence of bounded solutions*

We start with the application of Theorem 4.4. Hence, let $(E, \|\cdot\|)$ be a reflexive Banach space and let $\mathcal{L}(E)$ be the space of all linear continuous transformations in E. The Hausdorff measure of noncompactness (MNC) will be denoted by γ.

We are interested in the existence of a bounded solution to the semilinear differential inclusion

$$\dot{x}(t) + A(t)x(t) \in F\big(t, x(t)\big), \quad \text{for a.a. } t \in \mathbb{R} \tag{5.1}$$

with $A(t) \in \mathcal{L}(E)$ and a set-valued transformation F.

Our assumptions concerning the inclusion (5.1) will be the following:

(A1) $A:\mathbb{R}\to\mathcal{L}(E)$ is strongly measurable (cf. Definition 2.9) and Bochner integrable, on every compact interval $[a,b]$.

(A2) Assume that

$$\dot{x}+A(t)x=0 \tag{5.2}$$

admits a regular exponential dichotomy (cf. Remark 5.3 below; for more details see, e.g., [56]). Denote by G the principal Green's function for (5.2).

(F1) Let $F:\mathbb{R}\times E\multimap E$ be a u-Carathéodory set-valued map (cf. Definition 2.10) such that

$$\|F(t,x)\|\leqslant m(t),\qquad \text{for a.a. } t\in\mathbb{R},\ x\in E.$$

Here $m\in L^1_{\mathrm{loc}}(\mathbb{R})$ is such that, for a constant M,

$$\sup\left\{\int_t^{t+1}m(s)\,\mathrm{d}s\;\middle|\;t\in\mathbb{R}\right\}<M.$$

(F2) Assume that

$$\gamma\bigl(F(t,\Omega)\bigr)\leqslant g(t)h\bigl(\gamma(\Omega)\bigr),\qquad\text{for a.a. } t\in\mathbb{R}$$

and each bounded $\Omega\subset E$, where g,h, are positive functions, g is measurable, h is nondecreasing such that

$$L:=\sup\left\{\int_{\mathbb{R}}\|G(t,s)\|_{\mathcal{L}(E)}g(s)\,\mathrm{d}s\;\middle|\;t\in\mathbb{R}\right\}<\infty$$

and $qh(t)L<t$, for each $t>0$, with a constant $q=1$, if E is separable, and $q=2$, in the general case.

THEOREM 5.1. *Under the assumptions* (A1), (A2), (F1), (F2), *the semilinear differential inclusion* (5.1) *admits a bounded solution on* $\mathbb{R}$.

The main obstruction in the application of Theorem 4.4 will be the estimation of a suitably chosen MNC. For this purpose, we recall the following rule of taking the MNC under the sign of the integral (see [78, Corollary 4.2.5]).

LEMMA 5.1. *Let* $\{f_n\}\subset L^1([a,b],E)$ *be a sequence of functions such that*

(i) $\|f_n(t)\|\leqslant\nu(t)$, *for all* $n\in\mathbb{N}$ *and a.a.* $t\in[a,b]$, *where* $\nu\in L^1([a,b])$,

(ii) $\gamma(\{f_n(t)\})\leqslant c(t)$, *for a.a.* $t\in[a,b]$, *where* $c\in L^1([a,b])$.

Then we have the estimate

$$\gamma\left(\left\{\int_a^t f_n(s)\,\mathrm{d}s\right\}\right)\leqslant q\int_a^t c(s)\,\mathrm{d}s,$$

for each $t\in[a,b]$, *with* $q=1$, *if* E *is separable, and* $q=2$, *in general case.*

PROOF OF THEOREM 5.1. We carry out the proof in several steps.

(i) Let

$$Q := \left\{ x \in C(\mathbb{R}, E) \;\middle|\; \|x(t)\| \leqslant K, \text{ for each } t \in \mathbb{R},\ \|x(t_1) - x(t_2)\| \right.$$
$$\left. \leqslant K \int_{t_1}^{t_2} \|A(s)\|_{\mathcal{L}(E)}\, \mathrm{d}s + \int_{t_1}^{t_2} m(s)\, \mathrm{d}s, \text{ for all } t_1, t_2 \in \mathbb{R},\ t_1 \leqslant t_2 \right\}$$

with a constant K to be specified below. Clearly Q is a closed convex subset of $C(\mathbb{R}, E)$.

For a given $q \in Q$, we are interested in bounded solutions to the differential inclusion

$$\dot{x}(t) + A(t)x(t) \in F\bigl(t, q(t)\bigr), \quad \text{for a.a. } t \in \mathbb{R}. \tag{5.3}$$

Take $f \in N_F(q)$ (recall that such f exists, in view of Lemma 4.3), where N_F denotes the Nemytskiĭ operator. Since A admits an exponential dichotomy, we know that the problem

$$\dot{x}(t) + A(t)x(t) = f(t), \quad \text{for a.a. } t \in \mathbb{R},$$

has a unique, entirely bounded solution given by

$$x(f) = \int_{\mathbb{R}} G(t, s) f(s)\, \mathrm{d}s$$

(cf. [56,86]). Thus, problem (5.3) has a nonempty set of solutions $T(q)$. Using Lemmas 4.3 and 4.4, it is also clear that this set is closed convex, and since its compactness will become clear in the subsequent steps of the proof, it is in fact an R_δ-set.

(ii) We will show that, for each $q \in Q$, we actually have $T(q) \subset Q$. Let $x \in T(q)$. Then, for suitable $f \in N_F(q)$, we have

$$\begin{aligned}
\|x(t)\| &\leqslant \int_{\mathbb{R}} \|G(t, s)\|_{\mathcal{L}(E)} \|f(s)\|\, \mathrm{d}s \\
&\leqslant k \int_{-\infty}^{t} \mathrm{e}^{-\mu(t-s)} m(s)\, \mathrm{d}s + k \int_{t}^{\infty} \mathrm{e}^{-\mu(s-t)} m(s)\, \mathrm{d}s \\
&= k \sum_{j=0}^{\infty} \int_{j}^{j+1} \mathrm{e}^{-\mu\sigma} m(t-\sigma)\, \mathrm{d}\sigma + k \sum_{j=0}^{\infty} \int_{j}^{j+1} \mathrm{e}^{-\mu\sigma} m(t+\sigma)\, \mathrm{d}\sigma \\
&\leqslant k \sum_{j=0}^{\infty} \mathrm{e}^{-\mu j} \int_{j}^{j+1} m(t-\sigma)\, \mathrm{d}\sigma + k \sum_{j=0}^{\infty} \mathrm{e}^{-\mu j} \int_{j}^{j+1} m(t+\sigma)\, \mathrm{d}\sigma \\
&\leqslant k \sum_{j=0}^{\infty} \mathrm{e}^{-\mu j} \int_{t-j-1}^{t-j} m(s)\, \mathrm{d}s + k \sum_{j=0}^{\infty} \mathrm{e}^{-\mu j} \int_{t+j}^{t+j+1} m(s)\, \mathrm{d}s \\
&\leqslant 2kM \sum_{j=0}^{\infty} \mathrm{e}^{-\mu j} = 2kM\bigl(1 - \mathrm{e}^{-\mu}\bigr)^{-1} =: K.
\end{aligned} \tag{5.4}$$

In this estimation, we have used the fact that, by assumption (A2), there exist positive constants k, μ such that

$$\|G(t,s)\|_{\mathcal{L}(E)} \leqslant k\mathrm{e}^{-\mu|t-s|}.$$

Now, let $t_1, t_2 \in \mathbb{R}$. Then

$$\begin{aligned}\|x(t_1)-x(t_2)\| &\leqslant \int_{t_1}^{t_2} \|\dot{x}(s)\|\,\mathrm{d}s\\ &\leqslant \int_{t_1}^{t_2} \|A(s)\|_{\mathcal{L}(E)}\|x(s)\|\,\mathrm{d}s + \int_{t_1}^{t_2}\|f(s)\|\,\mathrm{d}s\\ &\leqslant K\int_{t_1}^{t_2}\|A(s)\|_{\mathcal{L}(E)}\,\mathrm{d}s + \int_{t_1}^{t_2} m(s)\,\mathrm{d}s.\end{aligned}$$

Consequently, $T(q) \subset Q$.

(iii) Let $\mathcal{M}$ be the power set of Q and define, for each $\Omega \in \mathcal{M}$, the real-valued MNC ψ by

$$\psi(\Omega) := \max_{D\in\mathcal{D}}(\Omega)\Bigl(\sup_{t\in\mathbb{R}}\gamma\bigl(D(t)\bigr)\Bigr),$$

where $\mathcal{D}(Q)$ denotes the collection of all denumerable subsets of Ω and $D(t) = \{d(t) \mid d \in D\} \subset E$. Then ψ is well-defined and from the corresponding properties of γ it is clear that ψ has monotone and nonsingular properties of measure of noncompactness (see Proposition 2.3 in Section 2.2). Finally, observe that ψ is regular in view of the Arzelà–Ascoli lemma.

We wish to show that the mapping T given in step (i) is condensing w.r.t. the MNC ψ.

Take $\Omega \in \mathcal{M}$. Considering $T(\Omega)$, we see that by the definition of ψ there exists a sequence $\{x_n\} \subset T(\Omega)$ such that

$$\psi\bigl(T(\Omega)\bigr) = \sup_{t\in\mathbb{R}}\gamma\bigl(\{x_n(t)\}\bigr).$$

Thus, for each $n \in \mathbb{N}$, there is $z_n \in \Omega$ and $f_n \in N_F(z_n)$ such that

$$x_n(t) = \int_{\mathbb{R}} G(t,s)f_n(s)\,\mathrm{d}s. \tag{5.5}$$

Let $\varepsilon > 0$ be fixed. Choose a number $a > 0$ such that $K\mathrm{e}^{-\mu a} < \varepsilon$. Analogously to the estimation (5.4), one shows that

$$\left\|\int_{-\infty}^{t-a} G(t,s)f_n(s)\,\mathrm{d}s + \int_{t+a}^{\infty} G(t,s)f_n(s)\,\mathrm{d}s\right\| \leqslant \varepsilon,$$

for every $n \in \mathbb{N}$. Using (5.5), we thus infer for an arbitrary $t \in \mathbb{R}$ that

$$\gamma\big(\{x_n(t)\}\big) \leqslant \varepsilon + \gamma\left(\left\{\int_{t-a}^{t+a} G(t,s) f_n(s)\,\mathrm{d}s\right\}\right).$$

Now, from assumption (F1), we get that

$$\big\|G(t,s) f_n(s)\big\| \leqslant \big\|G(t,s)\big\|_{\mathcal{L}(E)} m(s),$$

for a.a. $s \in \mathbb{R}$ and each $n \in \mathbb{N}$. Furthermore, using assumption (F2) and properties of γ (see Proposition 2.3 in Section 2.2), we see that the following estimate holds, namely

$$\begin{aligned}\gamma\big(\{G(t,s) f_n(s)\}\big) &\leqslant \big\|G(t,s)\big\|_{\mathcal{L}(E)} \gamma\big(\{f_n(s)\}\big)\\ &\leqslant \big\|G(t,s)\big\|_{\mathcal{L}(E)} g(s) h\big(\gamma\big(\{z_n(s)\}\big)\big)\\ &\leqslant \big\|G(t,s)\big\|_{\mathcal{L}(E)} g(s) h\big(\psi(\Omega)\big),\end{aligned}$$

for a.a. $s \in \mathbb{R}$. Hence, an application of Lemma 5.1 gives us

$$\gamma\big(\{x_n(t)\}\big) \leqslant \varepsilon + q h\big(\psi(\Omega)\big) \int_{t-a}^{t+a} \big\|G(t,s)\big\|_{\mathcal{L}(E)} g(s)\,\mathrm{d}s.$$

It follows that

$$\psi\big(T(\Omega)\big) \leqslant \varepsilon + q h\big(\psi(\Omega)\big) L,$$

and subsequently, since $\varepsilon > 0$ was arbitrary,

$$\psi\big(T(\Omega)\big) \leqslant q h\big(\psi(\Omega)\big) L. \tag{5.6}$$

Let us now assume that Ω is not relatively compact. Then $\psi(\Omega) > 0$ and so, by assumption (F2) and (5.6), we obtain

$$\psi\big(T(\Omega)\big) \leqslant \psi(\Omega).$$

Finally, observe that the estimate (5.6) also implies the quasi-compactness of the mapping T which subsequently justifies the compactness of the solution set to (5.3), as claimed.

Hence, we have verified all the assumptions of Theorem 4.4 (cf. also Definition 2.8) and we can establish the existence of a bounded solution to problem (5.1). □

REMARK 5.1. One can easily check that the assumption (F1) can be replaced by a weaker one, namely

$$\big\|F(t,x)\big\| \leqslant m(t) + K\|x\|, \quad \text{for a.a. } t \in \mathbb{R},\ x \in E, \tag{5.7}$$

where $K \geqslant 0$ is a sufficiently small constant and m is the same as above.

REMARK 5.2. If $A : E \to E$ is a linear, bounded operator whose spectrum does not intersect the imaginary axis, then the constant K in (5.7) can be easily taken as $K < 1/C(A)$, where

$$\sup_{t\in\mathbb{R}}\left|\int_{-\infty}^{\infty}\left\|G(t,s)\right\|\mathrm{d}s\right| \leqslant C(A), \quad G(t,s)=\begin{cases}\mathrm{e}^{-A(t-s)}P_-, & \text{for } t>s,\\ \mathrm{e}^{-A(t-s)}P_+, & \text{for } t<s\end{cases} \tag{5.8}$$

and P_-, P_+ stand for the corresponding spectral projections to the invariant subspaces of A.

REMARK 5.3. For $E = \mathbb{R}^n$ ($\Rightarrow$ (F2) holds automatically), condition (A2) is satisfied, provided there exists a projection matrix P ($P = P^2$) and constants $k > 0$, $\lambda > 0$ such that

$$\begin{cases}\left|X(t)PX^{-1}(s)\right| \leqslant k\exp\bigl(-\lambda(t-s)\bigr), & \text{for } s\leqslant t,\\ \left|X(t)(I-P)X^{-1}(s)\right| \leqslant k\exp\bigl(-\lambda(s-t)\bigr), & \text{for } t\leqslant s,\end{cases} \tag{5.9}$$

where $X(t)$ is the fundamental matrix of (5.2), satisfying $X(0) = I$, i.e., the unit matrix.

If A in (A1) is a piece-wise continuous and periodic, then it is well known that (5.9) takes place, whenever all the associated Floquet multiplies lie off the unit cycle. If A in (A1) is (continuous and) almost-periodic, then it is enough (see [93, p. 70]) that (5.9) holds only on a half-line $[t_0, \infty)$ or even on a sufficiently long finite interval.

Now, the information concerning the topological structure of solution sets in Section 4.2 will be employed for obtaining existence criteria, on the basis of general methods established in Section 4.1.

EXAMPLE 5.1. Consider the system

$$\begin{cases}\dot{x}_1 \in F_1(t,x_1,x_2)x_1 + F_2(t,x_1,x_2)x_2 + E_1(t,x_1,x_2),\\ \dot{x}_2 \in -F_2(t,x_1,x_2)x_1 + F_1(t,x_1,x_2)x_2 + E_2(t,x_1,x_2),\end{cases} \tag{5.10}$$

where $E_1, E_2, F_1, F_2 : [0,\infty) \times \mathbb{R}^2 \multimap \mathbb{R}^2$ are product-measurable u-Carathéodory maps.

Assume, furthermore, the existence of positive constants $\widetilde{E}_1$, $\widetilde{E}_2$, $\widetilde{F}_1$, $\widetilde{F}_2$, λ such that

$$\operatorname*{ess\,sup}_{t\in[0,\infty)}\Bigl[\sup_{|x_i|\leqslant D, i=1,2} F_1(t,x_1,x_2)\Bigr] \leqslant -\lambda, \tag{5.11}$$

$$\operatorname*{ess\,sup}_{t\in[0,\infty)}\Bigl[\sup_{|x_i|\leqslant D, i=1,2} \bigl|F_1(t,x_1,x_2)\bigr|\Bigr] \leqslant \widetilde{F}_1, \tag{5.12}$$

$$\operatorname*{ess\,sup}_{t\in[0,\infty)}\Bigl[\sup_{|x_i|\leqslant D, i=1,2} \bigl|F_2(t,x_1,x_2)\bigr|\Bigr] \leqslant \widetilde{F}_2, \tag{5.13}$$

$$\operatorname*{ess\,sup}_{t\in[0,\infty)}\Bigl[\sup_{|x_i|\leqslant D, i=1,2} \bigl|E_1(t,x_1,x_2)\bigr|\Bigr] \leqslant \widetilde{E}_1, \tag{5.14}$$

$$\operatorname*{ess\,sup}_{t\in[0,\infty)}\Big[\sup_{|x_i|\leqslant D, i=1,2}\big|E_2(t,x_1,x_2)\big|\Big]\leqslant\widetilde{E}_2, \tag{5.15}$$

where $D=1/\lambda(\widetilde{E}_1,\widetilde{E}_2)$. Observe that, under the assumptions (5.12)–(5.15), we have

$$\operatorname*{ess\,sup}_{t\in[0,\infty)}\big|\dot{x}_i(t)\big|\leqslant D',\quad i=1,2, \tag{5.16}$$

where $D'=(\widetilde{F}_1+\widetilde{F}_2)D+\max(\widetilde{E}_1,\widetilde{E}_2)$, so long as the solution $(x_1(t),x_2(t))$ of (5.10) satisfies

$$\sup_{t\in[0,\infty)}\big|x_i(t)\big|\leqslant D,\quad i=1,2. \tag{5.17}$$

Our aim is to prove, under the assumptions (5.11)–(5.15), the existence of a solution $x(t)=(x_1(t),x_2(t))$ satisfying

$$x(0)=0\quad\text{and}\quad\sup_{t\in[0,\infty)}\big|x_i(t)\big|\leqslant D,\quad i=1,2. \tag{5.18}$$

In order to apply Corollary 4.1 for this goal, define two sets

$$\begin{aligned}Q&:=\Big\{r(t)=\big(r_1(t),r_2(t)\big)\in C\big([0,\infty)\times[0,\infty),\mathbb{R}^2\big)\mid\sup_{t\in[0,\infty)}\big|r_i(t)\big|\leqslant D,\\&\qquad i=1,2\Big\},\\S&:=\big\{s(t)=\big(s_1(t),s_2(t)\big)\in C\big([0,\infty)\times[0,\infty),\mathbb{R}^2\big)\cap Q\mid\big|s_i(t)\big|\leqslant D't,\\&\qquad i=1,2\big\}\end{aligned}$$

(observe that $s(0)=0$), where Q is a closed convex subset of $C([0,\infty)\times[0,\infty),\mathbb{R}^2)$ and S is a bounded closed subset of Q.

For $q(t)=(q_1(t),q_2(t))\in Q$, consider still the family of systems

$$\begin{cases}\dot{x}_1=p_1(t)x_1+p_2(t)x_2+r_1(t),\\\dot{x}_2=-p_2(t)x_1+p_1(t)x_2+r_2(t),\end{cases} \tag{5.19}$$

where $p_1(t)\subset F_1(t,q(t))$, $p_2(t)\subset F_2(t,q(t))$, $r_1(t)\subset E_1(t,q(t))$, $r_2(t)\subset E_2(t,q(t))$ are measurable selections (see Proposition 2.5).

To show the solvability of (5.10) and (5.18) by means of Corollary 4.1, we need to verify that, for each $q\in Q$, the linearized system

$$\begin{cases}\dot{x}_1\in F_1\big(t,q(t)\big)x_1+F_2\big(t,q(t)\big)x_2+E_1\big(t,q(t)\big),\\\dot{x}_2\in -F_2\big(t,q(t)\big)x_1+F_1\big(t,q(t)\big)x_2+E_2\big(t,q(t)\big)\end{cases} \tag{5.20}$$

has an R_δ-set of solutions in S.

It is well known that the general solution $x(t,0,\xi)$ of (5.19), where $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, reads as follows:

$$\begin{aligned}
x_1(t,0,\xi) = {} & \left[\xi_1 \cos\left(\int_0^t p_2(s)\,\mathrm{d}s\right) + \xi_2 \sin\left(\int_0^t p_2(s)\,\mathrm{d}s\right)\right] \exp\int_0^t p_1(s)\,\mathrm{d}s \\
& + \int_0^t \left[r_1(s) \exp\int_s^t p_1(w)\,\mathrm{d}w \cos\left(\int_s^t p_2(w)\,\mathrm{d}w\right)\right] \mathrm{d}s \\
& + \int_0^t \left[r_2(s) \exp\int_s^t p_1(w)\,\mathrm{d}w \sin\left(\int_s^t p_2(w)\,\mathrm{d}w\right)\right] \mathrm{d}s, \\
x_2(t,0,\xi) = {} & \left[-\xi_1 \sin\left(\int_0^t p_2(s)\,\mathrm{d}s\right) + \xi_2 \cos\left(\int_0^t p_2(s)\,\mathrm{d}s\right)\right] \exp\int_0^t p_1(s)\,\mathrm{d}s \\
& - \int_0^t \left[r_1(s) \exp\int_s^t p_1(w)\,\mathrm{d}w \sin\left(\int_s^t p_2(w)\,\mathrm{d}w\right)\right] \mathrm{d}s \\
& + \int_0^t \left[r_2(s) \exp\int_s^t p_1(w)\,\mathrm{d}w \cos\left(\int_s^t p_2(w)\,\mathrm{d}w\right)\right] \mathrm{d}s.
\end{aligned}$$

Because of (5.11), (5.14) and (5.15), we get

$$\begin{aligned}
& \sup_{t\in[0,\infty)} \left| \int_0^t \left[r_i(s)\,\mathrm{d}s \exp\int_s^t p_1(w)\,\mathrm{d}w \cos\left(\int_s^t p_2(w)\,\mathrm{d}w\right)\right] \mathrm{d}s \right| \\
& \quad \leqslant \widetilde{E}_i \sup_{t\in[0,\infty)} \int_0^t \exp\left[-\int_s^t \left|p_1(w)\right| \mathrm{d}w\right] \mathrm{d}s \leqslant \frac{\widetilde{E}_i}{\lambda}, \\
& \sup_{t\in[0,\infty)} \left| \int_0^t \left[r_i(s)\,\mathrm{d}s \exp\int_s^t p_1(w)\,\mathrm{d}w \sin\left(\int_s^t p_2(w)\,\mathrm{d}w\right)\right] \mathrm{d}s \right| \\
& \quad \leqslant \widetilde{E}_i \sup_{t\in[0,\infty)} \int_0^t \exp\left[-\int_s^t \left|p_1(w)\right| \mathrm{d}w\right] \mathrm{d}s \leqslant \frac{\widetilde{E}_i}{\lambda},
\end{aligned}$$

for $i = 1, 2$, and subsequently we arrive at

$$\sup_{t\in[0,\infty)} \left|x_i(t,0,\xi)\right| \leqslant |\xi_1| + |\xi_2| + D, \quad i = 1, 2, \tag{5.21}$$

and $x(0,0,\xi) = \xi$.

According to Theorem 4.9 (see also Proposition 2.5 and estimate (5.21)), problem (5.17) ∩ (5.20) has the R_δ-set of solutions $x(t,0,0)$. Moreover, in view of the indicated implication ((5.17) ⇒ (5.16)), these solutions $x(t,0,0)$ belong obviously to S, for every $q \in Q$, as required.

Thus, it follows from Corollary 4.1 that problem (5.10) ∩ (5.18) has, under the assumptions (5.11)–(5.15), at least one solution.

If the inequality (5.11) or both inequalities (5.14) and (5.15) are sharp, then the same conclusion is true for $x(0) = 0$ in (5.18) replaced by $x(0) = \xi$, where $|\xi|$ is sufficiently small. For bigger values of $|\xi|$, the above assumptions can be appropriately modified as well.

THEOREM 5.2. *Consider the target problem*

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big), & \textit{for a.a. } t \in [0, \infty), \\ \lim\limits_{t\to\infty} x(t) = x_\infty \in \mathbb{R}^n \end{cases} \tag{5.22}$$

and assume that $F : [0, \infty) \times \mathbb{R}^n \multimap \mathbb{R}^n$ *is a product-measurable u-Carathéodory map. Let, furthermore, there exist a globally integrable function* $\alpha : [0, \infty) \to [0, \infty)$ *and a positive constant* β *such that, for every* $x \in \mathbb{R}^n$ *and for a.a.* $t \in [0, \infty)$, *we have* $|F(t, x)| \leqslant \alpha(t)(\beta + |x|)$, *where* $|F(t, x)| = \sup\{|y| \mid y \in F(t, x)\}$, *and* $\int_0^\infty \alpha(t)\,\mathrm{d}t < \infty$. *Then problem* (5.22) *admits a* (*bounded*) *solution, for every* $x_\infty \in \mathbb{R}^n$.

PROOF. It is convenient to consider, instead of problem (5.22), the equivalent problem

$$\begin{cases} \dot{x}(t) \in G\big(t, x(t)\big), & \text{for a.a. } t \in [0, \infty), \\ \lim\limits_{t\to\infty} x(t) = x_\infty \in \mathbb{R}^n, \end{cases} \tag{5.23}$$

where

$$G(t, x) = \begin{cases} F(t, x), & \text{for } |x| \leqslant D \text{ and } t \in [0, \infty), \\ F\left(t, D\dfrac{x}{|x|}\right), & \text{for } |x| \geqslant D \text{ and } t \in [0, \infty), \end{cases}$$

$$D \geqslant \big(|x_\infty| + AB\big) \exp A, \qquad A = \int_0^\infty \alpha(t)\,\mathrm{d}t < \infty.$$

Moreover, there certainly exists a positive constant γ such that

$$|x_0| + \big|G(t, x)\big| \leqslant |x_0| + A(B + D) \leqslant \gamma, \\ \text{for all } x \in \mathbb{R}^n \text{ and a.a. } t \in [0, \infty). \tag{5.24}$$

Besides problem (5.23), consider still a one-parameter family of linear problems

$$\begin{cases} \dot{x}(t) \in G\big(t, q(t)\big), & \text{for a.a. } t \in [0, \infty),\ q \in Q, \\ x \in Q \cap S, \end{cases} \tag{5.25}$$

where

$$S = \Big\{x \in C\big([0, \infty), \mathbb{R}^n\big) \ \Big|\ \lim_{t\to\infty} x(t) = x_\infty\Big\},$$

$$Q = \big\{q \in C\big([0, \infty), \mathbb{R}^n\big) \ \big|\ \big|q(t)\big| \leqslant |x_\infty| + A(B + D),\ \text{for } t \geqslant 0\big\}.$$

Consider the set

$$S_1 = \left\{ x \in Q \;\middle|\; \left|x(t) - x_\infty\right| \leqslant (B + D) \int_t^\infty \alpha(s)\,\mathrm{d}s, \text{ for } t \geqslant 0 \right\} \subset S.$$

It is evident that S_1 is a closed subset of S and all solutions to problem (5.25) belong to S_1.

At first, we assume that $G = g$ is single-valued. Then we have a single-valued continuous operator

$$T(q) = x_\infty + \int_\infty^t g\big(s, q(s)\big)\,\mathrm{d}s, \quad \text{for every } q \in Q.$$

Thus, to apply Corollary 4.2, only the condition $T(Q) \subset Q$ should be verified. But this follows immediately from (5.24), because

$$\begin{aligned} \sup_{t \in [0,\infty)} \left| x_\infty + \int_\infty^t G\big(s, q(s)\big)\,\mathrm{d}s \right| &\leqslant |x_\infty| + \int_0^\infty \big|G\big(t, q(t)\big)\big|\,\mathrm{d}t \\ &\leqslant |x_\infty| + (B + D) \int_0^\infty \alpha(t)\,\mathrm{d}t \\ &= |x_\infty| + A(B + D) < \infty. \end{aligned} \tag{5.26}$$

By Corollary 4.2, we obtain a solution to the problem with g as a right-hand side. This existence result can be used jointly with Theorem 4.12 which is needed to prove our statement in a general case. In fact, in view of the (just proved) existence result and Theorem 4.12 (cf. also Proposition 2.5), the map T which assigns to every $q \in Q$ the set of solutions to the linear problem (5.25), has nonempty, acyclic sets of values. Once more, we use Corollary 4.2, obtaining a solution to problem (5.22), and the proof is complete. □

5.2. *Solvability of boundary value problems with linear conditions*

Now, we shall deal with boundary value problems of the type

$$\begin{cases} \dot{x}(t) + A(t)x(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in [0, \tau], \\ Lx = \Theta, \end{cases} \tag{5.27}$$

where

(i) $A : [0, \tau] \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is a measurable linear operator such that $|A(t)| \leqslant \gamma(t)$, for all $t \in [0, \tau]$ and some integrable function $\gamma : [0, \tau] \to [0, \infty)$,

(ii) the associated homogeneous problem

$$\begin{cases} \dot{x}(t) + A(t)x(t) = 0, & \text{for a.a. } t \in [0, \tau], \\ Lx = 0 \end{cases}$$

has only the trivial solution,

(iii) $F : [0,\tau] \times \mathbb{R}^n \multimap \mathbb{R}^n$ is a u-Carathéodory mapping with nonempty, compact and convex values (cf. Definition 2.10),

(iv) there are two nonnegative Lebesgue-integrable functions $\delta_1, \delta_2 : [0,\tau] \to [0,\infty)$ such that

$$\big|F(t,x)\big| \leqslant \delta_1(t) + \delta_2(t)|x|, \quad \text{for a.a. } t \in [0,\tau] \text{ and all } x \in \mathbb{R}^n,$$

where $|F(t,x)| = \sup\{|y| \mid y \in F(t,x)\}$.

Applying Theorem 4.13 (cf. also Proposition 2.5) to replace condition (i) in Corollary 4.3 for (5.27), we can immediately give

PROPOSITION 5.1. *Consider problem* (5.27) *with* (i)–(iv) *above and let* $G : [0,\tau] \times \mathbb{R}^n \times \mathbb{R}^n \times [0,1] \to \mathbb{R}^n$ *be a product-measurable u-Carathéodory map* (*cf. Definition* 2.10) *such that*

$$G(t,c,c,1) \subset F(t,c), \quad \textit{for all } (t,c) \in [0,\tau] \times \mathbb{R}^n.$$

Assume, furthermore, that

(v) *there exists a* (*bounded*) *retract* Q *of* $C([0,\tau],\mathbb{R}^n)$ *such that* $Q \setminus \partial Q$ *is nonempty* (*open*) *and such that* $G(t,x,q(t),\lambda)$ *is Lipschitzian in* x *with a sufficiently small Lipschitz constant* (*see Theorem* 4.13), *for a.a.* $t \in [0,\tau]$ *and each* $(q,\lambda) \in Q \times [0,1]$,

(vi) *there exists a Lebesgue integrable function* $\alpha : [0,\tau] \to [0,\infty)$ *such that*

$$\big|G\big(t,x(t),q(t),\lambda\big)\big| \leqslant \alpha(t), \quad \textit{a.e. in } [0,\tau],$$

for any $(x,q,\lambda) \in \Gamma_T$ (*i.e. from the graph of* T), *where* T *denotes the set-valued map which assigns, to any* $(q,\lambda) \in Q \times [0,1]$, *the set of solutions of*

$$\begin{cases} \dot{x}(t) + A(t)x(t) \in G\big(t,x(t),q(t),\lambda\big), & \textit{for a.a. } t \in [0,1],\\ Lx = \Theta, \end{cases}$$

(vii) $T(Q \times \{0\}) \subset Q$ *holds and* ∂Q *is fixed point free w.r.t.* T, *for every* $(q,\lambda) \in Q \times [0,1]$.

Then problem (5.27) *has a solution.*

REMARK 5.4. Rescaling t in (5.27), the interval $[0,\tau]$ can be obviously replaced in Proposition 5.1 by any compact interval J, e.g., $J = [-m,m]$, $m \in \mathbb{N}$. Therefore, the second part of Proposition 4.5 can be still applied for obtaining an entirely bounded solution.

EXAMPLE 5.2. Consider problem (5.27). Assume that conditions (i)–(iv) are satisfied. Taking (for a product-measurable $F : [0,\tau] \times \mathbb{R}^n \multimap \mathbb{R}^n$)

$$G\big(t,q(t)\big) = F\big(t,q(t)\big), \quad \text{for } q \in Q,$$

where $Q = \{\mu \in C([0,\tau],\mathbb{R}^n) \mid \max_{t\in[0,\tau]} |\mu(t)| \leqslant D\}$ and $D > 0$ is a sufficiently big constant which will be specified below, we can see that (v) holds trivially. Furthermore, according to (iv), we get

$$\bigl|G\bigl(t,q(t)\bigr)\bigr| \leqslant \delta_1(t) + \delta_2(t)D, \quad \text{for a.a. } t \in [0,\tau], \tag{5.28}$$

i.e. (vi) holds as well with $\alpha(t) = \delta_1(t) + \delta_2(t)D$. At last, the associated linear problem

$$\begin{cases} \dot{x}(t) + A(t)x(t) \in F\bigl(t,q(t)\bigr), & \text{for a.a. } t \in [0,\tau], \\ Lx = \Theta \end{cases}$$

has, according to Theorem 4.13, for every $q \in Q$, an R_δ-set of solutions of the form

$$T(q) = \int_0^\tau H(t,s) f\bigl(s,q(s)\bigr)\,\mathrm{d}s,$$

where H is the related Green function and $f \subset F$ is a measurable selection (see again Proposition 2.5).

Therefore, in order to apply Proposition 5.1 for the solvability of (5.27), we only need to show (cf. (vii)) that $T(Q) \subset Q$ (and that ∂Q is fixed point free w.r.t. T, for every $q \in Q$, which is, however, not necessary here). Hence, in view of (5.28), we have that

$$\begin{aligned} \max_{t\in[0,\tau]} \bigl|T(q)\bigr| &= \max_{t\in[0,\tau]} \left| \int_0^\tau H(t,s) f\bigl(s,q(s)\bigr)\,\mathrm{d}s \right| \\ &\leqslant \max_{t\in[0,\tau]} \int_0^\tau \bigl|H(t,s)\bigr| \bigl(\delta_1(s) + \delta_2(s)D\bigr)\,\mathrm{d}s \\ &= \max_{t,s\in[0,\tau]} \bigl|H(t,s)\bigr| \left[\int_0^\tau \delta_1(t)\,\mathrm{d}t + D \int_0^\tau \delta_2(t)\,\mathrm{d}t \right], \end{aligned}$$

and subsequently the above requirement holds for

$$D \geqslant \frac{\max_{t,s\in[0,\tau]} |H(t,s)| \int_0^\tau \delta_1(t)\,\mathrm{d}t}{1 - \max_{t,s\in[0,\tau]} |H(t,s)| \int_0^\tau \delta_2(t)\,\mathrm{d}t},$$

provided

$$\int_0^\tau \delta_2(t)\,\mathrm{d}t < \frac{1}{\max_{t,s\in[0,\tau]} |H(t,s)|}.$$

(Observe that for D strictly bigger than the above quantity, ∂Q becomes fixed point free.)

5.3. *Existence of periodic and anti-periodic solutions*

Now, consider the following special cases of Floquet boundary value problems in a reflexive Banach space E:

$$\begin{cases} \dot{x}(t) + A(t)x(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in [0,\tau], \\ x(\tau) = Mx(0), & \text{where } M = \mathrm{id} \text{ or } M = -\,\mathrm{id}. \end{cases} \tag{5.29}$$

Besides (A1), (A2), assume still

$$A(t) \equiv A(t+\tau) \quad \text{and} \quad F(t,x) \equiv F(t+\tau, x) \text{ or } F(t,x) \equiv -F(t+\tau, -x)$$
$$\text{for some } \tau > 0, \tag{5.30}$$

and, instead of (F1), that only

$$\big\|F(t,x)\big\| \leqslant \alpha_0(t) + \alpha_1(t)\|x\| \tag{5.31}$$

holds for a u-Carathéodory map $F : [0,\tau] \times E \multimap E$, for a.a. $t \in [0,\tau]$ and every $x \in E$, where $\alpha_0, \alpha_1 \in L^1([0,\tau])$. Then one can check, as in the proof of Theorem 5.1, that the solution operator $T : Q \times [0,1] \multimap E$, where $Q = \{q \in C(\mathbb{R}, E) \mid q(t) \equiv q(t+\tau) \text{ or } q(t) \equiv -q(t+\tau)\}$, associated with the fully linearized problem

$$\begin{cases} \dot{x}(t) + A(t)x(t) \in \lambda F\big(t, q(t)\big), & \text{for a.a. } t \in [0,\tau],\ \lambda \in [0,1], \\ x(\tau) = Mx(0), & \text{where } M = \mathrm{id} \text{ or } M = -\,\mathrm{id}, \end{cases} \tag{5.32}$$

is condensing and that the set of (bounded, after τ-periodic or 2τ-periodic prolongation) solutions is convex and compact, provided an analogy of (F2) holds.

Therefore, taking $G(t,c,c,0) \equiv -A(t)c$, Corollary 4.5 can be simplified as follows.

COROLLARY 5.1. *Assume that conditions* (5.30) *and* (5.31) *hold, jointly with*

- (i) $A : [0,\tau] \to \mathcal{L}(E)$ *is strongly measurable* (*cf. Definition* 2.9) *and Bochner integrable on the interval* $[0,\tau]$.
- (ii) *The linear equation* $\dot{x}(t) + A(t)x(t) = 0$ *admits a regular exponential dichotomy. Denote by G the related principal Green's function.*
- (iii) $F : [0,\tau] \times E \multimap E$ *is a u-Carathéodory map with nonempty, compact and convex values.*

Assume, furthermore, that there exists a nonempty, bounded, closed, convex subset Q of $\{q \in C(\mathbb{R}, E) \mid q(t) \equiv q(t+\tau)$ or $q(t) \equiv -q(t+\tau)\}$ with nonempty interior $\operatorname{int} Q$ *such that*

- (iv) $\gamma(F(t,\Omega)) \leqslant g(t)h(\gamma(\Omega))$, *for a.a.* $t \in [0,\tau]$, *and each* $\Omega \subset \{q(t) \in E \mid t \in [0,\tau],\ q \in Q\}$, *where* g, h *are positive functions*, g *is measurable*, h *is nondecreasing such that*

$$L := \sup\left\{ \int_0^\tau \big\|G(t,s)\big\|_{\mathcal{L}(E)} g(s)\,\mathrm{d}s \;\middle|\; t \in [0,\tau] \right\} < \infty$$

and $gh(t)L < t$, *for each* $t \in [0, \tau]$, *with a constant* $q = 1$, *if* E *is still separable, and* $q = 2$, *in general case.*

(v) $\{0\} \subset \operatorname{int} Q$ *and the boundary* ∂Q *of* Q *is fixed point free w.r.t.* T, *for every* $(q, \lambda) \in Q \times (0, 1]$, *where* T *is the map assigning, to any* $(q, \lambda) \in Q \times [0, 1]$, *the set of solutions of the fully linearized problem* (5.32).

Then problem (5.29) *admits a solution.*

REMARK 5.5. Obviously, under the assumptions (i), (ii), (iii), (5.7), (5.30) and (F2) in Section 5.1, problem (5.29) admits a solution, provided K in (5.7) is sufficiently small.

For periodic and anti-periodic problems, Proposition 5.1 can be easily simplified as follows (cf. Remark 5.3).

COROLLARY 5.2. *Consider problem*

$$\begin{cases} \dot{x}(t) + A(t)x(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in [0, \tau], \\ x(0) = x(\tau), \end{cases}$$

where $F(t, x) \equiv F(t + \tau, x)$ *satisfies conditions* (iii) *and* (iv) *in Proposition* 5.1. *Let* $G : [0, \tau] \times \mathbb{R}^n \times \mathbb{R}^n \times [0, 1] \multimap \mathbb{R}^n$ *be a product-measurable u-Carathéodory map such that*

$$G(t, c, c, 1) \subset F(t, c), \quad \text{for all } (t, c) \in [0, \tau] \times \mathbb{R}^n.$$

Assume that A *is a piece-wise continuous* (*single-valued*) *bounded* τ*-periodic* $(n \times n)$*-matrix whose Floquet multipliers lie off the unit cycle*, *jointly with* (v)–(vii) *in Proposition* 5.1, *where* $Lx = x(0) - x(\tau)$ *and* $\Theta = 0$. *Then the inclusion* $\dot{x} + A(t)x \in F(t, x)$ *admits a* τ*-periodic solution.*

COROLLARY 5.3. *Consider problem*

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big), & \text{for a.a. } t \in [0, \tau], \\ x(0) = -x(\tau), \end{cases}$$

where $F(t, x) \equiv -F(t + \tau, -x)$ *satisfies conditions* (iii) *and* (iv) *in Proposition* 5.1. *Let* $G : [0, \tau] \times \mathbb{R}^n \times \mathbb{R}^n \times [0, 1] \multimap \mathbb{R}^n$ *be a product-measurable u-Carathéodory map such that*

$$G(t, c, c, 1) \subset F(t, c), \quad \text{for all } (t, c) \in [0, \tau] \times \mathbb{R}^n.$$

Assume that (v)–(vii) *in Proposition* 5.1 *hold*, *where* $Lx = x(0) + x(\tau)$ *and* $\Theta = 0$. *Then the inclusion* $\dot{x} \in F(t, x)$ *admits a* 2τ*-periodic solution.*

In the case of ODEs, Corollary 5.3 can be still improved, in view of Theorem 4.14, where $\lambda = 1$ and $\xi = 0$, as follows.

COROLLARY 5.4. *Consider problem*

$$\begin{cases} \dot{x}(t) = f\big(t, x(t)\big), & \text{for a.a. } t \in [0, \tau], \\ x(0) = -x(\tau), \end{cases}$$

where $f(t, x) \equiv -f(t + \tau, -x)$ *is a Carathéodory function. Let* $g : [0, \tau] \times \mathbb{R}^n \times \mathbb{R}^n \times [0, 1] \to \mathbb{R}^n$ *be a Carathéodory function such that*

$$g(t, c, c, 1) = f(t, c), \quad \text{for all } (t, c) \in [0, \tau] \times \mathbb{R}^n.$$

Assume that

(i) *there exists a bounded retract* Q *of* $C([0, \tau], \mathbb{R}^n)$ *such that* $Q \setminus \partial Q$ *is nonempty* (*open*) *and such that* $g(t, x, q(t), \lambda)$ *satisfies*

$$\big|g\big(t, x, q(t), \lambda\big) - g\big(t, y, q(t), \lambda\big)\big| \leqslant p(t)|x - y|, \quad x, y \in \mathbb{R}^n$$

for a.a. $t \in [0, \tau]$ *and each* $(q, \lambda) \in Q \times [0, 1]$, *where* $p : [0, \tau] \to [0, \infty)$ *is a Lebesgue integrable function with* (*see* (4.31))

$$\int_0^\tau p(t)\, \mathrm{d}t \leqslant \pi,$$

(ii) *there exists a Lebesgue integrable function* $\alpha : [0, \tau] \to [0, \infty)$ *such that*

$$\big|g\big(t, x(t), q(t), \lambda\big)\big| \leqslant \alpha(t), \quad \text{a.e. in } [0, \tau],$$

for any $(x, q, \lambda) \in \Gamma_T$, *where* T *denotes the set-valued map which assigns, to any* $(q, \lambda) \in Q \times [0, 1]$, *the set of solutions of*

$$\begin{cases} \dot{x}(t) = g\big(t, x(t), q(t), \lambda\big), & \text{for a.a. } t \in [0, \tau], \\ x(0) = -x(\tau), \end{cases}$$

(iii) $T(Q \times \{0\}) \subset Q$ *holds and* ∂Q *is fixed point free w.r.t.* T, *for every* $(q, \lambda) \in Q \times [0, 1]$.

Then the equation $\dot{x} = f(t, x)$ *admits a* 2τ*-periodic solution.*

REMARK 5.6. Since in Corollaries 5.2 and 5.3 the associated homogeneous problems (cf. (ii) at the beginning of Section 5.2) have obviously only the trivial solution, the requirement $T(Q \times \{0\}) \subset Q$ reduces to $\{0\} \subset Q$, provided $G(t, x, q, \lambda) = \lambda G(t, x, \lambda)$, $\lambda \in [0, 1]$.

REMARK 5.7. The requirement concerning a fixed point free boundary ∂Q of Q in Proposition 5.1, Corollaries 5.1, 5.2, 5.3 and 5.4 can be verified by means of bounding (Liapunov-like) functions (see [11,32–34] and cf. [22, Chapter III.8]).

EXAMPLE 5.3. Consider the anti-periodic problem

$$\begin{cases} \dot{x} \in F_1(t,x) + F_2(t,x), \\ x(a) = -x(b), \end{cases} \tag{5.33}$$

where $x = (x_1, \ldots, x_n)$, $F = F_1 + F_2 = (f_{11}, \ldots, f_{1n}) + (f_{21}, \ldots, f_{2n})$, $F_1, F_2 : [a,b] \times \mathbb{R}^n \multimap \mathbb{R}^n$ are globally upper semicontinuous multivalued functions with nonempty, convex, compact values which are bounded in $t \in [a,b]$, for every $x \in \mathbb{R}^n$, and linearly bounded in $x \in \mathbb{R}^n$, for all $t \in [a,b]$.

Assume, furthermore, that there exist positive constants R_i, $i = 1, \ldots, n$ such that

$$\bigl|f_{1i}\bigl(t, x(\pm R_i)\bigr)\bigr| > \max_{\substack{t\in[a,b]\\ x\in\overline{K}}} \bigl|f_{2i}(t,x)\bigr|, \quad i = 1, \ldots, n,\ t \in (a,b),$$

where $x(\pm R_i) = (x_1, \ldots, x_{i-1}, \pm R_i, x_{i+1}, \ldots, x_n)$, $|x_j| \leqslant R_j$ and $K = \{x \in \mathbb{R}^n \mid |x_i| < R_i, i = 1, \ldots, n\}$,

$$\begin{aligned} &\bigl[f_{1i}\bigl(a, x(\pm R_i)\bigr) + f_{2i}(a,y)\bigr] \cdot \bigl[f_{1i}\bigl(b, -x(\pm R_i)\bigr) + f_{2i}(b,z)\bigr] < 0, \\ &\quad i = 1, \ldots, n, \end{aligned}$$

where $x, y, z \in \overline{K}$,

$$F_1(t,\cdot) \text{ is Lipschitzian with a sufficiently small constant } L,$$

for every $t \in [a,b]$; in the single-valued case, when $F \in C([a,b] \times \mathbb{R}^n, \mathbb{R}^n)$, it is enough to take $L \leqslant \pi/(b-a)$ (see condition (i) in Corollary 5.4).

It can be checked (see [32]; cf. [22, Example 8.40 in Chapter III.8]) that all assumptions of Corollary 5.3 (in the single-valued case, of Corollary 5.4) are satisfied, and so problem (5.33) admits a solution.

6. Multiplicity results

6.1. *Several solutions of initial value problems*

Let us recall that in order to apply Theorem 4.3, the following main steps have to be taken:

(i) the R_δ-structure of the solution set to (4.2) must be verified,
(ii) the inclusion $\overline{T(Q)} \subset S$ or, most preferably, $\overline{T(Q)} \subset Q \cap S$ must be guaranteed, together with the retractibility of T in the sense of Definition 4.1,
(iii) $N(r|_{T(Q)} \circ T)$ must be computed.

For initial value problems, condition (i) can be easily verified, provided G is still product-measurable. In fact, since u-Carathéodory inclusions (cf. Definition 2.10) with product-measurable right-hand sides G possess (according to Theorem 4.9; cf. also Proposition 2.5), for each $q \in Q \subset C(I, \mathbb{R}^n)$, an R_δ-set of solutions $x(\cdot, x_0)$ with $x(0, x_0) = x_0$, for every $x_0 \in \mathbb{R}^n$, such a requirement can be, in Theorem 4.3 with $S := \{x \in \mathrm{AC}_{\mathrm{loc}}(I, \mathbb{R}^n) \mid$

$x(0, x_0) = x_0\}$, simply avoided. Moreover, if Q is still compact and such that $\overline{T(Q)} \subset Q \cap S$, then (see Remark 4.3) $\pi_1(Q)$ need not be abelian and finitely generated.

Thus, Theorem 4.3 simplifies, for initial value problems, as follows:

COROLLARY 6.1. *Let $G : I \times \mathbb{R}^n \times \mathbb{R}^n \multimap \mathbb{R}^n$ be a product-measurable u-Carathéodory mapping, where $I = [0, \infty)$ or $I = [0, \tau]$, $\tau \in (0, \infty)$. Assume, furthermore, that there exists a (nonempty) compact, connected subset $Q \subset C(I, \mathbb{R}^n)$ which is a neighbourhood retract of $C(I, \mathbb{R}^n)$ such that $|G(t, x(t), q(t))| \leqslant \mu(t)(|x| + 1)$ holds, for every $(t, x, q) \in I \times \mathbb{R}^n \times Q$. Let the initial value problem*

$$\begin{cases} \dot{x}(t) \in G\big(t, x(t), q(t)\big), & \textit{for a.a. } t \in I, \\ x(0) = x_0 \end{cases}$$

have, for each $q \in Q$, a nonempty set of solutions $T(q)$ such that $\overline{T(Q)} \subset Q \cap S$, where $S := \{x \in \mathrm{AC}_{\mathrm{loc}}(I, \mathbb{R}^n) \mid x(0) = x_0\}$. Then the original initial value problem

$$\begin{cases} \dot{x}(t) \in F\big(t, x(t)\big), & \textit{for a.a. } t \in I, \\ x(0) = x_0 \end{cases}$$

admits at least $N(T)$ solutions, provided $G(t, c, c) \subset F(t, c)$ holds a.e. on I, for any $c \in \mathbb{R}^n$.

EXAMPLE 6.1. Consider the scalar ($n = 1$) initial value problem with $x_0 = 0$ and $I = [0, \tau]$, $\tau > 0$. Letting

$$Q := \big\{q \in \mathrm{AC}\big([0, \tau], \mathbb{R}\big) \mid q(0) = 0 \text{ and } \delta_2 \leqslant \dot{q}(t) \leqslant \delta_1$$
$$\text{or } -\delta_1 \leqslant \dot{q}(t) \leqslant -\delta_2, \text{ for a.a. } t \in [0, \tau]\big\},$$

where $0 < \delta_2 < \delta_1$ are suitable constants, Q can be easily verified to be a disjoint (!) union of two convex, compact sets, and consequently Q is a compact ANR, i.e. also a neighbourhood retract of $C([0, \tau], \mathbb{R})$. Unfortunately, Q is disconnected which excludes the direct application of Corollary 6.1.

Nevertheless, e.g., the inclusion

$$\dot{x}(t) \in \delta \operatorname{Sgn}\big(x(t)\big), \quad \text{for a.a. } t \in [0, \tau], \ \delta > 0, \tag{6.1}$$

where

$$\operatorname{Sgn}(x) = \begin{cases} -1, & \text{for } x \in (-\infty, 0), \\ [-1, 1], & \text{for } x = 0, \\ 1, & \text{for } x \in (0, \infty), \end{cases}$$

admits obviously two classical solutions $x_1(t) = \delta t$ with $x_1(0) = 0$ and $x_2(t) = -\delta t$ with $x_2(0) = 0$, satisfying the given inclusion everywhere.

The linearized inclusion

$$\dot{x}(t) \in \delta\,\mathrm{Sgn}\big(q(t)\big), \quad \text{for a.a. } t \in [0,\tau],\ \delta > 0,$$

possesses, for each $q \in Q$, either the solution $x_1(t) = \delta t$ with $x_1(0) = 0$ or $x_2(t) = -\delta t$ with $x_2(0) = 0$, depending on $\mathrm{sgn}(q(t))$, provided $\delta_2 \leqslant \delta \leqslant \delta_1$. Observe that there are no more solutions, for each $q \in Q$. Thus, we also have $\overline{T(Q)} \subset Q \cap S$ (i.e. condition (ii)), where $S := \{x \in \mathrm{AC}_{\mathrm{loc}}(I, \mathbb{R}^n) \mid x(0) = 0\}$.

The only handicap is related to the mentioned disconnectedness of Q. However, since $T : Q \to Q$, where

$$T(q) = \begin{cases} \delta t, & \text{for } q \geqslant 0, \\ -\delta t, & \text{for } q \leqslant 0, \end{cases}$$

is obviously single-valued, the application of the multivalued Nielsen theory, in the proof of Theorem 4.3 (and subsequently of Corollary 6.1), can be replaced by the single-valued one, where $Q \in \mathrm{ANR}$ can be disconnected (see the definition of the Nielsen number at the beginning of Section 3.2). We can, therefore, conclude, on the basis of the appropriately modified Corollary 6.1, that the original inclusion (6.1) admits at least $N(T) = 2$ solutions $x(t)$ with $x(0) = 0$, as observed by the direct calculations. In fact, it must therefore have, according to Theorem 4.5, a nontrivial R_δ-set of infinitely many piece-wise linear solutions $x(t)$ with $x(0) = 0$. The computation of $N(T) = 2$ (i.e. condition (iii)) is trivial, because $Q = Q^+ \cup Q^-$, where

$$\begin{aligned} (\mathrm{AR} \ni)\ Q^+ &:= \big\{q \in \mathrm{AC}\big([0,\tau], \mathbb{R}\big) \mid q(0) = 0 \text{ and} \\ &\qquad \delta_2 \leqslant \dot{q}(t) \leqslant \delta_1,\ \text{for a.a. } t \in [0,\tau]\big\}, \\ (\mathrm{AR} \ni)\ Q^- &:= \big\{q \in \mathrm{AC}\big([0,\tau], \mathbb{R}\big) \mid q(0) = 0 \text{ and} \\ &\qquad -\delta_1 \leqslant \dot{q}(t) \leqslant -\delta_2,\ \text{for a.a. } t \in [0,\tau]\big\}, \end{aligned}$$

and so for the computation of the generalized Lefschetz numbers we have $\Lambda(T|_{Q^+}) = \Lambda(T|_{Q^-}) = 1$, where $T|_{Q^+} : Q^+ \to Q^+$ and $T|_{Q^-} : Q^- \to Q^-$.

The same is obviously true for the inclusion

$$\dot{x}(t) \in \big[\delta + f\big(t, x(t)\big)\big]\,\mathrm{Sgn}\big(x(t)\big), \quad \text{for a.a. } t \in [0,\tau],\ \delta > 0,$$

where $f : [0,\tau] \times \mathbb{R} \to \mathbb{R}$ is a Carathéodory and locally Lipschitz function in x, for a.a. $t \in [0,\tau]$, such that $\delta_2 \leqslant \delta + f(t,x) \leqslant \delta_1$, for some $0 < \delta_2 < \delta_1$, because again $T : Q \to Q$.

Of course, we could arrive at the same conclusion even without an explicit usage of the Nielsen theory arguments, just through double application (separately on Q^+ and Q^-) of Theorem 3.3 (i.e. of the Lefschetz theory arguments).

REMARK 6.1. In view of Example 6.1, it is more realistic to suppose in Corollary 6.1 that (at least, for $n = 1$) the solution operator T is single-valued and that Q can be disconnected and not necessarily compact. Naturally, the first requirement seems to be rather associated with differential equations than inclusions.

6.2. *Several periodic and bounded solutions*

Now, Theorem 4.3 will be applied to boundary value problems associated with semilinear differential inclusions.

Hence, consider the problem

$$\begin{cases} \dot{x} + A(t)x \in F(t, x), \\ Lx = \Theta. \end{cases} \tag{6.2}$$

Since the composed multivalued function $F(t, q(t))$, where $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ is a product-measurable u-Carathéodory mapping with nonempty, compact and convex values and $q \in C(J, \mathbb{R}^n)$, is, according to Proposition 2.5, measurable, we can also employ Theorem 4.13 to the associated linearized system

$$\begin{cases} \dot{x} + A(t)x \in F\big(t, q(t)\big), \\ Lx = \Theta, \end{cases} \tag{6.3}$$

provided

$$\big|F(t, x)\big| \leqslant \mu(t)\big(|x| + 1\big), \tag{6.4}$$

where $\mu : J \to [0, \infty)$ is a suitable (locally) Lebesgue integrable bounded function.

We can immediately give

THEOREM 6.1. *Consider boundary value problem* (6.2) *on a compact interval J. Assume that $A : J \to \mathbb{R}^{n^2}$ is a single-valued continuous $(n \times n)$-matrix and $F : J \times \mathbb{R}^n \multimap \mathbb{R}^n$ is a product-measurable u-Carathéodory mapping with nonempty, compact and convex values satisfying* (6.4). *Furthermore, let $L : C(J, \mathbb{R}^n) \to \mathbb{R}^n$ be a linear operator such that the homogeneous problem*

$$\begin{cases} \dot{x} + A(t)x = 0, \\ Lx = 0 \end{cases}$$

has only the trivial solution on J. Then the original problem (6.2) *has at least $N(r|_{T(Q)} \circ T(\cdot))$ solutions* (*for the definition of the Nielsen number N, see Definition* 3.2 *in Section* 3.2), *provided there exists a closed connected subset Q of $C(J, \mathbb{R}^n)$ with a finitely generated abelian fundamental group such that*

(i) *$T(Q)$ is bounded,*
(ii) *$T(q)$ is retractible onto Q with a retraction r in the sense of Definition* 4.1,
(iii) $\overline{T(Q)} \subset \{x \in \mathrm{AC}(J, \mathbb{R}^n) \mid Lx = \Theta\}$,

where $T(q)$ denotes the set of (*existing*) *solutions to* (6.3).

REMARK 6.2. In the single-valued case, we can obviously assume the unique solvability of the associated linearized problem. Moreover, Q need not then have a finitely generated abelian fundamental group (see Remark 4.3). In the multivalued case, the latter is true, provided Q is compact and $\overline{T(Q)} \subset Q$ (see again Remark 4.3).

Before presenting a nontrivial example, it will be convenient to have the following reduction property (see [6] and cf. [22, Lemma 6.6 in Chapter III.6]).

LEMMA 6.1 (Reduction). *Let X and its closed subset Y be ANR-spaces. Assume that $f : X \to X$ is a compact map, i.e. $\overline{f(X)}$ is compact, such that $f(X) \subset Y$. Denoting by $f' : Y \to Y$ the restriction of f, we have*

(i) $\mathrm{Fix}(f') = \mathrm{Fix}(f)$,

(ii) *the Nielsen relations coincide,*

(iii) $\mathrm{ind}(C, f') = \mathrm{ind}(C, f)$, *for any Nielsen class* $C \subset \mathrm{Fix}(f)$.

Thus, $N(f') = N(f)$.

Consider the u-Carathéodory system (the functions e, f, g, h have the same regularity as in Theorem 6.1)

$$\begin{cases} \dot{x} + ax \in e(t,x,y)y^{(1/m)} + g(t,x,y), \\ \dot{y} + by \in f(t,x,y)x^{(1/n)} + h(t,x,y), \end{cases} \tag{6.5}$$

where a, b are positive numbers and m, n are odd integers with $\min(m,n) \geqslant 3$. Let suitable positive constants E_0, F_0, G, H exist such that

$$\begin{aligned} \big|e(t,x,y)\big| &\leqslant E_0, & \big|f(t,x,y)\big| &\leqslant F_0, \\ \big|g(t,x,y)\big| &\leqslant G, & \big|h(t,x,y)\big| &\leqslant H, \end{aligned}$$

hold, for a.a. $t \in (-\infty, \infty)$ and all $(x,y) \in \mathbb{R}^2$.

Furthermore, assume the existence of positive constants e_0, f_0, δ_1, δ_2 such that

$$0 < e_0 \leqslant e(t,x,y), \tag{6.6}$$

for a.a. t, all x and $|y| \geqslant \delta_2$, jointly with

$$0 < f_0 \leqslant f(t,x,y), \tag{6.7}$$

for a.a. t, $|x| \geqslant \delta_1$ and all y.

As a constraint S, consider at first the periodic boundary condition

$$\big(x(0), y(0)\big) = \big(x(\omega), y(\omega)\big). \tag{6.8}$$

More precisely, we take $S = Q = Q_1 \cap Q_2 \cap Q_3$, where

$$Q_1 = \Big\{q(t) \in C\big([0,\omega], \mathbb{R}^2\big) \mid \\ \big\|q(t)\big\| := \max\Big\{\max_{t\in[0,\omega]}\big|q_1(t)\big|, \max_{t\in[0,\omega]}\big|q_2(t)\big|\Big\} \leqslant D\Big\},$$

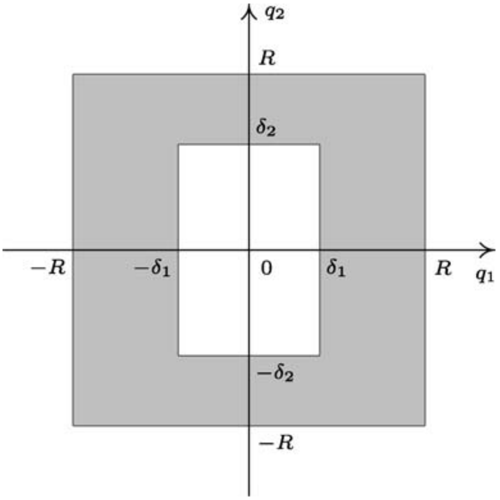


Fig. 3.

$$Q_2 = \Big\{q(t) \in C\big([0,\omega],\mathbb{R}^2\big) \mid \min_{t\in[0,\omega]}\big|q_1(t)\big| \geqslant \delta_1 > 0$$

$$\text{or } \min_{t\in[0,\omega]}\big|q_2(t)\big| \geqslant \delta_2 > 0\Big\},$$

$$Q_3 = \big\{q(t) \in C\big([0,\omega],\mathbb{R}^2\big) \mid q(0) = q(\omega)\big\},$$

the constants δ_1, δ_2, D will be specified below. For $(Q_1 \cap Q_2) \cap \mathbb{R}^2$, the situation is schematically sketched in Fig. 3.

Important properties of the set Q can be expressed as follows.

LEMMA 6.2. *The set Q defined above satisfies*:

(i) *Q is a closed connected subset of $C([0,\omega],\mathbb{R}^2)$*,
(ii) *$Q \in ANR$*,
(iii) *$\pi_1(Q) = \mathbb{Z}$*.

PROOF. Since Q is an intersection of closed sets Q_1, Q_2, Q_3, we conclude that Q is a closed subset of $C([0,\omega],\mathbb{R}^2)$ as well. The connectedness follows from the proof of (iii) below.

For (ii), it is enough to realize (see [69, Corollary 4.4 on p. 284]) that Q is the union of four closed, convex sets in the Banach space $C([0,\omega],\mathbb{R}^2)$, namely $Q = (Q_1 \cap Q_{21} \cap Q_3) \cup (Q_1 \cap Q_{22} \cap Q_3) \cup (Q_1 \cap Q_{23} \cap Q_3) \cup (Q_1 \cap Q_{24} \cap Q_3)$, where

$$Q_{21} = \Big\{q(t) \in C\big([0,\omega],\mathbb{R}^2\big) \mid \min_{t\in[0,\omega]} q_1(t) \geqslant \delta_1 > 0\Big\},$$

$$Q_{22} = \Big\{q(t) \in C\big([0,\omega],\mathbb{R}^2\big) \mid \min_{t\in[0,\omega]} q_2(t) \geqslant \delta_2 > 0\Big\},$$

$$Q_{23} = \Big\{ q(t) \in C\big([0,\omega],\mathbb{R}^2\big) \;\Big|\; \min_{t\in[0,\omega]} q_1(t) \leqslant -\delta_1 < 0 \Big\},$$

$$Q_{24} = \Big\{ q(t) \in C\big([0,\omega],\mathbb{R}^2\big) \;\Big|\; \min_{t\in[0,\omega]} q_2(t) \leqslant -\delta_2 < 0 \Big\}.$$

At last, we will show (iii). It is obvious that $\pi_1(A) = \mathbb{Z}$, where

$$A = \big\{ (x,y) \in \mathbb{R}^2 \;\big|\; \max\big(|x|,|y|\big) \leqslant D \text{ and } \big[|x| \geqslant \delta_1 \text{ or } |y| \geqslant \delta_2\big]\big\}.$$

At the same time, $A = Q \cap \mathbb{R}^2$, when regarding $\mathbb{R}^2$ as a subspace of constant functions of Q_3. For (iii), it is sufficient to show that A is a deformation retract of Q.

We define $\rho : Q \times [0,1] \to A$ by the formula

$$\rho(q,\lambda) = \big(\lambda q_1 + (1-\lambda)\overline{q_1}, \lambda q_2 + (1-\lambda)\overline{q_2}\big),$$

where $q = (q_1,q_2) \in Q$ and $\overline{q_1} = q_1(0)$, $\overline{q_2} = q_2(0)$. One can readily check that ρ is a deformation retraction, which completes the proof of our lemma. □

Besides (6.5) consider still its embedding into

$$\begin{cases} \dot{x} + ax \in \big[(1-\mu)e_0 + \mu e(t,x,y)\big]y^{1/m} + \mu g(t,x,y), \\ \dot{y} + by \in \big[(1-\mu)f_0 + \mu f(t,x,y)\big]x^{1/n} + \mu h(t,x,y), \end{cases} \tag{6.9}$$

where $\mu \in [0,1]$ and observe that (6.9) reduces to (6.5), for $\mu = 1$.

The associated linearized system to (6.9) takes, for $\mu \in [0,1]$, the form

$$\begin{cases} \dot{x} + ax \in \big[(1-\mu)e_0 + \mu e\big(t,q_1(t),q_2(t)\big)\big]q_2(t)^{1/m} + \mu g\big(t,q_1(t),q_2(t)\big), \\ \dot{y} + by \in \big[(1-\mu)f_0 + \mu f\big(t,q_1(t),q_2(t)\big)\big]q_1(t)^{1/n} + \mu h\big(t,q_1(t),q_2(t)\big), \end{cases} \tag{6.10}$$

or, equivalently,

$$\begin{cases} \dot{x} + ax = \big[(1-\mu)e_0 + \mu e_t\big]q_2(t)^{1/m} + \mu g_t, \\ \dot{y} + by = \big[(1-\mu)f_0 + \mu f_t\big]q_1(t)^{1/n} + \mu h_t, \end{cases} \tag{6.11}$$

where $e_t \subset e(t,q_1(t),q_2(t))$, $f_t \subset f(t,q_1(t),q_2(t))$, $g_t \subset g(t,q_1(t),q_2(t))$, $h_t \subset h(t,q_1(t),q_2(t))$ are measurable selections. These exist, because the u-Carathéodory functions e, f, g, h are weakly superpositionally measurable (see Proposition 2.5).

It is well known that problem (6.11) $\cap$ (6.8) has, for each $q(t) \in Q$ and every fixed quadruple of selections e_t, f_t, g_t, h_t, a unique solution $X(t) = (x(t),y(t))$, namely

$$X(t) = \begin{cases} x(t) = \displaystyle\int_0^\omega G_1(t,s)\big[\big((1-\mu)e_0 + \mu e_s\big)q_2(s)^{1/m} + \mu g_s\big]\,\mathrm{d}s, \\ y(t) = \displaystyle\int_0^\omega G_2(t,s)\big[\big((1-\mu)f_0 + \mu f_s\big)q_1(s)^{1/n} + \mu h_s\big]\,\mathrm{d}s, \end{cases}$$

where

$$G_1(t,s)=\begin{cases}\dfrac{\mathrm{e}^{-a(t-s+\omega)}}{1-\mathrm{e}^{-a\omega}}, & \text{for } 0\leqslant t\leqslant s\leqslant\omega,\\[2mm] \dfrac{\mathrm{e}^{-a(t-s)}}{1-\mathrm{e}^{-a\omega}}, & \text{for } 0\leqslant s\leqslant t\leqslant\omega,\end{cases}$$

$$G_2(t,s)=\begin{cases}\dfrac{\mathrm{e}^{-b(t-s+\omega)}}{1-\mathrm{e}^{-b\omega}}, & \text{for } 0\leqslant t\leqslant s\leqslant\omega,\\[2mm] \dfrac{\mathrm{e}^{-b(t-s)}}{1-\mathrm{e}^{-b\omega}}, & \text{for } 0\leqslant s\leqslant t\leqslant\omega.\end{cases}$$

In order to verify that $\overline{T_\mu(Q)}\subset S=Q$, where $T_\mu(\cdot)$ is the solution operator to (6.10) $\cap$ (6.8), it is just sufficient to prove that $T_\mu(Q)\subset Q$, $\mu\in[0,1]$, because $S=Q$ is closed. Hence, the Nielsen number $N(T_\mu)$ is well-defined, for every $\mu\in[0,1]$, provided only product-measurability of e, f, g, h and $T_\mu(Q)\subset Q$.

Since $X(0)=X(\omega)$, i.e. $T_\mu(Q)\subset Q_3$, it remains to prove that $T_\mu(Q)\subset Q_1$ as well as $T_\mu(Q)\subset Q_2$. Let us consider the first inclusion. In view of

$$\min_{t,s\in[0,\omega]} G_1(t,s)\geqslant\frac{\mathrm{e}^{-a\omega}}{1-\mathrm{e}^{-a\omega}}>0\quad\text{and}\quad\min_{t,s\in[0,\omega]} G_2(t,s)\geqslant\frac{\mathrm{e}^{-b\omega}}{1-\mathrm{e}^{-b\omega}}>0,$$

we obtain, for the above solution $X(t)$, that

$$\begin{aligned}\max_{t\in[0,\omega]}\left|x(t)\right| &\leqslant \max_{t\in[0,\omega]}\int_0^\omega \left|G_1(t,s)\right|\left[\left[(1-\mu)e_0+\mu e_s\right]q_2(s)^{1/m}+\mu g_s\right]\mathrm{d}s\\ &\leqslant\left[(e_0+E_0)D^{1/m}+G\right]\int_0^\omega G_1(t,s)\,\mathrm{d}s\\ &=\frac{1}{a}\left[(e_0+E_0)D^{1/m}+G\right]\end{aligned}$$

and

$$\begin{aligned}\max_{t\in[0,\omega]}\left|y(t)\right| &\leqslant \max_{t\in[0,\omega]}\int_0^\omega \left|G_2(t,s)\right|\left[\left[(1-\mu)f_0+\mu f_s\right]q_1(s)^{1/n}+\mu h_s\right]\mathrm{d}s\\ &\leqslant\left[(f_0+F_0)D^{1/n}+H\right]\int_0^\omega G_2(t,s)\,\mathrm{d}s\\ &=\frac{1}{b}\left[(f_0+F_0)D^{1/n}+H\right].\end{aligned}$$

Because of

$$\left\|X(t)\right\|=\max\left\{\max_{t\in[0,\omega]}\left|x(t)\right|,\max_{t\in[0,\omega]}\left|y(t)\right|\right\}$$

$$\leqslant \max\left\{\frac{1}{a}\big[(e_0+E_0)D^{1/m}+G\big], \frac{1}{b}\big[(f_0+F_0)D^{1/n}+H\big]\right\},$$

a sufficiently large constant D certainly exists such that $\|X(t)\| \leqslant R$, i.e. $T_\mu(Q) \subset Q_1$, independently of $\mu \in [0,1]$ and e_t, f_t, g_t, h_t.

For the inclusion $T_\mu(Q) \subset Q_2$, we proceed quite analogously.

Assuming that $q(t) \in Q_2$, we have

$$\text{either} \quad \min_{t\in[0,\omega]}\big|q_1(t)\big| \geqslant \delta_1 > 0 \quad \text{or} \quad \min_{t\in[0,\omega]}\big|q_2(t)\big| \geqslant \delta_2 > 0.$$

Therefore, we obtain for the above solution $X(t)$ that (see (6.6))

$$\begin{aligned}\min_{t\in[0,\omega]}\big|x(t)\big| &= \min_{t\in[0,\omega]}\int_0^\omega \big|G_1(t,s)\big|\big[\big[(1-\mu)e_0+\mu e_s\big]q_2(s)^{1/m}+\mu g_s\big]\,\mathrm{d}s\\ &\geqslant \big|e_0\delta_2^{1/m}-G\big|\int_0^\omega G_1(t,s)\,\mathrm{d}s = \frac{1}{a}\big|e_0\delta_2^{1/m}-G\big| > 0,\end{aligned}$$

provided $G < e_0\delta_2^{1/m}$, for $|q_2| \geqslant \delta_2$, or (see (6.7))

$$\begin{aligned}\min_{t\in[0,\omega]}\big|y(t)\big| &= \min_{t\in[0,\omega]}\int_0^\omega \big|G_2(t,s)\big|\big[\big[(1-\mu)f_0+\mu f_s\big]q_1(s)^{1/n}+\mu h_s\big]\,\mathrm{d}s\\ &\geqslant \big|f_0\delta_1^{1/n}-H\big|\int_0^\omega G_2(t,s)\,\mathrm{d}s = \frac{1}{b}\big|f_0\delta_1^{1/n}-H\big| > 0,\end{aligned}$$

provided $H < f_0\delta_1^{1/n}$, for $|q_1| \geqslant \delta_1$.

So, in order to prove that $X(t) \in Q_2$, we need to fulfill simultaneously the following inequalities:

$$\begin{cases}(1/a)\big|e_0\delta_2^{1/m}-G\big| \geqslant \delta_1 > (H/f_0)^n\\ (1/b)\big|f_0\delta_1^{1/m}-H\big| \geqslant \delta_2 > (G/e_0)^m.\end{cases} \tag{6.12}$$

Let us observe that the "amplitudes" of the multivalued functions g, h must be sufficiently small. On the other hand, if e_0 and f_0 are sufficiently large (for fixed quantities a, b, G, H), then we can easily find δ_1, δ_2 satisfying (6.12).

After all, if there exist constants δ_1, δ_2 obeying (6.12), then we arrive at $X(t) \in Q_2$, i.e., $T_\mu(Q) \subset Q_2$, independently of $\mu \in [0,1]$ and e_t, f_t, g_t, h_t. This already means that $T_\mu(Q) \subset Q$, $\mu \in [0,1]$, as required.

Now, since all the assumptions of Theorem 6.1 are satisfied, problem (6.9) $\cap$ (6.8) possesses at least $N(T_\mu(\cdot))$ solutions belonging to Q, for every $\mu \in [0,1]$. In particular, problem (6.5) $\cap$ (6.8) has $N(T_1(\cdot))$ solutions, but according to the invariance under homotopy, $N(T_1(\cdot)) = N(T_0(\cdot))$. So, it remains to compute the Nielsen number $N(T_0(\cdot))$ for the operator $T_0 : Q \to Q$, where

$$T_0(q) = \left(e_0\int_0^\omega G_1(t,s)q_2(s)^{1/m}\,\mathrm{d}s, f_0\int_0^\omega G_2(t,s)q_1(s)^{1/n}\,\mathrm{d}s\right). \tag{6.13}$$

Hence, besides (6.13), consider still its embedding into the one-parameter family of operators

$$T^{\nu}(q) = \nu T_0(q) + (1-\nu) r \circ T_0(q), \quad \nu \in [0,1],$$

where $r(q) := (r(q_1), r(q_2))$ and

$$r(q_i) = q_i(0), \quad \text{for } i = 1, 2.$$

One can readily check that $r: Q \to Q \cap \mathbb{R}^2$ is a retraction and $T_0(\overline{q}): Q \cap \mathbb{R}^2 \to Q$ is retractible onto $Q \cap \mathbb{R}^2$ with the retraction r in the sense of Definition 4.1. Thus, $r \circ T_0(\overline{q}): Q \cap \mathbb{R}^2 \to Q \cap \mathbb{R}^2$ has a fixed point $\widehat{q} \in Q \cap \mathbb{R}^2$ if and only if $\widehat{q} = T_0(\widehat{q})$. Moreover, $r \circ T_0(q): Q \to Q \cap \mathbb{R}^2$ has evidently a fixed point $\widehat{q} = Q \cap \mathbb{R}^2$ if and only if $\widehat{q} = T_0(\widehat{q})$. So, the investigation of fixed points for $T^0(q) = r \circ T_0(q)$ turns out to be equivalent with the one for $T^0(\overline{q}): Q \cap \mathbb{R}^2 \to Q \cap \mathbb{R}^2$.

Since, in view of invariance under homotopy, we have

$$N\big(T_1(\cdot)\big) = N\big(T_0(\cdot)\big) = N\big(T^1(\cdot)\big) = N\big(T^0(\cdot)\big),$$

where

$$T^0(q) = \left(\frac{e_0 \mathrm{e}^{-a\omega}}{1 - \mathrm{e}^{-a\omega}} \int_0^{\omega} \mathrm{e}^{as} q_2(s)^{1/m}\, \mathrm{d}s, \frac{f_0 \mathrm{e}^{-b\omega}}{1 - \mathrm{e}^{-b\omega}} \int_0^{\omega} \mathrm{e}^{bs} q_1(s)^{1/n}\, \mathrm{d}s \right)$$

and

$$T^0(\overline{q}) = \left(\frac{e_0}{a} \overline{q_2}^{(1/m)}, \frac{f_0}{b} \overline{q_1}^{(1/n)} \right),$$
$$\text{for } \overline{q} = (\overline{q_1}, \overline{q_2}) = \big(q_1(0), q_2(0)\big) \in Q \cap \mathbb{R}^2,$$

it remains to estimate $N(T^0(\cdot))$. It will be useful to do it by passing to a simpler finite-dimensional analogy, namely by the direct computation of fixed points of the operator

$$T^0(\overline{q}): Q \cap \mathbb{R}^2 \to Q \cap \mathbb{R}^2,$$

belonging to different Nielsen classes.

There are two fixed points $\widehat{q}_+ = (\widehat{q}_1, \widehat{q}_2)$ and $\widehat{q}_- = (-\widehat{q}_1, -\widehat{q}_2)$ in $Q \cap \mathbb{R}^2$, where

$$\widehat{q}_1 = \left(\frac{e_0}{a}\right)^{mn/(mn-1)} \left(\frac{f_0}{b}\right)^{1/(mn-1)},$$
$$\widehat{q}_2 = \left(\frac{e_0}{a}\right)^{m/(mn-1)} \left(\frac{f_0}{b}\right)^{mn/(mn-1)}.$$

These fixed points belong to different Nielsen classes, because any path u connecting them in $Q \cap \mathbb{R}^2$ and its image $T^0(u)$ are not homotopic in the space $Q \cap \mathbb{R}^2$, as it is schematically

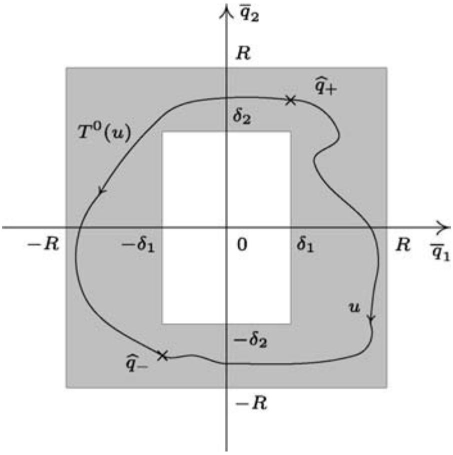


Fig. 4.

sketched in Fig. 4. Thus, according to the equivalent definition of the Nielsen number from the beginning of Section 3.2, $N(T^0(\overline{q})) = 2$. By means of the reduction property which is true here (see Lemma 6.1), we have, moreover,

$$N\bigl(T_1(\cdot)\bigr) = N\bigl(T^0(\cdot)\bigr) = N\bigl(T^0(\overline{q})\bigr) = 2$$

and so, according to Theorem 6.1, system (6.5) admits at least two solutions belonging to Q, provided suitable positive constants δ_1, δ_2 exist satisfying (6.12) and e, f, g, h are product-measurable.

In fact, system (6.5) possesses at least three solutions satisfying (6.8), when the sharp inequalities appear in (6.12), by which the lower boundary of Q becomes fixed point free. Indeed. Since

$$\Lambda\bigl(T_1(\cdot), Q\bigr) = \Lambda\bigl(T^0(\cdot), Q\bigr) = \lambda\bigl(T^0(\overline{q}), Q \cap \mathbb{R}^2\bigr)$$

holds for the generalized and ordinary Lefschetz numbers (see Section 3.1 and cf. [22, Chapter I.6]) and one can easily check that

$$\bigl|\lambda\bigl(T^0(\overline{q}), Q \cap \mathbb{R}^2\bigr)\bigr| = 2,$$

we obtain

$$\bigl|\Lambda\bigl(T_1(\cdot), Q\bigr)\bigr| = 2.$$

Furthermore, since for the self-map $T_1(\cdot)$ on the convex set $Q_1 \cap Q_3$ such that $\overline{T_1(Q_1 \cap Q_3)}$ is compact we have (see Remark 3.1)

$$\Lambda\bigl(T_1(\cdot), Q_1 \cap Q_3\bigr) = 1,$$

it follows from the additivity, contraction and existence properties of the fixed point index (see Proposition 3.2 in Section 3.3 and cf. [22, Chapter I.8]) that the mapping $T_1(\cdot)$ has the third coincidence point in $\overline{Q_1 \cap Q_3 \setminus Q}$ representing a solution of problem (6.5) $\cap$ (6.8) and belonging to $Q_1 \setminus Q$.

As we could see, problem (6.5) $\cap$ (6.8) admits at least two solutions in $Q_1 \cap Q_2$, for an arbitrary $\omega > 0$. Furthermore, because of rescaling (6.5), when replacing t by $t + (\omega/2)$, there are also two solutions of (6.5) satisfying $X(-\omega/2) = X(\omega/2)$, for an arbitrary $\omega > 0$, and belonging to $Q_1 \cap Q_2$.

Therefore, according to Proposition 4.5 and by obvious geometrical reasons, related to the appropriate subdomains of $Q_1 \cap Q_2$, system (6.5) possesses at least two entirely bounded solutions in $Q_1 \cap Q_2$.

Of course, because of replacing t by $(-t)$, the same result holds for (6.5) with negative constants a, b as well.

Finally, let us consider again system (6.5), where a, b, m, n are the same, but e, f, g, h are this time l.s.c. in (x, y), for a.a. $t \in (-\infty, \infty)$, multivalued functions with nonempty, convex, compact values and with the same estimates as above. Since each such mapping e, f, g, h has, under our regularity assumptions including the product-measurability, a Carathéodory selection (see, e.g., [22,73]), the same assertion must be also true in this new situation.

So, after summing up the above conclusions, we can give finally:

THEOREM 6.2. *Let suitable positive constants* δ_1, δ_2 *exist such that the inequalities*

$$\begin{cases} \dfrac{1}{|a|}\left|e_0\delta_2^{1/m} - G\right| \geqslant \delta_1 > \left(\dfrac{H}{f_0}\right)^n, \\ \dfrac{1}{|b|}\left|f_0\delta_2^{1/n} - H\right| \geqslant \delta_2 > \left(\dfrac{G}{e_0}\right)^m \end{cases} \tag{6.14}$$

are satisfied for constants e_0, f_0, G, H *estimating the product-measurable u-Carathéodory or l-Carathéodory multivalued functions* (*with nonempty, convex and compact values*) e, f, g, h *as above, for constants* a, b *with* $ab > 0$ *and for odd integers* m, n *with* $\min(m, n) \geqslant 3$. *Then system* (6.5) *admits at least two entirely bounded solutions. In particular, if multivalued functions* e, f, g, h *are still* ω*-periodic in* t, *then system* (6.5) *admits at least tree* ω*-periodic solutions, provided the sharp inequalities appear in* (6.14).

REMARK 6.3. Unfortunately, because of the invariance (w.r.t. the solution operator T_1) of the subdomains

$$\left\{q(t) \in C\left(\left[-\frac{\omega}{2}, \frac{\omega}{2}\right], \mathbb{R}^2\right) \,\middle|\, 0 < \delta_1 \leqslant q_1(t) \leqslant R \wedge 0 < \delta_2 \leqslant q_2(t) \leqslant R\right\}$$

and

$$\left\{q(t) \in C\left(\left[-\frac{\omega}{2}, \frac{\omega}{2}\right], \mathbb{R}^2\right) \,\middle|\, -R \leqslant q_1(t) \leqslant -\delta_1 < 0 \wedge -R \leqslant q_2(t) \leqslant -\delta_2 < 0\right\},$$

for each $\omega \in (-\infty, \infty)$, the same result can also be obtained, for example, by means of the fixed point index.

In order to avoid the handicap in Remark 6.3, let us still consider the planar system of integro-differential inclusions. For the sake of transparency, our presentation will be as simple as possible. Thus, the right-hand sides can apparently take a more general form and the multiplicity criteria in terms of inequalities can be improved.

Hence, let $x_i : [0, \omega] \to \mathbb{R}$, for $i = 1, 2$, $x = (x_1, x_2)$, $\varphi \in [0, \frac{\pi}{4}]$, $a > 0$ and consider the following system of integro-differential inclusions

$$\dot{x}_1 + a x_1 \in \sqrt[3]{p_2(x)} \cos\varphi - \sqrt[3]{p_1(x)} \sin\varphi + \varphi e, \tag{6.15}$$

$$\dot{x}_2 + a x_2 \in \sqrt[3]{p_1(x)} \cos\varphi + \sqrt[3]{p_2(x)} \sin\varphi + \varphi e, \tag{6.16}$$

where $e : [0, \omega] \times \mathbb{R}^2 \multimap \mathbb{R}$ is a product-measurable u-Carathéodory map with nonempty, convex and compact values with $|e(t, x)| \leqslant E$, for a.a. $t \in [0, \omega]$ and all $x \in \mathbb{R}^2$, and

$$p_i(x) = \frac{1}{\omega} \int_0^\omega x_i(s)\,\mathrm{d}s - B\left(\frac{1}{\omega} \int_0^\omega x_i(s)\,\mathrm{d}s - x_i\right),$$

with $B > 0$. For $\varphi = \frac{\pi}{4}$, the system takes the form

$$\dot{x}_1 + a x_1 \in \frac{\sqrt{2}}{2}\left(\sqrt[3]{p_2(x)} - \sqrt[3]{p_1(x)}\right) + \frac{\pi}{4} e, \tag{6.17}$$

$$\dot{x}_2 + a x_2 \in \frac{\sqrt{2}}{2}\left(\sqrt[3]{p_1(x)} + \sqrt[3]{p_2(x)}\right) + \frac{\pi}{4} e, \tag{6.18}$$

while, for $\varphi = 0$, it reduces to

$$\dot{x}_1 + a x_1 = \sqrt[3]{p_2(x)}, \qquad \dot{x}_2 + a x_2 = \sqrt[3]{p_1(x)}.$$

We shall be again looking for the lower estimate of the number of ω-periodic solutions to (6.17), (6.18).

Let us define sets $S = Q \subset C([0, \omega], \mathbb{R}^2)$ as follows. Function $q = (q_1, q_2)$ belongs to Q if the following conditions are satisfied:

(i) $q(0) = q(\omega)$ (ω-periodicity),

(ii) $|q(t)| \leqslant R$, for all $t \in [0, \omega]$ (boundedness),

(iii) $|q_1(t)| \geqslant \delta$ or $|q_2(t)| \geqslant \delta$, for all $t \in [0, \omega]$ (uniform boundedness of one component from below),

(iv) $q(t) = \overline{q} + \widetilde{q}(t)$, where $\overline{q} := \frac{1}{\omega} \int_0^\omega q(s)\,\mathrm{d}s$ is the integral average of q on $[0, \omega]$ (thus, $\frac{1}{\omega} \int_0^\omega \widetilde{q}(s)\,\mathrm{d}s = 0$) and $|\widetilde{q}(t)| \leqslant \varepsilon$, for all $t \in [0, \omega]$ (function q differs from its integral average by less than ε).

Values of a and ω are given, we shall specify the values of B, δ, R, E and ε in the subsequent parts.

Set Q is again a union of four closed, convex sets in the Banach space $C([0,\omega],\mathbb{R}^2)$, namely $Q = Q_1^+ \cup Q_2^+ \cup Q_1^- \cup Q_2^-$, where

$$Q_i^{\pm} = \Big\{ q(t) \in C\big([0,\omega],\mathbb{R}^2\big) \mid q \text{ satisfies (i), (ii), (iv) and} \\ \min_{t\in[0,\omega]} \pm q_i(t) \geqslant \delta > 0 \Big\}, \quad i = 1,2.$$

Thus, as in the proof of Lemma 6.2, it is a closed connected ANR-space such that $\pi_1(Q) = \mathbb{Z}$.

For homotopic parameter $\varphi = 0$, system (6.15), (6.16) reduces to a simpler case, which can be easily handled (in fact, we can explicitly compute two constant fixed points). For $\varphi = \frac{\pi}{4}$, the situation becomes non-trivial. We shall show that set Q is invariant under the solution operator, which takes a parameter $q \in Q$ to the solution x of the linearized inclusion. In this case, no obvious or easily detectable subset of Q can be recognized to be separately invariant.

In order to apply a slightly modified special case of Theorem 6.1 (cf. [18]), we use again the method of Schauder linearization. Let us take an arbitrary $q \in Q$. The system of fully linearized inclusions takes the form:

$$\dot{x}_1 + ax_1 \in \sqrt[3]{p_2(q)}\cos\varphi - \sqrt[3]{p_1(q)}\sin\varphi + \varphi e(t,q), \tag{6.19}$$

$$\dot{x}_2 + ax_2 \in \sqrt[3]{p_1(q)}\cos\varphi + \sqrt[3]{p_2(q)}\sin\varphi + \varphi e(t,q), \tag{6.20}$$

where

$$p_i(q) = \frac{1}{\omega}\int_0^{\omega} q_i(s)\,\mathrm{d}s - B\bigg(\frac{1}{\omega}\int_0^{\omega} q_i(s)\,\mathrm{d}s - q_i\bigg), \tag{6.21}$$

for $i = 1,2$.

Denoting $\overline{q} := \frac{1}{\omega}\int_0^{\omega} q(s)\,\mathrm{d}s$ the integral average of q on $[0,\omega]$, we can write $p\colon Q \subset C([0,\omega],\mathbb{R}^2) \to C([0,\omega],\mathbb{R}^2)$ in the form

$$p(q) = \overline{q} - B(\overline{q} - q). \tag{6.22}$$

For $B = 1$, operator p reduces to identity. For $B < 1$, the operator "shrinks" function q closer to its integral average. Indeed, if $q = \overline{q} + \widetilde{q}$, where $|\widetilde{q}(t)| \leqslant \varepsilon$, for all $t \in [0,\omega]$, then operator p takes q to $p(q) = \overline{q} - B\widetilde{q}$.

The fully linearized system (6.19), (6.20) possesses, for any $q \in Q$, and any Lebesgue integrable (single-valued) selection $e_0 \subset \{e(t,q)\}$, $t \in [0,\omega]$, a unique solution $x(t)$ which is given by the known convolution with the Green operator

$$x_i(t) = \int_0^{\omega} G(t,s) f_i(s)\,\mathrm{d}s, \tag{6.23}$$

where

$$f_1(s) := \sqrt[3]{p_2(q)}\cos\varphi - \sqrt[3]{p_1(q)}\sin\varphi + \varphi e_0, \tag{6.24}$$

$$f_2(s) := \sqrt[3]{p_1(q)}\cos\varphi + \sqrt[3]{p_2(q)}\sin\varphi + \varphi e_0. \tag{6.25}$$

Let us denote by T_φ the solution operator which takes $q \in Q$ to the solutions x, given by (6.23), of the linearized system (6.19), (6.20). We shall prove that Q is invariant under T_φ, namely that $T_\varphi(q) \subset Q$, for each $q \in Q$.

Let us take $q \in Q$ arbitrary. Operator p defined by (6.22) takes q to $p(q)$ such that $p(q) = \overline{q} + \widetilde{p}$, where $|\widetilde{p}(t)| \leqslant B\varepsilon$. Substituting this $p(q)$ into (6.24) and (6.25), we obtain $f_i(t) = F_i(t) + \widetilde{f}(t)$, where

$$F_1 := \sqrt[3]{\overline{q}_2}\cos\varphi - \sqrt[3]{\overline{q}_1}\sin\varphi, \qquad F_2 := \sqrt[3]{\overline{q}_1}\cos\varphi + \sqrt[3]{\overline{q}_2}\sin\varphi, \tag{6.26}$$

and $|\widetilde{f}_i(t)| \leqslant 3\sqrt[3]{B\varepsilon} + \frac{\pi E}{4}$ for all $t \in [0, \omega]$.

This estimate can be shown as follows. For $|\overline{q}_i| \geqslant 1$, one can get by direct calculation that $|\sqrt[3]{\overline{q}_i + \widetilde{p}_i} - \sqrt[3]{\overline{q}_i}| \leqslant \sqrt[3]{B\varepsilon}$, provided that $|\widetilde{p}_i| \leqslant B\varepsilon \leqslant 1$. For $|\overline{q}_i| \leqslant 1$, a careful examination of function $|\sqrt[3]{\overline{q}_i + \widetilde{p}_i} - \sqrt[3]{\overline{q}_i}|$ reveals that it is bounded from above by the value $2^{2/3}\sqrt[3]{B\varepsilon}$, provided again $|\widetilde{p}_i| \leqslant B\varepsilon \leqslant 1$. Altogether, f_i differs from F_i not more than by

$$2\frac{\sqrt{2}}{2}2^{2/3}\sqrt[3]{B\varepsilon} + \frac{\pi E}{4} \leqslant 3\sqrt[3]{B\varepsilon} + \frac{\pi E}{4},$$

as claimed.

The Green's function G in (6.23) takes the form

$$G(t,s) = \begin{cases} \dfrac{1}{1 - \mathrm{e}^{-a\omega}}\mathrm{e}^{-at}\mathrm{e}^{as}, & \text{for } 0 \leqslant s \leqslant t, \\ \dfrac{1}{1 - \mathrm{e}^{-a\omega}}\mathrm{e}^{-at}\mathrm{e}^{-a\omega}\mathrm{e}^{as}, & \text{for } t \leqslant s \leqslant \omega. \end{cases} \tag{6.27}$$

Substituting (6.26) and (6.27) into (6.23), we obtain $x_i(t) = \frac{F_i}{a} + \widetilde{x}_i(t)$, where $\widetilde{x}_i$ satisfies the inequality

$$\left|\widetilde{x}_i(t)\right| \leqslant \frac{3\sqrt[3]{B\varepsilon} + \frac{\pi}{4}E}{a},$$

for all $t \in [0, \omega]$. We can now take B and E small enough to fulfill

$$\frac{\sqrt{2}(3\sqrt[3]{B\varepsilon} + \frac{\pi}{4}E)}{a} \leqslant \frac{\varepsilon}{2}, \tag{6.28}$$

for example

$$B \leqslant \left(\frac{a}{12}\right)^3 \frac{\varepsilon^2}{\sqrt{2}} \quad \text{and} \quad E \leqslant \frac{a\varepsilon}{\pi\sqrt{2}}.$$

This means that function x differs from a constant function by less than $\frac{\varepsilon}{2}$, which implies that it differs from its integral average by less than ε. The above calculations ensure that the solution x satisfies condition (iv) of the definition of the parameter set Q, independently on $q \in Q$ and $\varphi \in [0, \frac{\pi}{4}]$. Condition (i) is trivially satisfied by the form of the Green function G.

We must further ensure that, for each $q \in Q$, function x satisfies conditions (ii) and (iii). Since both q and x differ from their integral averages by less than ε, let us first deal with constant functions. This is easy, because the solution mapping T_φ takes constant functions to functions that differ from a constant function by less than $\frac{\pi E}{4a}$ and it is a composition of

- reflection $(\overline{q}_1, \overline{q}_2) \to (\overline{q}_2, \overline{q}_1)$,
- re-scaling $(\overline{q}_1, \overline{q}_2) \to \frac{1}{a}(\sqrt[3]{\overline{q}_1}, \sqrt[3]{\overline{q}_2})$,
- rotation $(\overline{q}_1, \overline{q}_2) \to (\overline{q}_1 \cos\varphi - \overline{q}_2 \sin\varphi, \overline{q}_2 \cos\varphi + \overline{q}_1 \sin\varphi)$ by angle φ in the anticlockwise direction.

The re-scaling part of the composition ensures that constants R and δ can be specified so that the solution operator T_φ takes constant functions satisfying (ii) and (iii) to functions that again satisfy (ii) and (iii). Since functions in Q differ from their integral averages by less than ε, we need to find R, δ and ε such that the following conditions are satisfied:

$$\frac{1}{a}\sqrt[3]{R} \leqslant R - \varepsilon - \frac{\pi E}{4a} \quad \text{and} \quad \frac{\sqrt{2}}{2a}\sqrt[3]{\delta} \geqslant \delta + \sqrt{2}\varepsilon + \frac{\pi E}{4a}. \tag{6.29}$$

Inequalities (6.29) guarantee that x satisfies conditions (ii) and (iii) of the definition of the parameter set Q. Taking further $\varepsilon \leqslant \frac{\delta}{2}$ ensures that Q is a non-trivial ANR-space (leaving the "hole" inside).

Starting from $a > 0$ and $\omega > 0$, we have specified constants R, δ, B, E and ε such that set Q becomes invariant under the solution operator T_φ which takes any $q \in Q$ to the solutions x of the linearized problem (6.19), (6.20), for $\varphi \in [0, \frac{\pi}{4}]$. Moreover, observe that, for $\varphi = \frac{\pi}{4}$, there are no easily detectable subdomains of Q separately invariant under operator $T_{\pi/4}$. Figure 5 shows how operator $T_{\pi/4}$ treats constant functions in Q, for a particular choice of R and δ and helps understanding why we can not easily detect any subinvariant domains of Q.

Hence, a slightly modified special case of Theorem 6.1 (cf. [18]) ensures that system (6.15), (6.16) admits at least $N(T_{\pi/4})$ solutions. Since $T_{\pi/4}$ is compactly admissibly homotopic to T_0, we have $N(T_{\pi/4}) = N(T_0)$.

Let us further consider the retraction $r : Q \to Q$ which sends a function $q \in Q$ to its integral average $\overline{q}$. Let us define the homotopy $T^\mu : [0, 1] \times Q \to Q$ by

$$T^\mu(q) := \mu T_0(q) + (1 - \mu) r\bigl(T_0(q)\bigr).$$

This compact homotopy guarantees that $N(T^1) = N(T_0)$ equals to $N(T^0) = N(r \circ T_0)$. We can thus restrict ourselves to the computation of $N(r \circ T_0)$. Let us denote by $\overline{Q}$ the subset of Q consisting of constant functions. Since $r \circ T_0 : Q \to \overline{Q}$, all the fixed points of $r \circ T_0$ have to belong to $\overline{Q}$. Let us therefore deal with the restriction $L := r \circ T_0|_{\overline{Q}} : \overline{Q} \to \overline{Q}$, which can be explicitly written in the form

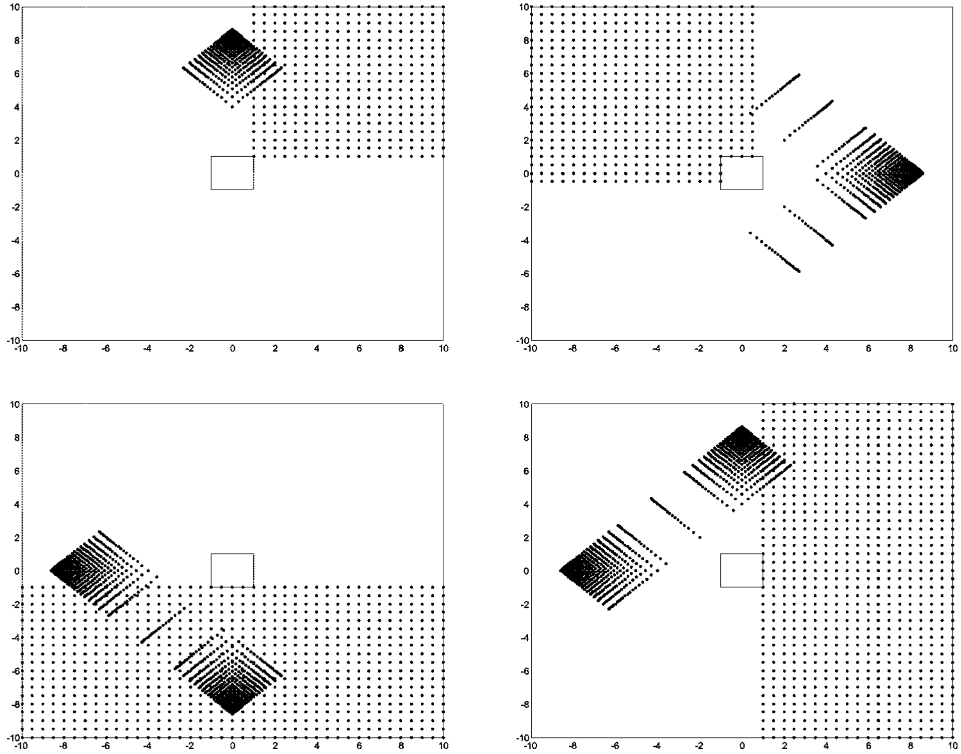


Fig. 5. Behaviour of $T_{\pi/4}$ on constant functions on Q for $R = 10$, $\delta = 1$ and $a = \frac{\sqrt{2}}{4}$. Rectangular grids of points represent constant functions $q \in Q$, the irregular grid represents their images under $T_{\pi/4}$. For simplicity, we take here $E = 0$, so that the images of constant functions become constant again. No easily detectable regions of the domain are subinvariant.

$$L(\overline{q}_1, \overline{q}_2) := \frac{1}{a}\left(\sqrt[3]{\overline{q}_2}, \sqrt[3]{\overline{q}_1}\right).$$

One can easily check by an explicit computation that L has two fixed-points in $\overline{Q}$ which belong to different Nielsen classes. Therefore, according to (reduction) Lemma 6.1, $N(L) = N(r \circ T_0) = 2$. This finally shows that system (6.17), (6.18) admits at least two ω-periodic solutions.

We are in the position to formulate the multiplicity criterion for ω-periodic solutions to system (6.17), (6.18).

THEOREM 6.3. *Let the following inequalities be satisfied*:

$$\delta < \frac{0.247}{a^{3/2}}, \quad \frac{\sqrt[3]{R}}{a} \leqslant R - \frac{5\delta}{8}, \quad E \leqslant \frac{a\delta}{2\sqrt{2}\pi}, \quad B \leqslant \left(\frac{a}{12}\right)^3 \frac{\delta^2}{4\sqrt{2}}, \tag{6.30}$$

and take $\varepsilon = \frac{\delta}{2}$. *Then system* (6.17), (6.18) *admits at least three* ω*-periodic solutions.*

PROOF. The assumptions guarantee that inequalities (6.28) and (6.29) are satisfied. Thus, two ω-periodic solutions have been already obtained by means of the Nielsen number, as above. The third can be proved quite analogously as in the proof of Theorem 6.2, by the additivity property of the fixed point index (see Proposition 3.2). □

REMARK 6.4. The inequalities in Theorem 6.3 are satisfied, e.g., for $a = \frac{\sqrt{2}}{4}$, $\delta = 1$, $R = 10$, $B \leqslant \frac{1}{2^4 3^3}$, and $E \leqslant \frac{1}{4\pi}$, as in Fig. 5, where $E = 0$.

REMARK 6.5. If the equality appears for δ in (6.30), then in the lack of periodicity, at least two entirely bounded solutions can be proved as in Theorem 6.2.

6.3. *Several anti-periodic solutions*

The following approach is via the Poincaré translation operator treated in Section 4.3. Hence, consider the system of differential inclusions

$$\dot{x} \in F(t, x), \tag{6.31}$$

where $F : \mathbb{R}^{n+1} \multimap \mathbb{R}^n$ is a u-Carathéodory mapping with nonempty, compact and convex values, satisfying (6.4). Then all solutions of (6.31) exist in the sense of Carathéodory, namely they are locally absolutely continuous and satisfy (6.31) a.e.

If $x(t, x_0) := x(t, 0, x_0)$ is a solution of (6.31) with $x(0, x_0) = x_0$, then we can define the Poincaré map (translation operator at the time $T > 0$) along the trajectories of (6.31) as follows:

$$\Phi_T : \mathbb{R}^n \multimap \mathbb{R}^n, \quad \Phi_T(x_0) := \bigl\{x(T, x_0) \mid x(t, x_0) \text{ satisfies } (6.31)\bigr\}. \tag{6.32}$$

The goal is to represent the admissible (see Theorem 4.17) map Φ_T in terms of an admissible pair (see Definition 2.5 in Section 2.2). We let $\varphi : \mathbb{R}^n \multimap C([0, T], \mathbb{R}^n)$, where

$$\varphi(X) := \bigl\{x \in C\bigl([0, T], \mathbb{R}^n\bigr) \bigm| x(0) = X \text{ and } x \text{ satisfies } (6.31)\bigr\}$$

and $C([0, T], \mathbb{R}^n)$ is the Banach space of continuous maps. According to Theorem 4.17, φ is an R_δ-mapping.

Now, we let

$$e_T : C\bigl([0, T), \mathbb{R}^n\bigr) \to \mathbb{R}^n, \quad e_T(x) = x(T),$$

where e_T is evidently continuous.

One can readily check that $\Phi_T = e_T \circ \varphi$. Moreover, $\mathcal{H} \circ \Phi_T = \mathcal{H} \circ \varphi \circ e_T$ is admissible for any homeomorphism $\mathcal{H} : \mathbb{R}^n \to \mathbb{R}^n$ (see Remark 4.15).

In fact, we have the diagram

$$\mathbb{R}^n \xleftarrow{p_\varphi} \Gamma_\varphi \xrightarrow{q_\varphi} C\bigl([0, a], \mathbb{R}^n\bigr) \xrightarrow{e_T} \mathbb{R}^n \xrightarrow{\mathcal{H}} \mathbb{R}^n,$$

where p_φ, q_φ are natural projections.

In what follows, we identify the Poincaré map Φ_T or its composition with $\mathcal{H}$, i.e. $\mathcal{H} \circ \Phi_T$, with the admissible pair $(p_\varphi, e_T \circ q_\varphi)$ or $(p_\varphi, \mathcal{H} \circ e_T \circ q_\varphi)$, i.e. we let

$$\Phi_T = (p_\varphi, e_T \circ q_\varphi) \quad \text{or} \quad \mathcal{H} \circ \Phi_T = (p_\varphi, \mathcal{H} \circ e_T \circ q_\varphi), \tag{6.33}$$

respectively.

Let $C(\mathcal{H} \circ \Phi_T)$ denote the set of coincidence points of the pair $(p_\varphi, \mathcal{H} \circ e_T \circ q_\varphi)$, while $\operatorname{Fix}(\mathcal{H} \circ \Phi_T) = \{X \in \mathbb{R}^n \mid X \in \mathcal{H}(e_T(q_\varphi(p_\varphi^{-1}(X))))\}$.

A pair for Φ_T can be easily shown to be homotopic to the identity map, so that the pair $(p_\varphi, \mathcal{H} \circ e_T \circ q_\varphi)$ is homotopic to $\mathcal{H}$ (see Theorem 4.17). We have a one-to-one correspondence between coincidence points and solutions. Since a coincidence point of $(p_\varphi, \mathcal{H} \circ e_T \circ q_\varphi)$ gives us in this way a solution $x(t)$ of (6.31) such that $x(0) = \mathcal{H}(x(T))$, the following proposition is self-evident.

PROPOSITION 6.1. *If* $\#C(p_\varphi, \mathcal{H} \circ e_T \circ q_\varphi) = \operatorname{card} C(\mathcal{H} \circ \Phi_T) \geqslant k$, *then system* (6.31) *has at least k solutions* $x_1(t), \ldots, x_k(t)$ *such that* $x_i(0) = \mathcal{H}(x_i(T))$, $i = 1, \ldots, k$.

The following lemma immediately follows from Theorem 4.17.

LEMMA 6.3. *Assume that* $Y \subset \mathbb{R}^n$ *is a compact connected ANR-space such that* $\Phi_s(Y) \subset Y$, *for every* $s \in [0, T]$ *and* $\mathcal{H}(Y) \subset Y$. *Then the pair* $(p_\varphi, \mathcal{H} \circ e_T \circ q_\varphi)$ *restricted to Y is admissibly homotopic to* $\mathcal{H}|_Y$.

As a consequence of Theorems 3.4 and 3.7 in Section 3.2, Proposition 6.1 and Lemma 6.3, we have the following

PROPOSITION 6.2. *Assume that* Φ_s, *for every* $s \in [0, T]$ *in* (6.32) *is a self-map on the torus* $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. *Then system* (6.31) *has at least* $N(\mathcal{H})$ *solutions* $x(t)$ *such that* $x(0) = \mathcal{H}(x(T))$ *on* $\mathbb{T}^n$, *where* $N(\mathcal{H})$ *denotes the Nielsen number of a homeomorphism* $\mathcal{H}: \mathbb{T}^n \to \mathbb{T}^n$.

PROOF. The proof follows directly from Theorems 3.4 and 3.7 in Section 3.2, Proposition 6.1, Lemma 6.3 and the properties of the Nielsen number mentioned in Section 3.2. □

If in particular $\mathcal{H} = \mathrm{id}$, then according to Remark 3.6,

$$N(\mathrm{id}) = |\lambda(\mathrm{id})| = |\chi(\cdot)|$$

holds on tori, and consequently the problem considered in Proposition 6.2 should have at least $|\chi(\cdot)|$ solutions, where λ is the Lefschetz number and χ denotes the Euler–Poincaré characteristic of $\mathbb{T}^n$. Since, unfortunately, $\chi(\cdot) = 0$ for tori, this is not a suitable case for applications. In other words, we are not able to establish several T-periodic solutions in this way.

On the other hand, as the simplest application of Proposition 6.2, we can give immediately

THEOREM 6.4. *Assume that*

$$F(t,\dots,x_j+1,\dots)\equiv F(t,\dots,x_j,\dots),\quad \textit{for } j=1,\dots,n, \tag{6.34}$$

where $x=(x_1,\dots,x_n)$ *and consider system* (6.31) *on the set* $[0,\infty)\times\mathbb{T}^n$, *where* $\mathbb{T}^n=\mathbb{R}^n/\mathbb{Z}^n$. *Then system* (6.31) *admits, for every positive constant* T, *at least* $N(\mathcal{H})$ *solutions* $x(t)$ *such that*

$$x(0)=\mathcal{H}\big(x(T)\big)\pmod 1,$$

where $\mathcal{H}$ *is a continuous self-map on* $\mathbb{T}^n$ *and* $N(\mathcal{H})$ *denotes the associated Nielsen number.*

As a consequence, we obtain for $\mathcal{H}=-\mathrm{id}$ easily:

COROLLARY 6.2. *If, in addition to the assumptions of Theorem* 6.4,

$$F(t+T,-x)\equiv -F(t,x),$$

then system (6.31) *admits at least* 2^n *anti-*T*-periodic* (*or* $2T$*-periodic*) *solutions* $x(t)$ *on* $\mathbb{T}^n$, *namely* $x(t+T)\equiv -x(t)\pmod 1$.

PROOF. According to Theorem 6.4, system (6.31) admits at least $N(-\mathrm{id})$ anti-T-periodic solutions on $\mathbb{T}^n$. On $\mathbb{T}^n$, the following formula holds (see Remark 3.6), $N(-\mathrm{id})=|\lambda(-\mathrm{id})|=|\det 2I|=2^n$, which completes the proof. □

7. Remarks and comments

7.1. *Remarks and comments to general methods*

Theorems 4.1, 4.2 and Corollaries 4.1, 4.2, 4.3 are taken from [19]. Corollaries 4.1 and 4.4 generalize many single-valued situations, e.g., in [52] and [63], where the parameter set Q was only convex. Corollary 4.3 was employed for the first time in [32,33]. Corollary 4.4 generalizes the single-valued case in [39].

Theorem 4.3, for the lower estimate of the number of solutions, comes from [23]. Its slightly modified version (cf. Remark 4.3) was presented in [5]. As far as we know, there are no other general methods, using the Nielsen number, as our Theorem 4.3.

Theorem 4.4, extending Theorem 4.1 to Banach spaces, was published in [14]. Corollary 4.5 is formally new. Intuitively clear Propositions 4.4 and 4.5 (see also Remark 4.6) are contained in [22, Chapter III.1].

In view of Remark 4.6, one can check by means of the Gronwall inequality (cf. [71]) that differential inclusion $\dot{x}\in F(t,x)$ admits an entirely bounded solution $x\in AC_{\mathrm{loc}}(\mathbb{R},E)$ with $x(0)=x_0$ such that

$$\|x(t)\|\leqslant\left(\|x_0\|+\int_{-\infty}^{\infty}r(t)\,\mathrm{d}t\right)\exp\int_{-\infty}^{\infty}r(t)\,\mathrm{d}t,\qquad t\in\mathbb{R},$$

provided E is a separable or reflexive Banach space and $F: \mathbb{R} \times E \multimap E$ is a u-Carathéodory mapping (cf. Definition 2.10) such that:

(i) $\mu(F(t, B)) \leqslant k(t)\mu(B)$, for bounded subsets $B \subset E$, $t \in \mathbb{R}$, where $\mu = \alpha$ or $\mu = \gamma$, and $k \in L^1_{\mathrm{loc}}(\mathbb{R})$,

(ii) $\|y\| \leqslant r(t)(1 + \|x\|)$, for every $(t, x) \in \mathbb{R} \times E$, $y \in F(t, x)$, where $r: \mathbb{R} \to [0, \infty)$ is an integrable function such that $\int_{-\infty}^{\infty} r(t)\,\mathrm{d}t < \infty$.

Many alternative continuation principles for ODEs can be found, e.g., in [45,58,61,62, 64,65,82,83,90], and the references therein.

Theorems 4.6, 4.8, 4.9 and Proposition 4.9, dealing with the topological structure of solution sets are taken from [19]. Theorem 4.9 extends Theorem 4.5 in [57] for arbitrary (possibly infinite) intervals. In [20,21], we have developed and applied a powerful inverse limit method for the investigation of the topological structure of the solution sets. In these papers, the structure was also systematically studied for less regular right-hand sides of multivalued ODEs than those in Section 4.2.

Theorems 4.10 and 4.11 concerning the topological dimensions of solution sets come from [66]. An extension to multivalued ODEs in Banach spaces was recently published in [45]. For further results concerning the solution sets to initial value problems, see, e.g., [42,55,58,59,67,74,78,97].

Unlike for initial value problems, there are only several results concerning the topological structure of solution sets to boundary value problems. One of the most important is Theorem 4.13 in [43] which was improved by us (cf. Remark 4.10) in [21]. Theorem 4.12 comes from [20], Theorem 4.14 from [46] and Theorems 4.15, 4.16 from [7]. In [22, Chapter III.3], we collected for the first time practically all results in this field (cf. also [59, Chapter 6]). Paper [43] contains also the information about the topological dimension of the solution sets.

Theorem 4.17 about admissibility of Poincaré's operators appeared on various levels of abstraction in many papers (cf. [22, Chapter III.4 and the related comments on pp. 592–593]). For some generalizations and extensions, see [2,10,28], and the references therein.

7.2. *Remarks and comments to existence results*

Theorem 5.1 from [14] can be regarded as an infinite-dimensional extension of our earlier results in [3,4], where under suitable additional restrictions also almost-periodic solutions were detected. In [15], the methods applied to Theorem 5.1 were modified for obtaining almost-periodic solutions with values in separable Banach spaces. Theorem 5.2 is contained in [20].

Proposition 5.1, for solvability of a rather general class of boundary value problems with linear conditions, was presented for the first time in [3].

There is an enormous amount of results about the existence of periodic and anti-periodic solutions (see, e.g., [22,62,64,65,70,74,78,81,82,90]). Our Corollary 5.1 is formally new. Corollaries 5.2, 5.3, 5.4 can be found in [22, Chapter III.5]. As pointed out in Remark 5.7, the requirement concerning a fixed point free boundary of a parameter set Q were satisfied in [11,32–34] by means of bound sets. An alternative approach for satisfying this requirement can be found, e.g., in [30,31], where canonical domains or upper and lower solutions technique were applied, in the single-valued case (cf. also [65,70,81,82,90]).

7.3. *Remarks and comments to multiplicity results*

Our multiplicity results are based on the application of the Nielsen theory in Section 3.2, eventually combined with the additivity property of the fixed point index in Section 3.3. For further Nielsen theories which can be also used here, see [12,22–26,38], and the references therein. Practically all results obtained in this way were collected in [13].

Corollary 6.1 is from [5]. Variants of Theorems 6.1 and 6.2 can be found in [5,6,9], and [23], while Theorem 6.3 is a multivalued generalization of a single-valued version in [17]. Theorem 6.4 and Corollary 6.2 were presented for the first time in [24].

As pointed out in Section 1, the delicate problem of application of the Nielsen theory to differential equations is associated with the name of J. Leray.

For further multiplicity results obtained by different methods, see, e.g., [22, Chapter III.6; 62; 82, Chapter 6], and the references therein.

Acknowledgements

The collaboration and fruitful discussions with the coauthors of the quoted papers, in particular with Prof. Lech Górniewicz (Toruń), are highly appreciated. The author is also indebted to Dr. Tomáš Fürst (Olomouc) for his critical reading of the manuscript.

References

[1] R.P. Agarwal and V. Lakshmikantham, *Uniqueness and Nonuniqueness Criteria for Ordinary Differential Equations*, World Scientific, Singapore (1993).

[2] J. Andres, *On the multivalued Poincaré operators*, Topol. Methods Nonlinear Anal. **10** (1997), 171–182.

[3] J. Andres, *Bounded, almost-periodic and periodic solutions of quasi-linear differential inclusions*, Differential Inclusions and Optimal Control, J. Andres, L. Górniewicz and P. Nistri, eds., Lecture Notes in Nonlinear Anal., Vol. 2, N. Copernicus Univ., Toruń (1998), 19–32.

[4] J. Andres, *Almost-periodic and bounded solutions of Carathéodory differential inclusions*, Differential Integral Equations **12** (1999), 887–912.

[5] J. Andres, *Multiple bounded solutions of differential inclusions. The Nielsen theory approach*, J. Differential Equations **155** (1999), 285–310.

[6] J. Andres, *A nontrivial example of application of the Nielsen fixed-point theory to differential systems: Problem of Jean Leray*, Proc. Amer. Math. Soc. **128** (2000), 2921–2931.

[7] J. Andres, *Ordinary differential equations in the lack of uniqueness*, Atti. Sem. Mat. Fis. Univ. Modena **49** (2001), 247–267.

[8] J. Andres, *Nielsen number, Artin braids, Poincaré operators and multiple nonlinear oscillations*, Nonlinear Anal. **47** (2001), 1017–1028.

[9] J. Andres, *Nielsen number and multiplicity results for multivalued boundary value problems*, Nonlinear Analysis and Differential Equations, M.R. Grossinho, M. Ramos, C. Rebelo and L. Sanchez, eds., Birkhäuser, Basel (2001), 175–187.

[10] J. Andres, *Poincaré's translation multioperator revisited*, Proceedings of the 3rd Polish Symposium on Nonlinear Analysis, W. Kryszewski and A. Nowakowski, eds., Lecture Notes in Nonlinear Anal., Vol. 3, N. Copernicus Univ., Toruń (2002), 7–22.

[11] J. Andres, *Using the integral manifolds to solvability of boundary value problems*, Set-Valued Mappings with Application in Nonlinear Analysis, R.P. Agarwal and D. O'Regan, eds., Ser. Math. Anal. Appl., Vol. 4, Taylor and Francis, Singapore (2002), 27–38.

[12] J. Andres, *Applicable fixed point principles*, Handbook of Topological Fixed Point Theory, R.F. Brown, M. Furi, L. Górniewicz and B. Jiang, eds., Springer, Berlin (2005), 687–739.

[13] J. Andres, *Nielsen number and differential equations*, Fixed Point Theory Appl. **2** (2005), 137–167.

[14] J. Andres and R. Bader, *Asymptotic boundary value problems in Banach spaces*, J. Math. Anal. Appl. **247** (2002), 437–457.

[15] J. Andres and A.M. Bersani, *Almost-periodicity problem as a fixed-point problem for evolution inclusions*, Topol. Methods Nonlinear Anal. **18** (2001), 337–349.

[16] J. Andres, J. Fišer and L. Jüttner, *On a multivalued version of the Sharkovskiĭ theorem and its application to differential inclusions*, Set-Valued Anal. **10** (2002), 1–14.

[17] J. Andres and T. Fürst, *An example of application of the Nielsen theory to integro-differential equations*, Proc. Amer. Math. Soc. **134** (2006), 1985–1993.

[18] J. Andres and T. Fürst, *Nontrivial applications of Nielsen theory to differential systems*, J. Differential Equations, to appear.

[19] J. Andres, G. Gabor and L. Górniewicz, *Boundary value problems on infinite intervals*, Trans. Amer. Math. Soc. **351** (1999), 4861–4903.

[20] J. Andres, G. Gabor and L. Górniewicz, *Topological structure of solution sets to multivalued asymptotic problems*, Z. Anal. Anwendungen **19** (2000), 35–60.

[21] J. Andres, G. Gabor and L. Górniewicz, *Acyclicity of solution sets to functional inclusions*, Nonlinear Anal. **49** (2002), 671–688.

[22] J. Andres and L. Górniewicz, *Topological Fixed Point Principles for Boundary Value Problems*, Kluwer, Dordrecht (2003).

[23] J. Andres, L. Górniewicz and J. Jezierski, *Noncompact version of the multivalued Nielsen theory and its application to differential inclusions*, Differential Inclusions and Optimal Control, J. Andres, L. Górniewicz and P. Nistri, eds., Lecture Notes in Nonlinear Anal., Vol. 2, N. Copernicus Univ., Toruń (1998), 33–50.

[24] J. Andres, L. Górniewicz and J. Jezierski, *A generalized Nielsen number and multiplicity results for differential inclusions*, Topol. Appl. **100** (2000), 193–209.

[25] J. Andres, L. Górniewicz and J. Jezierski, *Relative versions of the multivalued Lefschetz and Nielsen theorems and their application to admissible semi-flows*, Topol. Methods Nonlinear Anal. **16** (2000), 73–92.

[26] J. Andres, L. Górniewicz and J. Jezierski, *Periodic points of multivalued mappings with applications to differential inclusions on tori*, Topol. Appl. **127** (2003), 447–472.

[27] J. Andres, L. Górniewicz and P. Nistri, eds., *Differential Inclusions and Optimal Control*, Lecture Notes in Nonlinear Anal., Vol. 2, J. Schauder Center for Nonlinear Studies and N. Copernicus Univ., Toruń (1998).

[28] J. Andres and L. Jüttner, *Periodic solutions of discontinuous differential systems*, Nonlinear Anal. Forum **6** (2001), 391–407.

[29] J. Andres, L. Jüttner and K. Pastor, *On a multivalued version of the Sharkovskiĭ theorem and its application to differential inclusions, II*, Set-Valued Anal. **13** (2005), 47–68.

[30] J. Andres and B. Krajc, *Bounded solutions in a given set of Carathéodory differential systems*, J. Comput. Appl. Math. **113** (2000), 73–82.

[31] J. Andres and B. Krajc, *Periodic solutions in a given set of differential systems*, J. Math. Anal. Appl. **264** (2001), 495–509.

[32] J. Andres, L. Malaguti and V. Taddei, *Floquet boundary value problems for differential inclusions: A bound sets approach*, Z. Anal. Anwendungen **20** (2001), 709–725.

[33] J. Andres, L. Malaguti and V. Taddei, *Bounded solutions of Carathéodory differential inclusions: A bound sets approach*, Abstr. Appl. Anal. **9** (2003), 547–571.

[34] J. Andres, L. Malaguti and V. Taddei, *Boundary value problems in Banach spaces: A bound sets approach*, in preparation.

[35] J. Andres and K. Pastor, *On a multivalued version of the Sharkovskiĭ theorem and its application to differential inclusions, III*, Topol. Methods Nonlinear Anal. **22** (2003), 369–386.

[36] J. Andres and K. Pastor, *A version of Sharkovskii's theorem for differential equations*, Proc. Amer. Math. Soc. **133** (2005), 449–453.

[37] J. Andres and M. Väth, *Coincidence index for noncompact mappings on nonconvex sets*, Nonlinear Funct. Anal. Appl. **7** (2002), 619–658.

[38] J. Andres and M. Väth, *Two topological definitions of a Nielsen number for coincidences of noncompact maps*, Fixed Point Theory Appl. **1** (2004), 49–69.

[39] G. Anichini, G. Conti and P. Zecca, *Using solution sets for solving boundary value problems for ordinary differential equations*, Nonlinear Anal. **17** (1991), 465–472.

[40] J. Appell, E. De Pascale, N.H. Thái and P.P. Zabreiko, *Multi-Valued Superpositions*, Diss. Math., Vol. 345, PWN, Warsaw (1995).

[41] N. Aronszajn, *Le correspondant topologique de l'unicité dans la théorie des équations différentialles*, Ann. Math. **43** (1942), 730–738.

[42] J.-P. Aubin and A. Cellina, *Differential Inclusions*, Springer, Berlin (1984).

[43] A. Augustynowicz, Z. Dzedzej and B.D. Gel'man, *The solution set to BVP for some functional differential inclusions*, Set-Valued Anal. **6** (1998), 257–263.

[44] R. Bader, *Fixpunktindextheorie mengenwertiger Abbildungen und einige Anwendungen*, Ph.D. Thesis, University of Munich (1995).

[45] R. Bader, B.D. Gel'man, M. Kamenskii and V. Obukhovskii, *On the topological dimension of the solution sets for some classes of operator and differential inclusions*, Discuss. Math. Differ. Incl. Control Optim. **22** (2002), 17–32.

[46] D. Bielawski and T. Pruszko, *On the structure of the set of solutions of a functional equation with application to boundary value problems*, Ann. Polon. Math. **53** (1991), 201–209.

[47] K. Borsuk *Theory of Retracts*, Mathematical Monographs, Vol. 44, PWN, Warsaw (1967).

[48] S. Bochner, *Integration von Funktionen deren Werte die Elemente eines Vektorraumes sind*, Fund. Math. **20** (1933), 262–276.

[49] A. Bressan, *Directionally continuous selections and differential inclusions*, Funkcial. Ekvac. **31** (1988), 459–470.

[50] A. Bressan, *On the qualitative theory of lower semicontinuous differential inclusions*, J. Differential Equations **77** (1989), 379–391.

[51] R.F. Brown, *The Lefschetz Fixed Point Theorem*, Scott, Foresman and Co., Glenview, IL (1971).

[52] M. Cecchi, M. Furi and M. Marini, *About solvability of ordinary differential equations with asymptotic boundary conditions*, Boll. Un. Mat. Ital. C (6) **4** (1985), 329–345.

[53] S.-N. Chow and A. Lasota, *On boundary value problems for ordinary differential equations*, J. Differential Equations **14** (1973), 326–337.

[54] E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York (1955).

[55] G. Conti, V. Obukhovskii and P. Zecca, *On the topological structure of the solution set for a semilinear functional-differential inclusion in a Banach space*, Topology in Nonlinear Analysis Banach Center Publ., Vol. 35, K. Gęba and L. Górniewicz, eds., Inst. of Polish Acad. Sci. Math., Warsaw (1996), 159–169.

[56] J.L. Daleckiĭ and M.G. Krein, *Stability of Solutions of Differential Equations and Banach Space*, Transl. Math. Monogr., Vol. 43, Amer. Math. Soc., Providence, RI (1974).

[57] F.S. DeBlasi and J. Myjak, *On the solutions sets for differential inclusions*, Bull. Polish Acad. Sci. Math. **12** (1985), 17–23.

[58] K. Deimling, *Multivalued Differential Equations*, W. de Gruyter, Berlin (1992).

[59] R. Dragoni, J.W. Macki, P. Nistri and P. Zecca, *Solution Sets of Differential Equations in Abstract Spaces*, Pitman Res. Notes in Math., Vol. 342, Longman, Harlow (1996).

[60] A.F. Filippov, *Differential Equations with Discontinuous Right-Hand Sides*, Kluwer, Dordrecht (1999).

[61] M. Frigon and A. Granas, eds., *Topological Methods in Differential Equations and Inclusions*, Kluwer, Dordrecht (1995).

[62] S. Fučík, *Solvability of Nonlinear Equations and Boundary Value Problems*, Reidel, Dordrecht (1988).

[63] M. Furi and P. Pera, *A continuation method on locally convex spaces and applications to ordinary differential equations on noncompact intervals*, Ann. Polon. Math. **47** (1987), 331–346.

[64] M. Furi and P. Zecca, eds., *Topological Methods in the Theory of Ordinary Differential Equations*, Lecture Notes in Math., Vol. 1537, Springer, Berlin (1993).

[65] R.G. Gaines and J. Mawhin, *Coincidence Degree and Nonlinear Differential Equations*, Lecture Notes in Math., Vol. 568, Springer, Berlin (1977).

[66] B.D. Gelman, *On topological dimension of a set of solutions of functional inclusions*, Differential Inclusions and Optimal Control, J. Andres, L. Górniewicz and P. Nistri, eds., Lecture Notes in Nonlinear Anal., Vol. 2, N. Copernicus Univ., Toruń (1998), 163–178.

[67] L. Górniewicz, *Topological Fixed Point Theory of Multivalued Mappings*, Kluwer, Dordrecht (1999).

[68] L. Górniewicz, *On the Lefschetz fixed point theorem*, Handbook of Topological Fixed Point Theory, R.F. Brown et al., eds., Springer, Berlin (2005), 43–82.

[69] A. Granas and J. Dugundji, *Fixed Point Theory*, Springer, Berlin (2003).

[70] A. Granas, R. Guenther and J. Lee, *Nonlinear Boundary Value Problems for Ordinary Differential Equations*, Diss. Math., Vol. 244, Warsaw (1985).

[71] P. Hartman, *Ordinary Differential Equations*, Willey, New York (1964).

[72] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Math., Vol. 840, Springer, Berlin (1981).

[73] S. Hu and N.S. Papageorgiou, *Handbook of Multivalued Analysis, Vol. I: Theory*, Kluwer, Dordrecht (1997).

[74] S. Hu and N.S. Papageorgiou, *Handbook of Multivalued Analysis, Vol. II: Applications*, Kluwer, Dordrecht (2000).

[75] M. Hukuhara, *Sur les systèmes des équations différentialles ordinaires*, Japan J. Math. **5** (1928), 345–350.

[76] J. Jarník and J. Kurzweil, *On conditions on right hand sides of differential relations*, Časopis Pěst. Mat. **102** (1977), 334–349.

[77] B. Jiang, *Lectures on Nielsen Fixed Point Theory*, Contemp. Math., Vol. 14, Amer. Math. Soc., Providence, RI (1983).

[78] M.I. Kamenskii, V.V. Obukhovskii and P. Zecca, *Condensing Multivalued Maps and Semilinear Differential Inclusions in Banach Space*, W. de Gruyter, Berlin (2001).

[79] M. Kisielewicz, *Differential Inclusions and Optimal Control*, Kluwer, Dordrecht (1991).

[80] H. Kneser, *Über die Lösungen eine Systeme gewöhnlicher Differentialgleichungen, das der Lipschitzschen Bedingung nicht genügt*, S.B. Preuss. Akad. Wiss. Phys. Math. Kl. **4** (1923), 171–174.

[81] M.A. Krasnosel'skiĭ, *Translation Operator Along the Trajectories of Differential Equations*, Amer. Math. Soc., Providence, RI (1968).

[82] M.A. Krasnosel'skii and P.P. Zabreiko, *Geometrical Methods of Nonlinear Analysis*, Springer, Berlin (1984).

[83] W. Krawcewicz and J. Wu, *Theory of Degrees with Applications to Bifurcations and Differential Equations*, Willey, New York (1997).

[84] M. Kunze, *Non-Smooth Dynamical Systems*, Lecture Notes in Math., Vol. 1744, Springer, Berlin (2000).

[85] A. Lasota, *On the existence and uniqueness of solutions of a multipoint boundary value problem*, Ann. Polon. Math. **38** (1980), 205–310.

[86] J.L. Massera and J.J. Schaeffer, *Linear Differential Equations and Function Spaces*, Academic Press, New York (1966).

[87] T. Matsuoka, *The number and linking of periodic solutions of periodic systems*, Invent. Math. **70** (1983), 319–340.

[88] T. Matsuoka, *Waveform in the dynamical study of ordinary differential equations*, Japan J. Appl. Math. **1** (1984), 417–434.

[89] T. Matsuoka, *The number and linking of periodic solutions of non-dissipative systems*, J. Differential Equations **76** (1988), 190–201.

[90] J. Mawhin, *Topological Degree Methods in Nonlinear Boundary Value Problems*, CBMS Lecture Notes in Math., Vol. 40, Amer. Math. Soc., Providence, RI (1979).

[91] M.D.P. Monteiro Marques, *Differential Inclusions in Nonsmooth Mechanical Problems*, Birkhäuser, Basel (1993).

[92] W. Orlicz, *Zur Theorie der Differentialgleichung* $y' = f(x, y)$, Bull. Akad. Polon. Sci. Ser. A (1932), 221–228.

[93] K.J. Palmer, *Exponential dichotomies for almost-periodic equations*, Proc. Amer. Math. Soc., **101** (1987), 293–298.

[94] V.A. Pliss, *Nonlocal Problems in the Theory of Oscillations*, Nauka, Moscow (1964) (in Russian); Academic Press, New York (1966).

[95] A.N. Sharkovskii, *Coexistence of cycles of a continuous map of a line into itself*, Ukrainian Math. J. **16** (1964) 61–71 (in Russian).

[96] G.V. Smirnov, *Introduction to the Theory of Differential Inclusions*, Grad. Stud. Math., Vol. 41, Amer. Math. Soc., Providence, RI (2002).

[97] A.A. Tolstonogov, *Differential Inclusions in a Banach Space*, Kluwer, Dordrecht (2000).

[98] Ya.I. Umanskii, *On a property of the solution set of differential inclusions in a Banach space*, Differ. Uravn. **28** (1992), 1346–1351 (in Russian).

[99] M. Väth, *Volterra and Integral Equations of Vector Functions*, M. Dekker, New York (2000).

[100] I.I. Vrabie, *Compactness Methods for Nonlinear Evolutions*, second edition, Longman House, Burn Mill, Harlow (1995).

This page intentionally left blank

CHAPTER 2

Heteroclinic Orbits for Some Classes of Second and Fourth Order Differential Equations

Denis Bonheure

Université Catholique de Louvain, Institut de Mathématique Pure et Appliquée, Chemin du Cyclotron 2,
1348 Louvain-La-Neuve, Belgium
E-mail: bonheure@anma.ucl.ac.be

Luís Sanchez

Faculdade de Ciências da Universidade de Lisboa, CMAF, Avenida Professor Gama Pinto 2, 1649-003 Lisboa,
Portugal
E-mail: sanchez@ptmat.fc.ul.pt

Contents

1. Introduction . . . 105
Part 1. Second Order Equations and Systems . . . 109
2. Second order Hamiltonian systems . . . 109
  2.1. Heteroclinics for some scalar second order differential equations . . . 109
  2.2. Autonomous Hamiltonian systems . . . 117
  2.3. Periodic Hamiltonian systems: Multiplicity. Multibump solutions . . . 123
  2.4. Notes and further comments . . . 129
3. Wave profiles in reaction diffusion equations . . . 131
  3.1. Reduction to a first order equation . . . 133
  3.2. Fast solutions and heteroclinics . . . 137
  3.3. Heteroclinics in a combustion model . . . 142
  3.4. Notes and further comments . . . 144
Part 2. Fourth Order Equations . . . 145
4. The Extended Fisher–Kolmogorov equation: An overview . . . 145
  4.1. A shooting method . . . 147
  4.2. Minimization . . . 151
  4.3. Homotopy classes of multi-transition solutions . . . 155
  4.4. Notes and further comments . . . 157
5. Sign changing Lagrangians . . . 159
  5.1. Functionals with sign changing acceleration coefficient . . . 160

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 3
Edited by A. Cañada, P. Drábek and A. Fonda

© 2006 Elsevier B.V. All rights reserved

5.2. Functionals of Swift–Hohenberg type 171
5.3. Notes and further comments 180
6. A Ginzburg–Landau model for ternary mixtures: Connections between nonconsecutive equilibria 183
6.1. Local analysis of a saddle-focus equilibrium 186
6.2. Clipping 189
6.3. Existence of a minimizer 191
6.4. Notes and further comments 194
7. Multi-transition connections 194
7.1. Notes and further comments 199
References 199

Keywords: heteroclinic solutions, variational methods, minimization, Hamiltonian systems, Fisher–Kolmogorov equation, reaction diffusion equations, travelling waves, fourth order equations, extended-Fisher–Kolmogorov equation, Swift–Hohenberg equation

MSC: 34C37, 37J45

1. Introduction

In qualitative theory of differential equations, a prominent role is played by special classes of solutions, like periodic solutions or solutions to some kind of boundary value problem. When a system of ordinary differential equations has equilibria (i.e. constant solutions) whose stability properties are known, it becomes significant to study the connections between them by trajectories of solutions of the given system. These are called *homoclinic* or *heteroclinic* solutions (sometimes *pulses* or *kinks*) according to whether they just describe a loop based at one single equilibrium or they “start” and “end” at two distinct equilibria.

In addition to their intrinsic mathematical interest (for autonomous second order systems, for instance, homoclinics and heteroclinics act in the phase portrait as separatrices between regions where solutions behave differently; when they appear as connections between saddles they prevent structural stability, see, e.g., [38]), homoclinic and heteroclinic solutions appear in the context of a number of mathematical models for problems arising in Mechanics, Chemistry and Biology.

The study of existence of homoclinic and heteroclinic solutions has a long history. Besides phase plane analysis, whose applicability is confined to second order autonomous equations, the study of such solutions has been often made with resource to the geometric theory of ordinary differential equations and dynamical systems techniques. In the nineteenth century, Poincaré [64] already studies perturbed time periodic systems. Poincaré’s results have been the outbreak of many works. In particular, Melnikov’s theory provides instruments for the analysis of how homoclinics and heteroclinics are affected by perturbations on an Hamiltonian system. The main idea is that existence of loops or connections at some rest points can be proved by analyzing the intersection properties of the stable and unstable manifolds through those equilibrium points. In the 60’s, Melnikov [54] proves by analytical method the existence of homoclinics for non conservative perturbations, leading to chaos. Smale [81,82] then shows that in presence of a transverse homoclinic point, the Poincaré map admits a Bernouilli shift structure. Similar ideas are present in works of Birkhoff [15]. We refer to Moser [57] and Wiggins [92] for further developments.

However, starting mainly in the 80’s, a functional analytic approach added powerful tools to the research in this field. Variational methods, combining classical ideas with modern critical point theory, have thus provided a wealth of new results. A large number of contributions are devoted to this topic, the main developments being due to Ambrosetti, Bolotin, Coti Zelati, Ekeland, Rabinowitz and Séré [16,66,27,28,17,9,79,67]. The advantage of this approach comes from the fact that we can often bypass the question of transversality of the stable and unstable manifolds whose verification is delicate in practice, or obtain weaker nondegeneracy conditions. Moreover, it leads to results of a global nature. Some comparison with geometric methods and variational interpretation of Melnikov and Smale–Birkhoff theorems are studied by Ambrosetti and Badiale [7,8], see also [56,86].

This monograph is devoted mainly to the existence of heteroclinics in several types of differential equations and systems of the second and fourth orders. We divide our survey in two parts. Part 1 deals with second order equations and systems, while Part 2 concerns several important types of fourth order equations.

Part 1 starts with the analysis of classical conservative systems arising in Mechanics, like

$$u'' = \nabla V(u), \tag{1.1}$$

where $V \in C^1(\mathbb{R}^N, \mathbb{R})$ is a potential (according to the usual classical mechanics terminology, $-V$ should be called the potential) with several equilibria at the same (minimum) level. We have selected a basic set of results for this and similar problems, where V may also depend on time and be periodic in each variable. We present them with some detail. These results are due mainly to P.H. Rabinowitz and to T.O. Maxwel. Further results are mentioned without proofs. Roughly speaking, the common underlying idea is the following: suppose that V has two minima, say ξ and η; we look for heteroclinic connections between ξ and η as minimizers of the *action functional*

$$\int_{-\infty}^{+\infty} \left(\frac{|u'(t)|^2}{2} + V\bigl(u(t)\bigr) \right) \mathrm{d}t \tag{1.2}$$

in an appropriate class of functions u defined in $\mathbb{R}$ so that $u(-\infty) = \xi$ and $u(+\infty) = \eta$. Then u appears as a solution of (1.1) since this is the Euler–Lagrange equation of (1.2).

We then proceed to a different kind of problem. To state it in a simple form, we recall the Fisher–Kolmogorov type equation,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + g(u), \tag{1.3}$$

proposed as a model of diffusion in Biomathematics [34]. Here g is a positive, continuous function in [0, 1] such that $g(0) = g(1) = 0$. In the original model, $g(u) = u(1-u)$. A significant feature of the dynamics of (1.3) is the existence of *travelling wave fronts*, which are solutions of the form

$$u(x,t) = U(x - ct) \tag{1.4}$$

for some constant $c > 0$. The meaning of c is the speed of wave propagation. The *front profile* $U(s)$ solves then the ordinary differential equation

$$U'' + c\,U' + g(U) = 0, \tag{1.5}$$

where we look at c as a parameter. It is required that the profile satisfies $U(-\infty) = 1$, $U(+\infty) = 0$. Hence this is a problem of existence of heteroclinics for (1.5) and it is a classical one with a very rich literature. Its mathematical treatment began with Kolmogorov, Petrovski and Piskounov [45] and reached a highlight with the paper by Aronson and Weinberger [11]. However, it has continually attracted the interest of many mathematicians up to nowadays. We give an account of the most basic features of this problem dealing with more general equations than (1.5). We present a proof of the existence of a *threshold speed* c^* such that the heteroclinic exists if and only if $c \geqslant c^*$. Following an approach similar to that of Malaguti and Marcelli [48] and introducing some variational procedures like those developed in Arias, Campos, Robles-Pérez and Sanchez [10] we essentially reduce the problem to a first order ordinary differential equation and use the elementary theory of the Cauchy problem. We show that the same method is efficient in the treatment of equations in combustion theory, which differ from (1.3) in that $g(u)$ changes sign in $[0,1]$ while keeping its integral positive; in this case, it is curious that the heteroclinic appears for just one value of c.

In the second part, we consider a class of fourth order autonomous differential equations of the form

$$u'''' - \beta u'' + u^3 - u = 0, \tag{1.6}$$

where β is a real parameter. This equation admits three equilibrium states: $-1, 0$ and $+1$ and we focus on the existence of heteroclinic solutions going from -1 to $+1$. The set of bounded solutions of (1.6) has been the object of much research in the past ten years. In an impressive series of papers [59–62], Peletier and Troy have performed a systematic study of periodic, homoclinic, heteroclinic and chaotic solutions of the model equation (1.6) for the parameter range $\beta \geqslant 0$. In Section 4 we present an overview of results obtained concerning heteroclinic solutions. These include the pioneer works of Peletier and Troy and later results of Kalies and VanderVorst [43], Kalies, Kwapisz and VanderVorst [42] and van den Berg [87]. We give a first insight of methods that can be used to track heteroclinic orbits. It is convenient for simplicity to first restrict our attention to the simple model (1.6). The existence of heteroclinics for (1.6) for any $\beta \geqslant 0$ has been first proved by Peletier and Troy using a shooting method. We briefly describe their arguments in Section 4.1.

Equation (1.6) has also a variational structure. It is an elementary fact that heteroclinic solutions of (1.6) correspond to critical points of the functional

$$\mathcal{F}_\beta(u) := \int_{-\infty}^{+\infty} \left(\frac{1}{2}\left(u''^2 + \beta u'^2\right) + \frac{1}{4}\left(u^2 - 1\right)^2 \right) \mathrm{d}x$$

in an appropriate functional space. The existence of heteroclinics via variational arguments was first investigated by Peletier, Troy and VanderVorst [58] and Kalies and VanderVorst [43]. For $\beta \geqslant 0$, $\mathcal{F}_\beta$ is a positive functional. It is therefore natural to look for heteroclinics as minimizers of $\mathcal{F}_\beta$. Section 4.2 deals with a global minimization procedure which works fine for all $\beta \geqslant 0$.

When $0 \leqslant \beta \leqslant \sqrt{8}$, more heteroclinics are obtained via local minimization in well-chosen homotopy classes. Kalies, Kwapisz and VanderVorst [42] have defined precise types of functions that describe the complex structure of those solutions. In Section 4.3, we describe these homotopy classes and the profiles of the corresponding minimizers.

Many of the arguments used in the minimization process (both in the whole space and in homotopy classes) rely on the positivity of the Lagrangian

$$L\left(u, u', u''\right) = \frac{1}{2}\left(u''^2 + \beta u'^2\right) + \frac{1}{4}\left(u^2 - 1\right)^2$$

and therefore on the positivity of the parameter β. A first attempt to consider changing sign Lagrangian is made in [20,37]. The authors consider the more general functional

$$\mathcal{F}_g(u) := \int_{-\infty}^{+\infty} \left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u) \right) \mathrm{d}x \tag{1.7}$$

whose Euler–Lagrange equation is given by

$$u'''' - g(u)u'' - \frac{1}{2}g'(u)u'^2 + f'(u) = 0. \tag{1.8}$$

Here f is assumed to be a double-well potential with bottoms at ± 1 and g is neither necessarily constant nor positive. The main idea of [37] is to impose a condition on g to ensure a lower bound on the action $\mathcal{F}_g(u)$. In Section 5.1 and Section 5.2, we use the ideas of [19,20,37] to prove the existence of a global minimizer of $\mathcal{F}_g$ under convenient assumptions.

In Section 6, we discuss the case of an equation similar to (1.8) where f is a triple-well potential and address the following question. Does the dynamics possess a heteroclinic orbit connecting the extremal equilibria? Let us take a look at a classical mechanics analogy. Consider a moving particle in a potential characterized by three hills of equal height. To fix the ideas, suppose the tops of the peaks are located at -1, 0 and $+1$ and consider a motion starting from -1 at time $t \to -\infty$. As the potential energy is identical at the top of each hill, the law of energy conservation implies that the particle needs an infinite amount of time to reach the top of the second hill and therefore cannot pass through the middle-equilibrium. For the fourth order equation (1.8), the energy reads

$$E\left(u, u', u'', u'''\right) = u'''u' - \frac{1}{2}u''^2 + \frac{g(u)}{2}u'^2 + f(u).$$

For this equation of motion, a particle does not need to come at rest at the top of the second hill as the constant of motion can be satisfied with a non-zero u' due to the presence of the new terms $u'''u' - \frac{1}{2}u''^2$. In other words, the intersection of the zero energy manifold and the space $u = 0$ does not reduce to a point.

We answer positively the above question at least when the middle-equilibrium is of *saddle-focus* type. We mention that this question is relevant in the study of Ginzburg–Landau models of amphiphilic systems [36]. We refer to Section 6 for a brief description of these models.

We come back to equations with a double-well potential in Section 7. We present an existence result of multi-transition connections for Eq. (1.8). These solutions are obtained as local minimizers in classes of functions with prescribed profiles. We describe these subsets and explain briefly how the local minimization process works.

Each section is followed by some remarks and complementary results that we have chosen not to consider in details. We also give references for some extensions of results presented here in a simplified version, we refer to open questions, problems for which less is known or simply mention some related topics that did not find their way in this monograph.

Part 1. Second Order Equations and Systems

2. Second order Hamiltonian systems

2.1. *Heteroclinics for some scalar second order differential equations*

Recall that a good description of the various kinds of solutions of the mathematical pendulum equation

$$u'' + a\sin u = 0 \quad (a > 0) \tag{2.1}$$

is provided by the representation of the corresponding trajectories in the phase plane (u, u'). These are level curves of the energy, that is, they are a locus of points (u, u') of the form

$$\frac{u'^2}{2} + a(1 - \cos u) = k \tag{2.2}$$

for some constant $k \in \mathbb{R}^+$. That the energy function

$$E\bigl(u, u'\bigr) := \frac{u'^2}{2} + a(1 - \cos u)$$

is constant along solutions of (2.1) is an elementary consequence of multiplying Eq. (2.1) by u'.

When $k = 0$ in (2.2), we obtain the stable equilibria $u = 2k\pi$ $(k \in \mathbb{Z})$. If $0 < k < 2a$, the trajectories are closed curves corresponding to periodic solutions. For $k > 2a$, we obtain unbounded trajectories corresponding to solutions with periodic derivative. Finally, $k = 2a$ in (2.2) yields a locus consisting of the unstable equilibria ($u = (2k + 1)\pi$, $k \in \mathbb{Z}$) and the graphs of the functions

$$u' = \pm\sqrt{2a(1 + \cos u)}, \quad u \neq (2k + 1)\pi \text{ for all } k \in \mathbb{Z}. \tag{2.3}$$

For instance, the solutions having as trajectories the graph of this function in $]-\pi, \pi[$ have the property that

$$\lim_{t\to+\infty} u(t) = \pi, \qquad \lim_{t\to-\infty} u(t) = -\pi$$

and

$$\lim_{t\to\pm\infty} u'(t) = 0$$

or the same conditions with the roles of $+\infty$ and $-\infty$ reversed. Hence these trajectories connect two distinct (consecutive) unstable equilibria. They are called *heteroclinics* and

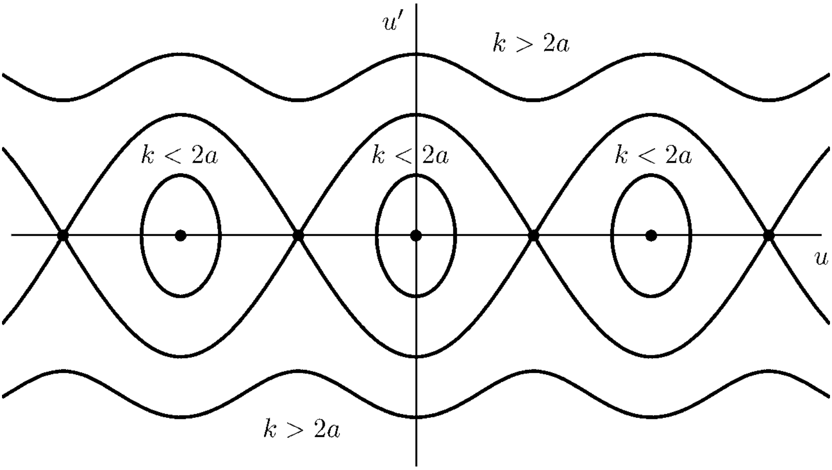


Fig. 1. The pendulum phase-plane.

any underlying solution is called a *heteroclinic solution* of (2.1). In this example it is apparent that they separate regions of the (u, u')-plane where the solutions of (2.1) have a different nature, see Fig. 1.

Of course, given the physical meaning of (2.1), it is sometimes preferable to depict trajectories not in a plane but in a cylinder (which is a plane where the points (u, u') and (v, v') are identified if and only if $u \equiv v \pmod{2\pi}$ and $u' = v'$. Then $(-\pi, 0)$ and $(\pi, 0)$ are in fact the same equilibrium and a trajectory connecting these points in the plane becomes a trajectory with equal limits at $\pm\infty$. We would then rather speak of a *homoclinic*. However, if one forgets the 2π-periodicity of the potential $a(1 - \cos u)$, or, for that matter, if one modifies it outside, say, the interval $]-\pi, \pi[$, the consideration of heteroclinics is meaningful.

Let us consider a more general autonomous scalar equation

$$u'' = f(u), \tag{2.4}$$

where $f \in C(\mathbb{R}, \mathbb{R})$ is a function such that

(A1) $f(\pm 1) = 0$;

(A2) there exists a primitive F of f such that $F(-1) = F(1) = 0$ and $F(u) > 0$ for all $u \in]-1, 1[$.

Hence Eq. (2.4) has two equilibria, $u = \pm 1$, at the (same) zero level of the potential. As for solutions of (2.4) energy is conserved, that is

$$\frac{u'^2}{2} - F(u) = K \tag{2.5}$$

for some constant K, it makes sense to look for heteroclinic solutions connecting -1 and 1, i.e. solutions such that

$$u(\pm\infty) := \lim_{x\to\pm\infty} u(x) = \pm 1 \quad \text{and} \quad u'(\pm\infty) := \lim_{x\to\pm\infty} u'(x) = 0$$

or the same properties with the roles of the $+\infty$ and $-\infty$ reversed. In fact, for such solutions we must have $K = 0$ in (2.5) and the corresponding phase plane trajectories are given explicitly by

$$u' = \pm\sqrt{2F(u)}, \quad -1 < u < 1.$$

These are equations with separable variables. With the $+$ sign, for example, integrating we obtain

$$\int_0^u \frac{\mathrm{d}v}{\sqrt{2F(v)}} = t + C, \tag{2.6}$$

for some constant C. In particular, the equilibria ± 1 are not reached in finite time whenever

$$\int_0^{\pm 1} \frac{\mathrm{d}v}{\sqrt{F(v)}}$$

diverges. This condition is obviously satisfied provided that f is locally Lipschitz. More generally, it also holds if F is at most quadratic near its minima, that is if there exists $c > 0$ so that

$$F(u) \leqslant c(u \pm 1)^2$$

in a neighborhood of -1 and $+1$, respectively.

Now we look for a heteroclinic of (2.4) from another angle. Instead of using elementary integration techniques, we show that such a solution can be characterized by a variational property. Needless to say, it may seem cumbersome to treat that simple problem in such an involved way, but since the variational method has an important role to play in the search of heteroclinics for systems and non-autonomous equations, it is worth to grasp the main ideas in an uncomplicated case.

Formally, (2.4) is the Euler–Lagrange equation of the functional

$$\mathcal{I}(u) := \int_{-\infty}^{+\infty} \left(\frac{u'^2}{2} + F(u) \right) \mathrm{d}t, \tag{2.7}$$

where $F(u)$ is the primitive of f as given in (A2). We look for the heteroclinics of (2.4) as minimizers of $\mathcal{I}$ in the functional space

$$\mathcal{E} := \left\{ u \in H^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}) \mid u(\pm\infty) = \pm 1 \right\}.$$

In fact it will be clear that we may confine ourselves to functions taking values in $[-1,1]$ by simply assuming that F is extended by 0 on $]-\infty,-1[\cup]+1,+\infty[$. We therefore make this hypothesis from now on.

FACT 1. *Let* $\varepsilon\in]0,1[$ *and*

$$\beta_\varepsilon:=\min\left\{F(z)\mid 1-\varepsilon\leqslant z\leqslant 1-\frac{\varepsilon}{2}\ \textit{or}\ -1+\frac{\varepsilon}{2}\leqslant z\leqslant -1+\varepsilon\right\}.$$

If $u\in\mathcal{E}$ *has the property that there exist* $t_1,t_2\in\mathbb{R}$ *such that* $u(t_1)=1-\frac{\varepsilon}{2}$ *and* $u(t_2)=1-\varepsilon$ (*or* $u(t_1)=-1+\frac{\varepsilon}{2}$ *and* $u(t_2)=-1+\varepsilon$), *then we have*

$$\mathcal{I}(u)\geqslant\left|\int_{t_1}^{t_2}\left(\frac{u'^2}{2}+F(u)\right)\mathrm{d}t\right|\geqslant\frac{\varepsilon\sqrt{\beta_\varepsilon}}{\sqrt{2}}.$$

PROOF. We may assume $t_1<t_2$ and $1-\varepsilon\leqslant u(t)\leqslant 1-\frac{\varepsilon}{2}$ or $-1+\frac{\varepsilon}{2}\leqslant u(t)\leqslant -1+\varepsilon$ for $t\in[t_1,t_2]$. Using the positivity of the integrand and Schwarz's inequality we write

$$\mathcal{I}(u)\geqslant\int_{t_1}^{t_2}\left(\frac{u'^2}{2}+F(u)\right)\mathrm{d}t\geqslant\frac{\varepsilon^2}{8(t_2-t_1)}+\beta_\varepsilon(t_2-t_1)$$

and the conclusion follows from the elementary inequality $\frac{a^2}{x}+b^2x\geqslant 2ab$ which holds for all $x\geqslant 0$. □

With these preliminaries we can state and prove:

THEOREM 2.1. *If* $f\in C([-1,1],\mathbb{R})$ *satisfies the assumptions* (A1), (A2) *and is extended by* 0 *outside the interval* $]-1,1[$, *the functional* $\mathcal{I}$ *defined by* (2.7) *attains a minimum in* $\mathcal{E}$. *A minimizer is a heteroclinic solution of* (2.4) *connecting* -1 *and* 1.

PROOF. Let $(u_n)_n\subset\mathcal{E}$ be a minimizing sequence that is $\mathcal{I}(u_n)\to\inf_{\mathcal{E}}\mathcal{I}$. By passing to $v_n=\sup(-1,\inf(u_n,1))$ and observing that $\mathcal{I}(v_n)\leqslant\mathcal{I}(u_n)$, we may assume without loss of generality $-1\leqslant u_n\leqslant 1$.

For each $\varepsilon>0$, we find an interval $[s_n,t_n]$ such that $u_n(s_n)=-1+\varepsilon$, $u_n(t_n)=1-\varepsilon$ and

$$-1+\varepsilon\leqslant u_n(t)\leqslant 1-\varepsilon\quad\text{for all } t\in[s_n,t_n].$$

If u_n takes values greater than $-1+\varepsilon$ in $]-\infty,s_n]$, we may choose $s_n'<s_n$ so that $u_n(s_n')=-1+\varepsilon$ and $-1\leqslant u_n\leqslant -1+\varepsilon$ in $]-\infty,s_n']$. In the same way we find $t_n'\geqslant t_n$ so that $u_n(t_n')=1-\varepsilon$ and $1-\varepsilon\leqslant u_n\leqslant 1$ in $[t_n',+\infty[$. Define a new function

$$U_n(t)=\begin{cases}u_n(t-s_n+s_n') & \text{if } t\leqslant s_n,\\ u_n(t) & \text{if } s_n\leqslant t\leqslant t_n,\\ u_n(t+t_n'-t_n) & \text{if } t\geqslant t_n.\end{cases}$$

By the translation invariance of the integrals and the fact that the integrand is positive it is clear that

$$\mathcal{I}(U_n) \leqslant \mathcal{I}(u_n)$$

so that $(U_n)_n \subset \mathcal{E}$ is again a minimizing sequence which in addition satisfies $-1 \leqslant U_n \leqslant -1+\varepsilon$ in $]-\infty, s_n]$ and $1-\varepsilon \leqslant U_n \leqslant 1$ in $[t_n, +\infty[$. The sequence $t_n - s_n$ is bounded since

$$\mathcal{I}(U_n) \geqslant \Big(\min_{-1+\varepsilon\leqslant z\leqslant 1-\varepsilon} F(z)\Big)(s_n - t_n).$$

Observe also that as $\mathcal{I}$ is translation invariant, we may assume that $s_n = 0$.

Now, since $\sup_n \|U_n\|_{L^\infty} \leqslant 1$ and $\sup_n \|U_n'\|_{L^2}$ is bounded, we may apply the diagonal procedure to extract a subsequence that we still denote by $(U_n)_n$, and we obtain a function $u \in H^1_{\mathrm{loc}}(\mathbb{R})$ such that

$$U_n \overset{C_{\mathrm{loc}}(\mathbb{R})}{\longrightarrow} u$$

i.e. uniformly in compact intervals and

$$U_n' \overset{L^2(\mathbb{R})}{\rightharpoonup} u'.$$

In addition, we have $t_n \to \bar{t} > 0$.

Combining weak lower semicontinuity of the L^2-norm and Fatou's lemma we infer

$$\begin{aligned}\mathcal{I}(u) &\leqslant \liminf_{n\to\infty} \int_{-\infty}^{+\infty} \frac{U_n'^2}{2}\,\mathrm{d}t + \liminf_{n\to\infty}\int_{-\infty}^{+\infty} F(U_n)\,\mathrm{d}t \\ &\leqslant \lim_{n\to\infty} \mathcal{I}(U_n) = \inf_{\mathcal{E}} \mathcal{I}.\end{aligned} \tag{2.8}$$

By uniform convergence, we have $-1 \leqslant u(t) \leqslant -1+\varepsilon$ for $t \leqslant 0$ and $1-\varepsilon \leqslant u(t) \leqslant 1$ for $t > \bar{t}$. On the other hand, the fact that $\int_{-\infty}^{+\infty} F(u)\,\mathrm{d}t < +\infty$ implies that

$$\liminf_{t\to-\infty} u(t) = -1 \quad \text{and} \quad \limsup_{t\to+\infty} u(t) = 1.$$

In fact we have $\lim_{t\to\pm\infty} u(t) = \pm 1$. Otherwise there exist $\delta > 0$ and infinitely many disjoint intervals $[t_1, t_2]$ in the conditions of Fact 1 (with $\varepsilon = \delta$), implying $\mathcal{I}(u) = +\infty$ and contradicting (2.8). It follows that $u \in \mathcal{E}$ and therefore

$$\mathcal{I}(u) = \min_{\mathcal{E}} \mathcal{I}.$$

The fact that u is a solution of (2.4) follows by the usual argument that gives the Euler–Lagrange equation of a functional. For each $\varphi \in C_c^1(\mathbb{R})$, and $\tau \in \mathbb{R}$, the function $u + \tau\varphi$ belongs to $\mathcal{E}$; we compute

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\bigg|_{\tau=0} \mathcal{I}(u+\tau\varphi) = \int_{-\infty}^{+\infty} \big(u'\varphi' + f(u)\varphi\big)\,\mathrm{d}t = 0$$

and, by the Du Bois–Reymond lemma, u satisfies (2.4).

Finally, since u satisfies (2.5) for some K and there exist sequences $t_n \to \pm\infty$ with $u'(t_n) \to 0$, we infer that $K = 0$ and we then conclude that $u'(\pm\infty) = 0$. $\square$

We next apply the above method to a less trivial situation: let us consider the second order non-autonomous differential equation

$$u'' = a(t)f(u), \tag{2.9}$$

where $f \in C(\mathbb{R},\mathbb{R})$ satisfies the assumptions (A1), (A2) and $a \in L^\infty(\mathbb{R},\mathbb{R})$ is such that

(A3) there exist $a_1, a_2 \in \mathbb{R}$ so that $0 < a_1 \leqslant a(t) \leqslant a_2$ for all $t \in \mathbb{R}$.

We look for a heteroclinic connection between the equilibria -1 and $+1$. In the absence of a conservation law, the variational argument appears as a natural device. So we now consider the functional

$$\mathcal{J}(u) := \int_{-\infty}^{+\infty} \left(\frac{u'^2}{2} + a(t)F(u)\right)\mathrm{d}t \tag{2.10}$$

and seek conditions that allow to minimize it in $\mathcal{E}$. If $a(t)$ is a T-periodic function then the proof of Theorem 2.1 can be easily adapted. While we cannot completely mimic the powerful modification arguments used in the autonomous case, the periodicity of a allows to localize a transition from a small neighborhood of -1 to a small one of $+1$. Indeed, keeping the notations of the proof of Theorem 2.1, the sequence of left endpoints of the intervals $[s_n, t_n]$ can be placed in the interval $[0, T]$ by time translations whose length is a multiple of T. The remaining of the proof is then similar.

If a does not possess any symmetry or periodicity property then we have to face a possible loss of compactness as the interval $[s_n, t_n]$ could escape to $+\infty$ or $-\infty$ so that the weak limit of u_n could either be one equilibrium or a homoclinic solution. The natural way to avoid such behaviours is to impose some coercivity assumption on the function a at $\pm\infty$. In some sense this penalizes transitions tending to $\pm\infty$.

Note that Fact 1 proved above has an obvious extension for this new functional: under condition (A3), it suffices to replace, in the right-hand side of the inequality, $\frac{\varepsilon\sqrt{\beta_\varepsilon}}{\sqrt{2}}$ with $\frac{\varepsilon\sqrt{a_1\beta_\varepsilon}}{\sqrt{2}}$, namely we have:

FACT 1′. *Under the assumptions of Fact* 1, *we have*

$$\mathcal{J}(u) \geqslant \left|\int_{t_1}^{t_2} \left(\frac{u'^2}{2} + a(t)F(u)\right)\mathrm{d}t\right| \geqslant \frac{\varepsilon\sqrt{a_1\beta_\varepsilon}}{\sqrt{2}}.$$

The following two facts are useful to ensure that the quasi-minimizers stay close to ±1 once they enter a sufficiently small neighborhood of ±1.

FACT 2. *Let* $\varepsilon \in \,]0,1[$ *and*

$$\mathcal{E}_{\bar t,1-\varepsilon} := \bigl\{u \in C\bigl([\bar t,+\infty[,\mathbb{R}\bigr) \bigm| u(\bar t) = 1-\varepsilon,\ u(+\infty) = 1,\ u' \in L^2\bigl([\bar t,+\infty[\bigr)\bigr\}.$$

Then, setting

$$\mathcal{J}_{[\bar t,+\infty[}(u) := \int_{\bar t}^{+\infty} \left(\frac{u'^2}{2} + a(t)F(u)\right) \mathrm{d}t,$$

there exists $v \in \mathcal{E}_{\bar t,1-\varepsilon}$ *so that* $1-\varepsilon \leqslant v(t) \leqslant 1$ *for all* $t \geqslant \bar t$ *and*

$$\mathcal{J}_{[\bar t,+\infty[}(v) \leqslant \frac{\varepsilon^2}{2} + a_2 \max_{z\in[1-\varepsilon,1]} F(z).$$

PROOF. It suffices to compute $\mathcal{J}_{[\bar t,+\infty[}(v)$ where

$$v(t) = \begin{cases} 1-\varepsilon+\varepsilon\bigl(t-\bar t\bigr), & \text{if } \bar t \leqslant t \leqslant \bar t+1,\\ 1, & \text{if } t \geqslant \bar t+1. \end{cases} \qquad \square$$

Analogously we have:

FACT 3. *Let* $\varepsilon \in \,]0,1[$ *and*

$$\mathcal{E}_{\bar t,-1+\varepsilon} := \bigl\{u \in C\bigl(]-\infty,\bar t],\mathbb{R}) \bigm| u(\bar t) = -1+\varepsilon,\ u(-\infty) = -1,\\ u' \in L^2\bigl(]-\infty,\bar t]\bigr)\bigr\}.$$

Then, setting

$$\mathcal{J}_{]-\infty,\bar t]}(u) := \int_{-\infty}^{\bar t} \left(\frac{u'^2}{2} + a(t)F(u)\right) \mathrm{d}t,$$

there exists $v \in \mathcal{E}_{\bar t,-1+\varepsilon}$ *such that* $-1 \leqslant v(t) \leqslant -1+\varepsilon$ *for all* $t \leqslant \bar t$ *and*

$$\mathcal{J}_{]-\infty,\bar t]}(v) \leqslant \frac{\varepsilon^2}{2} + a_2 \max_{z\in[-1,-1+\varepsilon]} F(z). \qquad \square$$

The family of functionals

$$\mathcal{I}_\alpha(u) := \int_{-\infty}^{+\infty} \left(\frac{u'^2}{2} + \alpha F(u)\right) \mathrm{d}t,$$

where $\alpha > 0$ is constant, is useful in what follows. According to Theorem 2.1, $\mathcal{I}_\alpha$ has a minimum in $\mathcal{E}$. Hence we may define $u_\alpha \in \mathcal{E}$ such that

$$\varphi(\alpha) := \min_{u \in \mathcal{E}} \mathcal{I}_\alpha(u) = \mathcal{I}_\alpha(u_\alpha).$$

FACT 4. *The function φ is a strictly increasing, continuous function in* $]0, +\infty[$.

PROOF. The first statement is obvious. The second is a consequence of the inequalities

$$\varphi(\alpha) + (\beta - \alpha)\int_{-\infty}^{+\infty} F(u_\beta)\,\mathrm{d}t \leqslant \varphi(\beta) \leqslant \varphi(\alpha) + (\beta - \alpha)\int_{-\infty}^{+\infty} F(u_\alpha)\,\mathrm{d}t$$

together with the fact that, when β is close to α, $\int_{-\infty}^{+\infty} F(u_\beta)\,\mathrm{d}t$ is uniformly bounded from above. □

We can now state the following theorem.

THEOREM 2.2. *Assume that $f \in C(\mathbb{R}, \mathbb{R})$, $a \in L^\infty(\mathbb{R}, \mathbb{R})$ satisfy* (A1)–(A3). *If, in addition,*

$$\lim_{|t| \to \infty} a(t) = a_2$$

and $a(t) < a_2$ in some set of nonzero measure, then (2.9) *has a heteroclinic solution from -1 to 1. This solution takes values in* $[-1, 1]$.

PROOF. Observe first that we clearly have $\inf_{\mathcal{E}} \mathcal{J} < \varphi(a_2)$ and using Fact 4, we may find $\alpha \in]0, a_2[$ so that $\inf_{\mathcal{E}} \mathcal{J} < \varphi(\alpha)$.

Fix $\bar{\varepsilon} \in]0, 1[$ and then choose $\varepsilon > 0$ sufficiently small to satisfy $\varepsilon < \bar{\varepsilon}/2$ and

$$\left(1 + \frac{a_2}{a_1}\right)\left(\frac{\varepsilon^2}{2} + a_2\mu_\varepsilon\right) < \min\left(\frac{\bar{\varepsilon}\sqrt{a_1\beta_{\bar{\varepsilon}}}}{\sqrt{2}}, \frac{\varphi(\alpha) - \inf_{\mathcal{E}} \mathcal{J}}{2}\right) \tag{2.11}$$

where $\beta_{\bar{\varepsilon}}$ is defined as in Fact 1 and

$$\mu_\varepsilon := \max_{z \in [-1, -1+\varepsilon] \cup [1-\varepsilon, 1]} F(z).$$

Let $u_n \in \mathcal{E}$ be a minimizing sequence for $\mathcal{J}$. Define the interval $[s_n, t_n]$ with respect to u_n and ε as in the proof of Theorem 2.1. According to the choice of ε, Fact 1′, Fact 2 and Fact 3 and modifying $u_n(t)$ in $]-\infty, s_n] \cup [t_n, +\infty[$ if necessary, we may assume that

$$-1 \leqslant u_n(t) \leqslant -1 + \bar{\varepsilon} \quad \text{if } t \leqslant s_n, \qquad 1 - \bar{\varepsilon} \leqslant u_n(t) \leqslant 1 \quad \text{if } t \geqslant t_n$$

and

$$\sup\bigl(\mathcal{J}_{]-\infty, s_n]}(u_n),\ \mathcal{J}_{[t_n, +\infty[}(u_n)\bigr) \leqslant \frac{\varepsilon^2}{2} + a_2\mu_\varepsilon.$$

As in the proof of Theorem 2.1, $t_n - s_n$ is bounded. We claim that s_n is bounded from above and t_n is bounded from below. Let us prove this for s_n, the proof for t_n being similar. If the claim is not true there exists $s_0 \in \mathbb{R}$ such that $a(t) \geqslant \alpha$ whenever $t \geqslant s_0$ so that $a(t) \geqslant \alpha$ whenever $t \geqslant s_n$ for large n. We then compute

$$\begin{aligned} \mathcal{J}(u_n) &\geqslant \int_{s_n}^{+\infty} \left(\frac{u_n'^2}{2} + \alpha F(u_n) \right) \mathrm{d}t \\ &\geqslant \varphi(\alpha) - \frac{a_2}{a_1} \left(\frac{\varepsilon^2}{2} + a_2 \mu_\varepsilon \right) \end{aligned}$$

and by our choice of ε in (2.11) we infer that

$$\mathcal{J}(u_n) \geqslant \frac{\varphi(\alpha) + \inf_{\mathcal{E}} \mathcal{J}}{2},$$

a contradiction with the assumption $\mathcal{J}(u_n) \to \inf_{\mathcal{E}} \mathcal{J}$.

It now follows that along some subsequence

$$u_n \overset{C_{\mathrm{loc}}(\mathbb{R})}{\longrightarrow} u, \quad u_n' \overset{L^2(\mathbb{R})}{\rightharpoonup} u', \quad s_n \to \bar{s}, \quad t_n \to \bar{t}$$

and in particular

$$-1 \leqslant u(t) \leqslant -1 + \varepsilon \quad \text{for all } t \leqslant \bar{s}, \qquad 1 - \varepsilon \leqslant u(t) \leqslant 1 \quad \text{for all } t \geqslant \bar{t}.$$

Arguing as in the proof of Theorem 2.1, it is easily shown that $u \in \mathcal{E}$, $\mathcal{J}(u) = \inf_{\mathcal{E}} \mathcal{J}$ and u is a solution of (2.9).

It remains to show that $u'(\pm\infty) = 0$. Assume, by contradiction, that there exist $\tau_n \to \infty$ and $\Delta > 0$ with $|u'(\tau_n)| \geqslant \Delta$. As from (2.9) we have $u''(\pm\infty) = 0$, we may assume that, for n sufficiently large, $|u'(t)| \geqslant \Delta/2$ for all $t \in [\tau_n, \tau_n + 1]$. But then $|u(\tau_n) - u(\tau_n + 1)| \geqslant \Delta/2$, which is impossible since $u(\pm\infty) = \pm 1$. □

2.2. *Autonomous Hamiltonian systems*

In this section we consider a basic situation for systems, adapted from the results of Rabinowitz [66]. We concentrate on the autonomous system

$$u'' = \nabla V(u), \tag{2.12}$$

where $u = (u_1, \ldots, u_N)$ and $V \in C^1(\mathbb{R}^N, \mathbb{R})$ is a non-negative potential with several isolated equilibria at minimum level. Precisely, we assume

(A4) $V \geqslant 0$, $\mathcal{M} := V^{-1}(0)$ contains at least two points and

$$\inf\bigl\{ |\xi_1 - \xi_2| \mid \xi_1, \xi_2 \in V^{-1}(0) \text{ and } \xi_1 \neq \xi_2 \bigr\} > 0;$$

(A5) for each $\varepsilon > 0$, $\inf\{V(u) \mid \operatorname{dist}(u, \mathcal{M}) \geqslant \varepsilon\} > 0$.

As in the previous section, we investigate the existence of heteroclinic connections between elements of $\mathcal{M}$. A solution $u(t)$ of (2.12) such that there exist $\xi, \eta \in \mathcal{M}$, $\xi \neq \eta$ and

$$u(-\infty) = \xi, \quad u(+\infty) = \eta, \quad u'(\pm\infty) = 0$$

is called a *heteroclinic solution of* (2.12) *from* ξ *to* η *or a heteroclinic connection between* ξ *and* η. We say also that u *starts* at ξ and *ends* at η.

Clearly, if $u(t)$ is a heteroclinic of (2.12) from ξ to η, for every $c \in \mathbb{R}$ $u(t+c)$ is also a heteroclinic from ξ to η, and $u(-t)$ is a heteroclinic from η to ξ.

Let us fix an element in $\mathcal{M}$ which we may assume without loss of generality to be 0. Then we search for heteroclinics from 0 to some $\xi \in \mathcal{M} \setminus \{0\}$. They appear as suitable minimizers of the functional

$$\mathcal{J}(u) := \int_{-\infty}^{+\infty} \left(\frac{|u'|^2}{2} + V(u) \right) \mathrm{d}t. \tag{2.13}$$

Given $\xi \in \mathcal{M} \setminus \{0\}$, it is natural to consider the class of vector functions

$$\Gamma(\xi) := \left\{ u \in H^1_{\mathrm{loc}}\left(\mathbb{R}, \mathbb{R}^N\right) \mid u(-\infty) = 0,\ u(+\infty) = \xi \right\}.$$

Let us set

$$c(\xi) := \inf_{u \in \Gamma(\xi)} \mathcal{J}(u).$$

The following fact is proved with a computation quite similar to the proof of Fact 3 in Section 2.1.

FACT 1. *Assume that* (A5) *holds. Let* $\varepsilon > 0$, *and* $z \in H^1((t_1, t_2), \mathbb{R}^N)$ *be such that*

$$\operatorname{dist}\big(u(t), \mathcal{M}\big) \geqslant \varepsilon \quad \textit{for all } t \in [t_1, t_2].$$

Then

$$\int_{t_1}^{t_2} \left(\frac{|u'|^2}{2} + V(u) \right) \mathrm{d}t \geqslant \sqrt{2\alpha_\varepsilon} \big| u(t_2) - u(t_1) \big|,$$

where $\alpha_\varepsilon := \inf\{V(u) \mid \operatorname{dist}(u, \mathcal{M} \geqslant \varepsilon)\} > 0$.

FACT 2. *Assume that* V *satisfies* (A4) *and* (A5). *Let* $u \in H^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^N)$ *be such that* $\mathcal{J}(u) < +\infty$. *Then* $u(-\infty)$, $u(+\infty)$ *exist and belong to* $\mathcal{M}$.

PROOF. As $\int_{-\infty}^{+\infty} V(u(t))\,\mathrm{d}t < +\infty$, we have

$$\liminf_{t \to -\infty} V\big(u(t)\big) = \liminf_{t \to +\infty} V\big(u(t)\big) = 0.$$

By assumption (A5), there exist sequences $t_n \to -\infty$, $s_n \to +\infty$ so that

$$\lim_{n\to\infty} \operatorname{dist}\bigl(u(t_n), \mathcal{M}\bigr) = \lim_{n\to\infty} \operatorname{dist}\bigl(u(s_n), \mathcal{M}\bigr) = 0.$$

Let us prove that $\lim_{t\to+\infty} u(t)$ exists (a similar argument applies at $-\infty$): if the limit does not exist we are able to find

$$0 < \varepsilon < \gamma := \frac{1}{3} \inf\bigl\{|\xi_1 - \xi_2| \mid \xi_1 \neq \xi_2 \text{ and } \xi_1, \xi_2 \in \mathcal{M}\bigr\} \tag{2.14}$$

and disjoint intervals $[\tau_n, \sigma_n]$ such that $\operatorname{dist}(u(\tau_n), \mathcal{M}) = \varepsilon, \operatorname{dist}(u(\sigma_n), \mathcal{M}) = 2\varepsilon$ and $\operatorname{dist}(u(t), \mathcal{M}) \geqslant \varepsilon$ for all $t \in [\tau_n, \sigma_n]$. But then, by virtue of Fact 1, we infer that

$$\begin{aligned}\mathcal{J}(u) &\geqslant \sum_n \int_{\tau_n}^{\sigma_n} \left(\frac{|u'|^2}{2} + V\bigl(u(t)\bigr) \right) \mathrm{d}t \\ &\geqslant \sum_n \sqrt{2\alpha_\varepsilon} \min(\varepsilon, \gamma - \varepsilon) = +\infty\end{aligned}$$

which is a contradiction. □

FACT 3. *If V has infinitely many minima, then*

$$\lim_{|\xi|\to+\infty} c(\xi) = +\infty.$$

PROOF. It suffices to prove that, given $M \geqslant 0$, there exists another number $N = N(M)$ so that $\mathcal{J}(u) \leqslant M$ and $u \in \Gamma(\xi)$ implies $|\xi| \leqslant N$. Fix $\varepsilon > 0$ as in (2.14). For u as stated, let us call $[a, b]$ a transition interval from $\xi_a \in \mathcal{M}$ to $\xi_b \in \mathcal{M}$ ($\xi_a \neq \xi_b$) if $|u(a) - \xi_a| = \varepsilon = |u(b) - \xi_b|$ and for all $t \in [a, b]$, $\operatorname{dist}(u(t), \mathcal{M}) \geqslant \varepsilon$. By definition of $\Gamma(\xi)$ we find (disjoint) transition intervals $[t_i, s_i]$, from ξ_i to η_i, $i = 1, \ldots, k$, so that $\xi_1 = 0$, $\xi_{j+1} = \eta_j$ and the $\{\xi_j\}$ are all distinct. Fact 1 implies that

$$\mathcal{J}(u) \geqslant \sum_{i=1}^{k} \sqrt{2\alpha_\varepsilon}\gamma = k\sqrt{2\alpha_\varepsilon}\gamma.$$

Hence the number k of such intervals has an upper bound, $k \leqslant M/(\sqrt{2\alpha_\varepsilon}\gamma)$. Since $u(t)$ reaches $B(\xi, \varepsilon)$, we may assume that $\eta_k = \xi$. The length of each transition interval has an upper bound that depends only on M, since

$$M \geqslant \int_{t_i}^{s_i} V\bigl(u(t)\bigr) \mathrm{d}t \geqslant \alpha_\varepsilon (s_i - t_i).$$

We then infer that

$$\bigl|u(t_i) - u(s_i)\bigr| \leqslant M\sqrt{\frac{2}{\alpha_\varepsilon}}.$$

Therefore

$$|\xi| < 2\varepsilon + 2\varepsilon(k-2) + (k-1)M\sqrt{\frac{2}{\alpha_\varepsilon}}$$

and we obtain an upper bound for $|\xi|$ depending on M. □

THEOREM 2.3. *Let* $V \in C^1(\mathbb{R}^N, \mathbb{R})$ *satisfy the assumptions* (A4) *and* (A5), *with* $0 \in \mathcal{M}$. *Then the system* (2.12) *has a heteroclinic solution starting at* 0, *and another one ending at* 0.

PROOF. By reversing time, it suffices to show that the system (2.12) possesses a heteroclinic starting at 0. By the assumption (A4) and Fact 3 (if $\mathcal{M}$ is not finite) it is clear that

$$c := \min_{\zeta \in \mathcal{M}\setminus\{0\}} c(\zeta)$$

exists, since a bounded domain intersects $\mathcal{M}$ in a finite set. Hence there exists $\xi \in \mathcal{M} \setminus \{0\}$ such that $c = c(\xi)$. Now let (u_n) be a minimizing sequence, so that

$$u_n \in \Gamma(\xi) \quad \text{and} \quad \mathcal{J}(u_n) \to c(\xi) = c.$$

By translation invariance we may assume (with ε as in (2.14))

$$\left|u_n(0)\right| = \varepsilon, \quad \left|u_n(t)\right| \leqslant \varepsilon \quad \text{for all } t \leqslant 0 \text{ and all } n \in \mathbb{N}. \tag{2.15}$$

From the boundedness of $\mathcal{J}(u_n)$ it follows immediately that $(u_n')_n$ is bounded for the L^2-norm, and the arguments of the proof of Fact 3 imply that $\|u_n\|_{L^\infty}$ is also bounded. Hence for each finite interval $[a,b]$, $\|u_n\|_{H^1(a,b)}$ is bounded and, by the diagonal procedure we may extract a subsequence, also labelled $(u_n)_n$, and find a function $u \in H^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^N)$ so that

$$u_n \overset{C_{\mathrm{loc}}(\mathbb{R},\mathbb{R}^N)}{\longrightarrow} u \quad \text{and} \quad u_n' \overset{L^2(\mathbb{R},\mathbb{R}^N)}{\rightharpoonup} u'.$$

The weak lower semicontinuity of the norm and Fatou's lemma imply

$$\mathcal{J}(u) \leqslant \liminf_{n\to\infty} \mathcal{J}(u_n) = c.$$

By Fact 2, $u(\pm\infty)$ exist and belong to $\mathcal{M}$. From (2.15), using local uniform convergence we have $u(-\infty) = 0$. We now prove that $u(+\infty) = \xi$. This follows from the following two claims.

CLAIM 1. $u(+\infty) \neq 0$. Suppose by contradiction that $u(+\infty) = 0$. Fix $\delta > 0$ so that $4\delta < \varepsilon$ and

$$2\delta^2 + \sup_{B_{2\delta}(0)} V < \frac{\varepsilon}{4}\sqrt{2\alpha_{\varepsilon/2}}, \tag{2.16}$$

where $\alpha_{\varepsilon/2}$ has been defined in Fact 1. There exists t_δ so that $u(t) \in B_\delta(0)$ for all $t \geqslant t_\delta$, and therefore $u_m(t_\delta) \in B_{2\delta}(0)$ for all m sufficiently large. As $|u_m(0)| = \varepsilon$, Fact 1 implies that

$$\mathcal{J}(u_m) \geqslant \frac{\varepsilon}{2}\sqrt{2\alpha_{\varepsilon/2}} + \int_{t_\delta}^{+\infty} \left(\frac{|u_m'|^2}{2} + V(u_m) \right) \mathrm{d}t. \tag{2.17}$$

Define a new function $U_m \in \Gamma(\xi)$ by

$$U_m(t) := \begin{cases} 0 & \text{if } t \leqslant t_\delta - 1, \\ (t - t_\delta + 1)u_m(t_\delta) & \text{if } t_\delta - 1 \leqslant t \leqslant t_\delta, \\ u_m(t) & \text{if } t \geqslant t_\delta. \end{cases}$$

We clearly have the estimate

$$\mathcal{J}(U_m) \leqslant \frac{1}{2}(2\delta)^2 + \sup_{B_{2\delta}(0)} V + \int_{t_\delta}^{+\infty} \left(\frac{|u_m'|^2}{2} + V(u_m) \right) \mathrm{d}t. \tag{2.18}$$

From (2.16)–(2.18) we conclude that

$$\mathcal{J}(u_m) - \mathcal{J}(U_m) \geqslant \frac{\varepsilon}{4}\sqrt{2\alpha_{\varepsilon/2}},$$

which yields a contradiction since

$$c \leqslant \limsup_{m\to\infty} \mathcal{J}(U_m) \leqslant c - \frac{\varepsilon}{4}\sqrt{2\alpha_{\varepsilon/2}} < c.$$

CLAIM 2. If $\eta \in \mathcal{M} \setminus \{0, \xi\}$, $u(+\infty) \neq \eta$. Suppose, by contradiction, that $u(+\infty) = \eta \in \mathcal{M} \setminus \{0, \xi\}$. We then fix $\delta > 0$ so that $2\delta < \varepsilon$ and

$$2\delta^2 + \sup_{B_{2\delta}(\eta)} V < \frac{\gamma}{2}\sqrt{2\alpha_\varepsilon}.$$

Let t_δ be such that $u_m(t_\delta) \in B_{2\delta}(\eta)$ for all m sufficiently large. We introduce the new function $U_m \in \Gamma(\eta)$ defined by

$$U_m(t) := \begin{cases} u_m(t) & \text{if } t \leqslant t_\delta, \\ (1 - t + t_\delta)u_m(t_\delta) + (t - t_\delta)\eta & \text{if } t_\delta \leqslant t \leqslant t_\delta + 1, \\ \eta & \text{if } t \geqslant t_\delta + 1. \end{cases}$$

As $u_m \in \Gamma(\xi)$, there exists $\underline{t}_m < \bar{t}_m$ with $t_\delta < \underline{t}_m$, so that $|u_m(\underline{t}_m) - \eta| = \varepsilon$, $\operatorname{dist}(u_m(\bar{t}_m), \mathcal{M}) = \varepsilon$ and $\operatorname{dist}(u_m(t), \mathcal{M}) \geqslant \varepsilon$ for all $t \in [\underline{t}_m, \bar{t}_m]$. Hence we have

$$\mathcal{J}(u_m) \geqslant \int_{-\infty}^{t_\delta} \left(\frac{|u_m'|^2}{2} + V(u_m) \right) \mathrm{d}t + \gamma\sqrt{2\alpha_\varepsilon}.$$

On the other hand, we infer that

$$\mathcal{J}(U_m) \leqslant 2\delta^2 + \sup_{B_{2\delta}(\eta)} V + \int_{-\infty}^{t_\delta} \left(\frac{|u_m'|^2}{2} + V(u_m) \right) \mathrm{d}t$$

and we therefore conclude that

$$\mathcal{J}(u_m) - \mathcal{J}(U_m) \geqslant \frac{\gamma}{2}\sqrt{2\alpha_\varepsilon}.$$

It now follows that

$$\limsup_{m\to\infty} \mathcal{J}(U_m) \leqslant c - \frac{\gamma}{2}\sqrt{2\alpha_\varepsilon} < c,$$

which again contradicts the definition of the level c.

Having shown that $u \in \Gamma(\xi)$, it follows that $c = \mathcal{J}(u)$ is a minimum. The usual elementary argument of the Calculus of Variations shows that u is a solution of (2.12). At last we must check that $u'(\pm\infty) = 0$. For the autonomous system , this is particularly simple. Indeed, since u is a solution of (2.12), it satisfies the energy identity

$$\frac{|u'|^2}{2} + V(u) = K$$

for some constant K, and it is easy to see that $K = 0$. □

In presence of symmetries something else can be said. Let us consider the important case where V is periodic in each coordinate. For definiteness, we assume the period is the same for all coordinates and that the minimizers of V are the translates of 0:

(A6) $V \in C^1(\mathbb{R}^N, \mathbb{R})$ is a potential periodic in each variable u_i with period 1, $\min_{\mathbb{R}^N} V = 0$ and $\mathcal{M} = V^{-1}(0) = \mathbb{Z}^N$.

Of course, the condition (A6) implies that (A4) and (A5) hold, so that Theorem 2.3 applies to this class of potentials.

THEOREM 2.4. *If* $V \in C^1(\mathbb{R}^N, \mathbb{R})$ *satisfies* (A6) *then for each* $\beta \in \mathcal{M}$ *there are at least* $2N$ *heteroclinic solutions of* (2.12) *starting from* β *and at least* $2N$ *heteroclinic solutions of* (2.12) *ending at* β.

IDEA OF THE PROOF. We may assume that $\beta = 0$ and we keep the notation introduced before. Whenever $c(\xi)$ is attained for some $\xi \in \mathcal{M} \setminus \{0\}$, the minimizer is a heteroclinic from 0 to ξ. Let G be the subgroup of $\mathbb{Z}^N$ spanned by the elements $\xi \in \mathcal{M} \setminus \{0\}$ so that $c(\xi)$ is attained. It follows from Theorem 2.3 that $G \neq 0$. If $G \neq \mathbb{Z}^N$, we can select $\zeta \in \mathbb{Z}^N \setminus G$ so that

$$c(\zeta) = \min_{\zeta' \in \mathbb{Z}^N \setminus G} c(\zeta').$$

Then we prove that $c(\zeta)$ is attained, by mimicking the proof of Theorem 2.3. The main difference is in the proof of Claim 2. When we take $\eta \in \mathcal{M} \setminus \{0, \zeta\}$ we have to allow the possibility that $\eta \in G$. But this implies that $\zeta - \eta \in \mathbb{Z}^N \setminus G$. The minimizing sequence (u_m) is then modified to a sequence in $\Gamma(\zeta - \eta)$, given by

$$U_m(t) := \begin{cases} 0 & \text{if } t \leqslant t_\delta - 1, \\ (t - t_\delta + 1)\big(u_m(t_\delta) - \eta\big) & \text{if } t_\delta - 1 \leqslant t \leqslant t_\delta, \\ u_m(t) - \eta & \text{if } t \geqslant t_\delta. \end{cases}$$

The appropriate choice of δ and the fact that $V(u_m(t) - \eta) = V(u_m(t))$ lead to the usual contradiction with the choice of ζ. Therefore we conclude that $G = \mathbb{Z}^N$. This yields N $\mathbb{Z}$-independent elements ξ so that there exist heteroclinics u form 0 to ξ. For each such ξ, $u(-t) - \xi$ is a heteroclinic from 0 to $-\xi$. □

2.3. *Periodic Hamiltonian systems: Multiplicity. Multibump solutions*

In this section we are interested in non-autonomous systems of the form

$$u'' = \nabla_u V(t, u), \tag{2.19}$$

where the potential is periodic in the time variable as well as in each spatial variable u_i ($i = 1, \ldots, N$). For simplicity, the period is supposed to be the same for all the variables. The results of the previous section carry over to this kind of systems provided the assumptions on the potential are adequately rephrased. Namely, $V(t, \cdot)$ must have equilibria at the minimum level of the potential that are independent of t.

Let us state the assumptions on V:

(A7) $V \in C^2(\mathbb{R} \times \mathbb{R}^N, \mathbb{R})$ is 1-periodic in each variable $t, u_i, i = 1, \ldots, n$;

(A8) $V(t, 0) = 0 < V(t, u)$ for all $t \in \mathbb{R},\ u \in \mathbb{R}^N \setminus \mathbb{Z}^N$.

Clearly, we may consider different given periods in each variable of V by rescaling. Note also that the assumptions (A7) and (A8) imply that $\mathcal{M} := \{u \in \mathbb{R}^N\colon\ V(t, u) = 0\} = \mathbb{Z}^N$ independently of t.

We easily recognize that Fact 1 and Fact 2 of the preceding section extend to the class of potentials we are considering by now. We only have to replace $V(u)$ by $V(t, u)$ and define

$$\alpha_\varepsilon := \inf\big\{V(t, u) \mid t \in \mathbb{R},\ \operatorname{dist}(u, \mathcal{M}) \geqslant \varepsilon\big\}. \tag{2.20}$$

As before, we make use of the functional $\mathcal{J}$ given by

$$\mathcal{J}(u) := \int_{-\infty}^{+\infty} \left(\frac{|u'|^2}{2} + V(t, u)\right) \mathrm{d}t. \tag{2.21}$$

For each given pair of elements $\xi, \eta \in \mathcal{M}$, consider the class of functions

$$\Gamma(\xi, \eta) := \big\{u \in C\big(\mathbb{R}, \mathbb{R}^N\big) \bigm| u(-\infty) = \xi,\ u(+\infty) = \eta \text{ and } u' \in L^2(\mathbb{R})\big\}.$$

Obviously, $\Gamma(\xi, \eta) \subset H^1_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^N)$. Set

$$c(\xi, \eta) := \inf_{u \in \Gamma(\xi, \eta)} \mathcal{J}(u). \tag{2.22}$$

We cannot guarantee, in general, that $c(\xi, \eta)$ is attained at a heteroclinic from ξ to η. However it is attained at a *heteroclinic chain* between ξ and η in the sense of the following statement which, according to Rabinowitz [70], was first obtained by Strobel in his Ph.D. dissertation.

THEOREM 2.5. *Assume* $V \in C^2(\mathbb{R} \times \mathbb{R}^N, \mathbb{R})$ *satisfies* (A7) *and* (A8). *For each pair* $\xi, \eta \in \mathcal{M}$ $(\xi \neq \eta)$ *there exists a finite family* $\{w_1, \dots, w_j\} \subset \mathcal{M}$ *with* $w_1 = \xi$, $w_j = \eta$, *and a corresponding family of heteroclinics* v^i *of* (2.19), *with*

$$\mathcal{J}\left(v^i\right) = c(w_i, w_{i+1}), \quad v^i \in \Gamma(w_i, w_{i+1}),\ i = 1, \dots, j-1$$

and

$$\sum_{i=1}^{j-1} \mathcal{J}\left(v^i\right) = c(\xi, \eta).$$

It is easily seen by (A7)–(A8) that the v^i's may be assumed to be *basic heteroclinics* by which we mean that none of the heteroclinic chains, with more than one element, joining w_i and w_{i+1}, achieve the value $\mathcal{J}(v^i)$.

Variational methods have proved to be fruitful in investigating the existence of multiple heteroclinics for Hamiltonian systems of the type (2.19). Different approaches have been proposed by several authors but all of them share a common feature: at the start these always require some nondegeneracy assumption, like stating that some solution of (2.19) is in a sense isolated. This kind of assumption naturally excludes autonomous systems whereas it is commonly conjectured it is generically fulfilled when V depends on t although it is quite complicate to check it on concrete examples.

We shall take advantage from the existence of a basic heteroclinic chain from ξ to η in order to describe, following [70], how true heteroclinics from ξ to η may be obtained. We shall see that we can impose these solutions to spend arbitrarily large amounts of time near the "vertices" w_i. Hence there are infinitely many such solutions. This is one of several multiplicity results where solutions are distinguished by means of their behaviour with respect to the equilibria of the system (2.19): one therefore refers to "*multibump*" solutions.

Before stating the theorem, we first introduce a nondegeneracy condition. Given $\xi \neq \eta \in \mathcal{M}$, set

$$\mathcal{S}(\xi, \eta) := \left\{q(0) \in \mathbb{R}^N \mid q \in \Gamma(\xi, \eta),\ \mathcal{J}(q) = c(\xi, \eta)\right\}.$$

As $q(0) \in \mathcal{S}(\xi, \eta)$ implies $q(k) \in \mathcal{S}(\xi, \eta)$ for any integer k, we infer that ξ and η belong to $\bar{\mathcal{S}}(\xi, \eta)$. Denote respectively by $C_\xi(\xi, \eta)$ and $C_\eta(\xi, \eta)$ the connected components of ξ and η in $\bar{\mathcal{S}}(\xi, \eta)$. We assume

$$C_{w_i}(w_i, w_{i+1}) = \{w_i\} \quad \text{and} \quad C_{w_{i+1}}(w_i, w_{i+1}) = \{w_{i+1}\}, \\ \text{for } i = 1, \ldots, j-1. \tag{N}$$

It is shown in [70] that when these conditions fail then

$$C_{w_i}(w_i, w_{i+1}) = C_{w_{i+1}}(w_i, w_{i+1}),$$

which can be interpreted as a strongly degenerate situation.

THEOREM 2.6. *Assume* $V \in C^2(\mathbb{R} \times \mathbb{R}^N, \mathbb{R})$ *satisfies* (A7) *and* (A8). *Let there exist a basic heteroclinic chain of* (2.19), *consisting in the basic heteroclinics* v^i *with vertices* $\xi = w_1, \ldots, w_j = \eta$, *as described in Theorem* 2.5. *Then if* (N) *holds, the system* (2.19) *has infinitely many heteroclinic solutions connecting* ξ *and* η. *The time a solution spends in a neighbourhood of the vertices may be prescribed to exceed any arbitrary positive number.*

PROOF OF THEOREM 2.6. If I is an interval in $\mathbb{R}$, we write

$$\mathcal{J}_I(u) := \int_I \left(\frac{|u'|^2}{2} + V(t, u) \right) \mathrm{d}t.$$

By the nondegeneracy condition (N), given $\rho > 0$ we can fix open neighbourhoods A_i of w_i, B_{i+1} of w_{i+1} with diameter smaller than ρ and so that

$$\partial(A_i \cup B_{i+1}) \cap \bar{S}(w_i, w_{i+1}) = \emptyset, \quad \text{for } 1 \leqslant i \leqslant j-1. \tag{2.23}$$

We now consider a vector $z \in \mathbb{Z}^{2j-2}$ with coordinates

$$z_1 < z_2' < z_2 < z_3' < \cdots < z_j'$$

and define

$$X_z = \bigl\{ q \in \Gamma(\xi, \eta) \bigm| q(z_i) \in \bar{A}_i, q\bigl(z_{i+1}'\bigr) \in \bar{B}_{i+1}, i = 1, \ldots, j-1 \bigr\}.$$

The solutions we look for will be solutions of the following minimization problem:

$$b_z = \inf_{u \in X_z} \mathcal{J}(u) \tag{2.24}$$

provided that ρ is sufficiently small and the differences between the successive coordinates of z are sufficiently large. By the way, observe that the coordinates of z may be translated by any integer without affecting the variational problem.

CLAIM 1. *If* $0 < r < \frac{1}{3}$ *and* $\rho > 0$ *is sufficiently small*, b_z *is attained at a function* $u_z \in X_z$ *such that* $|u_z(t) - w_i| \leqslant r$ *for all* $t \in I_i$, *where* $I_i = [z_i', z_i]$, $i = 2, \dots, j-1$, $I_1 =]-\infty, z_1]$ *and* $I_j = [z_j', \infty[$.

Taking a minimizing sequence $(u_n)_n \subset X_z$ for (2.24), the claim follows from arguments similar to those we used in the proof of Theorem 2.3 that enable us to extract a convergent subsequence, still denoted by $(u_n)_n$, such that

$$u_n \overset{C_{\mathrm{loc}}(\mathbb{R},\mathbb{R}^N)}{\longrightarrow} u_z \quad \text{and} \quad u_n' \overset{L^2(\mathbb{R},\mathbb{R}^N)}{\rightharpoonup} u_z',$$

together with the fact that we may assume that the minimizing sequence satisfies the inequalities in the above statement. To see this last fact with respect to I_1, let $K(s, t; x, y)$ denote the affine map

$$\tau \to x + \frac{\tau - s}{t - s}(y - x), \quad s \leqslant \tau \leqslant t,$$

and consider the modified sequence $(U_n)_n \subset X_z$ defined by

$$U_n(t) = \begin{cases} w_1(t) & \text{if } t \leqslant z_1 - 1, \\ K\big(z_1 - 1, z_1; w_1, u_n(z_1)\big)(t) & \text{if } z_1 - 1 < t < z_1, \\ u_n(t) & \text{if } t \geqslant z_1. \end{cases}$$

A straightforward computation yields $\mathcal{J}_{I_1}(U_n) = o(\rho)$ as $\rho \to 0$. However if u_n reaches the boundary of $B_r(w_1)$ at some instant $t \in I_1$, the adaptation of Fact 1 of Section 2.2 to the present situation allows to prove $\mathcal{J}_{I_1}(u_n)$ is bounded from below by a constant depending only on r. Hence, if ρ is chosen sufficiently small, we may substitute U_n for u_n and the inequality follows. The same argument enables us to deal with the remaining intervals. This proves Claim 1.

CLAIM 2. *Let* $0 < \sigma < \rho$. *If the differences* $z_i - z_i'$ *are large enough, there exist subintervals* $J_i = [a_i, a_i + 2] \subset I_i$, *such that* $|u_z(t) - w_i| \leqslant \sigma$ *for all* $t \in J_i$ *and* $2 \leqslant i \leqslant j - 1$. *In addition* $\mathcal{J}_{J_i}(u_z) = o(\sigma)$ *as* $\sigma \to 0$.

By Claim 1, taking $z_i - z_i'$ large, we cannot have $|u_z(t) - w_i| \geqslant \sigma/2$ for all $t \in I_i$, since otherwise

$$\alpha_{\sigma/2}\big(z_i - z_i'\big) \leqslant \mathcal{J}_{I_i}(u_z) = o(\rho) \quad \text{for } \rho \to 0,$$

where $\alpha_{\sigma/2}$ is defined according to (2.20). In fact, the same argument shows that a subinterval $L \subset I_i$ with length greater than a certain $l > 0$ must contain at least one instant t such that $|u_z(t) - w_i| < \sigma/2$.

Suppose the first statement of Claim 2 is false. Then, each subinterval $J_i = [a_i, a_i + 2] \subset I_i$ containing an instant t where $|u_z(t) - w_i| < \sigma/2$ also contains an instant t' where $|u_z(t') - w_i| \geqslant \sigma$. If k is the number of such disjoint intervals, applying again the arguments of Fact 1 of Section 2.2 and the computations in the proof of Claim 1, we deduce a bound on k. Namely, we have

$$k\frac{\sigma}{2}\sqrt{2\alpha_{\sigma/2}} < o(\rho), \quad \text{for } \rho \to 0.$$

But, by the above remark, if $z_i - z_i'$ grows to infinity, so does k. This shows that the first statement of Claim 2 is true.

The second statement of Claim 2 is checked by noting that

$$\mathcal{J}_{J_i}(u_z) \leqslant \mathcal{J}_{J_i}(\psi) = o(\sigma), \quad \text{as } \sigma \to 0,$$

where ψ is defined by

$$\psi(t) := \begin{cases} u_z(t) & \text{if } t \in]-\infty, a_i[\cup]a_i+2, \infty[, \\ K\big(a_i, a_i+2, u_z(a_i), u_z(a_i+2)\big)(t) & \text{if } t \in [a_i, a_i+2]. \end{cases}$$

Let us now complete the proof. It is clear that $u_z(t)$ is a solution of (2.19) for $t \neq z_i,\ z_i'$. To prove that it is indeed a solution, we have to show that $u(z_i) \in A_i$ and $u(z_{i+1}') \in B_{i+1}$, for $1 \leqslant i \leqslant j-1$. To this purpose, an auxiliary variational problem is introduced: define

$$\Lambda(w_i, w_{i+1}) := \big\{u \in \Gamma(w_i, w_{i+1}) \mid u(0) \in \partial(A_i \cup B_{i+1})\big\}$$

and

$$d(w_i, w_{i+1}) := \inf_{u \in \Lambda(w_i, w_{i+1})} \mathcal{J}(u).$$

It is easily seen that this infimum is attained at some $S \in \Lambda(w_i, w_{i+1})$ and

$$d(w_i, w_{i+1}) > c(w_i, w_{i+1}).$$

Indeed, if there holds an equality, then $S(0) \in \partial(A_i \cup B_{i+1}) \cap \bar{S}(w_i, w_{i+1})$, contradicting (2.23).

Arguing by contradiction, assume that for some $i \in \{1, \dots, j-1\}$, we have, say,

$$u_z\big(z_{i+1}'\big) \in \partial B_{i+1}. \tag{2.25}$$

The other case may be handled in a similar way. By integer translation, we may suppose $z_{i+1}' = 0$. Also, we can translate the time variable in the basic heteroclinic v^i thus obtaining a new basic heteroclinic V^i such that $V^i(t) \in \bar{A}_i$ for $t \leqslant z_i$ and $V^i(t)$ leaves $\bar{A}_i$ at some time in $[z_i, z_{i+1}]$. Taking large differences $z_i - z_i'$ we may assume $V^i(t) \in \bar{B}_{i+1}$ for $t \geqslant z_{i+1}'$ as well. Hence $V^i(t) \in \bar{A}_i$ for $t \in J_i$ and $V^i(t) \in \bar{B}_{i+1}$ for $t \in J_{i+1}$.

Let us call U_z the modification of u_z outside $]-\infty, a_i] \cup [a_{i+1}+2, \infty[$ obtained by gluing together the following pieces: $K(a_i, a_i+1; u_z(a_i), w_i)$; $K(a_i+1, a_i+2; w_i, V^i(a_i+2))$;

the restriction of V^i to the interval $[a_i+2, a_{i+1}]$; $K(a_{i+1}, a_{i+1}+1; V^i(a_{i+1}), w_{i+1})$ and $K(a_{i+1}+1, a_{i+1}+2; w_{i+1}, u_z(a_{i+1}+2))$.

Finally we compute the difference

$$\mathcal{J}(u_z)-\mathcal{J}(U_z)=\mathcal{J}_{[a_i,a_{i+1}+2]}(u_z)-\mathcal{J}_{[a_i,a_{i+1}+2]}(U_z)$$

in the following way. Define an auxiliary curve γ_z by gluing together the constant w_i on $]-\infty, a_i]$; $K(a_i, a_i+1; w_i, u_z(a_i+1))$; u_z restricted to $[a_i+1, a_{i+1}]$; $K(a_{i+1}, a_{i+1}+1; u_z(a_{i+1}), w_{i+1})$ and the constant w_{i+1} in $[a_{i+1}+1, \infty[$. Note that $\gamma_z \in \Lambda(w_i, w_{i+1})$ by our assumption (2.25). Then we write

$$\begin{aligned}\mathcal{J}_{[a_i,a_{i+1}+2]}(u_z)-\mathcal{J}_{[a_i,a_{i+1}+2]}(U_z) &= \big(\mathcal{J}_{[a_i,a_{i+1}+2]}(u_z)-\mathcal{J}_{[a_i,a_{i+1}+2]}(\gamma_z)\big)\\ &\quad+\big(\mathcal{J}_{[a_i,a_{i+1}+2]}(\gamma_z)-\mathcal{J}_{[a_i,a_{i+1}+2]}(U_z)\big).\end{aligned}$$

Using Claim 2, we may check that the first difference on the right-hand side is an $o(\sigma)$ as $\sigma \to 0$. On the other hand, we clearly have

$$\mathcal{J}_{[a_i,a_{i+1}+2]}(\gamma_z) \geqslant b_z$$

and

$$\mathcal{J}_{[a_i,a_{i+1}+2]}(U_z) \leqslant \mathcal{J}_{[a_i,a_{i+1}+2]}\big(V^i\big)+o(\sigma), \quad \text{as } \sigma \to 0.$$

Hence, we finally deduce that

$$\begin{aligned}\mathcal{J}(u_z)-\mathcal{J}(U_z) &= \mathcal{J}_{[a_i,a_{i+1}+2]}(u_z)-\mathcal{J}_{[a_i,a_{i+1}+2]}(U_z)\\ &\geqslant b(w_i, w_{i+1})-c(w_i, w_{i+1})+o(\sigma), \quad \text{as } \sigma \to 0,\end{aligned}$$

leading to a contradiction with the minimizing property of u_z.

Conclusion. From what precedes, we conclude that for each $0<r<\frac{1}{3}$ and each vector $(z_1, z_2', z_2, z_3', \dots, z_j') \in \mathbb{Z}^{2j-2}$ such that the distances $z_i - z_i'$, $2 \leqslant i \leqslant j-1$, are large enough, there exists a heteroclinic solution u_z of (2.19) with the particularity that

$$\big|u_z(t)-w_i\big| \leqslant r, \quad \text{if } t \in \big[z_i', z_i\big], \quad \text{for } i=2,\dots,j-1$$

and

$$\big|u_z(t)-\xi\big| \leqslant r \quad \text{if } t \leqslant z_1, \qquad \big|u_z(t)-\eta\big| \leqslant r \quad \text{if } t \geqslant z_j'. \qquad \square$$

Several methods adapted to the search of multibump solutions may be found in the literature. For instance, in [67] Rabinowitz considers a system with two equilibria 0 and ξ and, starting from a pair of heteroclinics, v from 0 to ξ and w from ξ to 0, constructs heteroclinics that oscillate an arbitrary number of times between neighbourhoods of the equilibria. The method uses the functional $u \to \mathcal{J}(v+u)$ defined in $H^1(\mathbb{R}, \mathbb{R}^N)$. The concentration of Palais–Smale sequences with respect to homoclinics and heteroclinics is a crucial tool.

In [72], Coti Zelati and Rabinowitz give a variant of the result of Strobel for potentials of the form $V(t,u) = a(t)W(u)$ where W is periodic, a is slowly oscillating and not necessarily periodic.

The almost periodic case has been considered by Alessio, Bertotti and Montechiari [3] for fairly general Lagrangian systems. In particular, they show that, in presence of a slowly oscillating extra term, chaotic multibump dynamics arises.

2.4. *Notes and further comments*

1. In Theorem 2.1, the minimizer u takes values in $[-1,1]$. If (2.4) has uniqueness or if $F(u)$ lies below some quadratic function $c(u \pm 1)^2$ in the neighborhood of ∓ 1, we have $u(t) \in \,]-1,1[$ for all $t \in \mathbb{R}$. The heteroclinic of (2.4) is essentially unique (up to translation) as a consequence of (2.5) with $k = 0$.

2. If (2.9) has uniqueness of solutions for the Cauchy problem, the heteroclinics assume values in $]-1,1[$. In addition, if f has only one zero in $]-1,1[$, it can be checked that the minimizers of the functional $\mathcal{J}$ defined by (2.10) are strictly increasing. On the other hand, the presence of the function $a(t)$ in the functional rules out an easier argument which in the autonomous case shows that the elements of a minimizing sequence may assumed to be increasing functions.

3. In comparison with Theorem 2.2, the next theorem illustrates what can be said about second order systems where the potential depends on time and has no particular symmetry properties. For definiteness, consider the system

$$u'' = \nabla_u V(t,u), \tag{2.26}$$

where $u : \mathbb{R} \to \mathbb{R}^N$, $V \in C^2(\mathbb{R} \times \mathbb{R}^N, \mathbb{R})$ and ∇_u is the gradient with respect to the variable $u \in \mathbb{R}^N$. Assume that

(A9) $V(t,u) \geqslant 0$ for all $t \in \mathbb{R}$ and all $u \in \mathbb{R}^N$, and there exist $\xi \neq \eta$ such that $V(t,u) = 0$ if and only if $u \in \{\xi, \eta\}$;

(A10) there exist constants $a_1, a_2 > 0$ and $\varepsilon > 0$ such that $a_1|u-z|^2 \leqslant V(t,u) \leqslant a_2|u-z|^2$ if $z \in \{\xi, \eta\}$ and $|u-z| < \varepsilon$;

(A11) $\liminf_{|u| \to \infty} V(t,u) > 0$ uniformly in $t \in \mathbb{R}$.

Then one can prove:

THEOREM 2.7. *If* $V \in C^2(\mathbb{R} \times \mathbb{R}^N, \mathbb{R})$ *satisfies the assumptions* (A9)–(A11) *and*

$$t \frac{\partial V}{\partial t}(t,u) > 0$$

whenever $t \neq 0$ *and* $u \notin \{\xi, \eta\}$, *then* (2.26) *has a heteroclinic solution from* ξ *to* η.

This theorem is a particular case of the results proved by Chen and Tzeng [26]. In their paper, further examples of existence, and results on multiplicity of heteroclinics, may be found. Thus in this respect the non-autonomous equation behaves quite differently from the autonomous one. We also mention that the system considered there may have an infinite set of equilibria.

4. It is, however, interesting to note that, in contrast to Theorems 2.2 and 2.7, Korman and Lazer [46] consider a special scalar equation

$$u'' = a(t)\big(|u|^{p-1}u - u\big), \quad p > 1, \tag{2.27}$$

and they find a heteroclinic from -1 to 1 assuming that $a(t)$ is even, $a'(t) < 0$ if $t < 0$ and $a(\infty) > 0$. They exploit the symmetry of (2.27), solve the approximate two point problem with boundary conditions $u(0) = 0$, $u(T) = 1$ and then they pass to the limit as $T \to +\infty$.

5. Many other properties of heteroclinics of Hamiltonian systems have been established by means of variational methods. Let us list a few of them:

(i) *Heteroclinics to periodics.* Consider a system of the form

$$u'' = \nabla_u V(t, u) + f(t), \tag{2.28}$$

where V is a function of class C^2, 1-periodic in t and in the spacial variables, and f is continuous, 1-periodic and of zero mean value. Assume in addition that the system is *reversible*, namely that V and f are *even* in t. Then it is known that (2.28) has periodic solutions at the minimum level of the action functional corresponding to the periodic boundary value problem for (2.28). In [68] Rabinowitz gives conditions for the existence of heteroclinics *connecting two such periodic solutions*. In [53] Maxwell proves that there are heteroclinic chains between any two periodic solutions, the "vertices" of the chain being also periodic solutions.

Calanchi and Serra [24] use a constrained minimization approach for the existence of connections between consecutive periodic solutions. For ordinary differential equation, Bosetto and Serra [21] obtain multi-bump heteroclinics between consecutive periodic motions without any reversibility assumption.

(ii) *Heteroclinics to almost periodic solutions.* Alessio, Carminati and Montecchiari [4] prove the existence of heteroclinic connections joining almost periodic solutions of a Lagrangian system.

(iii) *Heteroclinics with an endpoint at infinity.* For potentials $V \in C^1(\mathbb{R}^N, \mathbb{R})$ with one zero at the origin and vanishing at infinity like some power $|u|^{-\alpha}$, $\alpha > 0$, the system (2.12) has at least one "heteroclinic at infinity", i.e. a solution $u(t)$ such that $u(-\infty) = 0$, $|u(+\infty)| = +\infty$ and $u'(\pm\infty) = 0$. We refer to Serra [80]. A similar result holds for a potential with a finite singularity when $N \geqslant 3$.

(iv) *Almost periodic systems.* Bertotti and Montecchiari [14] and Alessio, Bertotti and Montecchiari [3] consider almost periodic Lagrangian systems. Under suitable conditions, they obtain infinitely many heteroclinic solutions connecting possibly degenerate equilibria.

(v) *Singular potentials*. Many authors consider homoclinic solutions to equilibrium points of Hamiltonian systems with a singular potential, see [85,69,13,25] and the references therein. In case of a potential in $\mathbb{R}^2$ singular at some unique point, the basic result is the existence of many homoclinics classified according to their winding number around the singularity. It seems that only few attention was paid to heteroclinic connections with such settings. In [25] an autonomous system in $\mathbb{R}^2$ is considered. The origin is a minimum of the potential and a strong force at the singularity ξ is assumed. The authors obtain a certain periodic solution $\bar{u}$ with index 1 with respect to ξ, homoclinics with arbitrarily large indexes with respect to ξ and a heteroclinic from 0 to $\bar{u}$ which is built upon an infinite sequence of homoclinics. We already mentioned the result of Serra [80] about heteroclinics at infinity which holds in some singular frameworks. But up to our knowledge, no multiplicity results were obtained for heteroclinic solutions in the spirit of those concerning homoclinics.

(vi) *Heteroclinic connections between minima at different levels of the potential*. V. Coti Zelati and Rabinowitz [29] give conditions for the existence of a heteroclinic from χ to η for a system of the form

$$u'' = a(t)\nabla V(u),$$

where a is periodic and bounded away from zero and χ, η are isolated minima of V at *different* levels.

(vii) *Spatial heteroclinics*. Consider the PDE

$$\Delta u = a(x, y) f(u) \quad \text{in } \mathbb{R}^2, \tag{2.29}$$

where a is periodic in both variables and f is a non-negative smooth function such that $f(0) = f(1) = 0$. There are solutions $u(x, y)$ of (2.29) periodic in y and "heteroclinic in x" from 0 to 1, i.e. approaching respectively 0 and 1 as $x \to -\infty$ and $x \to +\infty$. This and a lot more of related results can be consulted in Rabinowitz and Stredulinsky [73, 74] and Rabinowitz [71]. See also the papers by Alessio, Jeanjean and Montecchiari [5, 6].

3. Heteroclinics as front wave profiles in reaction diffusion equations

An efficient model for many chemical and biological phenomena is provided by the partial differential equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} H(u) = \frac{\partial}{\partial x}\left[p(u)\frac{\partial u}{\partial x}\right] + g(u), \tag{3.1}$$

where $H, p \in C^1(\mathbb{R})$, $p > 0$ and following the terminology used in [49], g is a *function of type* A in [0, 1] by which we mean that

$$\begin{aligned}&g \text{ is continuous in } [0,1], \quad g(0) = g(1) = 0 \quad \text{and}\\ &g(u) > 0 \text{ if } u \in]0,1[.\end{aligned} \tag{Type A}$$

Hence the constants $u = 0$ and $u = 1$ are solutions of (3.1).

The first term in the right-hand side represents density dependent diffusion, while the term $\frac{\partial}{\partial x} H(u)$ accounts for convection effects.

An important problem related to this equation is that of finding positive travelling wave solutions, that is, positive solutions of the form $u(t,x) = U(x - ct)$ for some $c > 0$. Here c is the propagation speed of the wave. It is in addition required that the wave front $U(s)$ is defined in $]-\infty, +\infty[$ and satisfies $U(-\infty) = 1$, $U(+\infty) = 0$. This amounts to look at the solutions of the second order ordinary differential equation

$$\big(p(u)u'\big)' + \big(c - H'(u)\big)u' + g(u) = 0 \tag{3.2}$$

satisfying the limit conditions

$$u(-\infty) = 1, \quad u(+\infty) = 0.$$

We are therefore close to the problem of finding a heteroclinic between the equilibria 1 and 0.

The simplest and most classical case corresponds to $H \equiv 0$, $p \equiv 1$, the associated equation being then

$$u'' + cu' + g(u) = 0. \tag{3.3}$$

This (or the corresponding evolution equation) is usually referred as Fisher's equation, although in Fisher's original model $g(u)$ is the function $u(1-u)$. Among the vast literature on the existence of heteroclinics for (3.3) at least two contributions should be singled out: the pioneering work of Kolmogorov, Petrovsky and Piskounoff [45] and the paper by Aronson and Weinberger [11]—the latter concerning also diffusion in n-dimensional space. It is shown in [11] that, for $g \in C^1[0,1]$, there exists a heteroclinic connection of (3.3) between 1 and 0 if and only if $c \geqslant c^*$, where c^* is a positive number such that

$$2\sqrt{g'(0)} \leqslant c^* \leqslant 2\sqrt{\sup_{0<u<1} \frac{g(u)}{u}}.$$

The lower bound is clear, since linearizing (3.3) around $u = 0$ shows that for $c < 2\sqrt{g'(0)}$ the origin cannot act as an attractor of *positive* solutions. This result has been recently extended by Malaguti and Marcelli [48,49] for the more general problem (3.2) (in fact for even more general equations).

In this monograph we shall concentrate on the existence of heteroclinics and some bounds for the minimal speed c^*. Many other interesting problems concerning the equation (3.3) or systems of equations of this type have been dealt with in the literature. Some examples will be mentioned at the end of the section.

3.1. *Reduction to a first order equation*

Maybe the simplest way to obtain existence results for heteroclinic solutions of (3.2) or (3.3) consists in studying an equivalent first-order problem. This device may be found in [76] and has been extensively used in [49,50]. We modify it slightly, in the spirit of the approach given in [10].

Let us consider the equation (3.2) written as

$$\bigl(p(u)u'\bigr)' + h(u)u' + g(u) = 0 \tag{3.4}$$

and assume

(B1) $p \in C^1([0,1])$ and $p(u) > 0$ for all $u \in [0,1]$;
(B2) $h \in C[0,1]$ and $h(u) > 0$, for all $u \in [0,1]$;
(B3) g is a function of type A in $[0,1]$.

We first notice that monotone solutions of (3.4) such that $0 < u(s) < 1$ have no critical points. Indeed, if $u'(t_0) = 0$ and $0 < u(t_0) < 1$, then (3.4) implies $u''(t_0) < 0$. On the other hand, we observe that any decreasing solution $u(s)$ of (3.4) such that u is defined in $]-\infty, +\infty[$, $u(-\infty) = 1$ and $u(+\infty) = 0$ has the property

$$\lim_{s \to \pm\infty} u'(s) = 0, \tag{3.5}$$

so that u is indeed a heteroclinic between the equilibria 1 and 0. To prove (3.5) first remark that $\limsup_{s \to \pm\infty} u'(s) = 0$ by the boundedness of $u(s)$. Suppose, by contradiction, that $\liminf_{s \to +\infty} u'(t) < 0$, the limit at $-\infty$ is handled in a similar way. Let $t_n \to +\infty$ be a sequence so that $u'(t_n) \to 0$. Integrating (3.4) in $[0, t_n]$ shows that $\int_0^{t_n} g(u(s))\,\mathrm{d}s$ is bounded, whence by the positivity of g, we deduce that $\int_0^{+\infty} g(u(s))\,\mathrm{d}s$ is finite. Now let $t_n \to +\infty$ and $s_n \to +\infty$ be sequences so that $t_n < s_n < t_{n+1}$, $u'(t_n) \to 0$ and $u'(s_n) \to -\delta < 0$. From (3.4) we derive, for some $\tau_n \in [t_n, s_n]$

$$\begin{aligned}&\bigl[p\bigl(u(s_n)\bigr)u'(s_n) - p\bigl(u(t_n)\bigr)u'(t_n)\bigr] + h\bigl(u(\tau_n)\bigr)\bigl(u(s_n) - u(t_n)\bigr)\\&\quad + \int_{t_n}^{s_n} g\bigl(u(s)\bigr)\,\mathrm{d}s = 0.\end{aligned}$$

In the above equation the term in brackets has limit $-p(0)\delta$, while the remaining terms tend to 0, leading to a contradiction.

Now let $u = U(t)$ be a monotone decreasing heteroclinic of (3.4) and let $]t_-, t_+[$ be the maximal interval where

$$0 < U(t) < 1.$$

Then we have $U'(t) < 0$ for $t \in]t_-, t_+[$ and thus we can define $t(u)$, the inverse function of $u = U(t)$. Set

$$\varphi(u) := p\big[U\big(t(u)\big)\big]U'\big(t(u)\big). \tag{3.6}$$

Then φ is a C^1 function in $]0, 1[$ which can be continuously extended to $[0, 1]$ with $\varphi(0) = \varphi(1) = 0$. Moreover, φ satisfies

$$\varphi'(u)\varphi(u) + h(u)\varphi(u) + p(u)g(u) = 0.$$

Observe that the boundedness of h and the positivity of p and g implies that

$$\big|\varphi(u)\big| \leqslant ku$$

for some positive constant k. From Gronwall's inequality we then infer this rules out the possibility that $t_+ < +\infty$. Since $\varphi < 0$, we find that $\psi(u) := \varphi(u)^2$ is a solution of the first order differential equation

$$\psi'(u) = 2h(u)\sqrt{\psi(u)} - 2p(u)g(u) \tag{3.7}$$

in $[0, 1]$, together with the endpoint conditions

$$\psi(0) = \psi(1) = 0. \tag{3.8}$$

In particular, ψ is of type A in $[0, 1]$.

Conversely, if (3.7) has a solution ψ of type A in $[0, 1]$, we use it to define the solution $u(t)$ of the Cauchy problem

$$u' = -\frac{\sqrt{\psi(u)}}{p(u)}, \qquad u(0) = \frac{1}{2}. \tag{3.9}$$

The domain of this solution is $]t_-, t_+[$, where

$$t_- = -\int_{1/2}^{1} \frac{p(u)\,\mathrm{d}u}{\sqrt{\psi(u)}}, \qquad t_+ = \int_0^{1/2} \frac{p(u)\,\mathrm{d}u}{\sqrt{\psi(u)}}.$$

The boundedness of h and (3.7) imply that in a neighborhood of $u = 0$, we have

$$\sqrt{\psi(u)} \leqslant ku$$

for some positive constant k. Therefore it follows from assumption (B1) that $t_+ = +\infty$. If we assume in addition that

(B4) there exists $\ell > 0$ such that for all $u \in [0, 1]$, $g(u) \leqslant \ell(1 - u)$, then we infer from (3.7) that there exists some constant $k_1 > 0$ so that $\psi(u) \leqslant k_1(1-u)^2$ in a neighborhood of $u = 1$. Hence under the condition (B4) we also have $t_- = -\infty$.

A straightforward computation shows that $u(t)$ satisfies (3.4) in $]t_-, t_+[$, and it is clear that

$$\lim_{t\to t_-} u(t) = 1, \qquad \lim_{t\to t_+} u(t) = 0, \qquad \lim_{t\to t_\pm} u'(t) = 0.$$

Therefore we have shown:

PROPOSITION 3.1. *Assume* (B1)–(B4) *hold and let* φ *be the function defined by* (3.6). *Then* $u = U(t)$ *is a* (*strictly*) *decreasing heteroclinic solution of* (3.4) *between* 1 *and* 0 *if and only if* $\varphi(u)^2$ *is a solution of type A of* (3.7) *in* [0, 1].

REMARK. Under weaker hypotheses, namely if we allow a singularity at zero, i.e. $p(0) = 0$, or in the absence of (B4), t_- or t_+, or both, may be finite. In these cases, we still obtain a heteroclinic by trivially extending $U(t)$ with the value 1 to the left and with the value 0 to the right.

Proposition 3.1 reduces the existence of heteroclinics for (3.4) to the existence of type A solutions of (3.7). Therefore it is useful to have criteria that ensure existence of such solutions.

PROPOSITION 3.2. *Assume* (B1)–(B3) *hold.*

(i) *Suppose that* $s(u)$ *is a* C^1 *function in* [0, 1] *such that* $s(0) = 0$, $s(u) > 0$ *if* $u \in]0, 1[$ *and for all* $u \in [0, 1]$,

$$s'(u) \leqslant 2h(u)\sqrt{s(u)} - 2p(u)g(u). \tag{3.10}$$

Then Eq. (3.7) *has a solution of type A.*

(ii) *Equation* (3.7) *has at most one solution of type A.*

PROOF. (i) The assumption (3.10) which means s is a lower solution of the initial value problem

$$\begin{aligned} &\psi' = 2h(u)\sqrt{\psi} - 2p(u)g(u), \\ &\psi(0) = 0, \end{aligned} \tag{3.11}$$

implies, as is well known (see, e.g., [91]), that (3.11) has a solution $\psi(u)$ such that $\psi(u) \geqslant s(u)$. If $\psi(1) = 0$ then ψ is a type A solution of (3.7) and we are done.

If $\psi(1) > 0$, we consider the solution $\bar{\psi}$ of the initial value problem

$$\begin{aligned} &\bar{\psi}' = 2h(u)\sqrt{|\bar{\psi}|} - 2p(u)g(u), \\ &\bar{\psi}(1) = 0. \end{aligned} \tag{3.12}$$

It is clear that we may assume $\bar{\psi} \geqslant 0$ in $[0, 1]$, since 0 is a lower solution for the initial value problem (3.12) in $[0, 1]$. We claim that $0 < \bar{\psi}(u) < \psi(u)$ for all $u \in]0, 1[$. In fact, if u_0 is the largest zero of $\bar{\psi}$ in $]0, 1[$, then (3.12) implies $\bar{\psi}'(u_0) < 0$, which is impossible. If $u_1 \in]0, 1[$ is such that $\bar{\psi}(u_1) = \psi(u_1)$, then $\bar{\psi} \equiv \psi$, which is also impossible. Hence the claim is proved and by continuity, $\bar{\psi}(0) = 0$ so that $\bar{\psi}$ is the desired solution of type A.

(ii) Assume by contradiction that ψ_1 and ψ_2 are distinct type A solutions of (3.7). By uniqueness of the solution of the Cauchy problem, these two solutions are ordered, say $\psi_1(u) < \psi_2(u)$ for all $u \in]0, 1[$. But then (3.7) shows that $\psi_2 - \psi_1$ is increasing and this contradicts $\psi_1(1) = \psi_2(1) = 0$. □

The results we have established have straightforward consequences:

PROPOSITION 3.3.
(i) *The decreasing heteroclinic of* (3.4), *if it exists, is unique up to translation.*
(ii) *Consider two equations*

$$\bigl(p_i(u)u'\bigr)' + h_i(u)u' + g_i(u) = 0, \quad i = 1, 2, \tag{$*$}_i$$

where p_i, h_i, g_i satisfy assumptions (B1)–(B4). *Then if $h_1 \leqslant h_2$, $p_1 \geqslant p_2$ and $g_1 \leqslant g_2$, and* $(*)_1$ *has a decreasing heteroclinic, then so does* $(*)_2$.

REMARK. If $c_1 < c_2$ the heteroclinics u_1, u_2 of (3.3) with respectively $c = c_1$, c_2 satisfy $\psi_2 < \psi_1$ by the above ordering argument.

EXAMPLES. (1) Suppose H' in (3.2) is continuous. Then there exists a number $c^* \in \mathbb{R}$ such that (3.2) admits a decreasing solution if and only if $c \geqslant c^*$. In particular, for (3.3) we have the classical estimate

$$0 < c^* \leqslant 2\sqrt{\sup_{0<u<1} \frac{g(u)}{u}}. \tag{3.13}$$

In fact, if there exists $M > 0$ so that for all $u \in [0, 1]$, $g(u) \leqslant Mu$, then we can choose a constant $\beta > 0$ so that $s(u) = \beta u^2$ is a lower solution of (3.11) with $h(u) \equiv c$, $p(u) \equiv 1$, provided that $c^2 \geqslant 4M$.

(2) Consider the autonomous case

$$u'' + cu' + g(u) = 0$$

in which there exists $M > 0$ and $\alpha \geqslant 2$ so that

$$g(u) \leqslant Mu^\alpha(1 - u).$$

(This type of nonlinear term appears in many applications.) It is easy to see that $s = \beta u^\alpha(1 - u)^2$ is a lower solution of (3.11) with $h(u) \equiv c$, $p(u) \equiv 1$, if for all $u \in [0, 1]$,

$$\alpha\beta u^{\alpha/2-1} + \bigl(2M - \beta(\alpha + 2)\bigr)u^{\alpha/2} \leqslant 2c\sqrt{\beta}.$$

Elementary but tedious computations show that such a number $\beta > 0$ exists if $c \geqslant \sqrt{2M}\big(\frac{\alpha-1}{\alpha+2}\big)^{(\alpha-1)/2}$. We thus obtain

$$c^* \leqslant \sqrt{2M}\left(\frac{\alpha-1}{\alpha+2}\right)^{(\alpha-1)/2},$$

an estimate which improves, for small values of α (e.g., $\alpha \leqslant 4$) that allowed by (3.13).

3.2. *Fast solutions and heteroclinics*

In this section we consider the simple model

$$u'' + cu' + g(u) = 0 \tag{3.14}$$

with g of type A. Following [10] we say that a solution $u(t)$ of (3.14) is a *fast solution* if it is defined in some interval $[t_0, +\infty]$ and satisfies the integrability condition

$$\int_{t_0}^{+\infty} \mathrm{e}^{ct} u'(t)^2 \,\mathrm{d}t < +\infty.$$

In order to look for this kind of solutions we introduce the Hilbert space

$$H_c = \left\{ u \in H^1_{\mathrm{loc}}(0, +\infty) \;\middle|\; \int_0^{+\infty} \mathrm{e}^{ct} u'(t)^2 \,\mathrm{d}t < +\infty \text{ and } u(+\infty) = 0 \right\}$$

with the norm

$$\|u\| = \left(\int_0^{+\infty} \mathrm{e}^{ct} u'(t)^2 \,\mathrm{d}t \right)^{1/2}.$$

Note that $u(+\infty)$ exists whenever the above integral is finite, since by Schwarz's inequality

$$\big|u(T) - u(S)\big| \leqslant \left(\frac{|\mathrm{e}^{-cS} - \mathrm{e}^{-cT}|}{c} \int_S^T \mathrm{e}^{ct} u'(t)^2 \,\mathrm{d}t \right)^{1/2}.$$

On the other hand, taking limits in the above inequality as $T \to +\infty$ it turns out that for all $u \in H_c$

$$\|u\|_{L^\infty(S,+\infty)} \leqslant \frac{\mathrm{e}^{-cS/2}}{\sqrt{c}} \|u\|. \tag{3.15}$$

PROPOSITION 3.4. *For all $u \in H_c$, we have*

$$\frac{c^2}{4} \int_0^{+\infty} \mathrm{e}^{ct} u(t)^2 \,\mathrm{d}t \leqslant \int_0^{+\infty} \mathrm{e}^{ct} u'(t)^2 \,\mathrm{d}t. \tag{3.16}$$

PROOF. In Hardy's inequality (see [40])

$$\int_0^{+\infty} \frac{v(s)^2}{s^2}\,\mathrm{d}s \leqslant 4\int_0^{+\infty} v'(s)^2\,\mathrm{d}s,$$

where v is absolutely continuous in $[0,+\infty[$ and $v(0)=0$, make the change of variables $v(s)=u(t)$, $s=\mathrm{e}^{-ct}$ to obtain

$$c\int_{-\infty}^{+\infty} \mathrm{e}^{ct}u(t)^2\,\mathrm{d}t \leqslant \frac{4}{c}\int_{-\infty}^{+\infty} \mathrm{e}^{ct}u'(t)^2\,\mathrm{d}t.$$

Then choose $u(t)$ to be constant for $t\leqslant 0$ and ignore, in the left-hand side of the estimate, the integral over $]-\infty,0]$ which is positive. □

Assume that g satisfies
(B5) $\sup_{0<u<1}\frac{g(u)}{u}<+\infty$
and extend $g(u)$ to $]-\infty,\infty[$ with the value 0 outside $[0,1]$. Solutions of (3.14) defined in $[0,+\infty[$ and taking values in $[0,1]$ can thus be identified with critical points of the functional $\mathcal{F}\colon H_c\to\mathbb{R}$ defined by

$$\mathcal{F}(u)=\int_0^{+\infty}\mathrm{e}^{ct}\left(\frac{u'^2}{2}-G(u)\right)\mathrm{d}t,$$

where $G(s)=\int_0^s g(\tau)\,\mathrm{d}\tau$. In fact, one can check that under the assumption (B5), $\mathcal{F}$ is a functional of class C^1 in H_c and (3.14) (with the extended g) is the Euler–Lagrange equation for $\mathcal{F}$.

Let us introduce one more assumption:
(B6) there exists a constant k such that $k\in\,]0,c^2/4[$ and

$$G(u)\leqslant\frac{ku^2}{2},\qquad \text{for } 0\leqslant u\leqslant 1.$$

PROPOSITION 3.5. *Assume* (B3), (B5) *and* (B6) *hold. Then* $\mathcal{F}$ *attains a minimum in the set*

$$M=\bigl\{u\in H_c \mid u(0)=1\bigr\}.$$

The minimizer u *is a solution of* (3.14) *in* $[0,+\infty[$ *such that for all* $t\geqslant 0$, $u'(t)\leqslant 0$ *and* $0\leqslant u(t)\leqslant 1$.

PROOF. We write $\mathcal{F}$ in the form

$$\mathcal{F}(u)=\int_0^{+\infty}\mathrm{e}^{ct}\left(\frac{u'^2}{2}-k\frac{u^2}{2}\right)\mathrm{d}t+\int_0^{+\infty}\mathrm{e}^{ct}H(u)\,\mathrm{d}t, \tag{3.17}$$

where by (B6),

$$H(u) = k\frac{u^2}{2} - G(u) \geqslant 0.$$

Let $(u_n)_n \subset M$ be a minimizing sequence for $\mathcal{F}$. Since $G(u) = G(\max(0, \min(1, u)))$, it is clear that arguing as in the proof of Theorem 2.1, we may assume $0 \leqslant u_n \leqslant 1$. Because of (3.16), the first integral in (3.17) is the square of an equivalent norm in H_c. Hence $(u_n)_n$ is bounded and, passing to a subsequence we may assume that there exists $u \in H_c$ so that $u_n \rightharpoonup u$ in H_c and $u_n \to u$ uniformly in compact intervals. In particular $u \in M$ and, by the lower semicontinuity of the norm and Fatou's lemma, we infer from (3.17) that

$$\mathcal{F}(u) \leqslant \liminf_{n\to\infty} \mathcal{F}(u_n).$$

Hence u is a minimizer of $\mathcal{F}$ in M. Since $0 \leqslant u \leqslant 1$, $u(t)$ is a solution of (3.14) with $u(0) = 1$. Writing (3.14) as

$$\left(\mathrm{e}^{ct}u'\right)' + \mathrm{e}^{ct}g(u) = 0 \tag{3.18}$$

we see that $\mathrm{e}^{ct}u'$ is decreasing, hence for all $t \geqslant 0$ we have $u'(t) \leqslant 0$. □

THEOREM 3.6. *Assume* (B3), (B5) *and* (B6) *hold. Then* (3.14) *has a heteroclinic connecting* 1 *and* 0.

PROOF. Let U be the solution found in the preceding proposition. If $U'(0) = 0$, then extending U with value 1 on $]-\infty, 0]$ we obtain a heteroclinic solution. Otherwise, we have $U'(t) < 0$, $0 < U(t) \leqslant 1$ for all $t \geqslant 0$ and we may define $\varphi(u) = U'(t(u))$ for $0 \leqslant u \leqslant 1$ as in (3.6). We thus obtain a solution $\psi(u) = \varphi(u)^2$ of (3.7) (with $p \equiv 1$, $h \equiv c$), such that $\psi(0) = 0$ and $\psi(u) > 0$ if $0 < u \leqslant 1$. It now follows from Proposition 3.2 that (3.7) has a solution of type A and Proposition 3.1 allows to conclude. □

We then have the straightforward corollary:

COROLLARY 3.7. *For a given function g of type A satisfying* (B5), *the threshold speed of Eq.* (3.14) *satisfies*

$$c^* \leqslant 2 \sup_{0<u<1} \sqrt{\frac{2G(u)}{u^2}}.$$

To end this section we illustrate the use of variational and comparison arguments to obtain a heteroclinic for a nonautonomous equation (see also [50] for further results relative to the non-autonomous case). We consider

$$u'' + p(t)u' + g(u) = 0, \tag{3.19}$$

where p is a continuous function in $\mathbb{R}$ such that

(B7) there exist $0 < c < d < +\infty$ such that for any $t \in \mathbb{R}$, $c \leqslant p(t) \leqslant d$.

As above, we make use of a weighted space of functions:

$$X_\tau = \left\{ u \in H^1_{\mathrm{loc}}(\tau, +\infty) \;\middle|\; \int_\tau^{+\infty} \mathrm{e}^{P(t)} u'(t)^2 \,\mathrm{d}t < +\infty \text{ and } u(+\infty) = 0 \right\},$$

where τ is taken negative and $P(t) = \int_0^t p(s)\,\mathrm{d}s$. Arguing as in the proof of Proposition 3.4 and assuming p satisfy (B7), we derive the inequality

$$\frac{c^2}{4} \int_\tau^{+\infty} \mathrm{e}^{P(t)} u(t)^2 \,\mathrm{d}t \leqslant \int_\tau^{+\infty} \mathrm{e}^{P(t)} u'(t)^2 \,\mathrm{d}t, \tag{3.20}$$

which is valid for any $u \in X_\tau$. On the other hand, it is easily seen that for any $u \in X_\tau$ we have the estimate

$$\|u\|_{L^\infty(0,+\infty)} \leqslant C \left(\int_0^{+\infty} \mathrm{e}^{P(t)} u'(t)^2 \,\mathrm{d}t \right)^{1/2} \tag{3.21}$$

for some positive constant C.

THEOREM 3.8. *Assume* (B3) *holds*, $p \in C(\mathbb{R})$ *satisfies* (B7), g *is Lipschitz in* $[0, 1]$ *and there exists* $k < c^2/4$ *such that*

$$g(u) \leqslant ku, \quad \textit{for } u \in [0, 1]. \tag{3.22}$$

Then, *given* $a \in\,]0, 1[$, (3.19) *has a strictly decreasing heteroclinic connecting* 1 *and* 0 *and satisfying* $u(0) = a$.

PROOF. We outline the proof in four steps.

Step 1. We claim that the linear equation

$$\varphi'' + p(t)\varphi' + k\varphi = 0 \tag{3.23}$$

has a solution φ_τ such that $\varphi_\tau(\tau) = 1$, $0 < \varphi_\tau < 1$ in $]\tau, +\infty[$ and $\varphi_\tau \in X_\tau$.

Indeed, it suffices to take φ_τ as the minimizer of the quadratic functional

$$\mathcal{F}_\tau(u) = \int_\tau^{+\infty} \mathrm{e}^{P(t)} \left(\frac{u'^2}{2} - k\frac{u^2}{2} \right) \mathrm{d}t$$

in the convex set $\{u \in X_\tau \colon\ u(\tau) = 1\}$. It is easy to see, on the basis of (3.20) which remains valid with τ substituted by any $\xi \geqslant \tau$, that in fact $\varphi_\tau > 0$. Observe also that, multiplying the self-adjoint form of (3.23) by φ_τ and integrating by parts, we infer that $\mathrm{e}^{P(t)}\varphi'(t)\varphi(t)$ has limit as $t \to +\infty$. It then follows from (3.20) and the fact that $\varphi_\tau > 0$ that this limit is negative. This in turn implies that $\varphi_\tau'(\tau) < 0$ and $\varphi_\tau < 1$ in $]\tau, +\infty[$. Hence this step is completed.

Step 2. We claim $\lim_{\tau\to-\infty} \varphi_\tau(0) = 0$.

Given $\beta > 0$, we first evaluate $\mathcal{F}_\tau$ on the test function $u(t) = \mathrm{e}^{-(\beta/2)P(t-\tau)}$ and obtain

$$\mathcal{F}_\tau(\varphi_\tau) = \inf_{X_\tau} \mathcal{F}_\tau \leqslant \frac{\beta^2 d^2}{8} \int_0^{+\infty} \mathrm{e}^{P(\tau+t)-\beta P(t)}\,\mathrm{d}t. \tag{3.24}$$

From the inequality (3.20), we come out with the estimate

$$\mathcal{F}_\tau(\varphi_\tau) \geqslant \frac{1}{2}\left(1 - \frac{4k}{c^2}\right) \int_\tau^{+\infty} \mathrm{e}^{P(t)} \varphi_\tau'(t)^2\,\mathrm{d}t. \tag{3.25}$$

It then follows from the inequalities (3.24), (3.25) and $k < c^2/4$ that

$$\int_0^{+\infty} \mathrm{e}^{P(t)} \varphi_\tau'(t)^2\,\mathrm{d}t \leqslant C \int_0^{+\infty} \mathrm{e}^{P(\tau+t)-\beta P(t)}\,\mathrm{d}t$$

for some $C \geqslant 0$ and consequently we deduce from Step 1 and (3.21) that there exists a number $L \geqslant 0$ such that for all $\tau < 0$,

$$0 < \varphi_\tau(0) \leqslant L \left(\int_0^{+\infty} \mathrm{e}^{P(\tau+t)-\beta P(t)}\,\mathrm{d}t \right)^{1/2}.$$

If we take $\beta > d/c$, we may apply the dominated convergence theorem to the integral in the right-hand side and conclude that

$$\lim_{\tau\to-\infty} \varphi_\tau(0) = 0.$$

Step 3. Let $\tau < 0$ be such that $\varphi_\tau(0) < a$. From (3.22) it follows that φ_τ is an upper solution of (3.19) with respect to the two-point boundary conditions $u(\tau) = 1$, $u(0) = 0$. Since 0 is a lower solution, we obtain a solution v of (3.19) in $[\tau, 0]$ such that $v(\tau) = 1$, $v(0) = 0$, $0 \leqslant v \leqslant \varphi_\tau$. This in turn is a lower solution in $[\tau, 0]$ with respect to the boundary conditions $u(\tau) = 1$, $u(0) = a$. Since 1 is an upper solution, there is a solution u_τ of (3.19) such that

$$u_\tau(\tau) = 1, \qquad u_\tau(0) = a, \qquad v \leqslant u_\tau \leqslant 1 \quad \text{in } [\tau, 0].$$

We claim that the solution u_τ satisfies for all $t \geqslant 0$,

$$u_\tau(t) \geqslant \varphi_\tau(t) > 0.$$

Otherwise, since $u_\tau(\tau) = \varphi(\tau)$ and $u_\tau(0) > \varphi_\tau(0)$, there exist t_0, t_1 such that $\tau \leqslant t_0 < 0$, $t_0 < t_1 \leqslant +\infty$, $u_\tau(t_i) = \varphi_\tau(t_i)$ for $i = 0, 1$, $u_\tau > \varphi_\tau$ in $]t_0, t_1[$. Multiplying the self-adjoint forms of (3.19) and (3.23) respectively by φ_τ and u_τ, and integrating by parts, we obtain

$$\left[\mathrm{e}^{P(t)}\left(u_\tau'(t)\varphi_\tau(t) - u_\tau(t)\varphi_\tau'(t)\right)\right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \mathrm{e}^{P(t)}\left(g(u_\tau)\varphi_\tau - k u_\tau\varphi_\tau\right) \mathrm{d}t = 0. \tag{3.26}$$

On the other hand, since g satisfies a Lipschitz condition in $[0, 1]$, we must have $u_\tau'(t_0) > \varphi_\tau'(t_0)$ and $u_\tau'(t_1) \leqslant \varphi_\tau'(t_1)$. But then using (3.22), we conclude that the left-hand side of (3.26) is strictly negative, yielding a contradiction.

Step 4. From the last two steps it follows that we may define a sequence $\tau_n \to -\infty$ and a corresponding sequence $u_n = u_{\tau_n}$ of solutions of (3.19) such that

$$u_n(\tau_n) = 1, \qquad u_n(0) = a \qquad \text{and} \qquad 0 < u_n < 1 \quad \text{in }]\tau_n, +\infty[.$$

A diagonal argument, based on simple estimates on the sequence $(u_n)_n$, shows that some subsequence converges to a solution u of (3.19) in $]-\infty, +\infty[$ such that $u(0) = a$ and $0 < u < 1$. It is easy to see that $u' < 0$, $u(-\infty) = 1$, $u(+\infty) = 0$ and indeed u is the desired heteroclinic solution. □

3.3. *Heteroclinics in a combustion model*

In this section we consider the differential equation

$$u'' + cu' + f(u) = 0, \tag{3.27}$$

where $f \in C([0, 1], \mathbb{R})$ has the following behaviour:

(B8) $f(0) = f(1) = 0$;
(B9) there exists $a \in]0, 1[$ such that $f(u) < 0$ if $u \in]0, a[$;
(B10) $\int_0^1 f(t)\,\mathrm{d}t > 0$.

Equation (3.27) with this type of nonlinearity arises as the front wave equation for some combustion models. We refer the reader to [12,33,52] for a study in depth of related problems. Here we confine ourselves to look for solutions of (3.27) which are monotone heteroclinics connecting the equilibria 1 and 0.

Note that by virtue of assumptions (B9) and (B10), Eq. (3.27) has at least one more equilibrium between 0 and 1. The behavior of solutions of (3.27) is quite different from those of (3.14). The main result describing the existence of heteroclinics in terms of the parameter c is the following.

THEOREM 3.9. *Assume* (B8)–(B10) *hold and suppose in addition that* $f(u) > 0$ *for* $\xi < u < 1$, *where* ξ *is the smallest number in* $]0, 1[$ *with the property that* $\int_0^\xi f(t)\,\mathrm{d}t = 0$. *Then there exists a unique* $c^* > 0$ *such that* (3.27) *has a heteroclinic solution* $u(t)$, *connecting* 1 *and* 0, *with the property that* $u'(t) < 0$ *for all* $t \in \mathbb{R}$. *Moreover, this heteroclinic is unique up to translation.*

PROOF. In order to prove this theorem we proceed as in Section 3.1, obtaining the desired heteroclinic by means of a solution to the first-order problem (3.9) where $\psi(u)$ is a positive solution of

$$\psi'(u) = 2c\sqrt{\psi(u)} - 2f(u) \tag{3.28}$$

in [0, 1], $\psi(0) = 0 = \psi(1)$ and $p \equiv 1$. Let us therefore study the Cauchy problem

$$\begin{aligned} &\psi' = 2c\sqrt{\psi_+} - 2f(u),\\ &\psi(0) = 0, \end{aligned} \tag{3.29}$$

where $\psi_+(u) = \max(\psi(u), 0)$.

Step 1. Given $c \geqslant 0$, (3.29) has a unique solution in [0, 1].

If $c = 0$ the solution is $\psi_0(u) = -2F(u) := -2\int_0^u f(t)\,\mathrm{d}t$. Now assume $c > 0$. It is clear that any solution ψ of (3.29) satisfies $\psi(u) > \psi_0(u)$ if $0 < u \leqslant 1$. In particular we have $\psi(u) > 0$ in $]0, a[$. Write $\psi = \theta - 2F(u)$, so that in a right neighborhood of zero θ is a solution of

$$\begin{aligned} &\theta' = 2c\sqrt{\theta - 2F(u)},\\ &\theta(0) = 0. \end{aligned} \tag{3.30}$$

If ψ_1, ψ_2 are distinct positive solutions of (3.29), then they are ordered in a neighborhood of 0 that is $\psi_1(u) < \psi_2(u)$ and so are the corresponding solutions of (3.30): $\theta_1(u) < \theta_2(u)$. Since these are strictly increasing near zero we obtain, for their inverse functions $u_i(\theta)$ $(i = 1, 2)$ in some interval $0 < \theta \leqslant b$,

$$\frac{\mathrm{d}u_i}{\mathrm{d}\theta} = \frac{1}{2c\sqrt{\theta - 2F(u_i)}}$$

and $u_i(0) = 0$. However, since $-2F$ is strictly increasing in a neighborhood of zero, we may assume that $-2F(u_2) < -2F(u_1)$ and therefore $u_2 - u_1$ is increasing in $[0, b]$. This contradicts $u_1(0) = 0 = u_2(0)$. Hence the solution ψ of (3.29) is unique at least as long as it is strictly positive. Assume that $\psi(\tau) = 0$ for some $0 < \tau < 1$. We clearly have $\tau > \xi$ and, in view of our assumptions, $\psi'(\tau) < 0$. But then (3.29) implies $\psi(u) = -2\int_\tau^u f(t)\,\mathrm{d}t$ for $\tau \leqslant u \leqslant 1$.

Step 2. There exists $c > 0$ such that the solution $\psi_c(u)$ of (3.29) satisfies $\psi_c(1) > 0$.

It suffices to note that, if we set

$$M = \sup_{0<u<1} \frac{f(u)}{u}$$

and choose $c > 1 + M$, then u^2 is a lower solution of (3.29).

Step 3. Uniqueness and the fact that solutions of (3.29) are uniformly bounded when c runs over bounded sets, imply that the solution $\psi_c(u)$ of (3.29) depends continuously on c. By assumption *(B10)*, we infer that $\psi_0(1) < 0$. By Step 2 we conclude that there exists $c^* > 0$ such that $\psi_{c^*}(1) = 0$. Moreover, by uniqueness and a standard comparison theorem, $\psi_c(1)$ is strictly increasing as a function of c, so that c^* is unique.

To complete the proof, if suffices to observe that a solution of (3.29) with $\psi(1) = 0$ satisfies $\psi(u) > 0$ if $0 < u < 1$ and is therefore a solution of (3.28). But this is a straightforward consequence of the argument used at the end of Step 1. ☐

3.4. *Notes and further comments*

1. Elementary approaches to the existence of the travelling front solutions, for an equation or a system with two equilibria, can be found in [2,77,78] (in [78] a class of nonautonomous systems is considered). Front waves and their minimal speed are studied in [32] for two-dimensional systems with four equilibria.

2. Comparison techniques are used in [44] to give analytic approximation of the front solutions.

3. We might have enlarged the scope of the results described here by considering the presence of singularities, namely $p(0) = 0$ in (3.1). Such singularities are important in applications. In [51] Malaguti and Marcelli study the problem

$$\big(p(u)u'\big)' + cu' + g(u) = 0, \tag{3.31}$$

where g is of type A and p is a C^1 function with $p(0) = 0$ and $p'(0) > 0$. In addition to decreasing heteroclinics between 1 and 0 there appear *sharp type* solutions, by which one understands decreasing solutions $u(t)$ of (3.31) in $]-\infty, 0]$ such that $u(-\infty) = 1$, $u(0) = 0$ and $u'(0) = -c/p'(0)$. It turns out that a threshold speed c^* exists such that (3.31) has a decreasing heteroclinic from 1 to 0 if $c > c^*$, a sharp type solution if $c = c^*$ and no solution of either type if $c < c^*$.

4. Hamel and Nadirashvili [39] use front wave solutions of

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + g(u),$$

with g concave of type A, to construct an infinite dimensional manifold of entire solutions.

5. A wealth of results concerning wavefronts and their admissible speeds can be found in the recent monograph by Gilding and Kersner [35].

Part 2. Fourth Order Equations

4. The Extended Fisher–Kolmogorov equation: An overview

The class of fourth order differential equations considered in this survey is related to semilinear evolution equations of the form

$$\frac{\partial u}{\partial t} + \frac{\partial^4 u}{\partial x^4} - \beta \frac{\partial^2 u}{\partial x^2} + u^3 - u = 0, \tag{4.1}$$

where β is a real parameter. This equation serves as a model in studies of pattern formation in many physical, chemical or biological systems. It also arises in the study of phase transitions.

When $\beta > 0$, it is related to the so-called *Extended Fisher–Kolmogorov* equation which was proposed in 1988 by Dee and van Saarloos [30] as a higher-order model equation for *bistable* systems. The term bistable indicates that the states $u = \pm 1$ are stable for the homogenized equation $-u' = u^3 - u$. When suitably scaled, solutions of (4.1) solve the equation

$$\frac{\partial v}{\partial t} + \gamma \frac{\partial^4 v}{\partial x^4} - \frac{\partial^2 v}{\partial x^2} + v^3 - v = 0, \tag{4.2}$$

where the positive parameter γ is related to β by the formula $\beta = 1/\sqrt{\gamma}$ and $v(x,t) = u(\sqrt{\beta} x, t)$. When $\gamma = 0$ in (4.2), the equation is a second order evolution equation (of Fisher type) which enters into the framework of Section 3. It is often referred to as the Fisher–Kolmogorov equation.

In this section, we summarize recent results concerning the existence of heteroclinic solutions for the stationary Extended Fisher–Kolmogorov equation

$$u'''' - \beta u'' + u^3 - u = 0. \tag{4.3}$$

Heteroclinic solutions of (4.3) connecting -1 to $+1$ in the phase-space satisfy the following conditions

$$\lim_{x \to \pm\infty} \big(u, u', u'', u'''\big)(x) = (\pm 1, 0, 0, 0). \tag{4.4}$$

Of course, we can also consider connections from $+1$ to -1 by reversing the role of $\pm\infty$ in the above condition.

The parameter β plays a central role in the analysis of the behaviour of solutions of (4.3). Indeed, the nature of the equilibria ± 1 changes at two critical values $\beta = \pm\sqrt{8}$. The linearization of (4.3) at $u = \pm 1$ reads

$$v'''' - \beta v'' + 2v = 0, \tag{4.5}$$

where v stands respectively for $u-1$ and $u+1$. The eigenvalues of the associated characteristic equation

$$\lambda^4-\beta\lambda^2+2=0$$

are

$$\lambda=\pm\sqrt{\frac{\beta\pm\sqrt{\beta^2-8}}{2}}. \tag{4.6}$$

When $\beta\geqslant\sqrt{8}$, the four eigenvalues are real so that $u=\pm 1$ are *saddle-nodes*. For $\beta\in(-\sqrt{8},\sqrt{8})$, they are all complex with non vanishing real part. The equilibria are then called *saddle-foci*. When β passes below $-\sqrt{8}$, the eigenvalues become purely imaginary and therefore $u=\pm 1$ are *centers*.

The behaviour of the solutions of the linearization of (4.3) around the equilibria provides important informations for the solutions of the nonlinear equation. Indeed, when no eigenvalue has a vanishing real part, it is well known that under some smoothness assumptions, the nonlinear flow and the flow defined by the linearization are conjugate in a neighborhood of the equilibria, see [41]. Consequently, when $\beta\geqslant-\sqrt{8}$, the solutions of (4.3) inherit some properties of the small solutions of (4.5) when they are close to $u=\pm 1$ (with small derivatives up to third order). For example, we easily obtain a qualitative description of the shape of any heteroclinic at $\pm\infty$. Indeed, when ± 1 are saddle-nodes, the solutions of (4.5) that have limit 0 at $+\infty$ or $-\infty$ are monotone while in the saddle-foci case, they oscillate around zero.

We focus on the model equation (4.3) to present different approaches which have been considered in the literature. We assume throughout this section that β is positive. We first look at the shooting method developed by Peletier and Troy. Their results are given without proof as the arguments are not used further in this monograph. We just mention some central ideas showing again the important role of the parameter β.

As previously mentioned, Eq. (4.3) can be written

$$\frac{1}{\beta^2}u''''-u''+u^3-u=0 \tag{4.7}$$

after a suitable rescaling of u. For large β, Eq. (4.7) is a fourth-order perturbation of the Fisher–Kolmogorov equation

$$-u''+u^3-u=0$$

which can be analyzed using elementary methods. It is easily seen that $u^{\pm}(x)=\pm\tanh(\frac{x+a}{\sqrt{2}})$, $a\in\mathbb{R}$, are the only heteroclinics of the Fisher–Kolmogorov equation and these are monotone. We therefore expect that for large β the heteroclinic solutions of (4.3) are monotone. It turns out that a topological shooting method adapted to track monotone heteroclinics works fine for instance for $\beta\geqslant\sqrt{8}$. For this range of β, it is also remarkable that the uniqueness (up to translations and symmetry) of the heteroclinic solution of the Fisher–Kolmogorov equation extends to the fourth-order equation, see Note 1 below and [89].

For $|\beta| < \sqrt{8}$, the oscillatory behaviour of the solutions of (4.3) close to the equilibria makes the shooting method much more tricky. However, two families of heteroclinics can be singled out thanks to a careful analysis.

The remaining of the section is concerned with variational methods and more particularly with minimization. We expose in Section 4.2 a global minimization process. Here we develop the arguments in detail as they will be confronted with a more general framework in Section 5. We conclude our overview of the model equation (4.3) with the nice local minimization method developed by Kalies, Kwapisz and VanderVorst. Their method handles perfectly oscillatory graphs and is therefore efficient when $0 < \beta < \sqrt{8}$. In Section 4.3, we define an homotopy type which allows to consider convenient subsets for local minimization and describe the multi-transition profiles of the local minimizers. We do not present the results in detail but we emphasize that parts of the arguments are developed and used later on in Sections 6 and 7.

4.1. *A shooting method*

One of the basic tools that can be used to analyze the solutions of (4.3) is a topological shooting method developed by Peletier and Troy in [59–62], see also [63]. The main idea of a classical shooting method is to look the way solutions change with respect to initial conditions (taken as parameters) at some fixed initial point. The success of the method for second order ordinary differential equation comes from the number of parameters that is usually reduced to one. Here we reduce the number of parameters by restricting our attention to odd solutions. The reversibility of (4.3) and the oddness of the nonlinear term $u^3 - u$ allows indeed to look at the problem

$$\begin{cases} u''''(x) - \beta u''(x) + u^3(x) - u(x) = 0, \quad x \geqslant 0, \\ u(0) = u''(0) = 0, \\ \lim\limits_{x\to+\infty} \big(u, u', u'', u'''\big)(x) = (+1, 0, 0, 0). \end{cases} \tag{4.8}$$

If u solves (4.8), then the odd extension of u,

$$u^*(x) = \begin{cases} u(x) & \text{for } x \geqslant 0, \\ -u(-x) & \text{for } x < 0, \end{cases}$$

is an odd heteroclinic solution of (4.3) connecting -1 to $+1$. To find a solution of (4.8), we consider the Cauchy problem

$$\begin{cases} u''''(x) - \beta u''(x) + u^3(x) - u(x) = 0, \quad x \geqslant 0, \\ u(0) = u''(0) = 0, \\ u'(0) = \mu, \\ u'''(0) = \nu, \end{cases} \tag{S}$$

where $\mu \in \mathbb{R}^+$ and $\nu \in \mathbb{R}$. According to classical theory of ordinary differential equations, this Cauchy problem defines a unique local solution $u(\mu, \nu, \cdot)$ for every $\mu,\ \nu \in \mathbb{R}$. The shooting method then consists in finding the right initial conditions $\mu,\ \nu \in \mathbb{R}$ such that $u(\mu, \nu, \cdot)$ is globally defined and solves (4.8). A priori, we thus have to handle a two dimensional topological shooting. However, we can take benefit of a first integral to link the parameters μ and ν. In fact it is easily checked that the Hamiltonian

$$H\bigl(u,\, u',\, u'',\, u'''\bigr) = u'''u' - \frac{1}{2}u''^2 - \frac{\beta}{2}u'^2 + \frac{(u^2-1)^2}{4} \tag{4.9}$$

is constant along solutions, i.e. for every solution u of (4.3), there exists a constant E such that for all $x \in \mathbb{R}$, $H(u(x),\, u'(x),\, u''(x),\, u'''(x)) = E$. This first integral is referred to as the energy of the solution (by analogy with second order equations). Observe that any solution of (4.8) belongs to the level of energy $E = 0$. Indeed, this follows from the limit value

$$\lim_{x\to+\infty} H\bigl(u(x),\, u'(x),\, u''(x),\, u'''(x)\bigr) = 0.$$

We then infer that any solution of (4.8) satisfies

$$u'''(0) = \frac{\beta}{2}u'(0) - \frac{1}{4u'(0)}$$

taking also into account that $u'(0) \neq 0$ for such a solution. We therefore impose the conditions $\mu > 0$ and

$$\nu = \nu(\mu) := \frac{\beta}{2}\mu - \frac{1}{4\mu}$$

in the Cauchy problem (S). Summing up, the shooting method amounts to finding $\mu \in \mathbb{R}^+$ such that $u(x) := u(\mu, \nu(\mu), x)$ satisfies

$$\lim_{x\to+\infty} \bigl(u(x), u'(x), u''(x), u'''(x)\bigr) = (+1, 0, 0, 0).$$

For $\beta \geqslant \sqrt{8}$, the method is efficient to obtain a monotone heteroclinic.

THEOREM 4.1. *Let* $\beta \in [\sqrt{8}, +\infty[$. *Then there exists a monotone solution of* (4.8). *Moreover, its odd extension is a heteroclinic solution of* (4.3) *connecting* -1 *to* $+1$.

The proof goes as follows. For $\mu > 0$, we define

$$\xi(\mu) := \sup\bigl\{x > 0 \bigm| u'\bigl(\mu, \nu(\mu), \cdot\bigr) > 0 \text{ on } (0, x)\bigr\}$$

and

$$\mu^* := \sup\bigl\{\hat{\mu} \mid u\bigl(\mu, \nu(\mu), \xi(\mu)\bigr) < 1 \text{ for } \mu \in]0, \hat{\mu}[\bigr\}.$$

It turns out that $\xi(\mu^*) = +\infty$ and the orbit of $u(\mu^*, \nu(\mu^*), \cdot)$ tends to $(+1, 0, 0, 0)$ in the phase-space. We refer to [59,63] for a complete proof.

We think it is worth pointing the role of the assumption $\beta \geqslant \sqrt{8}$ in the preceding theorem. As we already mentioned, the roots of the characteristic equation associated to the linearization of (4.3) around the equilibria are real if $\beta \geqslant \sqrt{8}$. The squares of these roots are

$$\tau_\pm = \frac{\beta \pm \sqrt{\beta^2 - 8}}{2}.$$

Now observe that the linear operator $D^4 - \beta D^2 + 2I$ can be factorized as $(D^2 - \tau_+ I)(D^2 - \tau_- I)$. Letting $v = 1 - u$, Eq. (4.3) written in terms of v yields

$$v'''' - \beta v'' + 2v = v^2(3 - v). \tag{4.10}$$

Therefore, as long as $\beta \geqslant \sqrt{8}$, the solutions (v, w) of the second order system

$$\begin{cases} v'' - \tau_+ v = w, \\ w'' - \tau_- w = v^2(3 - v), \end{cases} \tag{4.11}$$

lead to solutions v of (4.10). This formulation of (4.3) as a system turns out to be very powerful to obtain estimates on v (and therefore on u) that allow to conclude that $\xi(\mu^*) = +\infty$. These estimates rely on a repeated application of the *Strong Maximum Principle* (see [65]) which can be used once we know the sign of the right hand sides in each equation of the system. The Strong Maximum Principle is also the main tool to prove the uniqueness of the heteroclinic of Theorem 4.1, see Note 1.

When $0 < \beta < \sqrt{8}$, the preceding argument involving the Maximum Principle is no more at hand. Anyway, Theorem 4.1 cannot hold in this parameter range. Indeed, as stressed in the introduction of this section, the linearization around the equilibria display oscillatory solutions so that any solution of (4.8) oscillates around $+1$ in its tail, i.e. when $x \to +\infty$. Therefore, the shooting method must now analyze carefully the location and value of the successive local extrema of $u(\mu, \nu(\mu), \cdot)$. One of the greatest difficulties when dealing with oscillatory solution graphs is to prove the convergence at infinity. The following result is the key to overcome this problem.

PROPOSITION 4.2. *Let* $\beta \geqslant 0$ *and assume* u *is a solution of* (4.3) *such that* $H(u, u', u'', u''') = 0$. *Suppose that for some* $a \in \mathbb{R}$ *and* $M > 1$,

$$\frac{1}{\sqrt{3}} \leqslant u(x) \leqslant M \quad \text{for } x > a.$$

Then

$$\lim_{x \to +\infty} \big(u, u', u'', u'''\big)(x) = (1, 0, 0, 0).$$

Basically, the proposition states that the equilibrium attracts all the solutions that stay in a strip around it. The lower bound of the strip has the particularity that the nonlinear term $u^3 - u$ is monotone increasing for all $u \geqslant \frac{1}{\sqrt{3}}$. This makes the functional

$$\mathcal{G}(u) := \frac{1}{2}u''^2 + \frac{1}{4}\big(u^2 - 1\big)^2$$

convex in the strip. This property is extensively used to prove Proposition 4.2, see [60,63].

Using sharp estimates on the critical points of $u(\mu, \nu(\mu), \cdot)$ for which we refer to [60, 63], the following can be proved. We denote respectively by ξ_k and η_k the local maxima and the local minima of the solutions of (4.8).

THEOREM 4.3. *Let β satisfy $0 < \beta < \sqrt{8}$. There exist two solutions u_1, u_2 of (4.8) that have the following properties*:

$$\frac{1}{\sqrt{3}} < u_1(\eta_k) < 1 < u_1(\xi_k) < \sqrt{2} \quad \textit{for } k = 1, 2, \ldots,$$

$$-1 < u_2(\eta_1) < 0, \quad 1 < u_2(\xi_1) < \sqrt{2} \quad \textit{and}$$

$$\frac{1}{\sqrt{3}} < u_2(\eta_k) < 1 < u_2(\xi_k) < \sqrt{2} \quad \textit{for } k = 2, 3, \ldots.$$

The odd extensions of u_1, u_2 are heteroclinic solutions of (4.3).

The solution u_1 is usually called the *principal heteroclinic* as it is monotone increasing in some interval $[-T, T]$ and it oscillates around -1 in $(-\infty, -T)$ and around $+1$ in $(T, +\infty)$. Also this heteroclinic is the first of two families of odd heteroclinic solutions which have $2n + 1$ zeros on $\mathbb{R}$. The two families differ by the amplitude of the oscillations. The first family consists of so-called *multi-transition* solutions as all the successive local extrema between the zeros are outside the region $[-1, 1]$. Thus the profiles of these solutions display $2n + 1$ jumps from -1 to $+1$ and two oscillatory tails around -1 and $+1$. In the second family, the amplitude of the oscillations is smaller than 1 so that the corresponding solutions are *single-transition* heteroclinics.

THEOREM 4.4. *Let β satisfy $0 < \beta < \sqrt{8}$. For each $n \in \mathbb{N}_0$, there exist two solutions u_1, u_2 of* (4.8) *that have the following properties*:

$$\textit{if } n \textit{ is even,} \quad \begin{cases} \textit{for } k \leqslant n/2, & u_1(\eta_k) < -1, \quad u_1(\xi_k) > 1, \\ & -1 < u_2(\eta_k) < 0 < u_2(\xi_k) < 1, \\ \textit{for } k > n/2, & 1/\sqrt{3} < u_i(\eta_k), \quad 1 < u_i(\xi_k), \quad i = 1, 2, \end{cases}$$

$$\textit{if } n \textit{ is odd,} \quad \begin{cases} \textit{for } k \leqslant (n-1)/2, & u_1(\eta_k) < -1, \quad u_1(\xi_k) > 1, \\ & -1 < u_2(\eta_k) < 0 < u_2(\xi_k) < 1, \\ \textit{for } k = (n+1)/2, & u_1(\eta_k) < -1, \quad u_1(\xi_k) > 1, \\ & -1 < u_2(\eta_k) < 0, \quad u_2(\xi_k) > 1, \\ \textit{for } k > (n+1)/2 & 1/\sqrt{3} < u_i(\eta_k), \quad u_i(\xi_k) > 1, \quad i = 1, 2. \end{cases}$$

Moreover, for all $n \geqslant 1$, $\|u_i\|_\infty < \sqrt{2}$ *and* $u_i'(0) \neq 0$ *for* $i = 1, 2$.

We refer to [60,63] for the proof.

4.2. *Minimization*

Heteroclinic solutions of (4.3) are critical points of the action functional

$$\mathcal{F}_\beta(u) = \int_{\mathbb{R}} \left(\frac{1}{2}\bigl(u''^2 + \beta u'^2\bigr) + \frac{1}{4}\bigl(u^2 - 1\bigr)^2 \right) \mathrm{d}x. \tag{4.12}$$

This functional is well defined for functions u having first and second square integrable derivatives and being such that the potential is integrable. Taking into account conditions (4.4) which are satisfied by heteroclinics connecting -1 to $+1$, we can define $\mathcal{F}_\beta$ in the space

$$\bigl\{u : \mathbb{R} \to \mathbb{R} \mid u + 1 \in H^2\bigl(\mathbb{R}^-\bigr),\ u - 1 \in H^2\bigl(\mathbb{R}^+\bigr)\bigr\}.$$

Indeed, if $u + 1 \in H^1(\mathbb{R}^-)$, then $\lim_{x \to -\infty} u(x) = -1$ see for example [22] and u is bounded in every compact interval $[-T, 0]$. A similar observation can be made on $\mathbb{R}^+$ so that we infer

$$\begin{aligned}\int_{\mathbb{R}} \frac{1}{4}\bigl(u^2 - 1\bigr)^2 \mathrm{d}x &= \int_{\mathbb{R}^-} \frac{1}{4}\bigl(u^2 - 1\bigr)^2 \mathrm{d}x + \int_{\mathbb{R}^+} \frac{1}{4}\bigl(u^2 - 1\bigr)^2 \mathrm{d}x \\ &\leqslant \frac{C}{4}\bigl(\|u + 1\|^2_{L^2(\mathbb{R}^-)} + \|u - 1\|^2_{L^2(\mathbb{R}^+)}\bigr)\end{aligned}$$

for some positive constant $C = C(\|u\|_{L^\infty(\mathbb{R})})$ and therefore $\mathcal{F}_\beta(u)$ is finite.

For $\beta \geqslant 0$, $\mathcal{F}_\beta$ is positive and hence bounded from below. The a priori simplest way to find critical points of $\mathcal{F}_\beta$ is therefore to search for minimizers. The next theorem confirms the efficiency of a minimization approach. To avoid a loss of compactness due to the invariance under translations, we look for a minimizer in the space

$$\mathcal{E} := \bigl\{u : \mathbb{R} \to \mathbb{R} \mid u(0) = 0,\ u + 1 \in H^2\bigl(\mathbb{R}^-\bigr),\ u - 1 \in H^2\bigl(\mathbb{R}^+\bigr)\bigr\}. \tag{4.13}$$

Of course, this is not restrictive.

THEOREM 4.5. *For all* $\beta \geqslant 0$, *the functional* $\mathcal{F}_\beta$ *has a global minimizer* $\varphi_\beta \in \mathcal{E}$ *which is a heteroclinic solution of* (4.3). *Furthermore, any minimizer is odd and positive in* $]0, +\infty[$.

This theorem is the analogous of Theorem 2.1. Remember that for the functional $\mathcal{J}$ of Theorem 2.1, minimizing sequences $(u_n)_n$ can be taken in such a way that $-1 \leqslant u_n \leqslant 1$. Indeed, $v_n = \sup(-1, \inf(u_n, 1))$ is still a minimizing sequence. Using similar modification arguments it is easily seen that the minimizers of $\mathcal{J}$ are monotone.

One of the first differences that we encounter when considering the functional $\mathcal{F}_\beta$ in $\mathcal{E}$, comes from the fact that these modifications which keep functions in $H^1_{\mathrm{loc}}(\mathbb{R})$ do not necessarily produce functions of class C^1 and therefore, in general, the modified functions do not belong to $H^2_{\mathrm{loc}}(\mathbb{R})$. By the way, as already mentioned, it is not true in general that minimizers of $\mathcal{F}_\beta$ in $\mathcal{E}$ are monotone. When dealing with $\mathcal{J}$, modification arguments are also used to ensure that quasi-minimizers stay close to ± 1 outside a fixed compact interval which can always be centered around zero by translation invariance.

To substitute these rather simple arguments, we take benefit of the symmetry of the functional $\mathcal{F}_\beta$. The original proof of Theorem 4.5 is due to Peletier, Troy and VanderVorst [58]. We give here a proof which is closer to Kalies and VanderVorst [43].

PROOF OF THEOREM 4.5. We introduce the spaces

$$\mathcal{E}^+ := \{u : \mathbb{R}^+ \to \mathbb{R} \mid u(0) = 0,\ u - 1 \in H^2(\mathbb{R}^+)\},$$
$$\mathcal{E}^- := \{u : \mathbb{R}^- \to \mathbb{R} \mid u(0) = 0,\ u + 1 \in H^2(\mathbb{R}^-)\}$$

and define the restricted functional $\mathcal{F}^\pm_\beta : \mathcal{E}^\pm \to \mathbb{R}$ by

$$\mathcal{F}^\pm_\beta(u) := \int_{\mathbb{R}^\pm} L(u, u', u'')\,\mathrm{d}x,$$

where $L(u, u', u'')$ is the Lagrangian defined by

$$L(u, u', u'') := \frac{1}{2}(u''^2 + \beta u'^2) + \frac{1}{4}(u^2 - 1)^2. \tag{4.14}$$

Let us consider the values

$$c := \inf_{\mathcal{E}} \mathcal{F}_\beta \quad \text{and} \quad c^\pm := \inf_{\mathcal{E}^\pm} \mathcal{F}^\pm_\beta.$$

Since $\mathcal{F}_\beta$ is symmetric, it is easily seen that for all $u^+ \in \mathcal{E}^+$, $\mathcal{F}^+_\beta(u^+) = \mathcal{F}^-_\beta(u^-)$ where $u^- \in \mathcal{E}^-$ is defined by $u^-(x) = -u^+(-x)$. Therefore, $c^+ = c^- = c/2$.

Step 1. The variational problem

$$\inf\{\mathcal{F}^+_\beta(u) \mid u \in \mathcal{E}^+\}$$

has a positive solution.

Let $(v_n)_n \subset \mathcal{E}^+$ be a minimizing sequence, i.e. $\mathcal{F}^+_\beta(v_n) \to c^+$. For each $n \geqslant 0$, we define

$$x_n := \sup\{x \geqslant 0 \mid v_n(x) = 0\}.$$

Since $\lim_{x\to+\infty} v_n(x) = 1$, $x_n < +\infty$ for all $n \geqslant 0$. We now consider the positive sequence $(v_n^+)_n \in \mathcal{E}^+$ where

$$v_n^+(x) := v_n(x + x_n)$$

for $x \geqslant 0$. We observe that

$$\int_0^{x_n} L\big(v_n, v_n', v_n''\big)\,\mathrm{d}x \geqslant 0$$

so that $\mathcal{F}_\beta^+(v_n^+) \leqslant \mathcal{F}_\beta^+(v_n)$ which implies that $(v_n^+)_n$ is also a minimizing sequence for $\mathcal{F}_\beta^+$.

As the sequence $\mathcal{F}_\beta^+(v_n^+)$ is uniformly bounded, we deduce a uniform estimate for $\|v_n^+ - 1\|_{H^2(\mathbb{R}^+)}$. Indeed, the L^2-bounds for $(v_n^+)'$ and $(v_n^+)''$ follow easily from the bound on $\mathcal{F}_\beta^+(v_n^+)$ while we infer from the positivity of v_n^+ that

$$\int_0^{+\infty} \frac{(v_n^+ - 1)^2}{4}\,\mathrm{d}x \leqslant \int_0^{+\infty} \frac{(v_n^{+2} - 1)^2}{4}\,\mathrm{d}x \leqslant \mathcal{F}_\beta^+\big(v_n^+\big).$$

We now deduce that going to a subsequence if necessary, there exists $v^+ \in H^2(\mathbb{R}^+) + 1$ such that

$$v_n^+ - 1 \overset{H^2(\mathbb{R}^+)}{\rightharpoonup} v^+ - 1$$

and

$$v_n^+ \xrightarrow{C^1_{\mathrm{loc}}(\mathbb{R}^+)} v^+.$$

As the two first terms in $\mathcal{F}_\beta^+$ are the square of seminorms and Fatou's Lemma is applicable to the last one, it follows that

$$\mathcal{F}_\beta^+\big(v^+\big) \leqslant \liminf_{n\to+\infty} \mathcal{F}_\beta^+\big(v_n^+\big) = c^+.$$

As the convergence is uniform on compact intervals, we conclude that $v^+(0) = 0$ so that $v^+ \in \mathcal{E}^+$ and $\mathcal{F}_\beta^+(v^+) = c^+$. Observe that v^+ is positive on $]0, +\infty[$ otherwise we could proceed as above to construct a positive function having smaller action.

Step 2. If $v \in \mathcal{E}^+$ is such that $(\mathcal{F}_\beta^+)'(v) = 0$, then $v''(0) = 0$, $v^* \in \mathcal{E}$ defined by

$$v^*(x) := \begin{cases} v(x) & \text{if } x \geqslant 0, \\ -v(-x) & \text{if } x < 0, \end{cases}$$

is a minimizer of $\mathcal{F}_\beta$ in $\mathcal{E}$ and v^* is a heteroclinic solution of (4.3).

We first compute

$$\big(\mathcal{F}_\beta^+\big)'(v)(h) = \int_0^{+\infty} \big(v''h'' + \beta v'h' + \big(v^3 - v\big)h\big)\,\mathrm{d}x \tag{4.15}$$

for all $h \in H^2(\mathbb{R}^+) \cap H_0^1(\mathbb{R}^+)$. Starting with $h \in C_c^2(\mathbb{R}^+)$ and using Du Bois–Reymond Lemma, we easily find that in fact $v \in C^4(\mathbb{R}^+)$ and v solves (4.3) for $x \geqslant 0$. As Eq. (4.3) expresses v'''' in terms of v'' and $v-1$, we deduce that $v'''' \in L^2(\mathbb{R}^+)$. Indeed, $v'' \in L^2(\mathbb{R}^+)$ and

$$\int_0^{+\infty} \big(v^3(x) - v(x)\big)^2 \,\mathrm{d}x \leqslant \big(\|v\|_\infty^2 + \|v\|_\infty\big)\|v-1\|_{L^2}^2 < \infty.$$

It now follows by interpolation that

$$\|v'''\|_{L^2}^2 \leqslant C\big(\|v''\|_{L^2} + \|v''''\|_{L^2}\big) < \infty.$$

Hence $v - 1 \in H^4(\mathbb{R}^+)$. Observe also that the L^2-integrability of the derivatives implies that

$$\lim_{x\to+\infty} v(x) = +1 \quad \text{and} \quad \lim_{x\to+\infty} v^{(n)}(x) = 0 \quad \text{for } n = 1, 2, 3. \tag{4.16}$$

Integrating (4.15) by parts, we obtain for all $h \in H^2(\mathbb{R}^+) \cap H_0^1(\mathbb{R}^+)$,

$$\big[v''h'\big]_0^{+\infty} - \big[v'''h\big]_0^{+\infty} - \beta\big[v'h\big]_0^{+\infty} = 0$$

as

$$\int_0^{+\infty} \big(v'''' - \beta v'' + v^3 - v\big)(x)h(x)\,\mathrm{d}x = 0.$$

Taking (4.16) into account, we now deduce that

$$v''(0)h'(0) = 0$$

for all $h \in H^2(\mathbb{R}^+) \cap H_0^1(\mathbb{R}^+)$ which of course implies that $v''(0) = 0$. The function $v^* : \mathbb{R} \to \mathbb{R}$ defined by

$$v^*(x) = \begin{cases} v(x) & \text{if } x \geqslant 0, \\ -v(-x) & \text{if } x < 0, \end{cases}$$

is therefore of class C^4 and straightforward arguments allow to conclude that v^* is a critical point of $\mathcal{F}_\beta$ which is a heteroclinic solution of (4.3), actually of class C^∞. Also, $\mathcal{F}_\beta(v^*) = 2\mathcal{F}_\beta^+(v) = 2c^+ = c$ so that v^* is a minimizer of $\mathcal{F}_\beta$ in $\mathcal{E}$.

Step 3. If $u \in \mathcal{E}$ minimizes $\mathcal{F}_\beta$, then u is odd.

Let us write $u^\pm = u_{|\mathbb{R}^\pm}$. As $\mathcal{F}_\beta(u) = c$, we have $\mathcal{F}_\beta^+(u^+) = \mathcal{F}_\beta^-(u^-) = c/2$ otherwise the odd extension of u^+ or u^- would have a lower action than c. From Step 2, we know that $u(0) = u''(0) = 0$. Hence $v^+(x) := -u^-(-x)$ satisfies $\mathcal{F}_\beta^+(v^+) = c^+$ and $v^+(0) = (v^+)''(0) = 0$. Thus both u^+ and v^+ are minimizers of $\mathcal{F}_\beta^+$ in $\mathcal{E}^+$. As

$(v^+)'(0) = (u^-)'(0) = (u^+)'(0)$ and $(v^+)'''(0) = (u^-)'''(0) = (u^+)'''(0)$, the functions v^+ and u^+ solve the same Cauchy problem

$$\begin{aligned}
&u''''(x) - \beta u''(x) + u^3(x) - u(x) = 0, \quad x \geqslant 0,\\
&u(0) = 0, \quad u'(0) = \big(u^+\big)'(0),\\
&u''(0) = 0, \quad u'''(0) = \big(u^+\big)'''(0).
\end{aligned}$$

By uniqueness, this implies $u^+(x) = v^+(x)$ for all $x \in \mathbb{R}^+$, i.e. $u^+(x) = -u^-(-x)$ for all $x \in \mathbb{R}^+$. □

4.3. *Homotopy classes of multi-transition solutions*

When $\beta < \sqrt{8}$, the equilibrium solution ± 1 of (4.3) are saddle-foci. It is known that heteroclinic connections between such kind of equilibria can exhibit a complex structure, see [23, 31,60]. In [43], Kalies and VanderVorst construct for (4.3) so-called *multi-bump* solutions i.e solutions with multiple oscillations separated by large distances. The usual methods used to obtain such solutions are rather tricky and require a careful study of the stable and unstable manifolds or a certain kind of nondegeneracy condition on a primary connection whose well separated copies are then glued together, see [23,28,79]. In [42], Kalies, Kwapisz and VanderVorst introduce a direct method to find multi-transition solutions. As we already mentioned, the term multi-transition refers to the fact that the graph of such solutions consists of multiple jumps from one equilibrium to the other. One jump is then called a transition. Also such solutions are qualitatively different from multi-bump solutions as the distance between two successive transition is not necessarily large. The method of Kalies et al. consists in minimizing the action functional (4.12) in specific subspaces of functions having the desired number of transitions.

When projected in the configuration plane (u, u'), a heteroclinic orbit yields a curve connecting the points $(\pm 1, 0)$. Choosing two oriented loops e_1 and e_2 around respectively $(-1, 0)$ and $(+1, 0)$ (see Fig. 2), an homotopy type can be associated to any curve connecting $(\pm 1, 0)$ (see the precise definition below). This homotopy type records the number of transitions between these points and the oscillations around them between the transitions. Hence, every orbit connecting $(-1, 0)$ to $(+1, 0)$ can be represented by a word of the form

$$e_1^{\theta_{2m}} \cdot e_2^{\theta_{2m-1}} \cdot \cdots \cdot e_1^{\theta_2} \cdot e_2^{\theta_1},$$

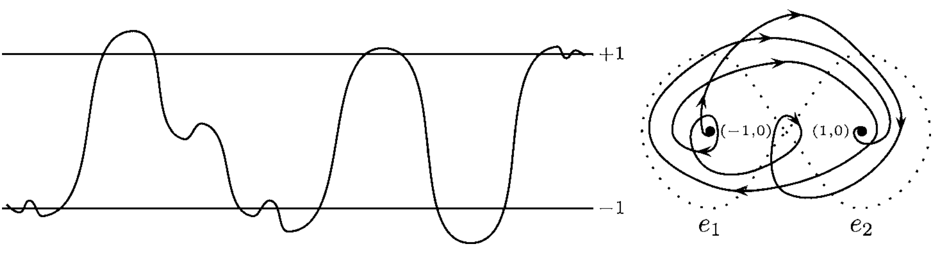


Fig. 2. An orbit with homotopy type $e_1 e_2 e_1^2 e_2$.

where $\theta(u) = (\theta_1, \dots, \theta_{2m}) \in \mathbb{N}^{2m}$. The vector $\theta(u)$ is called the *winding* vector associated to u. Observe that the word starts with the first visit at $(+1, 0)$ and stops recording after the last visit around $(-1, 0)$. Also the vector $g(u) = 2\theta(u)$ contains exactly the number of crossings that u makes with ± 1 between the transitions. The vector $g = 0$ is therefore associated to functions with a single transition. For each vector $g \in 2\mathbb{N}^{2m} \cup \{0\}$, the homotopy class $M(g)$ consisting of functions of $\mathcal{E}$ having winding vector $g/2$. The following more precise definition is introduced in [42].

DEFINITION 4.6. A function $u \in \mathcal{E}$ is in $M(g)$ if there are nonempty sets $\{A_i\}_{i=0}^{2m+1}$ such that

(i) $u^{-1}(\pm 1) = \bigcup_{i=0}^{2m+1} A_i$,
(ii) $\#A_i = g_i$ for $i = 1, \dots, 2m$,
(iii) $\max A_i < \min A_{i+1}$ for $i = 0, \dots, 2m$,
(iv) $u(A_i) = (-1)^{i+1}$,
(v) $\{\max A_0\} \cup (\bigcup_{i=1}^{2m} A_i) \cup \{\min A_{2m+1}\}$ consists of transverse crossings of ± 1.

Under these conditions $M(g)$ is an open subset of $\mathcal{E}$. As the functional $\mathcal{F}_\beta$ is positive, we can define

$$\inf_{u \in M(g)} \mathcal{F}_\beta(u). \tag{4.17}$$

If the infimum is achieved by a function in the interior of $M(g)$, then a local minimizer of $\mathcal{F}_\beta$ with the corresponding properties solves the Euler–Lagrange equation (4.3). The next theorem is a particular case of Theorem 1.3 in [42].

THEOREM 4.7. *Let* $0 < \beta < \sqrt{8}$. *Then for all* $g \in 2\mathbb{N}^{2m} \cup \{0\}$, *the functional* $\mathcal{F}_\beta$ *defined by* (4.12) *has a local minimizer* u_g *in the homotopy class* $M(g)$. *Moreover, the function* u_g *is a heteroclinic solution of* (4.3) *displaying* $2m + 1$ *transitions between* ± 1.

The proof of Theorem 4.7 is carried out in [42]. The main difficulty is to show that minimizing sequences in $M(g)$ have weak limits in the interior of $M(g)$. Indeed, minimizing sequences can approach the boundary of $M(g)$ so that the limit function could gain or lose complexity. For example, tangential crossings of ± 1 could appear due to a coalescence of two or more crossings or to a new spurious oscillation around one equilibrium. The oscillatory behaviour of solutions in a neighborhood of a saddle-focus equilibrium plays then a crucial role to control the minimizing sequences. Efficient tools were developed by Kalies et al. to adjust functions of the boundary of $M(g)$. Basically these tools rely on a cut and paste technique which allows to delete spurious oscillations or replace pieces of functions tangent to ± 1 with pieces of small oscillating orbits around $(\pm 1, 0, 0, 0)$ in the phase-space.

Another source of troubles in the minimization process comes from a possible lack of compactness when passing to a weak limit. Indeed the distance between two transitions or two crossings of either -1 or $+1$ could grow to infinity. Here again the oscillatory properties of the orbits close to a saddle-focus prevent from these losses of complexity in the limit.

4.4. *Notes and further comments*

1. As already mentioned, in case $\beta \in [\sqrt{8}, +\infty[$, the heteroclinic solution obtained in Theorem 4.1 is unique (up to translation and symmetry). The first result in this direction is due to Peletier and Troy [59]. They prove that the heteroclinic obtained via the shooting method is unique in the class of monotone antisymmetric functions and they conjecture that it is actually unique in the class of all functions. Kwapisz [47] and van den Berg [89] have then confirmed the conjecture. The proof of Kwapisz relies on the use of a twist-map while the arguments of van den Berg are based on the analysis of the phase-space and more precisely of its projection into the configuration plane (u, u'). The analysis of van den Berg applies to the model equation

$$u'''' - \beta u'' + f'(u) = 0, \tag{4.18}$$

where f is a double-well potential of class C^2. Considering the set of bounded functions

$$\mathcal{B}(a,b) := \big\{u \in C^4(\mathbb{R}) \bigm| u(x) \in [a,b] \text{ for all } x \in \mathbb{R}\big\},$$

and the value

$$\omega(a,b) = \max\Big\{0, \max_{u\in[a,b]} -f''(u)\Big\},$$

his key result states that when $\beta \geqslant 2\sqrt{\omega(a,b)}$, bounded solutions of (4.18) in $\mathcal{B}(a,b)$ do not cross in the configuration plane (u,u'). Observing then that for $\beta \geqslant \sqrt{8}$, any bounded solution u of (4.3) satisfies $\|u\|_\infty \leqslant 1$, he deduces that the uniqueness property in the configuration plane holds for the parameter range $\beta \geqslant \sqrt{8}$. The a priori bound is obtained by applying the maximum principle twice to the factorization (4.11). The uniqueness of the heteroclinic solution then follows from the energy ordering of the bounded solutions in the configuration plane (u,u'). As stressed in Section 4.1, the Hamiltonian

$$H\big(u,u',u'',u'''\big) = u'''u' - \frac{1}{2}u''^2 - \frac{\beta}{2}u'^2 + f(u)$$

is constant along solutions. It turns out that for $\beta \geqslant 2\sqrt{\omega(a,b)}$ and solutions of (4.18) in $\mathcal{B}(a,b)$, the energy $E = H(u,\, u',\, u'',\, u''')$ is a parameter that orders the solutions in the configuration plane.

It is worth mentioning that the results of [89] also imply that the unique heteroclinic of Theorem 4.1 is asymptotically stable in the space of bounded uniformly continuous functions.

2. The global minimization approach of Section 4.2 can be extended to the functional

$$\mathcal{J}(u) := \int_{-\infty}^{\infty} \Big(\frac{1}{2}\big(u''^2 + g(u)u'^2\big) + f(u)\Big)\,\mathrm{d}x,$$

where $g:\mathbb{R}\to\mathbb{R}$ is a positive function and $f:\mathbb{R}\to\mathbb{R}$ is a double well potential with nondegenerate minima at the same level of energy, see Section 5. If f and g are even then the equivalent of Theorem 4.5 holds. When this is not the case, the conclusion that any minimizer is odd is false and we cannot ensure that minimizers are still single-transition solutions.

3. The results of Section 4.3 hold in a more general framework. The potential $\frac{(u^2-1)^2}{4}$ may be replaced by a function $f\in C^2(\mathbb{R})$ that has exactly two nondegenerate global minima at $u=\pm 1$ and grows superquadratically at $\pm\infty$. If the parameter β satisfies $\beta^2<4f''(\pm 1)$ then $u=\pm 1$ are saddle-focus equilibria for the equation

$$u'''' - \beta u'' + f'(u) = 0.$$

If f is even then the equivalent of Theorem 4.7 holds, while without this symmetry assumption it is not clear that the infima are attained in the interior of each homotopy class. However, when $g_i=2$ for all i or if the g_i's are large enough (i.e. the profile is then similar to multi-bump solutions) then local minimizers exist in the corresponding class $M(g)$, see [42].

4. Not much is known about the role of the heteroclinic solutions for the evolution problem

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{\partial^4 u}{\partial x^4} - \beta\dfrac{\partial^2 u}{\partial x^2} + u^3 - u = 0, & 0<x<L,\ t>0,\\ u(t,0) = u(t,L) = 0, & t>0,\\ \dfrac{\partial^2 u}{\partial x^2}(t,0) = \dfrac{\partial^2 u}{\partial x^2}(t,L) = 0, & t>0,\\ u(0,x) = u_0(x), & 0<x<L. \end{cases}$$

When $\beta\geqslant\sqrt{8}$, the stability of the heteroclinic has been proved by van den Berg [89].

Another class of solutions, namely travelling wave solutions, is important for the dynamics of the evolution equation. These solutions are waves that evolve at constant speed. Thus they can be written $u(t,x)=U(x-ct)$ for some constant c and solve the equation

$$U'''' - \beta U'' - cU' + f'(U) = 0.$$

For this last equation, the heteroclinic solutions connecting -1 to $+1$ or 0 to -1 or $+1$ are of particular interest. The problem of existence of such solutions is not completely solved. Partial results exist for similar models with an asymmetric nonlinear term like $f_a'(u)=(u+a)(u^2-1)$, $a\neq 0$. The case $a=0$ can also be treated when dealing with heteroclinics connecting 0 to -1 or $+1$. For further details, we refer to [75,88,90].

5. It was observed in [58] that the arguments used in the proof of Theorem 4.5 can be carried over to treat sixth order bistable equations of the form

$$u^{(6)} + Au^{(4)} + Bu^{(2)} + u - u^3 = 0$$

provided that $A^2 < 4B$. The associated functional is then coercive and weakly lower-semicontinuous.

5. Sign changing Lagrangians

The success of the arguments of Section 4.2 depends on the positivity and the symmetry of the Lagrangian

$$L(u, u', u'') = \frac{1}{2}\big(u''^2 + \beta u'^2\big) + \frac{1}{4}\big(u^2 - 1\big)^2.$$

The symmetry of the potential allows to consider a restricted functional on $\mathbb{R}^+$ which makes the minimization process very simple. However it is easily seen that a more tedious proof works fine without exploiting the symmetry of the functional. On the other hand, the positivity of the Lagrangian seems much more crucial. Indeed, it gives a simple lower bound on the functional. Another great advantage of the Lagrangian L is that the potential $f(u) = \frac{1}{4}(u^2 - 1)^2$ is superquadratic at $\pm\infty$ and has nondegenerate minima i.e. $f''(\pm 1) \neq 0$. This last property implies that for u close to $+1$ (respectively -1), $f(u)$ behaves like the square of the L^2-norm of $u - 1$ (respectively $u + 1$).

In this section, we consider a larger class of Lagrangians

$$L_g\big(u, u', u''\big) := \frac{1}{2}\big(u''^2 + g(u)u'^2\big) + f(u). \tag{5.1}$$

Here the function g is not assumed to be positive and f is a positive two-well potential with bottoms at ± 1. We consider in a convenient space (which is made precise below) the functional

$$\mathcal{F}(u) := \int_{-\infty}^{+\infty} \left(\frac{1}{2}\big(u''^2 + g(u)u'^2\big) + f(u)\right) \mathrm{d}x \tag{5.2}$$

whose Euler–Lagrange equation is given by

$$u'''' - g(u)u'' - \frac{1}{2}g'(u)u'^2 + f'(u) = 0. \tag{5.3}$$

The main idea is then to impose a condition on g to ensure a lower bound on the action $\mathcal{F}(u)$.

The results of Section 5.1 are based on [20,37]. The main assumption is as follows:

(C1) there exist a function $\tilde{g} \in C(\mathbb{R})$ and some $k < 1$ such that for all $u \in \mathbb{R}$,

$$g(u) \geqslant \tilde{g}(u) \quad \text{and } \big|\tilde{G}(u)\big| \leqslant k\sqrt{8f(u)},$$

where

$$\tilde{G}(u) := \int_0^u \tilde{g}(s)\,\mathrm{d}s.$$

This condition can be seen as a good balance between the possible negativity of g and the positivity of the potential. Though assumption (C1) allows g to take negative values, it rules out a negative constant. Indeed, $\tilde{g}$ must satisfy

$$\int_{-1}^{0} \tilde{g}(s)\,\mathrm{d}s = \int_{0}^{+1} \tilde{g}(s)\,\mathrm{d}s = 0.$$

Some regularity assumptions have to be made on f and g but it should be stressed that the approach does not require the nondegeneracy of the equilibria ±1 as implicitly assumed when considering the model potential in $\mathcal{F}_\beta$. This rather weak assumption on the potential does not allow to define the functional $\mathcal{F}$ in an affine translate of $H^2(\mathbb{R})$ so that a slight modification of the functional settings has to be performed.

Section 5.2 reviews the results of [19]. Here more assumptions are in order but on the other hand g can be negative everywhere. Mainly, we only deal with symmetric functionals with an even potential having nondegenerate minima. Also, we do assume that the functional is bounded from below. We prove that this is the case when g^- is small.

5.1. *Functionals with sign changing acceleration coefficient*

We consider a two-well potential $f \in C^1(\mathbb{R})$ such that

(C2) for some $0 < a < 1$ and $\alpha > 0$,

$$\frac{f(u)}{(u-1)^2} \leqslant \alpha, \quad \text{for } |u-1| < a,$$
$$\frac{f(u)}{(u+1)^2} \leqslant \alpha, \quad \text{for } |u+1| < a;$$

(C3) $f(u) = 0$ if and only if $u = \pm 1$;

(C4) $f(u) \geqslant 0$ for all $u \in \mathbb{R}$ and

$$\liminf_{|u|\to\infty} f(u) > 0.$$

We also assume that $g \in C^1(\mathbb{R})$ satisfies (C1).

Notice that the assumption (C2) is automatically satisfied when ±1 are nondegenerate minima. In this case, if f is of class C^2 in a neighborhood of ±1, $g(\pm 1) < 0$ or (C1) holds with $g = \tilde{g}$ then ±1 are saddle-foci. Indeed, by l'Hospital's rule, we obtain

$$\frac{\tilde{g}^2(\pm 1)}{4f''(\pm 1)} = \lim_{u\to\pm 1} \frac{\tilde{G}^2(u)}{8f(u)} < 1,$$

so that

$$g(\pm 1)^2 \leqslant \tilde{g}(\pm 1)^2 < 4f''(\pm 1).$$

In order to find a solution u of (5.3) that satisfies

$$\lim_{x\to\pm\infty} u(x) = \pm 1,$$

we minimize the functional $\mathcal{F}$ in a convenient functional space. As the assumptions on f are quite weak and g can vanish close to ± 1, a function u that satisfies $\mathcal{F}(u) < \infty$ does not necessarily belong to an affine translate of $H^2(\mathbb{R})$ as it was the case for heteroclinic solutions of (4.3). In fact it is sufficient to search a minimizer of $\mathcal{F}$ in the set

$$\mathcal{H} := \Big\{ u \in C^1(\mathbb{R}),\ u'' \in L^2(\mathbb{R}),\ u' \in L^\infty(\mathbb{R}) \text{ and } \lim_{x\to\pm\infty} u(x) = \pm 1 \Big\}. \tag{5.4}$$

We first observe that under assumption (C1), the functional $\mathcal{F}$ is bounded from below in $\mathcal{H}$. This justifies our minimization procedure.

LEMMA 5.1. *If $f, g \in C(\mathbb{R})$ satisfy* (C1), *then there exists a constant $s > 0$ such that for all $u \in \mathcal{H}$*

$$\mathcal{F}(u) \geqslant s \int_{-\infty}^{+\infty} \left(\frac{u''^2}{2} + f(u) \right) \mathrm{d}x.$$

PROOF. Let k be given by the assumption (C1). For $c \in\,]k, 1[$, we compute

$$\begin{aligned} \mathcal{F}(u) &\geqslant \int_{-\infty}^{+\infty} \left(\frac{1}{2}\big(u''^2 + \tilde{g}(u)u'^2\big) + f(u) \right) \mathrm{d}x \\ &\geqslant \int_{-\infty}^{+\infty} \left(\frac{1}{2}\big(1 - c^2\big)u''^2 + \frac{1}{2}\left(cu'' - \frac{\tilde{G}(u)}{2c}\right)^2 + \left(f(u) - \frac{\tilde{G}(u)^2}{8c^2} \right) \right) \mathrm{d}x, \end{aligned}$$

where we have performed an integration by parts and used the fact that u' is bounded and $\tilde{G}(u(\pm\infty)) = \tilde{G}(\pm 1) = 0$ to obtain

$$-\int_{-\infty}^{+\infty} \tilde{G}(u)u''\,\mathrm{d}x = \int_{-\infty}^{+\infty} \tilde{g}(u)u'^2\,\mathrm{d}x.$$

Hence by our assumption, we obtain the inequality

$$\mathcal{F}(u) \geqslant \int_{-\infty}^{+\infty} \left(\frac{1}{2}\big(1 - c^2\big)u''^2 + \left(1 - \frac{k^2}{c^2}\right) f(u) \right) \mathrm{d}x$$

so that the conclusion follows. □

We now state the main theorem of the section.

THEOREM 5.2. *Suppose that $f, g \in C^1(\mathbb{R})$ satisfy* (C1)–(C4). *Then, there exists a minimizer u of $\mathcal{F}$ in $\mathcal{H}$ which is a solution of* (5.3).

To prove this theorem, we need sharp estimates on minimizing sequences. The key arguments are summarized in the next proposition.

PROPOSITION 5.3. *Suppose that* $f, g \in C^1(\mathbb{R})$ *satisfy* (C1)–(C4). *Then there exist* $L > 0$, $T > 0$ *and a sequence* $(u_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ *such that* $\mathcal{F}(u_n) \to \inf_{\mathcal{H}} \mathcal{F}$ *and for all* $n \in \mathbb{N}$,

(i) $\|u_n\|_{C^1} \leqslant L$,

(ii) $|u_n(x) + 1| \leqslant a$ *for all* $x \leqslant -T$ *and* $|u_n(x) - 1| \leqslant a$ *for all* $x \geqslant T$.

Observe that the first property of the minimizing sequence $(u_n)_n$ allows to choose a converging subsequence in C^1_{loc}. The second property then prevents from the lack of compactness of the functional space $\mathcal{H}$ due to the invariance of $\mathcal{F}$ under translations. The proof of Theorem 5.2 easily follows from Proposition 5.3.

PROOF OF THEOREM 5.2. Let $(u_n)_n \subset \mathcal{H}$ be a minimizing sequence satisfying the conclusions of Proposition 5.3.

Step 1. *Convergence.* According to Lemma 5.1, the sequence u_n'' is uniformly bounded in L^2. Together with the uniform bound on u_n in C^1 (property (i) of the minimizing sequence), this implies that $(u_n)_n$ has a subsequence (still written $(u_n)_n$ for simplicity) such that for some function u

$$u_n \xrightarrow{C^1_{\text{loc}}(\mathbb{R})} u, \quad u_n'' \overset{L^2(\mathbb{R})}{\rightharpoonup} u''.$$

We now infer from Lemma 5.1 and Fatou's Lemma that

$$\int_0^{+\infty} f\big(u(x)\big)\,\mathrm{d}x < +\infty. \tag{5.5}$$

Step 2. $u \in \mathcal{H}$. Observe first that $u' \in L^\infty(\mathbb{R})$. Indeed, as u_n' converges uniformly to u' on every compact subset of $\mathbb{R}$, it follows from the uniform bound on $\|u_n'\|_\infty$ that u' is bounded.

We next prove that

$$\lim_{x\to+\infty} u(x) = 1.$$

From the uniform convergence on compact sets, it is clear that $|u(x) - 1| \leqslant a$ for $x \geqslant T$. We therefore deduce that

$$1 - a \leqslant \liminf_{x\to+\infty} u(x) \leqslant \limsup_{x\to+\infty} u(x) \leqslant 1 + a.$$

Assume by contradiction that $\limsup_{x\to+\infty} u(x) > 1$. This means that for some $0 < \varepsilon < a$, there exist infinitely many disjoint intervals $[a_i, b_i] \subset \mathbb{R}^+$, $i \in \mathbb{N}$ such that

$$1 + \frac{\varepsilon}{2} \leqslant u(x) \leqslant 1 + \varepsilon \quad \text{for all } x \in [a_i, b_i].$$

We can suppose without loss of generality that $u(b_i) - u(a_i) = \pm\varepsilon/2$. As $\|u'\|_\infty \leqslant L$, we infer that

$$\frac{\varepsilon}{2} = \left|u(b_i) - u(a_i)\right| = \left|\int_{a_i}^{b_i} u'(s)\,\mathrm{d}s\right| \leqslant L(b_i - a_i)$$

which implies that

$$(b_i - a_i) \geqslant \frac{\varepsilon}{2L}.$$

Now, let $m_\varepsilon > 0$ be such that $f(u) \geqslant m_\varepsilon$ for all $u \in [1 + \frac{\varepsilon}{2}, 1 + \varepsilon]$. We then compute

$$\int_0^{+\infty} f\big(u(x)\big)\,\mathrm{d}x \geqslant \sum_{i=0}^{\infty} \int_{a_i}^{b_i} f\big(u(x)\big)\,\mathrm{d}x \geqslant \sum_{i=0}^{\infty} \frac{\varepsilon\, m_\varepsilon}{2L} = +\infty$$

which contradicts (5.5). If $\liminf_{x\to+\infty} u(x) < 1$, we derive the same contradiction so that

$$\liminf_{x\to+\infty} u(x) = \limsup_{x\to+\infty} u(x) = 1.$$

As we can argue similarly if $\limsup_{x\to-\infty} u(x) > -1$ or $\liminf_{x\to-\infty} u(x) < -1$, we also infer that

$$\lim_{x\to-\infty} u(x) = -1.$$

Summing-up, we come to the conclusion that $u \in \mathcal{H}$.
Conclusion. To see that u is a minimizer of $\mathcal{F}$, arguing as in Lemma 5.1 we compute

$$\begin{aligned}
\mathcal{F}(u_n) &= \int_{-\infty}^{\infty} \left(\frac{1}{2}\big(u_n''^2 + \tilde{g}(u_n)u_n'^2\big) + f(u_n)\right)\mathrm{d}x \\
&\quad + \int_{-\infty}^{\infty} \frac{1}{2}\big(g(u_n) - \tilde{g}(u_n)\big)u_n'^2\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{1}{2}\left(u_n'' - \frac{\tilde{G}(u_n)}{2}\right)^2 \mathrm{d}x + \int_{-\infty}^{\infty} \left(f(u_n) - \frac{\tilde{G}(u_n)^2}{8}\right)\mathrm{d}x \\
&\quad + \int_{-\infty}^{\infty} \frac{1}{2}\big(g(u_n) - \tilde{g}(u_n)\big)u_n'^2\,\mathrm{d}x.
\end{aligned} \tag{5.6}$$

As all the terms of the right-hand side are positive, we deduce that

$$\sup_{n\in\mathbb{N}} \left\| u_n'' - \frac{\tilde{G}(u_n)}{2} \right\|_{L^2} < +\infty.$$

Going to a subsequence if necessary, we may assume that $u_n'' - \frac{\tilde{G}(u_n)}{2}$ converges weakly in L^2 and therefore

$$\int_{-\infty}^{\infty} \frac{1}{2}\left(u'' - \frac{\tilde{G}(u)}{2}\right)^2 \mathrm{d}x \leqslant \liminf_{n\to\infty} \int_{-\infty}^{\infty} \frac{1}{2}\left(u_n'' - \frac{\tilde{G}(u_n)}{2}\right)^2 \mathrm{d}x.$$

Now observe that the integrands of the two last terms of (5.6) converge for each $x \in \mathbb{R}$ and are positive. Fatou's lemma is then applicable and we obtain

$$\int_{-\infty}^{+\infty} \left(f(u) - \frac{\tilde{G}(u)^2}{8}\right) \mathrm{d}x \leqslant \liminf_{n\to\infty} \int_{-\infty}^{+\infty} \left(f(u_n) - \frac{\tilde{G}(u_n)^2}{8}\right) \mathrm{d}x$$

and

$$\int_{-\infty}^{+\infty} \frac{1}{2}\big(g(u) - \tilde{g}(u)\big)u'^2 \,\mathrm{d}x \leqslant \liminf_{n\to\infty} \int_{-\infty}^{+\infty} \frac{1}{2}\big(g(u_n) - \tilde{g}(u_n)\big)u_n'^2 \,\mathrm{d}x.$$

From Step 2, we know that $u \in \mathcal{H}$. The equality (5.6) thus also holds with u_n replaced by u and we finally deduce that

$$\mathcal{F}(u) \leqslant \liminf_{n\to\infty} \mathcal{F}(u_n) = \inf_{\mathcal{H}} \mathcal{F}$$

which by Step 2 implies

$$\mathcal{F}(u) = \inf_{\mathcal{H}} \mathcal{F}.$$

This completes the proof. □

We devote the remaining of the section to the proof of Proposition 5.3. We first need some preliminary results. Our next purpose is to complete the functional settings by proving that the first derivative of any function $u \in \mathcal{H}$ that satisfies $\mathcal{F}(u) < \infty$, vanishes at $\pm\infty$. To this aim, we next state the following useful estimate.

LEMMA 5.4. *Given an interval* $[a, b] \subset \mathbb{R}$ *and a function* $u \in H^2(a, b)$ *such that* $u(a) = A$, $u(b) = B$, $u'(a) = A_1$, $u'(b) = B_1$, *the following inequality holds*:

$$\int_a^b u''^2 \,\mathrm{d}x \geqslant \frac{4}{b-a}\left((B_1 - A_1)^2 + 3\left(\frac{B-A}{b-a} - A_1\right)\left(\frac{B-A}{b-a} - B_1\right)\right),$$

and equality holds if and only if u is a third degree polynomial.

PROOF. Denote by P the third degree polynomial that coincides in H^2 with u at points a and b. Writing $u = P + w$, we compute

$$\int_a^b u''^2 \,\mathrm{d}x = \int_a^b P''^2 \,\mathrm{d}x + \int_a^b w''^2 \,\mathrm{d}x + 2\int_a^b P''w'' \,\mathrm{d}x. \tag{5.7}$$

Integrating $P''w''$ by parts and using the fact that $w(a) = w(b) = w'(a) = w'(b) = 0$, we see that the last integral in (5.7) is actually zero. We thus obtain the inequality

$$\int_a^b u''^2 \,\mathrm{d}x \geqslant \int_a^b P''^2 \,\mathrm{d}x$$

and the conclusion now follows by computing the integral of P''^2. □

LEMMA 5.5. *Let $u \in \mathcal{H}$ be such that $\mathcal{F}(u) < \infty$. Then*

$$\lim_{|x|\to\infty} u'(x) = 0.$$

PROOF. Let $\varepsilon > 0$ and $u \in \mathcal{H}$ be given. Suppose by contradiction that our conclusion is false. Assume for example that $u'(x_n) \geqslant \varepsilon$ for some sequence $x_n \to +\infty$, $n \in \mathbb{N}$. Let $\delta > 0$ be such that

$$\delta < \frac{s\varepsilon^3}{16\mathcal{F}(u)},$$

where s is given by Lemma 5.1. Then, as $\lim_{x\to+\infty} u(x) = 1$, there exists $R > 0$ such that for all $x \geqslant R$, $|u(x) - 1| \leqslant \delta/2$. Let $x_0 > R$ be such that $u'(x_0) = \varepsilon$.

We claim that we can find $x_1 > x_0$ such that $u'(x_1) = \varepsilon/2$, $u(x_1) \leqslant u(x_0) + \delta$ and $x_1 - x_0 \leqslant 2\delta/\varepsilon$. Indeed, if $x > x_0$, we have $|u(x) - u(x_0)| \leqslant \delta$ and as $\lim_{x\to+\infty} u(x) = 1$, there exists $x > x_0$ such that $u'(x) \leqslant \varepsilon/2$. We can therefore choose x_1 such that $u'(x) \geqslant \varepsilon/2$ for all $x \in [x_0, x_1]$. We then have

$$\frac{\varepsilon}{2}(x_1 - x_0) \leqslant \int_{x_0}^{x_1} u'(s)\,\mathrm{d}s = u(x_1) - u(x_0) \leqslant \delta.$$

Now, letting $m = \frac{u(x_1)-u(x_0)}{x_1-x_0}$, we infer from Lemma 5.4 that

$$\int_{x_0}^{x_1} u''^2 \,\mathrm{d}x \geqslant \frac{2\varepsilon}{\delta}\left(\left(\frac{\varepsilon}{2}\right)^2 + 3(m-\varepsilon)\left(m - \frac{\varepsilon}{2}\right)\right) \geqslant \frac{\varepsilon^3}{8\delta}.$$

Hence, we obtain a contradiction with the choice of δ since Lemma 5.1 implies that

$$2\mathcal{F}(u) \geqslant s\int_{x_0}^{x_1} u''^2 \,\mathrm{d}x \geqslant \frac{s\varepsilon^3}{8\delta}.$$

Similar arguments hold in the other cases, in particular if $\lim_{x\to-\infty} u'(x) \neq 0$. □

We now turn to the proof of Proposition 5.3.

PROOF. Proof of Proposition 5.3 We divide the proof in two parts. In the first part we prove the a priori bound (property (i)) and in the second one we prove the localization of at least one minimizing sequence (property (ii)).

Let $(u_n)_n \subset \mathcal{H}$ be a minimizing sequence of $\mathcal{F}$. As $\mathcal{F}(u_n)$ is a converging sequence, $\mathcal{F}(u_n)$ is uniformly bounded. Notice that we can assume without loss of generality that $\mathcal{F}(u_n) \leqslant C \leqslant \inf_{\mathcal{H}} \mathcal{F} + 1$.

Part 1. There exists $L > 0$ such that for all $n \in \mathbb{N}$, $\|u_n\|_{C^1} \leqslant L$.

Let s be given by Lemma 5.1. According to the assumptions on f there exist $K > 0$ and $b > 0$ such that

$$\frac{3K^2 s}{8b^3} > C \quad \text{and} \quad s\, f(u) > \frac{C}{b} \quad \text{for all } |u| \geqslant \frac{K}{2}.$$

We claim that $\|u_n\|_\infty \leqslant K$. Otherwise, either the set $\{x\colon\ |u_n(x)| \geqslant K/2\}$ has measure greater than b and

$$\mathcal{F}(u_n) \geqslant s \int_{-\infty}^{+\infty} f(u_n)\, \mathrm{d}x > s\, b \frac{C}{b\, s} = C,$$

a contradiction, or we can pick up an interval (c, d) such that $d - c < b$, $|u_n(c)| = |u_n|_\infty$, $|u_n(d)| = \frac{K}{2}$, $|u_n(x)| \geqslant \frac{K}{2}$, for all $x \in (c, d)$. It then follows using Lemma 5.4 that

$$\begin{aligned}
\mathcal{F}(u_n) &\geqslant s \int_c^d \frac{u_n''^2}{2}\, \mathrm{d}x \\
&\geqslant \frac{2s}{d-c}\left(u_n'(d)^2 + 3\frac{u_n(d) - u_n(c)}{d-c}\left(\frac{u_n(d) - u_n(c)}{d-c} - u_n'(d)\right)\right) \\
&\geqslant \frac{3s}{2(d-c)}\left(\frac{u_n(d) - u_n(c)}{d-c}\right)^2 \\
&\geqslant \frac{3K^2 s}{8b^3} > C,
\end{aligned}$$

leading again to a contradiction. Hence the bound on $\|u_n\|_\infty$ is established.

Next, choosing M such that

$$M > 4K \quad \text{and} \quad M^2 > \frac{8C}{s},$$

we show that u_n' cannot attain the value M. Assume by contradiction that $u_n'(x_0) = M$. Then there exists $x_1 \in (x_0, x_0 + 1)$ such that $u_n'(x_1) = \frac{M}{2}$. Hence, denoting again $m = \frac{u_n(x_1) - u_n(x_0)}{x_1 - x_0}$, it turns out that

$$\mathcal{F}(u_n) \geqslant s \int_{x_0}^{x_1} \frac{u_n''^2}{2}\, \mathrm{d}x$$

$$\geqslant \frac{2s}{x_1 - x_0}\left(\left(\frac{M}{2}\right)^2 + 3(m-M)\left(m - \frac{M}{2}\right)\right)$$
$$\geqslant \frac{M^2 s}{8} > C,$$

which is impossible. We show in a similar way that u_n' cannot attain the value $-M$. We thus deduce that the conclusion of Part 1 holds with $L = \max(K, M)$.

Part 2. There exist $T > 0$ and a minimizing sequence $(v_n)_n \subset \mathcal{H}$ such that for all $n \in \mathbb{N}$,

$$\left|v_n(x) + 1\right| \leqslant a \quad \text{for all } x \leqslant -T$$

and

$$\left|v_n(x) - 1\right| \leqslant a \quad \text{for all } x \geqslant T.$$

Let $\varepsilon > 0$ be given. For each $n \in \mathbb{N}$, we define

$$x_1 := \sup\left\{x \in \mathbb{R} \text{ such that } \left|u_n(x) + 1\right| \leqslant \varepsilon \text{ and } \left|u_n'(x)\right| \leqslant \varepsilon\right\},$$

and

$$x_2 := \inf\left\{x \in \mathbb{R} \text{ such that } \left|u_n(x) - 1\right| \leqslant \varepsilon \text{ and } \left|u_n'(x)\right| \leqslant \varepsilon\right\}.$$

Observe that as $u_n \in \mathcal{H}$ and $\mathcal{F}(u_n) < \infty$, x_1 and x_2 are real numbers. Basically, this part of the proof consists in showing that the length $x_2 - x_1$ can be controlled uniformly in $n \in \mathbb{N}$. Then choosing $\varepsilon > 0$ sufficiently small, pieces of orbits close to ± 1 on $]-\infty, x_1[$ and $]x_2, +\infty[$ can be glued to $u_{n|[x_1,x_2]}$ without increasing the action above $\mathcal{F}(u_n)$. This rather simple argument all the same requires some technical adjustments. We first focus on the estimate of the length $x_2 - x_1$.

Step 1. For each $0 < \varepsilon < 1$, there exists $T_\varepsilon > 0$ such that for all $n \in \mathbb{N}$, $x_2 - x_1 \leqslant 2T_\varepsilon$.

Let us define $N_\varepsilon^\pm :=]\pm 1 - \varepsilon, \pm 1 + \varepsilon[$, $N_\varepsilon := N_\varepsilon^- \cup N_\varepsilon^+$ and consider the set

$$Z := \left\{x \in [x_1, x_2] \,\middle|\, u_n(x) \in N_\varepsilon\right\}.$$

Observe that Z is a union of intervals I_i on which $|u_n'| \geqslant \varepsilon$. In the sequel, we assume that these intervals I_i are of maximal length. As $|u_n'(x)| \geqslant \varepsilon$ on any interval I_i, we infer that

$$|I_i|\varepsilon \leqslant \left|\int_{I_i} u_n'(x)\,\mathrm{d}x\right| \leqslant 2\varepsilon$$

so that $|I_i| \leqslant 2$. Further except maybe for the last one, each interval I_i is followed by an interval $J_i = [c_i, d_i]$ that we also suppose to be of maximal length and which is so that one of the following conditions holds for all $x \in [c_i, d_i]$:

$$\begin{array}{llll} \text{(a)} & u_n(x) \geqslant 1+\varepsilon, & u_n'(c_i) \geqslant \varepsilon, & u_n'(d_i) \leqslant -\varepsilon, \\ \text{(b)} & -1+\varepsilon \leqslant u_n(x) \leqslant 1-\varepsilon, & \left|u_n'(c_i)\right| \geqslant \varepsilon, & \left|u_n'(d_i)\right| \geqslant \varepsilon, \\ \text{(c)} & u_n(x) \leqslant -1-\varepsilon, & u_n'(c_i) \leqslant -\varepsilon, & u_n'(d_i) \geqslant \varepsilon. \end{array}$$

Consider an interval $J_i = [c_i, d_i]$ such that (a) or (c) hold. We then obtain the estimate

$$2\varepsilon \leqslant \left|u_n'(d_i) - u_n'(c_i)\right| = \left|\int_{c_i}^{d_i} u_n''(x)\,\mathrm{d}x\right| \leqslant \|u_n''\|_{L^2(c_i,d_i)}(d_i - c_i)^{1/2}$$

and we thus infer that

$$\int_{c_i}^{d_i} \left(\frac{1}{2}\left(u_n''\right)^2 + f(u_n)\right)\mathrm{d}x \geqslant \frac{2\varepsilon^2}{d_i - c_i} + r_\varepsilon(d_i - c_i) \geqslant 2\varepsilon\sqrt{2r_\varepsilon}, \tag{5.8}$$

where $r_\varepsilon := \inf\{f(u) \mid u \notin N_\varepsilon\}$. Notice that for an interval J_i of type (b) with $u_n'(c_i)u_n'(d_i) < 0$, the same inequality holds. At last, assume that J_i is of type (b) and such that $u_n'(c_i)u_n'(d_i) > 0$. Then we obtain

$$2 - 2\varepsilon = \left|\int_{c_i}^{d_i} u_n'(x)\,\mathrm{d}x\right| \leqslant L(d_i - c_i)$$

and

$$\int_{c_i}^{d_i} \left(\frac{1}{2}\left(u_n''\right)^2 + f(u_n)\right)\mathrm{d}x \geqslant r_\varepsilon(d_i - c_i) \geqslant \frac{r_\varepsilon(2-2\varepsilon)}{L}, \tag{5.9}$$

where L is the uniform bound on $\|u_n'\|_\infty$ obtained in Part 1. Let us denote by k the number of intervals J_i. Taking the estimates (5.8) and (5.9) into account, we infer from Lemma 5.1 that

$$\begin{aligned} C \geqslant \mathcal{F}(u_n) &\geqslant s\sum_{i=1}^{k}\int_{c_i}^{d_i}\left(\frac{1}{2}\left(u_n''\right)^2 + f(u_n)\right)\mathrm{d}x \\ &\geqslant sk\min\left(2\varepsilon\sqrt{2r_\varepsilon}, \frac{r_\varepsilon(2-2\varepsilon)}{L}\right) \end{aligned} \tag{5.10}$$

so that k is uniformly bounded with respect to $n \in \mathbb{N}$.

We are now able to conclude our first step. Indeed, setting $\tilde{Z} = [x_1, x_2] \setminus Z$, we have

$$C \geqslant \mathcal{F}(u_n) \geqslant s\int_{\tilde{Z}} f\left(u_n(x)\right)\mathrm{d}x \geqslant sr_\varepsilon\left|\tilde{Z}\right|$$

and since Z is the union of at most $k+1$ intervals of length smaller than 2, we finally deduce that

$$x_2 - x_1 = |Z| + \bigl|\tilde{Z}\bigr| \leqslant 2(k+1) + \frac{C}{s\, r_\varepsilon} =: 2T_\varepsilon.$$

Step 2. Modification of u_n in $]-\infty, x_1]$ and $[x_2, +\infty[$.

We consider the modification in $]-\infty, x_1]$. Define

$$\mathcal{F}_{]-\infty,x_1]}(u) := \int_{-\infty}^{x_1} \Bigl(\frac{1}{2}\bigl(u''^2 + g(u)u'^2\bigr) + f(u)\Bigr)\, \mathrm{d}x$$

having as domain

$$\mathcal{H}^n_{]-\infty,x_1]} := \bigl\{u \in \mathcal{H} \mid u = u_n \text{ on } [x_1, +\infty[\bigr\}.$$

Using integration by parts and arguing as in Lemma 5.1, we see that for any function $u \in \mathcal{H}^n_{]-\infty,x_1]}$

$$\mathcal{F}_{]-\infty,x_1]}(u) \geqslant s \int_{-\infty}^{x_1} \Bigl(\frac{u''^2}{2} + f(u)\Bigr)\, \mathrm{d}x + \frac{1}{2}\tilde{G}\bigl(u(x_1)\bigr)u'(x_1). \tag{5.11}$$

We also consider the set

$$\mathcal{D}^n_{]-\infty,x_1]} := \bigl\{u \in \mathcal{H}^n_{]-\infty,x_1]} \bigm| \text{ for all } x \leqslant x_1,\ u(x) \in [-1-a, -1+a]\bigr\}.$$

Let us fix again the notation $N_a^- :=]-1-a, -1+a[$. We then define

$$\begin{aligned}
&\gamma := \sup\bigl\{g(u) \mid u \in N_a^-\bigr\},\\
&\eta := \inf\bigl\{f(u) \mid u \in N_a^- \setminus N_{a/2}^-\bigr\} > 0,\\
&\delta := \sup\bigl\{\bigl|\tilde{G}(u)\bigr| \mid u \in N_a^-\bigr\} \geqslant 0.
\end{aligned}$$

We first derive a lower estimate on the action of any function $u \in \mathcal{H}^n_{]-\infty,x_1]}$ whose graph does not stay in the strip $]-\infty, x_1] \times N_a^-$.

CLAIM 1. *If* $u \in \mathcal{H}^n_{]-\infty,x_1]} \setminus \mathcal{D}^n_{]-\infty,x_1]}$ *and* $\|u'\|_\infty \leqslant L$ *then*

$$\mathcal{F}_{]-\infty,x_1]}(u) \geqslant \frac{s\eta a}{2L} - \delta\varepsilon. \tag{5.12}$$

If there exists $x \leqslant x_1$ such that $u(x) \notin N_a^-$, either there exist $s_1 \leqslant s_2 \leqslant x_1$ so that

$$u(s_1) = -1 + \frac{a}{2}, \quad u(s_2) = -1 + a$$

and for all $x \in [s_1, s_2]$,

$$u(x) \in \left[-1 + \frac{a}{2}, -1 + a\right]$$

or there exist $s_3 \leqslant s_4 \leqslant x_1$ so that

$$u(s_3) = -1 - a, \qquad u(s_4) = -1 - \frac{a}{2}$$

and for all $x \in [s_3, s_4]$,

$$u(x) \in \left[-1 - a, -1 - \frac{a}{2}\right].$$

Let us for instance consider the first possibility, the second being similar. We then have

$$L(s_2 - s_1) \geqslant \int_{s_1}^{s_2} u'(x)\,\mathrm{d}x = u(s_2) - u(s_1) = \frac{a}{2}$$

and

$$\int_{-\infty}^{x_1} \left(\frac{u''^2}{2} + f(u)\right)\mathrm{d}x \geqslant \int_{s_1}^{s_2} f(u)\,\mathrm{d}x \geqslant \frac{\eta a}{2L}.$$

On the other hand, we have

$$\left|\tilde{G}\big(u(x_1)\big)u'(x_1)\right| \leqslant \delta\varepsilon,$$

so that Claim 1 now follows from (5.11).

CLAIM 2. *There exists* $R > 0$ *so that for all* $\varepsilon > 0$,

$$\inf_{\mathcal{D}^n_{]-\infty,x_1]}} \mathcal{F}_{]-\infty,x_1]} \leqslant R\varepsilon^2.$$

For a function $u \in \mathcal{D}^n_{]-\infty,x_1]}$, we have by virtue of (C2)

$$\mathcal{F}_{]-\infty,x_1]}(u) \leqslant \int_{-\infty}^{x_1} \left(\frac{1}{2}\big(u''^2 + \gamma u'^2\big) + \alpha(u+1)^2\right)\mathrm{d}x.$$

Let us define the functional

$$\mathcal{J}_{]-\infty,x_1]}(u) := \int_{-\infty}^{x_1} \left(\frac{1}{2}\big(u''^2 + \gamma u'^2\big) + \alpha(u+1)^2\right)\mathrm{d}x$$

on $\mathcal{H}^n_{]-\infty,x_1]}$. Let P be the third degree polynomial satisfying $P(x_1-1)=-1$, $P'(x_1-1)=0$, $P(x_1)=u_n(x_1)$ and $P'(x_1)=u_n'(x_1)$. The function $v:]-\infty,x_1]\to\mathbb{R}$ defined by

$$v(x):=\begin{cases}0 & \text{if } x<x_1-1\\ P(x) & \text{if } x_1-1\leqslant x\leqslant x_1.\end{cases}$$

belongs to $\mathcal{H}^n_{]-\infty,x_1]}$ and in fact to $\mathcal{D}^n_{]-\infty,x_1]}$ if ε is taken sufficiently small. A straightforward computation then shows that there is a constant $C(\gamma,\alpha)$ such that

$$\mathcal{J}_{]-\infty,x_1]}(v)\leqslant C(\gamma,\alpha)\big(\big(u_n(x_1)+1\big)^2+u_n'(x_1)^2\big)\leqslant 2C(\gamma,\alpha)\varepsilon^2$$

so that we obtain the estimate

$$\inf_{\mathcal{D}^n_{]-\infty,x_1]}}\mathcal{F}_{]-\infty,x_1]}\leqslant\inf_{\mathcal{D}^n_{]-\infty,x_1]}}\mathcal{J}_{]-\infty,x_1]}\leqslant 2C(\gamma,\alpha)\varepsilon^2.$$

Conclusion of Step 2. Let us choose $\varepsilon>0$ sufficiently small in order to have

$$\mathcal{F}_{]-\infty,x_1]}(u)\geqslant\frac{s\eta a}{4L}$$

for $u\in\mathcal{H}^n_{]-\infty,x_1]}\setminus\mathcal{D}^n_{]-\infty,x_1]}$ satisfying $\|u'\|_\infty\leqslant L$ and

$$\inf_{\mathcal{D}^n_{]-\infty,x_1]}}\mathcal{F}_{]-\infty,x_1]}<\frac{s\eta a}{4L}.$$

If $u_n\notin\mathcal{D}^n_{]-\infty,x_1]}$, we infer from the above estimates that we can replace u_n by $v_n\in\mathcal{D}^n_{]-\infty,x_1]}$ such that $\mathcal{F}(v_n)\leqslant\mathcal{F}(u_n)$.

If $|u_n(x)-1|\nleqslant a$ for $x\geqslant x_2$ we proceed in the same way to modify u_n for $x\geqslant x_2$.

We are now in position to complete the proof of Part 2. Indeed, we deduce from Step 1 and Step 2 that there exist $T>0$ and a minimizing sequence $(v_n)_n$ such that for all $n\in\mathbb{N}$, there exist $x_1<x_2$ satisfying $x_2-x_1\leqslant 2T$ and

$$\big|v_n(x)+1\big|\leqslant a\quad\text{for all } x\leqslant x_1,$$

$$\big|v_n(x)-1\big|\leqslant a\quad\text{for all } x\geqslant x_2.$$

Translating v_n if necessary, we can assume $[x_1,x_2]\subset[-T,T]$ so that the conclusion of the second part follows. Observe that since $\mathcal{F}(v_n)\leqslant\mathcal{F}(u_n)$, the sequence $(v_n)_n$ satisfies the a priori bound derived in the first part. This ends the proof. □

5.2. *Functionals of Swift–Hohenberg type*

When β is negative, the model equation

$$u''''-\beta u''+u^3-u=0$$

is related to the *Swift–Hohenberg* equation

$$\frac{\partial u}{\partial t} - \kappa u + \left(1 + \frac{\partial^2}{\partial x^2}\right)^2 u + u^3 = 0, \tag{5.13}$$

where $\kappa \in \mathbb{R}$. This equation was proposed by Swift and Hohenberg [84] as a model for studies of Rayleigh–Bénard convection. When $\kappa > 1$, after rescaling, (5.13) can be written as

$$\frac{\partial u}{\partial t} + (\kappa - 1)^{3/2}\left(\frac{\partial^4 u}{\partial x^4} - \beta \frac{\partial^2 u}{\partial x^2} + u^3 - u\right) = 0$$

with $\beta = -2/\sqrt{\kappa - 1}$. The compatibility condition (C1) between f and g allows g to take negative values but it then implies that g takes positive values as well. Therefore, the stationary Swift–Hohenberg equation is not covered by the result of the preceding section.

When stronger assumptions are imposed on the potential, we can deal with negative functions g. Namely we assume that $f, g \in C^1(\mathbb{R})$ are even functions such that $f(1) = 0$ and

(C5) for some $k > 0$, $\beta < \sqrt{8k}$ and all $u \geqslant 0$,

$$f(u) \geqslant k(u-1)^2 \quad \text{and} \quad g(u) \geqslant -\beta.$$

We look then at the minimizers of the functional $\mathcal{F}$ defined by (5.2) in the space

$$\mathcal{E} = \left\{u \mid u(0) = 0,\ u + 1 \in H^2(\mathbb{R}^-),\ u - 1 \in H^2(\mathbb{R}^+)\right\}.$$

Our first observation is that under the assumption (C5), for small $\beta > 0$,

$$\inf_{u \in \mathcal{E}} \mathcal{F}(u) > -\infty.$$

LEMMA 5.6. *Let f and $g \in C(\mathbb{R})$ be even functions such that* (C5) *holds. Then there exists $\beta_1 > 0$ such that for $\beta \leqslant \beta_1$, we have*

$$\inf_{u \in \mathcal{E}} \mathcal{F}(u) \geqslant 0.$$

PROOF. Let us first recall the following interpolation inequality, see for example [1]. For all $\ell > 0$, there exists $C > 0$ such that for any interval I with $|I| \geqslant \ell$ and all $u \in H^2(I)$,

$$\|u'\|^2_{L^2(I)} \leqslant C\left(\|u\|^2_{L^2(I)} + \|u''\|^2_{L^2(I)}\right). \tag{5.14}$$

Let u be any function in $\mathcal{E}$ and consider an interval $[a, b]$ where u is nonnegative and $u(a) = u(b) = 0$. Suppose that $b - a \geqslant 1$. Then using the inequality (5.14) applied to $u - 1$, we obtain

$$\begin{aligned}\int_a^b \left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right) \mathrm{d}x &\geqslant \int_a^b \left(\frac{1}{2}\left(u''^2 - \beta u'^2\right) + k(u-1)^2\right) \mathrm{d}x \\ &\geqslant \frac{1-\beta C}{2} \left\|u''\right\|^2_{L^2(a,b)} \\ &\quad + \left(k - \frac{\beta C}{2}\right) \|u-1\|^2_{L^2(a,b)}. \end{aligned} \tag{5.15}$$

If $b - a < 1$, we compute

$$\left\|u'\right\|_{L^2(a,b)} \leqslant \frac{b-a}{\sqrt{2}} \left\|u''\right\|_{L^2(a,b)}$$

and

$$\begin{aligned}\int_a^b \left(\frac{1}{2}\left(u''^2 - g(u)u'^2\right) + f(u)\right) \mathrm{d}x &\geqslant \int_a^b \left(\frac{1}{2}\left(u''^2 - \beta u'^2\right) + f(u)\right) \mathrm{d}x \\ &\geqslant \frac{1}{2}\left(1 - \frac{\beta}{2}\right) \left\|u''\right\|^2_{L^2(a,b)}. \end{aligned}$$

In both cases, we obtain

$$\int_a^b \left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right) \mathrm{d}x \geqslant 0$$

if $\beta \leqslant \min(2, 1/C, 2k/C)$. A similar argument holds on intervals $[a, b]$ where u is non positive. At last, if u is nonnegative on the interval $[T, +\infty[$, we deduce as in (5.15) that

$$\int_a^{+\infty} \left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right) \mathrm{d}x \geqslant 0.$$

The conclusion follows. □

Next, assuming that the functional $\mathcal{F}$ is bounded from below, we prove the existence of a minimizer. We deduce from the previous lemma that the result holds at least when $g \geqslant -\beta$ with $\beta > 0$ small.

THEOREM 5.7. *Let $f \in C^1(\mathbb{R})$ and $g \in C^1(\mathbb{R})$ be even functions such that $f(1) = 0$ and* (C5) *holds. Assume further that*

$$\inf_{\mathcal{E}} \mathcal{F} > -\infty.$$

Then $\mathcal{F}$ has a minimizer which is a heteroclinic solution of Eq. (5.3) *having exactly one zero. Moreover, any minimizer is odd.*

PROOF. *Step* 1. We first prove the existence of an odd minimizing sequence $(u_n)_n$ such that $u_n(x) > 0$ for all $x > 0$. Let $(v_n)_n \subset \mathcal{E}$ be a minimizing sequence. Observe that if v_n is not odd, the appropriate symmetrization of v_n restricted to $\mathbb{R}^+$ or $\mathbb{R}^-$ decreases the action. Assume now that v_n has a positive zero and define

$$x_n := \sup\bigl\{x > 0 \bigm| v_n(x) = 0\bigr\}.$$

We claim that

$$\mathcal{F}_{[-x_n,x_n]}(v_n) := \int_{-x_n}^{x_n} \left(\frac{1}{2}\bigl(v_n''^2 + g(v_n)v_n'^{\,2}\bigr) + f(v_n)\right) \mathrm{d}x \geqslant 0.$$

Indeed, suppose that $\mathcal{F}_{[-x_n,x_n]}(v_n) = -C < 0$ and take $j \in \mathbb{N}$ so that $\mathcal{F}(v_n) - 2jC < \inf_{\mathcal{E}} \mathcal{F}$. Then defining the odd function $u_n^* \in \mathcal{E}$ by

$$u_n^*(x) := \begin{cases} v_n(x - 2ix_n) & \text{if } x \in \bigl[2ix_n, (2i+1)x_n\bigl[, \ i = 0, \dots, j, \\ -v_n(2ix_n - x) & \text{if } x \in \bigl[(2i-1)x_n, 2ix_n\bigl[, \ i = 1, \dots, j, \\ v_n\bigl(x - 2(j+1)x_n\bigr) & \text{if } x \in \bigl[(2j+1)x_n, +\infty\bigl[, \end{cases}$$

we obtain a contradiction because

$$\mathcal{F}\bigl(u_n^*\bigr) = \mathcal{F}(v_n) + 2j\mathcal{F}_{[-x_n,x_n]}(v_n) = \mathcal{F}(v_n) - 2jC < \inf_{\mathcal{E}} \mathcal{F}.$$

Now, let $u_n \in \mathcal{E}$ be the odd function defined by $u_n(x) = v_n(x + x_n)$ for $x \geqslant 0$. This function u_n vanishes only at 0 and since $\mathcal{F}_{[-x_n,x_n]}(v_n) \geqslant 0$, the sequence $(u_n)_n$ is also a minimizing sequence.

Step 2. As there exists an odd minimizing sequence, it is sufficient to minimize the restricted functional

$$\mathcal{F}^+(u) := \int_0^{+\infty} \left(\frac{1}{2}\bigl(u''^2 + g(u)u'^2\bigr) + f(u)\right) \mathrm{d}x$$

in the space of functions

$$\mathcal{E}^+ = \bigl\{u : \mathbb{R}^+ \to \mathbb{R} \bigm| u(0) = 0,\ u - 1 \in H^2\bigl(\mathbb{R}^+\bigr)\bigr\}.$$

We first claim that $u_n - 1$ is uniformly bounded in $H^2(\mathbb{R}^+)$. To prove this, we compute for $u \in \mathcal{E}^+$,

$$\begin{aligned} &\int_0^{+\infty} \left(\frac{1}{2}\bigl(u''^2 - \beta u'^2\bigr) + k(u-1)^2\right) \mathrm{d}x \\ &\quad = \varepsilon\|u - 1\|_{H^2}^2 + \frac{1 - 2\varepsilon}{2} \int_0^{+\infty} \left(u''^2 - \beta_\varepsilon u'^2 + \frac{1}{4}\beta_\varepsilon^2 (u-1)^2\right) \mathrm{d}x \\ &\qquad + \left(k - \varepsilon - \frac{(\beta + 2\varepsilon)^2}{8(1 - 2\varepsilon)}\right) \int_0^{+\infty} (u-1)^2 \,\mathrm{d}x, \end{aligned}$$

where $\beta_\varepsilon := \frac{\beta+2\varepsilon}{1-2\varepsilon}$. Notice that we may choose ε small enough in order to have $k-\varepsilon-\frac{(\beta+2\varepsilon)^2}{8(1-2\varepsilon)} \geqslant 0$. For the second term of the right-hand side, we have

$$\begin{aligned}\int_0^{+\infty}\left(u''^2-\beta_\varepsilon u'^2+\frac{\beta_\varepsilon^2}{4}(u-1)^2\right)\mathrm{d}x\\ =\int_0^{+\infty}\left(u''+\frac{\beta_\varepsilon}{2}(u-1)\right)^2\mathrm{d}x-\beta_\varepsilon u'(0)\geqslant -\beta_\varepsilon u'(0).\end{aligned}$$

As u_n is positive, we infer that

$$\begin{aligned}\mathcal{F}^+(u_n)&\geqslant\int_0^{+\infty}\left(\frac{1}{2}\big(u_n''^2-\beta u_n'^2\big)+k(u_n-1)^2\right)\mathrm{d}x\\ &\geqslant\varepsilon\|u_n-1\|_{H^2}^2-\beta_\varepsilon u_n'(0).\end{aligned}$$

We now deduce from the continuous injection of $H^2(0,\infty)$ into $\mathcal{C}^1([0,\infty[)$ that for some $C>0$

$$\mathcal{F}^+(u_n)\geqslant\varepsilon\|u_n-1\|_{H^2(0,\infty)}^2-C\|u_n-1\|_{H^2(0,\infty)}.$$

Hence, the claim follows from the uniform bound on $\mathcal{F}(u_n)$.

We now infer that at least for a subsequence (that we still denote by u_n)

$$u_n-1\overset{H^2(\mathbb{R}^+)}{\rightharpoonup}u-1\quad\text{and}\quad u_n\overset{C^1_{\mathrm{loc}}(\mathbb{R}^+)}{\longrightarrow}u$$

for some function $u\in\mathcal{E}^+$. In order to see that u is a minimizer, we write

$$\begin{aligned}\mathcal{F}^+(u_n)&=\frac{1}{2}\int_0^{+\infty}\left(u_n''^2-\beta u_n'^2+\frac{\beta^2}{4}(u_n-1)^2\right)\mathrm{d}x\\ &\quad+\frac{1}{2}\int_0^{+\infty}\big((g(u_n)+\beta)u_n'^2\big)\,\mathrm{d}x+\int_0^{+\infty}\left(f(u_n)-\frac{\beta^2}{8}(u_n-1)^2\right)\mathrm{d}x\\ &=\frac{1}{2}\int_0^{+\infty}\left(u_n''+\frac{\beta}{2}(u_n-1)\right)^2\mathrm{d}x-\frac{\beta}{2}u_n'(0)\\ &\quad+\frac{1}{2}\int_0^{+\infty}\big((g(u_n)+\beta)u_n'^2\big)\,\mathrm{d}x+\int_0^{+\infty}\left(f(u_n)-\frac{\beta^2}{8}(u_n-1)^2\right)\mathrm{d}x.\end{aligned}$$

Observe then that in the last equality, the first integral is convex and that Fatou's Lemma is applicable to the last two so that

$$\mathcal{F}^+(u)\leqslant\lim_{n\to+\infty}\mathcal{F}^+(u_n)=\inf_{\mathcal{E}^+}\mathcal{F}^+.$$

Now, as $u \in \mathcal{E}^+$, we conclude that $\mathcal{F}^+(u) = \inf_{\mathcal{E}^+} \mathcal{F}^+$ and still denoting by u its odd extension on $\mathbb{R}$, we have $\mathcal{F}(u) = \inf_{\mathcal{E}} \mathcal{F}$.

The proof that any minimizer is odd follows the lines of the third step of the proof of Theorem 4.5. □

In the following theorem, we observe that when ±1 are saddle-foci, the minimum of $\mathcal{F}$ is non-negative.

THEOREM 5.8. *Let f and $g \in C^2(\mathbb{R})$ be even functions such that $f(1) = 0$, $g(1)^2 < 4f''(1)$ and assume that* (C5) *holds. Then if u is a minimizer in $\mathcal{E}$ of $\mathcal{F}$, $\mathcal{F}(u) \geqslant 0$.*

In order to prove Theorem 5.8, we first consider the following lemma which states that if $+1$ is a saddle-focus stationary point, then the solutions of (5.3) that converge to $+1$ in the phase-space, do oscillate around the equilibrium in their tails. To simplify the notation, we assume the equilibria has been translated to 0.

LEMMA 5.9. *Let f and $g \in C^2(\mathbb{R})$ and $f(0) = 0$. Let 0 be a saddle-focus equilibria for the linearization of Eq.* (5.3), *i.e. $g(0)^2 < 4f''(0)$. Then, there exists $\Delta > 0$ such that if $\hat{u}$ is a nontrivial solution of* (5.3) *that satisfies*

$$\lim_{x \to +\infty} \big(\hat{u}(x), \hat{u}'(x), \hat{u}''(x), \hat{u}'''(x)\big) = (0, 0, 0, 0), \tag{5.16}$$

$\hat{u}$ changes sign in any interval $[x_0, x_0 + \Delta]$ for sufficiently large x_0.

PROOF. Let $\hat{u}$ be a solution of (5.3) that satisfies (5.16). Consider the linearization of (5.3)

$$z'''' - g(0)z'' + f''(0)z = 0. \tag{5.17}$$

Notice that the characteristic values of this linear equation read $\pm\rho \pm i\omega$. We then choose $\Delta = 2\pi/\omega$. Let $x_0 \in \mathbb{R}$ and $z_0 \in \mathbb{R}^4$. By the choice of Δ, the solution z of (5.17) with initial condition

$$\big(z(x_0), z'(x_0), z''(x_0), z'''(x_0)\big) = z_0$$

satisfies

$$\max_{[x_0, x_0+\Delta]} z, \quad \max_{[x_0, x_0+\Delta]} (-z) \geqslant c|z_0|$$

and

$$\|z\|_{C^2([x_0, x_0+\Delta])} \leqslant M|z_0|$$

for some $c > 0$ and $M > 0$ depending only on Δ, $g(0)$ and $f''(0)$. On the other hand, we can also find $N > 0$ so that the solutions of

$$\begin{aligned}&w'''' - g(0)w'' + f''(0)w = h(x),\\&\big(w(x_0), w'(x_0), w''(x_0), w'''(x_0)\big) = 0,\end{aligned}$$

satisfy

$$\|w\|_{\mathcal{C}^2([x_0,x_0+\Delta])} \leqslant N\|h\|_{L^\infty(x_0,x_0+\Delta)}.$$

Next, we take $\delta > 0$ so that $c - \frac{MN\delta}{1-N\delta} > 0$.

To fix the ideas we denote by $u(x; x_0, u_0)$ the solution of (5.3) with initial conditions $x_0 \in \mathbb{R}$, $u_0 := (u(x_0), u'(x_0), u''(x_0), u'''(x_0))$. Let $\tilde{u}(x; \lambda) = u(x; x_0, \lambda\hat{u}_0)$, with $\hat{u}_0 = (\hat{u}(x_0), \hat{u}'(x_0), \hat{u}''(x_0), \hat{u}'''(x_0))$ and define

$$\begin{aligned}&p(x) = -g\big(\tilde{u}(x; \lambda)\big),\\&q(x) = -g'\big(\tilde{u}(x; \lambda)\big)\tilde{u}'(x; \lambda),\end{aligned}$$

and

$$r(x) = -g'\big(\tilde{u}(x; \lambda)\big)\tilde{u}''(x; \lambda) - \frac{1}{2}g''\big(\tilde{u}(x; \lambda)\big)\tilde{u}'^2(x; \lambda) + f''\big(\tilde{u}(x; \lambda)\big).$$

Let us fix now x_0 large enough so that $|\hat{u}_0|$ is small enough to have, for $0 \leqslant \lambda \leqslant 1$,

$$\sup_{x_0 \leqslant x \leqslant x_0+\Delta} \big|p(x) + g(0)\big|, \quad \sup_{x_0 \leqslant x \leqslant x_0+\Delta} \big|q(x)\big|, \quad \sup_{x_0 \leqslant x \leqslant x_0+\Delta} \big|r(x) - f''(0)\big| \leqslant \delta.$$

Observe also that $|\hat{u}_0| \neq 0$ as a nonzero solution cannot reach 0 in the phase-space in a finite time. We then write

$$\hat{u}(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\lambda} u(x; x_0, \lambda\hat{u}_0)\,\mathrm{d}\lambda.$$

The function

$$\varphi(x; x_0, \hat{u}_0, \lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda} u(x; x_0, \lambda\hat{u}_0)$$

satisfies the Cauchy problem

$$\begin{aligned}&\varphi'''' + p(x)\varphi'' + q(x)\varphi' + r(x)\varphi = 0,\\&\big(\varphi(x_0), \varphi'(x_0), \varphi''(x_0), \varphi'''(x_0)\big) = \hat{u}_0,\end{aligned}$$

and we can write

$$\varphi = w + z.$$

Here w solves the equation

$$\begin{aligned} & w'''' - g(0)w'' + f''(0)w \\ & \quad = -\big(g(0) + p(x)\big)\varphi''(x) - q(x)\varphi'(x) + \big(f''(0) - r(x)\big)\varphi(x), \end{aligned}$$

together with the initial values

$$\big(w(x_0), w'(x_0), w''(x_0), w'''(x_0)\big) = 0$$

and z is a solution of

$$\begin{aligned} & z'''' - g(0)z'' + f''(0)z = 0, \\ & \big(z(x_0), z'(x_0), z''(x_0), z'''(x_0)\big) = \hat{u}_0. \end{aligned}$$

We next choose $\overline{x} \in [x_0, x_0 + \Delta]$ so that

$$z(\overline{x}) \geqslant c|\hat{u}_0|$$

and compute

$$\hat{u} = z + \int_0^1 w \, \mathrm{d}\lambda.$$

Notice that

$$\|w\|_{\mathcal{C}^2([x_0,x_0+\Delta])} \leqslant N\delta\|\varphi\|_{\mathcal{C}^2([x_0,x_0+\Delta])}.$$

We then obtain the estimates,

$$\|w\|_{\mathcal{C}^2([x_0,x_0+\Delta])} \leqslant \frac{N\delta}{1 - N\delta}\|z\|_{\mathcal{C}^2([x_0,x_0+\Delta])} \leqslant \frac{MN\delta}{1 - N\delta}|\hat{u}_0|.$$

At last, we have

$$\hat{u}(\overline{x}) \geqslant c|\hat{u}_0| - \|w\|_\infty \geqslant \left(c - \frac{MN\delta}{1 - N\delta}\right)|\hat{u}_0| > 0.$$

Arguing in a similar way, we find $\underline{x} \in [x_0, x_0 + \Delta]$ so that $\hat{u}(\underline{x}) < 0$. □

We now come back to the proof of Theorem 5.8.

PROOF OF THEOREM 5.8. Let $u \in \mathcal{E}$ be such that $\mathcal{F}(u) = \inf_{\mathcal{E}} \mathcal{F} = -C$ for some $C > 0$. As u is a minimizer, it satisfies Eq. (5.3) and we also have

$$\lim_{x \to +\infty} \big(u(x), u'(x), u''(x), u'''(x)\big) = (1, 0, 0, 0).$$

Observe that Theorem 5.7 implies that u is odd so that

$$\mathcal{F}(u) = 2\int_0^{+\infty}\left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right)\mathrm{d}x = -C.$$

As a consequence of Lemma 5.9, u oscillates around $+1$ at $+\infty$. Hence, we can find $T > 0$ large enough so that

$$\int_0^T\left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right)\mathrm{d}x \leqslant -C/4,$$

and $u'(T) = 0$. Defining the odd function $v^* \in \mathcal{E}$ as follows:

$$v^*(x) = \begin{cases} u(x) & \text{if } x \in [0, T), \\ u(2T - x) & \text{if } x \in [T, 2T), \\ -u(x - 2T) & \text{if } x \in [2T, 3T], \\ -u(4T - x) & \text{if } x \in [3T, 4T], \\ u(x - 4T) & \text{if } x \geqslant 4T, \end{cases}$$

we have

$$\begin{aligned}\mathcal{F}(v^*) = 8&\int_0^T\left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right)\mathrm{d}x \\ &+ 2\int_0^{+\infty}\left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right)\mathrm{d}x < -2C.\end{aligned}$$

This contradicts the definition of u. Hence, we have $\inf_{\mathcal{E}}\mathcal{F} \geqslant 0$. □

We next state an obvious corollary of Theorem 5.8.

COROLLARY 5.10. *Let f and $g \in C^2(\mathbb{R})$ be even functions such that $f(1) = 0$ and* (C5) *holds. Assume, moreover, that $g(1)^2 < 4f''(1)$. If there exists $u \in \mathcal{E}$ such that $\mathcal{F}(u) < 0$, then* $\inf_{\mathcal{E}}\mathcal{F} = -\infty$.

PROOF. Suppose by contradiction that the conclusion of the corollary is false. Then Theorem 5.7 implies the existence of a minimizer which has a negative action. This contradicts Theorem 5.8. □

Let us consider the stationary Swift–Hohenberg equation

$$u'''' + \beta u'' + u^3 - u = 0. \tag{5.18}$$

We know from Theorem 5.7 and Lemma 5.6 that the corresponding functional

$$\mathcal{F}_\beta(u) = \int_{-\infty}^{+\infty}\left(\frac{1}{2}\left(u''^2 - \beta u'^2\right) + \frac{1}{4}\left(u^2 - 1\right)^2\right)\mathrm{d}x$$

has a minimum in $\mathcal{E}$ for β lower than some small positive constant. If β becomes too large, the minimum no longer exists. Indeed, it is shown in [55] that taking the function $u(x) = A(\beta)\sin(\omega(\beta)x)$ where $A(\beta) = \sqrt{\frac{\beta^2+4}{3}}$ and $\omega(\beta) = \sqrt{\frac{\beta}{2}}$, we have

$$\int_0^{2\pi/\omega(\beta)} \left(\frac{1}{2}\left(u''^2 - \beta u'^2\right) + \frac{1}{4}\left(u^2-1\right)^2\right) \mathrm{d}x < 0 \tag{5.19}$$

if $\beta > \beta^* = \sqrt{2\sqrt{6}-4} = 0.9481\ldots$. It follows that the functional is unbounded from below whenever $\beta > \beta^*$. Therefore, there exists $0 < \beta_0 \leqslant \beta^*$ such that for $\beta > \beta_0$, the functional is unbounded from below. We give in the next theorem a characterization of β_0, which shows also that the minimum of $\mathcal{F}_\beta$ exists whenever $\beta \in (-\infty, \beta_0]$.

THEOREM 5.11. *Let β_0 be defined by*

$$\beta_0 = \sup\left\{\beta > 0 \;\middle|\; \inf_{\mathcal{E}} \mathcal{F}_\beta \geqslant 0\right\}.$$

Then $\beta_0 \in\,]0, \beta^*\,]$ and for all $\beta \leqslant \beta_0$, $\mathcal{F}_\beta$ has a minimum which is an odd heteroclinic solution of (5.18) *with exactly one zero. Moreover, for all $\beta > \beta_0$, $\mathcal{F}_\beta$ is unbounded from below.*

PROOF. For $u \geqslant 0$, we have $(u^2-1)^2/4 \geqslant (u-1)^2/4$ so that Theorem 5.7 is valid for $\beta < \sqrt{2}$. It follows from Lemma 5.6 and Theorem 5.7 that for small $\beta > 0$, there exists $u \in \mathcal{E}$ such that $\mathcal{F}_\beta(u) = \inf_{\mathcal{E}} \mathcal{F}_\beta$. Also we deduce from (5.19) that $\beta_0 \leqslant \beta^*$.

Notice now that $\inf_{\mathcal{E}} \mathcal{J}_{\beta_0} \geqslant 0$. If it were not the case, the same would hold true for β near enough β_0, which contradicts the definition of β_0. Next, as $\beta_0 \leqslant \beta^* < \sqrt{2}$, we infer from Theorem 5.7 that for any $\beta \leqslant \beta_0$, $\mathcal{F}_\beta$ has an odd minimizer having exactly one zero.

It follows also from Corollary 5.10 that if $\beta_0 < \beta < \sqrt{2}$, $\mathcal{F}_\beta$ is unbounded from below. In case $\sqrt{2} \leqslant \beta$ the conclusion follows from (5.19). □

Numerical computations [87] seems to indicate that heteroclinic solutions of (5.18) exist at least until $\beta \approx 2.32$. Theorem 5.11 shows that for $\beta > \beta_0$, these critical points are not global minima.

5.3. *Notes and further comments*

1. The solution u obtained in Theorem 5.2 can be seen as a heteroclinic solution in a weak sense as it does not satisfy in general the usual conditions at ±∞, namely

$$\lim_{x\to\pm\infty} \left(u, u', u'', u'''\right)(x) = (\pm 1, 0, 0, 0).$$

However the limits for u and u' hold and it is easily seen that u is in the manifold of zero energy which is the case for classical heteroclinics. Indeed, the conservation of the Hamiltonian

$$H\big(u, u', u'', u'''\big) = u'''u' - \frac{1}{2}u''^2 - \frac{1}{2}g(u)u'^2 + f(u),$$

implies $H(u(x)\, u'(x), u''(x), u'''(x)) = E$ for any $x \in \mathbb{R}$ and some $E \in \mathbb{R}$. Assume $E \neq 0$ and let $(x_n)_n \subset \mathbb{R}^+$ be such that $x_n \to +\infty$ and $(u''(x_n))_n \subset \mathbb{R}$ is a bounded sequence. Then we have

$$\int_0^{x_n} \left(u'''u' - \frac{1}{2}u''^2 - \frac{1}{2}g(u)u'^2 + f(u) \right) \mathrm{d}x = Ex_n.$$

Integrating by parts the first term of the left hand side, we notice that the integral is bounded contradicting the equality.

If in addition to the hypotheses of Theorem 5.2 we suppose that f satisfies for some $0 < b < 1/2$ and $\beta > 0$,

$$\frac{f(u)}{(u-1)^2} \geqslant \beta, \quad \text{for all } u \in (1-b, 1+b),$$
$$\frac{f(u)}{(u+1)^2} \geqslant \beta, \quad \text{for all } u \in (-1-b, -1+b),$$

then it is easily seen that a minimizer u of $\mathcal{F}$ in $\mathcal{H}$ satisfies $u + 1 \in L^2(\mathbb{R}^-)$ and $u - 1 \in L^2(\mathbb{R}^+)$. Now, as we also have $u'' \in L^2(\mathbb{R})$, we infer by interpolation that $u' \in L^2(\mathbb{R})$ and thus the minimizer u actually satisfies $u + 1 \in H^2(\mathbb{R}^-)$ and $u - 1 \in H^2(\mathbb{R}^+)$. Using Eq. (5.3), we conclude that $u'''' \in L^2(\mathbb{R})$ which in turn by interpolation implies $u - 1 \in H^4(\mathbb{R}^+)$ and $u + 1 \in H^4(\mathbb{R}^-)$. The limits at $\pm\infty$ for u'' and u''' follow now easily.

Notice that the above additional assumptions hold if -1 and $+1$ are nondegenerate minima of f.

2. Theorem 5.7 can be extended to a nonsymmetric functional, assuming for instance that for some $k_1, k_2 > 0$, $\beta_1, \beta_2 \in \mathbb{R}$,

$$f(u) \geqslant k_1(u-1)^2, \quad g(u) \geqslant -\beta_1 \quad \text{for } u \geqslant 0,$$

$$f(u) \geqslant k_2(u+1)^2, \quad g(u) \geqslant -\beta_2 \quad \text{for } u < 0$$

and $\beta_i < \sqrt{8k_i}$ for $i = 1, 2$. However, the proof is much more involved and requires the cutting method we introduce in the following section.

3. A sharp estimate of the critical parameter β_0 introduced in Theorem 5.11 is still missing. A very rough estimate can be obtained via Lemma 5.6 and an estimate of the best constant for the interpolation inequality used therein. However, this estimate is far from the numerical observation of van den Berg who claims $\beta_0 \approx 0.92$, see Conjecture 7 in [87].

4. Let us define $\mathcal{E}_0^+$ as the cone of positive functions in $\mathcal{E}^+$, i.e.

$$\mathcal{E}_0^+ := \{u \in \mathcal{E}^+ \mid u(x) > 0 \text{ for } x > 0\}$$

and $\mathcal{F}_\beta^+ : \mathcal{E}^+ \to \mathbb{R}$ the restricted functional

$$\mathcal{F}_\beta^+(u) = \int_0^{+\infty} \left(\frac{1}{2}\left(u''^2 - \beta u'^2\right) + \frac{1}{4}\left(u^2 - 1\right)^2\right) \mathrm{d}x.$$

It follows from Theorem 5.11 that $\mathcal{F}_\beta^+$ has a minimizer in $\mathcal{E}^+$ for every $\beta \leqslant \beta_0$. Moreover this minimizer belongs to the interior of $\mathcal{E}_0^+$. Indeed, as mentioned in Section 4.1, the Hamiltonian

$$H\left(u, u', u'', u'''\right) = u'''u' - \frac{1}{2}u''^2 - \frac{\beta}{2}u'^2 + \frac{(u^2 - 1)^2}{4}$$

is constant along solutions. As for any minimizer φ, $H(\varphi, \varphi', \varphi'', \varphi''') = 0$ and $\varphi''(0) = 0$, we deduce from the conservation of the Hamiltonian that $\varphi'(0) > 0$. On the other hand, we deduce from Step 2 of the proof of Theorem 5.7 that

$$\inf_{\mathcal{E}_0^+} \mathcal{F}_\beta^+ > -\infty$$

for every $\beta < \sqrt{2}$. Hence it is natural to ask if $\mathcal{F}_\beta^+$ has a local minimizer in the interior of $\mathcal{E}_0^+$ for $\beta_0 \leqslant \beta < \sqrt{2}$. To our knowledge, this question remains open. However, we can observe that an argument of continuity implies that the answer is positive at least for β close to β_0. Indeed, it is easily seen that $\mathcal{F}_\beta^+$ has a local minimizer in $\mathcal{E}_0^+$ for every $\beta < \sqrt{2}$. For $\beta \leqslant \beta_0$, we can prove that the minimizers cannot be on the boundary of $\mathcal{E}_0^+$ so that

$$\mathcal{F}_{\beta_0}^+(\varphi_{\beta_0}) < \inf_{\partial\mathcal{E}_0^+} \mathcal{F}_{\beta_0}^+, \tag{5.20}$$

where φ_{β_0} is a minimizer of $\mathcal{F}_{\beta_0}^+$. On the other hand, we have for $\beta_1 < \sqrt{2}$ and $\beta \in [\beta_0, \beta_1]$,

$$\mathcal{F}_\beta^+(u) = \mathcal{F}_{\beta_1}^+(u) + \frac{\beta_1 - \beta}{2} \int_0^{+\infty} u'^2 \, \mathrm{d}x \geqslant -C + \frac{\beta_1 - \beta}{2} \int_0^{+\infty} u'^2 \, \mathrm{d}x.$$

We now infer from this last estimate that any function $u \in \mathcal{E}_0^+$ satisfying

$$\mathcal{F}_\beta^+(u) = \inf_{\partial\mathcal{E}_0^+} \mathcal{F}_\beta^+$$

is a priori bounded in $\mathcal{E}^+$ for β in any compact subinterval of $[\beta_0, \beta_1[$. It follows that the inequality (5.20) holds true for β in a right neighborhood of β_0, i.e.

$$\mathcal{F}_\beta^+(\varphi_{\beta_0}) < \inf_{\partial\mathcal{E}_0^+} \mathcal{F}_\beta^+.$$

Since φ_{β_0} is in the interior of $\mathcal{E}_0^+$, this means that the infimum of $\mathcal{F}_\beta^+$ in $\mathcal{E}_0^+$ cannot be achieved on the boundary.

6. A Ginzburg–Landau model for ternary mixtures: Connections between nonconsecutive equilibria

A good example of phase transition phenomena is provided by the mixing–demixing transitions of a fluid. It is well known that oil and water do not mix so that a binary fluid composed of oil and water has two separated homogeneous phases usually called oil-rich and water-rich phases. Assuming that the essence of the transitions can be described in terms of the concentration difference between oil and water, we introduce a scalar order parameter ϕ which locally measures this difference. A Ginzburg–Landau model [36] then yields for the free energy a functional of the form

$$\mathcal{F}(\phi) = \int_{R^3} \bigl(g_0(\nabla\phi)^2 + f(\phi) - \mu\phi\bigr)\,\mathrm{d}r. \tag{6.1}$$

The function f is the free-energy density and μ is the chemical potential difference between oil and water. The free-energy density f is usually approximated by an even (due to the symmetry under the interchange of the two components) fourth order polynomial

$$f(\phi) = a_2\phi^2 + a_4\phi^4.$$

Thermodynamic stability of homogeneous phases requires that $a_4 > 0$. When $a_2 < 0$, there is coexistence of water-rich and oil-rich phases at $\mu = 0$, i.e. the two phases are at the same level of energy.

The addition of an amphiphile into the fluid can provoke the formation of wealth of complex self-assembled structures. An amphiphilic molecule consists of a polarizable head which prefers the highly polarizable water environment and a hydrocarbon tail which prefers the oil. The energy of the amphiphile is lowest when it can find or create surfaces between oil and water at which it can adsorb the other two components into various structures. Adding a small amount of amphiphile to the fluid will not in general modify the two-phase coexistence. The added amphiphile will partition itself between the two phases. On the other hand, if for example the amphiphile is more present in the water phase, a third phase can be made to appear by changing an external field such as the temperature or the chemical potential of a fourth component. This new phase contains more of the amphiphile and less of water and oil than the other two. Its density will be intermediate between that of the water-rich and the oil-rich phases and will be therefore physically located between them. Hence it is called the middle-phase.

Let us come back for a while to the free energy functional (6.1) of the binary fluid oil-water. Thermodynamic quantities and correlation functions can be obtained by approximation methods in which the functional $\mathcal{F}$ is minimized in the space of functions which spatially connect the two phases. The minimizer ϕ is then the interfacial profile between the oil-rich and the water-rich phase at coexistence.

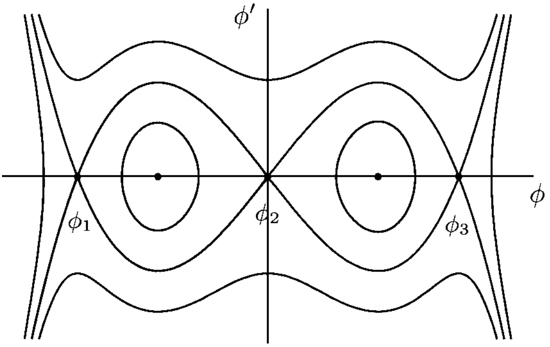


Fig. 3. The 3-stable system phase plane.

If we want to adapt the Ginzburg–Landau model (6.1) to describe the oil-water-amphiphile mixtures and keep a single order parameter which still measures the density difference between oil and water, we need to consider three-phase coexistence at $\mu = 0$ so that f must have three minima at $\phi_1 < \phi_2 < \phi_3$ corresponding respectively to the bulk phases oil-rich, middle and water-rich. Choosing their common value to be zero and assuming only one spatial dependence, the interfacial profile, $\phi(x)$, between the oil-rich phase which extends to $x \to -\infty$ and the water-rich phase which extends to $x \to +\infty$, then solves the equation

$$2g_0\phi'' - f'(\phi) = 0 \tag{6.2}$$

and the first integral

$$g_0\phi'(x)^2 - f\big(\phi(x)\big) = 0. \tag{6.3}$$

From a simple phase-plane analysis, see Fig. 3, it is easily seen that the only trajectory starting form ϕ_1 at $x \to -\infty$ and going to ϕ_3 at $x \to +\infty$ spends an infinite amount of time in ϕ_2, hence Eq. (6.2) has no solution ϕ satisfying $\phi(-\infty) = \phi_1$ and $\phi(+\infty) = \phi_3$. In the ternary mixture problem, this corresponds to an infinite thickness of middle phase, ϕ_2, between the oil-rich and water-rich phases. In other words, the model predicts the middle phase will always wet the interface between the oil- and water-rich phases, a prediction contrary to the results of experiment.

A simple way proposed by Gompper and Schick [36] to overcome this consequence of the model and yet consider a scalar order-parameter theory is to add a second order term in the Lagrangian, considering therefore the functional

$$\mathcal{F}(\phi) = \int_{R^3} \big(c\big(\nabla^2\phi\big)^2 + g(\phi)(\nabla\phi)^2 + f(\phi) - \mu\phi\big)\,\mathrm{d}r.$$

The function g, which quantifies the properties of the amphiphile, is negative close to the middle-phase as it tends to create interfaces and positive in the oil and water phases. The parameter c is positive and stabilizes the system.

The aim of this section is to prove that this last functional does not suffer the defect of the classical Ginzburg–Landau model (6.1) at least under some hypotheses on the nature of the middle-phase equilibrium. In the sequel, we only consider a scalar order-parameter ϕ with one spatial direction.

We consider a potential $f \in C^2(\mathbb{R})$ and a function $g \in C^2(\mathbb{R})$ satisfying (C1), (C4) and

(C2′) for some $0 < a < 1/2$ and $\alpha > 0$,

$$\frac{f(u)}{(u-1)^2} \leqslant \alpha, \quad \text{for } |u-1| < a,$$

$$\frac{f(u)}{(u+1)^2} \leqslant \alpha, \quad \text{for } |u+1| < a;$$

(C3′) $f(u) = 0$ if and only if $u = 0$ or $u = \pm 1$.

We also introduce the additional condition

$$g(0)^2 < 4f''(0).$$

This last condition implies that the trivial solution is a saddle-focus equilibrium for the linear equation

$$u'''' - g(0)u'' + f''(0)u = 0.$$

As observed in Section 5.1, if $g(0) < 0$ or (C1) holds with $g = \tilde{g}$, we just need f to be C^2 in a neighborhood of 0 and $f''(0) \neq 0$ as the inequality $g(0)^2 < 4f''(0)$ then holds.

Under the previous assumptions, it makes sense to consider the functional

$$\mathcal{F}(u) = \int_{-\infty}^{+\infty} \left(\frac{1}{2}\big(u''^2 + g(u)u'^2\big) + f(u) \right) \mathrm{d}x \tag{6.4}$$

in the space

$$\mathcal{H} := \left\{ u \in C^1(\mathbb{R}),\ u'' \in L^2(\mathbb{R}),\ u' \in L^\infty(\mathbb{R}) \text{ and } \lim_{x \to \pm\infty} u(x) = \pm 1 \right\}$$

and the associated Euler–Lagrange equation is given by

$$u'''' - g(u)u'' - \frac{1}{2}g'(u)u'^2 + f'(u) = 0. \tag{6.5}$$

Arguing as in Section 5, we observe that the functional (6.4) is bounded from below in $\mathcal{H}$.

LEMMA 6.1. *If f, $g \in C(\mathbb{R})$ satisfy* (C1), *then there exists a constant $s > 0$ such that for all $u \in \mathcal{H}$*

$$\mathcal{F}(u) \geqslant s \int_{-\infty}^{+\infty} \left(\frac{u''^2}{2} + f(u) \right) \mathrm{d}x.$$

The main theorem of the section goes as follows.

THEOREM 6.2. *Suppose that* $f, g \in C^2(\mathbb{R})$ *satisfy* (C1), (C2′), (C3′) *and* (C4). *If* $g(0)^2 < 4f''(0)$, *there exists a minimizer* u *of* $\mathcal{F}$ *in* $\mathcal{H}$ *which is a solution of* (6.5).

In comparison with Theorem 5.2, the additional condition $g(0)^2 < 4f''(0)$ allows to consider a potential f with a third bottom at 0. We do not know if Theorem 6.2 holds without this assumption.

The idea of the proof of Theorem 6.2 is identical to that of Theorem 5.2. We thus search for a control on the time it takes for a quasi-minimizer to travel from a neighborhood of $(-1, 0)$ to a neighborhood of $(+1, 0)$ in the uu'-plane. Observe that the presence of the third zero of the potential rules out some of the arguments used in the proof of Theorem 5.2. On the other hand, we easily obtain a control on time intervals from $(-1, 0)$ to $(0, 0)$ and $(0, 0)$ to $(+1, 0)$. To complete the arguments, we just need a control on the time quasi-minimizers spend close to 0. This control is obtained thanks to the additional condition that 0 is a saddle-focus. Basically, we show in Section 6.1 that due to the nature of the equilibrium, the local minimizers of (6.4) close to 0 (in the phase space) change sign in every interval of length larger than a fixed constant. We then introduce in Section 6.2 a *clipping* procedure which roughly speaking ensures that the quasi-minimizers are monotonic close to 0. This technique was first used by Kalies et al. [42]. The local analysis close to 0 then implies that quasi-minimizers do not spend much time close to 0. We give the proof of Theorem 6.2 in Section 6.3.

6.1. *Local analysis of a saddle-focus equilibrium*

We now focus on the local minimizers close to a saddle-focus equilibrium. To fix the ideas and to simplify the notation, we assume that f is a potential for which 0 is a nondegenerate global minimum and g is such that 0 is a saddle focus equilibrium of the linear equation

$$u'''' - g(0)u'' + f''(0)u = 0, \tag{6.6}$$

i.e. $g(0)^2 < 4f''(0)$. Basically, the following lemma shows that the minimizers of the functional

$$\mathcal{F}_{[a,b]}(u) := \int_a^b \left(\frac{1}{2}\left(\left(u''^2\right) + g(u)u'^2\right) + f(u)\right) \mathrm{d}x \tag{6.7}$$

on the set

$$\mathcal{H}_{a,b} := \left\{v \in H^2(a, b) \;\middle|\; \left(u(a), u'(a)\right) = y_0 \text{ and } \left(u(b), u'(b)\right) = y_1\right\}$$

are small (for the C^3-norm) whenever y_0 and y_1 are small. It then follows that the oscillatory behaviour of the solutions of the linearization (6.6) around the equilibrium extends to these minimizers. The following lemma is adapted from Theorem 4.1 [42] where it is assumed that g is positive.

LEMMA 6.3. *Let f and $g \in C^2(\mathbb{R})$ be such that $f(u) \geqslant 0$ for all $u \in \mathbb{R}$, $f(0) = f'(0) = 0$ and assume* (C1) *holds. Assume moreover that $f''(0) \neq 0$. Then, there exist $\delta_0 > 0$ and $S > 1$ such that if $b - a \geqslant 1$, $\|y_0\| \leqslant \delta_0$, $\|y_1\| \leqslant \delta_0$ and u minimizes $\mathcal{F}_{[a,b]}$ on $\mathcal{H}_{a,b}$, we have*

$$\|u\|_{C^3([a,b])} \leqslant S \max\bigl(\|y_0\|, \|y_1\|\bigr). \tag{6.8}$$

PROOF. As 0 is a nondegenerate minimum of f, there exist $\delta_0 > 0$, $\eta > 0$ and $\zeta > 0$ such that $|f'(u)| \leqslant 2\eta|u|$ and $\zeta u^2 \leqslant f(u) \leqslant \eta u^2$ for $|u| \leqslant \delta_0$. Notice also that using integration by parts and arguing as in Lemma 5.1, we see that there exists $s > 0$ such that for any function $u \in H^2(a, b)$,

$$\mathcal{F}_{[a,b]}(u) \geqslant s \int_a^b \left(\frac{u''^2}{2} + f(u)\right) \mathrm{d}x + \frac{1}{2}\bigl(\tilde{G}\bigl(u(b)\bigr)u'(b) - \tilde{G}\bigl(u(a)\bigr)u'(a)\bigr), \tag{6.9}$$

so that $\mathcal{F}_{[a,b]}$ is bounded from below on $\mathcal{H}_{a,b}$. The existence of a minimizer follows by standard arguments. Moreover, if u is a minimizer, u solves (6.5) on $[a, b]$ and satisfies the boundary conditions $(u(a), u'(a)) = y_0$ and $(u(b), u'(b)) = y_1$.

CLAIM 1. *There exists $C_1 > 0$ and $\delta_1 > 0$ such that if $\|y_0\| \leqslant \delta \leqslant \delta_1$ and $\|y_1\| \leqslant \delta \leqslant \delta_1$, then $\inf_{\mathcal{H}_{a,b}} \mathcal{F}_{[a,b]} \leqslant C_1\delta^2$.*

Define $P(x)$ as follows:

$$P(x) = \begin{cases} P_0(x) & \text{if } a \leqslant x \leqslant a + \frac{1}{2}, \\ 0 & \text{if } a + \frac{1}{2} < x \leqslant b - \frac{1}{2}, \\ P_1(x) & \text{if } b - \frac{1}{2} < x \leqslant b, \end{cases}$$

where P_i, $i = 0, 1$, are the third degree polynomials satisfying $(P_0(a), P_0'(a)) = y_0$, $(P_1(b), P_1'(b)) = y_1$ and $(P_0(a + \frac{1}{2}), P_0'(a + \frac{1}{2})) = (P_1(b - \frac{1}{2}), P_1'(b - \frac{1}{2})) = (0, 0)$. Observe that there exists $0 < \delta_1 \leqslant \delta_0$ such that if $\|y_0\| \leqslant \delta \leqslant \delta_1$ and $\|y_1\| \leqslant \delta \leqslant \delta_1$, then $\|P\|_{L^\infty} \leqslant \delta_0$. It then follows from an easy computation that

$$\inf_{\mathcal{H}_{a,b}} \mathcal{F}_{[a,b]} \leqslant \mathcal{F}_{[a,b]}(P) \leqslant C_1\delta^2,$$

where $C_1 > 0$ essentially depends on η and $\|g\|_{L^\infty(-\delta_0,\delta_0)}$.

CLAIM 2. *There exists $\delta_2 > 0$ such that if u is a minimizer in $\mathcal{H}_{a,b}$ with $\|y_0\| \leqslant \delta_2$ and $\|y_1\| \leqslant \delta_2$, then $\|u\|_\infty \leqslant \delta_0$.*

The ideas we use to prove the claim are already included in the proof of Proposition 5.3. Observe that the minimizers of $\mathcal{F}_{[a,b]}$ in $\mathcal{H}_{a,b}$ are a priori bounded in $C^1([a, b])$. Indeed, this follows easily arguing as in the first part of the proof of Proposition 5.3. We denote by L the C^1-bound. Consider next the set

$$\mathcal{D}_{a,b} := \{u \in \mathcal{H}_{a,b} \mid \text{ for all } x \in [a,b], |u(x)| \leqslant \delta_0\}.$$

Let us fix the notation $N_{\delta_0} := [-\delta_0, \delta_0]$. We then define

$$\mu := \min\{f(u) \mid u \in N_{\delta_0} \setminus N_{\delta_0/2}\} > 0,$$
$$\nu := \max\{|\tilde{G}(u)| \mid u \in N_{\delta_0}\} \geqslant 0.$$

As in Part 2—Step 2 of the proof of Proposition 5.3, we can derive a lower estimate on the action of functions $u \in \mathcal{H}_{a,b}$ whose graphs do not stay in the strip $[a,b] \times N_{\delta_0}$. Indeed, suppose $u \in \mathcal{H}_{a,b} \setminus \mathcal{D}_{a,b}$ minimizes $\mathcal{F}_{[a,b]}$ in $\mathcal{H}_{a,b}$ with $\|y_0\| \leqslant \delta \leqslant \delta_1$ and $\|y_1\| \leqslant \delta \leqslant \delta_1$. Then

$$\mathcal{F}_{[a,b]}(u) \geqslant \frac{s\mu\delta_0}{2L} - \nu\delta. \tag{6.10}$$

On the other hand, we infer from Claim 1 that

$$\mathcal{F}_{[a,b]}(u) \leqslant C_1\delta^2.$$

Choosing $\delta_2 > 0$ small enough and $0 < \delta \leqslant \delta_2$, the estimate (6.10) yields a contradiction so that $u \in \mathcal{D}_{a,b}$ and the claim follows.

CLAIM 3. *There exists $C_2 > 0$ such that if u is a minimizer in $\mathcal{H}_{a,b}$ with $\|y_0\| \leqslant \delta \leqslant \delta_2$ and $\|y_1\| \leqslant \delta \leqslant \delta_2$, then $\|u\|_{C^3} \leqslant C_2\delta$.*

Assuming that $\|y_0\| \leqslant \delta \leqslant \delta_2$ and $\|y_1\| \leqslant \delta \leqslant \delta_2$, we deduce from the estimate (6.9) that

$$\|u\|_{L^2} \leqslant C\delta \quad \text{and} \quad \|u''\|_{L^2} \leqslant C\delta.$$

Indeed, the L^2-bound for u'' follows easily while for the bound on u, it is sufficient to observe that f is bounded from below by the parabola ζu^2 for $|u| \leqslant \delta_0$. We now deduce by interpolation that

$$\|u'\|_{L^2} \leqslant C(\|u\|_{L^2} + \|u''\|_{L^2}) \leqslant C\delta.$$

The constant C can be chosen independent of the length $b-a$ as far as $b-a \geqslant 1$, see [1]. Using the continuous injection of $H^2(a,b)$ into $C^1([a,b])$, we deduce that

$$\|u\|_{C^1} \leqslant C\delta.$$

Observe that we can still choose a constant C that does not depend on the length of $[a,b]$ as $b-a \geqslant 1$. The differential equation (6.5) then yields the estimate

$$\|u''''\|_{L^2} \leqslant C\delta$$

so that by interpolation we also have

$$\|u'''\|_{L^2} \leqslant C\delta.$$

Now, the bound in C^3 follows from the bound in H^4 and the continuous injection in C^3. □

LEMMA 6.4. *Suppose that the assumptions of Lemma* 6.3 *hold and assume furthermore that* $g(0)^2 < 4f''(0)$. *Then there exist* $\delta_0 > 0$ *and* $\tau_0 > 0$ *such that if* $b - a \geqslant 1$, $\|y_0\| \leqslant \delta_0$, $\|y_1\| \leqslant \delta_0$, $\max(\|y_0\|, \|y_1\|) > 0$ *and* u *minimizes* $\mathcal{F}_{[a,b]}$ *on* $\mathcal{H}_{a,b}$, u *changes sign on every subinterval of* $[a, b]$ *having length larger than* τ_0.

The proof is included in Theorem 4.2 in [42]. A slight modification of Lemma 5.9 also leads to the conclusion. It is easily seen that the assumption on the convergence of the solution to the equilibrium in Lemma 5.9 can be replaced by a condition ensuring that the solution remains close (for the C^3-norm) to 0. Here, thanks to Lemma 6.3, the smallness of the boundary conditions on u and u' suffices to obtain such a control.

6.2. *Clipping*

Next, we recall the *clipping* procedure as introduced in [42]. When minimizing a functional in a certain space, we often want to be able to modify locally any function by another one being in the same space, having better properties and lower action. When dealing with a second order equation and its associated functional, we usually only have to worry about keeping functions continuous so that pieces of graph can be easily discarded. As our functional $\mathcal{F}$ requires square integrable second order derivative, things are more complicated. For example, any modification has to keep the functions at least C^1. Let us describe admissible cutoffs.

DEFINITION 6.5. Let $u \in C^1[a, b]$. If $u(\alpha) = u(\beta)$ and $u'(\alpha) = u'(\beta)$ for some $\alpha < \beta$ in $[a, b]$, we say that the interval $[\alpha, \beta]$ can be *clipped out* meaning that we can define a C^1 function $\hat{u}$ on the interval $[a, b - (\beta - \alpha)]$ which coincides with u and the $\beta - \alpha$ translate of $u_{|[\beta,b]}$ respectively on the intervals $[a, \alpha]$ and $[\alpha, b - (\beta - \alpha)]$. The function $\hat{u}$ is defined by

$$\hat{u}(x) := \begin{cases} u(x) & \text{if } x \in [a, \alpha], \\ u(x + \beta - \alpha) & \text{if } x \in \bigl[\alpha, b - (\beta - \alpha)\bigr]. \end{cases}$$

We say that $\hat{u}$ is a *clip* of u.

Observe that in case the Lagrangian

$$L_g\bigl(u, u', u''\bigr) = \frac{1}{2}\bigl(u''^2 + g(u)u'^2\bigr) + f(u)$$

is positive, the clipping process has the nice property that if $u \in \mathcal{H}$ and $\hat{u}$ is a clip of u, then

$$\mathcal{F}(\hat{u}) \leqslant \mathcal{F}(u).$$

When $L_g(u, u', u'')$ changes sign, this is no more true. However, we show below in Lemma 6.7 that under the assumption (C1), the clipping process decreases the action.

The following lemma gives the basic tool for clipping functions. It generalizes Lemma 3.1 in [42].

LEMMA 6.6. *Let $s_1 < s_2 < s_3 < s_4$ and let $u \in C^1(s_1, s_4)$ be such that*

$$u(s_1) = u(s_3), \qquad u(s_2) = u(s_4), \qquad \bigl(u'(s_1) - u'(s_3)\bigr)\bigl(u'(s_2) - u'(s_4)\bigr) \leqslant 0.$$

Assume, moreover, that

$$u(s_1) < u(s) < u(s_2) \quad \textit{for all } s \in]s_1, s_2[,$$

and

$$u(s_3) < u(s) < u(s_4) \quad \textit{for all } s \in]s_3, s_4[.$$

Then there exist $\alpha \in [s_1, s_2]$, $\beta \in [s_3, s_4]$ such that the interval $[\alpha, \beta]$ can be clipped out.

PROOF. Consider the set

$$E := \bigl\{(x, y) \in [s_1, s_2] \times [s_3, s_4] \bigm| u(x) = u(y)\bigr\}.$$

It follows from degree arguments that there exists a connected set $H \subset E$ that contains (s_1, s_3) and (s_2, s_4). Next we define the function $\varphi : H \to \mathbb{R}$ by

$$\varphi(x, y) = u'(x) - u'(y).$$

Since

$$\varphi(s_1, s_3)\varphi(s_2, s_4) \leqslant 0,$$

the continuity of φ leads to the conclusion. □

A typical example where Lemma 6.6 applies is displayed in Fig. 4.

As we already mentioned, the clipping procedure plays a key role to control and modify minimizing sequences. We therefore need to check that any clip has a lower action than the original function. This is the case when (C1) holds.

LEMMA 6.7. *Suppose that $f \in C(\mathbb{R})$ is a nonnegative function and $g \in C(\mathbb{R})$ satisfies* (C1). *If $u \in H^2(a, b)$ and $\hat{u} \in H^2(a, b - (\beta - \alpha))$ is a clip of u, then*

$$\mathcal{F}_{[a, b-(\beta-\alpha)]}(\hat{u}) \leqslant \mathcal{F}_{[a,b]}(u),$$

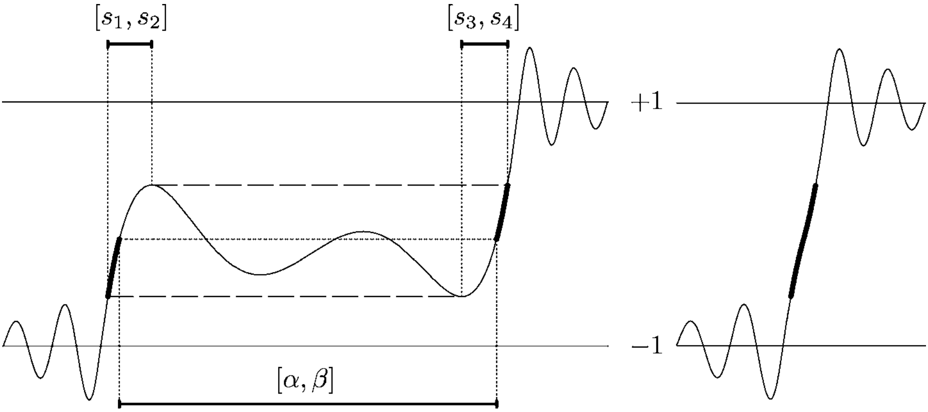


Fig. 4. *The clipping process*. Lemma 6.6 applies with $s_1 < s_2 < s_3 < s_4$ so that a subinterval $[\alpha, \beta]$ can be clipped out.

where $\mathcal{F}_{[a,b]}$ *is defined by* (6.7).

PROOF. We compute

$$\mathcal{F}_{[a,b-(\beta-\alpha)]}(\hat{u}) = \mathcal{F}_{[a,\alpha]}(u) + \mathcal{F}_{[\beta,b]}(u) = \mathcal{F}_{[a,b]}(u) - \mathcal{F}_{[\alpha,\beta]}(u)$$

and arguing as in Lemma 5.1, we infer that

$$\mathcal{F}_{[\alpha,\beta]}(u) \geqslant s \int_{\alpha}^{\beta} \left(\frac{u''^2}{2} + f(u) \right) \mathrm{d}x + \frac{1}{2}\big(\tilde{G}\big(u(\beta)\big)u'(\beta) - \tilde{G}\big(u(\alpha)\big)u'(\alpha)\big).$$

Since $u(\alpha) = u(\beta)$, $u'(\alpha) = u'(\beta)$, we obtain the inequality

$$\mathcal{F}_{[\alpha,\beta]}(u) \geqslant 0$$

so that the result follows. □

6.3. *Existence of a minimizer*

As for Theorem 5.2, the proof of Theorem 6.2 relies on the control and the localization of a minimizing sequence.

PROPOSITION 6.8. *Suppose that* $f, g \in C^2(\mathbb{R})$ *satisfy* (C1), (C2′), (C3′) *and* (C4). *Then there exists* $L > 0$, $T > 0$ *and a sequence* $(u_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ *such that* $\mathcal{F}(u_n) \to \inf_{\mathcal{H}} \mathcal{F}$ *and for all* $n \in \mathbb{N}$,

(i) $\|u_n\|_{C^1} \leqslant L$,

(ii) $|u_n(x)+1| \leqslant a$ *for all* $x \leqslant -T$ *and* $|u_n(x)-1| \leqslant a$ *for all* $x \geqslant T$.

The proof of property (i) follows exactly the lines of the proof of Proposition 5.3, Part 1. We divide the proof of the second statement in three main steps. In Step 1, we approximate the minimum of $\mathcal{F}$ with functions that stay close to ± 1 at $\pm\infty$. Step 2 estimates the time for a function $u \in \mathcal{H}$ to travel in the (u, u')-plane from neighborhoods of $(\pm 1, 0)$ to a neighborhood of $(0, 0)$. Finally, in Step 3, we show that functions in $\mathcal{H}$ that stay close to 0 with small velocity can be replaced, using the clipping procedure, by functions that spend in such a neighborhood a time which is a priori bounded.

PROOF. The proof of the first statement has been worked out in the proof of Proposition 5.3. We therefore restrict our attention to the proof of the conclusion (ii).

Let $(u_n)_n \subset \mathcal{H}$ be a minimizing sequence for $\mathcal{F}$. We can assume without loss of generality that $\mathcal{F}(u_n) \leqslant C \leqslant \inf_{\mathcal{H}} \mathcal{F} + 1$. Let $\varepsilon > 0$ be fixed and such that $S\varepsilon \leqslant \delta_0$, where $\delta_0 > 0$ and $S > 1$ are given by Lemma 6.3 and Lemma 6.4. For each $n \in \mathbb{N}$, we define

$$x_1 := \sup\bigl\{x \mid \bigl|u_n(x)+1\bigr| \leqslant \varepsilon \text{ and } \bigl|u_n'(x)\bigr| \leqslant \varepsilon\bigr\},$$

and

$$x_4 := \inf\bigl\{x \mid \bigl|u_n(x)-1\bigr| \leqslant \varepsilon \text{ and } \bigl|u_n'(x)\bigr| \leqslant \varepsilon\bigr\}.$$

Step 1. Modification of u_n in $]-\infty, x_1]$ and $[x_4, \infty[$. For all $n \in \mathbb{N}$, there exists a function $v_n \in \mathcal{H}$ such that for all $x \leqslant x_1$, $|v_n(x)+1| \leqslant a$, for all $x \geqslant x_4$, $|v_n(x)-1| \leqslant a$ and $\mathcal{F}(v_n) \leqslant \mathcal{F}(u_n)$.

This step is similar to Part 2—Step 2 of the proof of Proposition 5.3 so that we skip it.

Step 2. Estimates on time intervals. Define for all $n \in \mathbb{N}$,

$$x_2 := \inf\bigl\{x \geqslant x_1 \mid \bigl|v_n(x)\bigr| \leqslant \varepsilon \text{ and } \bigl|v_n'(x)\bigr| \leqslant \varepsilon\bigr\}$$

and

$$x_3 := \sup\bigl\{x \leqslant x_4 \mid \bigl|v_n(x)\bigr| \leqslant \varepsilon \text{ and } \bigl|v_n'(x)\bigr| \leqslant \varepsilon\bigr\}.$$

Notice that x_2 and x_3 need not exist. Arguing as in Part 2—Step 1 of the proof of Proposition 5.3, we obtain a bound T_1 on $x_2 - x_1$ and $x_4 - x_3$, or on $x_4 - x_1$ if x_2 and x_3 do not exist.

Step 3. Modification in $[x_2, x_3]$. If x_2 and x_3 do not exist, this step can of course be skipped.

CLAIM 1. *For all* $n \in \mathbb{N}$, *there exists a function* w_n *which minimizes* $\mathcal{F}$ *in*

$$\mathcal{H}_{2,3} := \bigl\{w \in \mathcal{H} \mid w = v_n \text{ on } \mathbb{R} \setminus [x_2, x_3]\bigr\}.$$

Indeed, we can replace $v_n$by a minimizer of $\mathcal{F}$ in $\mathcal{H}_{2,3}$. Such a function w_n exists and satisfies $\mathcal{F}(w_n) \leqslant \mathcal{F}(v_n)$. Further, it solves the differential equation (6.5) on $[x_2, x_3]$ together with the boundary conditions

$$w_n(x_2) = v_n(x_2), \quad w_n'(x_2) = v_n'(x_2), \quad w_n(x_3) = v_n(x_3), \quad w_n'(x_3) = v_n'(x_3).$$

CLAIM 2. *For each $n \in \mathbb{N}$, if $x_3 - x_2 \geqslant \max(1, 8\tau_0)$, where τ_0 is given from Lemma* 6.4, *the function w_n can be replaced by a clip $\hat{w}_n \in \mathcal{H}$ so that after clipping the interval* $[x_1, x_4]$ *is transformed into an interval of length smaller than* $(x_2 - x_1) + \max(1, 8\tau_0) + (x_4 - x_3)$. *Further, we have $\mathcal{F}(\hat{w}_n) \leqslant \mathcal{F}(w_n)$.*

Define

$$x_2' := \max\bigl\{x \leqslant x_2 \bigm| |w_n(x)| = \delta_0\bigr\} \quad \text{and} \quad x_3' := \min\bigl\{x \geqslant x_3 \bigm| |w_n(x)| = \delta_0\bigr\}.$$

It follows from Claim 1 and Lemma 6.3 that $|w_n(x)| \leqslant S\varepsilon \leqslant \delta_0$ for all $x \in [x_2, x_3]$ and therefore also for all $x \in [x_2', x_3']$.

Suppose first that $w_n(x_2') = -\delta_0$ and $w_n(x_3') = \delta_0$. Define

$$\begin{aligned}
s_2 &:= \min\bigl\{x \in [x_2', x_3'] \bigm| w_n'(x) = 0 \text{ and } w_n(x) \geqslant 0\bigr\},\\
s_4 &:= \max\bigl\{x \in [x_2', x_3'] \bigm| w_n(x) = w_n(s_2)\bigr\},\\
s_3 &:= \max\bigl\{x \in [s_2, s_4] \bigm| w_n'(x) = 0\bigr\}
\end{aligned}$$

and take

$$s_1 := \max\bigl\{x \in [x_2', s_2] \bigm| w_n(x) = w_n(s_3)\bigr\}.$$

Observe that Lemma 6.4 ensures the existence of s_2 and moreover $s_2 \in [x_2', x_2 + 2\tau_0]$. Also, $s_3 \in [x_3 - 2\tau_0, x_3']$. Further, we can apply Lemma 6.6 to the function w_n on $[s_1, s_4]$ and discard the restriction of w_n to some interval $[\alpha, \beta]$ containing $[s_2, s_3]$. We denote by $\hat{w}_n$ the clip of w_n. Namely, we define

$$\hat{w}_n(x) := \begin{cases} w_n(x) & \text{if } x \in [a, \alpha],\\ w_n(x + \beta - \alpha) & \text{if } x \in \bigl(\alpha, b - (\beta - \alpha)\bigr). \end{cases}$$

Letting $x_4^* := x_4 - (\beta - \alpha)$, we have

$$x_4^* - x_1 \leqslant (x_4 - x_3) + 4\tau_0 + (x_2 - x_1).$$

If $w_n(x_2') = \delta_0$ and $w_n(x_3') = -\delta_0$, we use the same argument.

Assume now that $w_n(x_2') = -\delta_0$ and $w_n(x_3') = -\delta_0$. Let $q \in [x_2', x_3']$ be such that $\max_{x \in [x_2', x_3']} w_n(x) = w_n(q)$. Notice that Lemma 6.4 implies $w_n(q) > 0$ and hence we can apply the preceding argument to each of the intervals $[x_2', q]$ and $[q, x_3']$. Here, denoting by x_4^{**} the point into which x_4 is transformed, we have after clipping, $x_4^{**} - x_1 \leqslant (x_4 - x_3) + 8\tau_0 + (x_2 - x_1)$.

The case $w_n(x_2') = \delta_0$ and $w_n(x_3') = \delta_0$ is handled similarly.

Conclusion of Step 3. For each $n \in \mathbb{N}$, there exists a function $\hat{w}_n \in \mathcal{H}$ ($\hat{w}_n = w_n$ if $x_3 - x_2 \leqslant \max(1, 8\tau_0)$) which satisfies $\mathcal{F}(\hat{w}_n) \leqslant \mathcal{F}(w_n)$ and

$$\begin{aligned} \left|\hat{w}_n(x) + 1\right| &\leqslant a \quad \text{for all } x \leqslant x_1, \\ \left|\hat{w}_n(x) - 1\right| &\leqslant a \quad \text{for all } x \geqslant x_4', \end{aligned}$$

for some $x_4' \in \mathbb{R}$ such that

$$x_4' - x_1 \leqslant 2T_1 + \max(1, 8\tau_0),$$

where T_1 is defined in Step 2.

Conclusion. Taking

$$T = T_1 + \frac{1}{2}\max(1, 8\tau_0)$$

and using a time-translation if necessary, we finally construct a minimizing sequence $(z_n)_n \subset \mathcal{H}$ that satisfies the second statement of Proposition 6.8. Notice that in case x_2 and x_3 are not defined, we can choose

$$T = \frac{T_1}{2} \geqslant \frac{1}{2}(x_4 - x_1). \qquad \square$$

Now that we have at hand a minimizing sequence satisfying Proposition 6.8, the proof of Theorem 6.2 follows the lines of the proof of Theorem 5.2.

6.4. *Notes and further comments*

1. As discussed in the first note of Section 5, the minimizer obtained in Theorem 6.2 is a heteroclinic solution in a weak sense.

2. Other Ginzburg–Landau models for ternary mixtures have been proposed in the literature, including two-order-parameter and three-order-parameter models, see [36]. In these models, a second scalar parameter ψ which describes the local concentration of amphiphile is introduced. Simple cases have been studied numerically by physicists but up to our knowledge the corresponding type of functional has not been treated mathematically.

7. Multi-transition connections

We consider again the case of a bi-stable equation. We already mentioned in Section 4.3 that when ±1 are saddle-focus equilibria, the extended Fisher–Kolmogorov equation (4.3) has infinitely many solutions that can be classified according to their homotopy type. In this section, we focus on multi-transition solutions which correspond to solutions of type $g = (g_1, \dots, g_n)$, $n \in \mathbb{N}_0$, with $g_i = 2$ for all $i = 1, \dots, n$.

Considering again the functional

$$\mathcal{F}(u) = \int_{-\infty}^{+\infty} \left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right) \mathrm{d}x \tag{7.1}$$

whose Euler–Lagrange equation is given by

$$u'''' - g(u)u'' - \frac{1}{2}g'(u)u'^2 + f'(u) = 0, \tag{7.2}$$

we assume that the function f is a positive symmetric double-well potential with bottoms at ± 1 and g is an even function which is not necessarily constant. Since f and g are even, arguing as in Section 4.2, we can restrict our attention to odd solutions. We thus look at the critical points of the functional

$$\mathcal{F}^+(u) = \int_0^{+\infty} \left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right) \mathrm{d}x \tag{7.3}$$

in the space

$$\mathcal{E}^+ = \left\{u \mid u - 1 \in H^2\left(\mathbb{R}^+\right), u(0) = 0\right\}. \tag{7.4}$$

If $\mathcal{F}^+$ has a minimizer u then $u''(0) = 0$ and extending u on $\mathbb{R}$ by

$$u^*(x) = \begin{cases} -u(-x) & \text{if } x < 0, \\ u(x) & \text{if } x \geqslant 0 \end{cases} \tag{7.5}$$

we obtain an odd solution of (7.2).

We expect multi-transition solutions when the equilibria ± 1 are saddle-foci, i.e. when $g(1)^2 < 4f''(1)$. We obtain these solutions by odd extension of local minima of the functional $\mathcal{F}^+$ in appropriate subsets of $\mathcal{E}^+$. Basically, these subsets correspond to classes of functions having the desired number of transitions. We define for each $n \geqslant 0$, the subset $\mathcal{E}_n^+ \subset \mathcal{E}^+$ consisting of functions whose odd extensions on $\mathbb{R}$ make $2n + 1$ transitions. More precisely, a function $u \in \mathcal{E}^+$ belongs to the subclass $\mathcal{E}_n^+$ if there exist $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = \infty$ such that

$$u(x)(-1)^{i+n} > 0 \quad \text{for } x \in (x_i, x_{i+1}),$$
$$\max_{(x_i, x_{i+1})} u(x)(-1)^{i+n} > 1.$$

We prove that $\mathcal{F}^+$ has a local minimum in each of these subspaces in the following situation.

THEOREM 7.1. *Let f and $g \in C^2(\mathbb{R})$ satisfy* (C1), (C3) *and* (C4). *Assume moreover that $g(1)^2 < 4f''(1)$. Then, for every $n \in \mathbb{N}$, $\mathcal{F}^+$ has in each subspace $\mathcal{E}_n^+$, a local minimizer u_n whose odd extension on $\mathbb{R}$ is a heteroclinic solution of* (1.8) *having exactly* $2n+1$ *zeros.*

For a function $u \in \mathcal{E}_n^+$, we denote by I_i the intervals (x_i, x_{i+1}) for $i = 0, \dots, n$ where by convention $x_0 = 0$ and $x_{n+1} = \infty$. The main idea in the proof of Theorem 7.1 is to show the existence of a minimizing sequence $(u_p)_p \subset \mathcal{E}_n^+$ that has the following properties:

(a) there exists $I > 0$ such that for all u_p, $|I_i| \leqslant I$ for all $i = 0, \dots, n-1$;

(b) for all $\varepsilon > 0$, there exists $T > 0$ such that for all u_p, $|u_p(x) - 1| \leqslant \varepsilon$ for $x \geqslant T$.

These two properties are closely related as they prevent from a loss of compactness when extracting a weakly converging subsequence. We do not enter into the details of the proof of the existence of a minimizing sequence having properties (a) and (b). Indeed, the details are rather technical and do not bring new arguments with respect to the previous sections. The control on the length of the intervals I_i is obtained thanks to the clipping procedure discussed in Section 6.2 and the oscillatory behaviour of minimizers close to the equilibria ± 1 described in Section 6.1. The scheme of the arguments is as follows. Let us suppose to fix the ideas that u_p is positive on the interval I_i. As the assumptions of Lemma 5.1 hold, the second derivative and the potential are the significant terms in $\mathcal{F}^+$. An analysis similar to that of Section 5.1 then shows that the only way for the I_i's to grow to infinity is that the u_p's stay on growing intervals in a neighborhood of $+1$. But in this case, arguing as in Step 3 of the proof of Proposition 6.8, we can locally replace u_p by a minimizer close to $+1$ and clip this minimizer to obtain a function which stays close to $+1$ in an a priori bounded interval. Once the minimizing sequence satisfies (a), the property (b) follows from the arguments of Section 5.1.

We now turn to the proof of Theorem 7.1. The main difficulty is to ensure that the minimizing sequence does not loose or gain complexity in the limit.

PROOF OF THEOREM 7.1.

Step 1. *Convergence*. Let $(u_p)_p \subset \mathcal{E}_n^+$ be a minimizing sequence for $\mathcal{F}^+$ that satisfies the properties (a) and (b). Arguing as in the previous sections, we infer that u_p converges in C^1_{loc} to some function $u \in \mathcal{E}^+$ which is such that

$$\mathcal{F}^+(u) \leqslant \inf_{\mathcal{E}_n^+} \mathcal{F}^+.$$

We denote the extremities of the intervals $I_i^{u_p}$ by x_i^p, $i = 0, \dots, n$. It is clear by uniform convergence that up to a subsequence, for all $i = 1, \dots, n$, x_i^p converges to some $x_i < T$. Remember that by convention, we set $x_0^p = x_0 = 0$ and $x_{n+1}^p = x_{n+1} = \infty$. We call I_i the intervals (x_i, x_{i+1}), $i = 0, \dots, n$. We also deduce from the convergence in $C^1([0, T])$ and the convergence in $C^1_{\text{loc}}([T, +\infty[)$ that

$$u(x)(-1)^{i+n} \geqslant 0 \quad \text{for } x \in (x_i, x_{i+1}),$$

$$\max_{(x_i, x_{i+1})} u(x)(-1)^{i+n} \geqslant 1.$$

Step 2. *Elimination of the zeros of u after x_n*. If u has zeros after x_n, we first modify it to keep only one of those zeros. So, suppose that u vanishes at least two times after x_n. We then define

$$a_1 := \min\{x > x_n \mid u(x) = 0\} \quad \text{and} \quad a_2 := \max\{x > x_n \mid u(x) = 0\}.$$

Since $u(a_1) = u(a_2) = u'(a_1) = u'(a_2) = 0$ by convergence in C^1_{loc}, the interval $[a_1, a_2]$ can be clipped out and the clip has only one zero after x_n. Of course this modification decreases the action.

Assume now that u vanishes at some point $\xi > x_n$. We then have $u'(\xi) = 0$. Now as $u(x_n) = u(\xi)$, there exists at least one critical point y between x_n and ξ such that $u(y) > 0$. Here, we have two possibilities, either y can be taken in such a way that $u(y) \leqslant 1$ or [0, 1] does not contain any critical value of $u_{|[x_n,\xi]}$.

Suppose first that we can find $y \in (x_n, \xi)$ such that $0 < u(y) \leqslant 1$ and $u'(y) = 0$. As $u \in \mathcal{E}^+$,

$$\lim_{x\to\infty} \big(u(x), u'(x)\big) = (1, 0).$$

Hence, by an argument similar to that of the proof of Lemma 6.4, we see that $u(x)$ oscillates around 1 for x large enough. Therefore, we can clip out an interval containing $[y, \xi]$ in such a way that the clip of u does not vanish after x_n.

In the second case, we can find $y \in (x_n, \xi)$ such that $u(y) > 1$ and if $x \in (x_n, \xi)$ satisfies $u'(x) = 0$, then $u(x) > 1$. We now define $v \in \mathcal{E}^+$ by

$$v(x) := \begin{cases} -u(x + x_1) & \text{if } 0 \leqslant x \leqslant \xi - x_1, \\ u(x + x_1) & \text{if } x > \xi - x_1. \end{cases}$$

Observe that since $\min_{[x_n - x_1, \xi - x_1]} v(x) < -1$ and v is negative in $(x_n - x_1, \xi - x_1)$, v has the right number of transitions. Also, v does not vanish after $\xi - x_1$. On the other hand, we deduce from Lemma 5.1 that

$$\int_{x_0}^{x_1} \left(\frac{1}{2}\big(u''^2 + g(u)u'^2\big) + f(u)\right) \mathrm{d}x \geqslant 0$$

so that $\mathcal{F}^+(v) \leqslant \mathcal{F}^+(u)$.

Step 3. *Elimination of the zeros of v in the bumps*. We still denote by $0 = x_0 < x_1 < \cdots < x_n$, the extremities of the intervals I_i associated to v (actually, these are the intervals I_i which have possibly been translated in Step 2). Suppose that there exists $\xi \in v^{-1}(0)$ so that $\xi \neq x_i$ for any $i = 0, \ldots, n$. Hence, ξ lies in the interior of an interval I_i. To fix the ideas, we assume that v is nonnegative therein and denoting by $v(\bar{x}_i)$ the maximum of v over this interval we assume that ξ is at the left of $\bar{x}_i$. Next, define $\xi_1 = \min\{x > x_i \mid v(x) = 0\}$ and $\xi_2 = \max\{x \in [\xi_1, \bar{x}_i] \mid v(x) = 0\}$. It is easily seen that an interval containing $[\xi_1, \xi_2]$ can be clipped out so that the zeros can be deleted.

Step 4. *Elimination of the tangencies with* ±1. The last condition that we have to check to be sure that $v \in \mathcal{E}_n^+$ is that

$$\max_{x\in I_i}\left|v(x)\right| > 1$$

for all $i = 0, \dots, n$. Assume that this condition fails to be true in one of the intervals I_i. In this interval, we thus have $\max_{x\in I_i} |v(x)| = 1$. Let $\tau \in I_i$ be such that $|v(\tau)| = 1$ and $v'(\tau) = 0$. To fix the ideas, assume that $v(\tau) = 1$, the second case being treated in the same way. Let τ_0 and δ_0 be given by Lemma 6.4. As the action of the function 1 is zero, we can modify v without increasing its action by stretching the point τ to an interval of arbitrary length and gluing the function 1 to both extremities, see [42]. Now, we take a_1, (respectively a_2) at the left (respectively at the right) of τ in such a way that $0 < \max_{i=1,2} \|(v(a_i) - 1, v'(a_i))\| \leqslant \delta_0$ and stretch τ to an interval of length τ_0. We still call v the function obtained after gluing 1 at τ and $\tau + \tau_0$. It follows from Lemma 6.4 that the minimizers of

$$\int_{a_1}^{a_2+\tau_0}\left(\frac{1}{2}\left(u''^2 + g(u)u'^2\right) + f(u)\right)\mathrm{d}x$$

on the set of functions $u \in H^2(a_1, \tau_0 + a_2)$ that satisfy $u(a_1) = v(a_1)$, $u'(a_1) = v'(a_1)$, $u(a_2 + \tau_0) = v(a_2)$ and $u'(a_2 + \tau_0) = v'(a_2)$ oscillate around 1. If we replace v locally by a minimizer, we obtain a new function w such that

$$\max_{x\in I_i}\left|w(x)\right| > 1$$

and $\mathcal{F}^+(w) \leqslant \mathcal{F}^+(v)$.

Conclusion. It follows from the previous steps that we can construct $w \in \mathcal{E}_n^+$ such that $\mathcal{F}^+(w) \leqslant \mathcal{F}^+(u)$. Consequently, we have $\mathcal{F}^+(w) = \inf_{\mathcal{E}_n} \mathcal{F}^+$. Now observe that for all $h \in H^2(\mathbb{R}^+)$ such that $h(0) = 0$, $\mathcal{F}^+(w) \leqslant \mathcal{F}^+(w + th)$ for t sufficiently small. Indeed, assume that there exists a sequence $(t_n)_n$ tending to 0 such that $\mathcal{F}^+(w) > \mathcal{F}^+(w + t_n h)$. If w is in the interior of the class $\mathcal{E}_n^+$, this is obviously a contradiction. In the case where w is on the boundary of $\mathcal{E}_n^+$ i.e. if for some points x_i, $w(x_i) = w'(x_i) = 0$, then even for t small, $w + th$ can have more than one zero close to the x_i's so that it does not belong necessarily to $\mathcal{E}_n^+$. However for t small enough, $w + th$ has the right number of transitions and the oscillations close to the points x_i can be erased using the clipping procedure. Therefore, for n large enough, modifying $w + t_n h$ close to the x_i's if necessary, we obtain a function in $\mathcal{E}_n^+$ whose action is strictly smaller than $\mathcal{F}^+(w)$. This contradicts the definition of w and therefore w is a critical point of $\mathcal{F}^+$ in $\mathcal{E}_n^+$. Finally, we infer from the conservation of the Hamiltonian

$$H\left(u, u', u'', u'''\right) = u'''u' - \frac{1}{2}u''^2 - \frac{1}{2}g(u)u'^2 + f(u)$$

that each minimizer is actually in the interior of $\mathcal{E}_n^+$, i.e. each crossing with zero is transverse. □

7.1. *Notes and further comments*

1. Theorem 7.1 can be seen as a partial extension of the results of [42] to Lagrangians that can take either signs. To deal with such Lagrangians, we only require an a priori lower bound on the action along admissible functions. In Theorem 7.1, this lower bound follows from assumption (C1). The conclusion of Theorem 7.1 holds also with the settings of Theorem 5.7, see [18]. In this case we explicitly assume that the functional is bounded from below.

2. Multi-transition solutions can also be obtained in the case of multi-well potentials. We can for example consider the framework of Section 6 assuming then that each equilibrium is of saddle-focus type.

3. The abundance of local minima suggests the existence of infinitely many other critical points of minimax type. However the lack of a compactness property of the Palais–Smale sequences makes the investigation of such solution very complicate. It would seem rather natural that a solution obtained from a minimax principle based on deformations from one local minimum to another one, behaves like the single-transition solutions obtained in Theorem 4.4 for $\beta \geqslant 0$. Even for this range of β, the variational nature of the single-transition solutions is unknown. D. Smets and J.B. van den Berg have used in [83] a modified functional and mountain-pass arguments to catch a homoclinic solution of the Swift–Hohenberg equation for almost every $\beta \in\,]-\sqrt{8}, 0[$.

References

[1] R.A. Adams, *Sobolev Spaces*, Pure and Applied Mathematics, Vol. 65, Academic Press, New York (1975).

[2] S. Ahmad and A.C. Lazer, *An elementary approach to traveling front solutions to a system of N competition-diffusion equations*, Nonlinear Anal. **16** (1991), 893–901.

[3] F. Alessio, M.L. Bertotti and P. Montecchiari, *Multibump solutions to possibly degenerate equilibria for almost periodic Lagrangian systems*, Z. Angew. Math. Phys. **50** (1999), 860–891.

[4] F. Alessio, C. Carminati and P. Montecchiari, *Heteroclinic motions joining almost periodic solutions for a class of Lagrangian systems*, Discrete Contin. Dynam. Systems **5** (1999), 569–584.

[5] F. Alessio, L. Jeanjean and P. Montecchiari, *Stationary layered solutions in $\mathbb{R}^2$ for a class of non autonomous Allen–Cahn equations*, Calc. Var. Partial Differential Equations **11** (2000), 177–202.

[6] F. Alessio, L. Jeanjean and P. Montecchiari, *Existence of infinitely many stationary layered solutions in $\mathbb{R}^2$ for a class of periodic Allen–Cahn equations*, Comm. Partial Differential Equations **27** (2002), 1537–1574.

[7] A. Ambrosetti and M. Badiale, *Homoclinics: Poincaré–Melnikov type results via a variational approach*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), 233–252.

[8] A. Ambrosetti and M. Badiale, *Variational perturbative methods and bifurcation of bound states from the essential spectrum*, Proc. Roy. Soc. Edinburgh Sect. A **128** (1998), 1131–1161.

[9] A. Ambrosetti and V. Coti Zelati, *Multiplicité des orbites homoclines pour des systèmes conservatifs*, C. R. Acad. Sci. Paris Sér. I Math. **314** (1992), 601–604.

[10] M. Arias, J. Campos, A. Robles-Pérez and L. Sanchez, *Fast and heteroclinic solutions for a second order ODE related to Fisher–Kolmogorov's equation*, Calc. Var. Partial Differential Equations **21** (2004), 319–334.

[11] D.G. Aronson and H.F. Weinberger, *Multidimensional nonlinear diffusion arising in population genetics*, Adv. in Math. **30** (1978), 33–76.

[12] H. Berestycki, B. Nicolaenko and B. Scheurer, *Traveling wave solutions to combustion models and their singular limits*, SIAM J. Math. Anal. **16** (1985), 1207–1242.

[13] M.L. Bertotti and L. Jeanjean, *Multiplicity of homoclinic solutions for singular second-order conservative systems*, Proc. Roy. Soc. Edinburgh Sect. A **126** (1996), 1169–1180.

[14] M.L. Bertotti and P. Montecchiari, *Connecting orbits for some classes of almost periodic Lagrangian systems*, J. Differential Equations **145** (1998), 453–468.

[15] G.D. Birkhoff, *Nouvelles recherches sur les systemes dynamiques*, Mem. Pont. Acad. Sci. Novi Lycaei **1** (1935), 85–216.

[16] S.V. Bolotin, *Existence of homoclinic motions*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. (1983), 98–103.

[17] S.V. Bolotin, *Variational methods for constructing chaotic motions in the dynamics of a rigid body*, Prikl. Mat. Mekh. **56** (1992), 230–239.

[18] D. Bonheure, *Multi-transition kinks and pulses for fourth order equations with a bistable nonlinearity*, Ann. Inst. H. Poincaré Anal. Non Linéaire **21** (2004), 319–340.

[19] D. Bonheure, P. Habets and L. Sanchez, *Minimizers for fourth order symmetric bistable equation*, Atti Sem. Mat. Fis. Univ. Modena e Reggio Emilia **52** (2004), 213–227.

[20] D. Bonheure, L. Sanchez, M. Tarallo and S. Terracini, *Heteroclinic connections between nonconsecutive equilibria of a fourth order differential equation*, Calc. Var. Partial Differential Equations **17** (2003), 341–356.

[21] E. Bosetto and E. Serra, *A variational approach to chaotic dynamics in periodically forced nonlinear oscillators*, Ann. Inst. H. Poincaré Anal. Non Linéaire **17** (2000), 673–709.

[22] H. Brezis, *Analyse fonctionnelle*, Collection mathématiques appliquées pour la maîtrise [Collection of Applied Mathematics for the Master's Degree], Masson, Paris (1983).

[23] B. Buffoni and E. Séré, *A global condition for quasi-random behavior in a class of conservative systems*, Comm. Pure Appl. Math. **49** (1996), 285–305.

[24] M. Calanchi and E. Serra, *Homoclinic solutions to periodic motions in a class of reversible equations*, Calc. Var. Partial Differential Equations **9** (1999), 157–184.

[25] P. Caldiroli and L. Jeanjean, *Homoclinics and heteroclinics for a class of conservative singular Hamiltonian systems*, J. Differential Equations **136** (1997), 76–114.

[26] C.-N. Chen and S.-Y. Tzeng, *Existence and multiplicity results for heteroclinic orbits of second order Hamiltonian systems*, J. Differential Equations **158** (1999), 211–250.

[27] V. Coti Zelati, I. Ekeland and E. Séré, *A variational approach to homoclinic orbits in Hamiltonian systems*, Math. Ann. **288** (1990), 133–160.

[28] V. Coti Zelati and P.H. Rabinowitz, *Homoclinic orbits for second order Hamiltonian systems possessing superquadratic potentials*, J. Amer. Math. Soc. **4** (1991), 693–727.

[29] V. Coti Zelati and P.H. Rabinowitz, *Heteroclinic solutions between stationary points at different energy levels*, Topol. Methods Nonlinear Anal. **17** (2001), 1–21.

[30] G.T. Dee and W. van Saarloos, *Bistable systems with propagating fronts leading to pattern formation*, Phys. Rev. Lett. **60** (1988), 2641–2644.

[31] R.L. Devaney, *Homoclinic orbits in Hamiltonian systems*, J. Differential Equations **21** (1976), 431–438.

[32] N. Fei and J. Carr, *Existence of travelling waves with their minimal speed for a diffusing Lotka–Volterra system*, Nonlinear Anal. Real World Appl. **4** (2003), 503–524.

[33] P.C. Fife and J.B. McLeod, *The approach of solutions of nonlinear diffusion equations to travelling front solutions*, Arch. Ration. Mech. Anal. **65** (1977), 335–361.

[34] R.A. Fisher, *The wave of advance of advantageous genes*, Ann. Eugenics **7** (1937), 355–369.

[35] B.H. Gilding and R.A. Kersner, *Travelling Waves in Nonlinear Diffusion–Convection Reaction*, Progress in Nonlinear Differential Equations and their Applications, Vol. 60, Birkhäuser, Basel (2004).

[36] G. Gompper and M. Schick, *Self-assembling amphiphilic systems*, Phase Transitions and Critical Phenomena, Vol. 16, C. Domb and J.L. Lebowitz, eds., Academic Press, London (1994), 1–176.

[37] P. Habets, L. Sanchez, M. Tarallo and S. Terracini, *Heteroclinics for a class of fourth order conservative differential equations*, Proceedings of Equadiff 10, Prague (2001), 203–216 (CD ROM).

[38] J.K. Hale and H. Koçak, *Dynamics and Bifurcations*, Texts in Applied Mathematics, Vol. 3, Springer, New York (1991).

[39] F. Hamel and N. Nadirashvili, *Travelling fronts and entire solutions of the Fisher–KPP equation in $\mathbb{R}^N$*, Arch. Ration. Mech. Anal. **157** (2001), 91–163.

[40] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities*, Cambridge Mathematical Library, Cambridge University Press, Cambridge (1988), Reprint of the 1952 edition.

[41] P. Hartman, *Ordinary Differential Equations*, Wiley, New York (1964).

[42] W.D. Kalies, J. Kwapisz and R.C.A.M. VanderVorst, *Homotopy classes for stable connections between Hamiltonian saddle-focus equilibria*, Comm. Math. Phys. **193** (1998), 337–371.

[43] W.D. Kalies and R.C.A.M. VanderVorst, *Multi-transition homoclinic and heteroclinic solutions of the extended Fisher–Kolmogorov equation*, J. Differential Equations **131** (1996), 209–228.

[44] W. Kelley, *Analytic approximations of travelling wave solutions*, Appl. Anal. **55** (1994), 287–299.

[45] A. Kolmogorov, I. Petrovski and N. Piscounov, *Etude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique*, Bull. Univ. Moskou Ser. Internat. Sec. A **1** (1937), 1–25.

[46] P. Korman, A.C. Lazer and Y. Li, *On homoclinic and heteroclinic orbits for Hamiltonian systems*, Differential Integral Equations **10** (1997), 357–368.

[47] J. Kwapisz, *Uniqueness of the stationary wave for the extended Fisher–Kolmogorov equation*, J. Differential Equations **165** (2000), 235–253.

[48] L. Malaguti and C. Marcelli, *Existence of bounded trajectories via upper and lower solutions*, Discrete Contin. Dynam. Systems **6** (2000), 575–590.

[49] L. Malaguti and C. Marcelli, *Travelling wavefronts in reaction–diffusion equations with convection effects and non-regular terms*, Math. Nachr. **242** (2002), 148–164.

[50] L. Malaguti and C. Marcelli, *Existence and multiplicity of heteroclinic solutions for a non-autonomous boundary eigenvalue problem*, Electron. J. Differential Equations (2003), 21 pp. (electronic).

[51] L. Malaguti and C. Marcelli, *Sharp profiles in degenerate and doubly degenerate Fisher–KPP equations*, J. Differential Equations **195** (2003), 471–496.

[52] J. Mallet-Paret, *Travelling waves in spatially discrete dynamical systems of diffusive type*, Dynamical Systems, J.W. Macki and P. Zecca, eds., Lecture Notes in Mathematics, Vol. 1822, Springer, Berlin (2003), 231–298.

[53] T. O. Maxwell, *Heteroclinic chains for a reversible Hamiltonian system*, Nonlinear Anal. **28** (1997), 871–887.

[54] V.K. Mel'nikov, *On the stability of a center for time-periodic perturbations*, Trudy Moskov. Mat. Obšč. **12** (1963), 3–52.

[55] V.J. Mizel, L.A. Peletier and W.C. Troy, *Periodic phases in second-order materials*, Arch. Ration. Mech. Anal. **145** (1998), 343–382.

[56] P. Montecchiari, M. Nolasco and S. Terracini, *A global condition for periodic Duffing-like equations*, Trans. Amer. Math. Soc. **351** (1999), 3713–3724.

[57] J. Moser, *Stable and Random Motions in Dynamical Systems: With Special Emphasis on Celestial Mechanics*, Ann. Math. Studies, Vol. 77, Princeton University Press, Princeton, NJ (1973).

[58] L.A. Peletier, W.C. Troy and R.C.A.M. Van der Vorst, *Stationary solutions of a fourth-order nonlinear diffusion equation*, Differentsial'nye Uravneniya **31** (1995), 327–337.

[59] L.A. Peletier and W.C. Troy, *Spatial patterns described by the extended Fisher–Kolmogorov (EFK) equation: kinks*, Differential Integral Equations **8** (1995), 1279–1304.

[60] L.A. Peletier and W.C. Troy, *A topological shooting method and the existence of kinks of the extended Fisher–Kolmogorov equation*, Topol. Methods Nonlinear Anal. **6** (1995), 331–355.

[61] L.A. Peletier and W.C. Troy, *Chaotic spatial patterns described by the extended Fisher–Kolmogorov equation*, J. Differential Equations **129** (1996), 458–508.

[62] L.A. Peletier and W.C. Troy, *Spatial patterns described by the extended Fisher–Kolmogorov equation: periodic solutions*, SIAM J. Math. Anal. **28** (1997), 1317–1353.

[63] L.A. Peletier and W.C. Troy, *Spatial Patterns*, Progress in Nonlinear Differential Equations and their Applications, Vol. 45, Birkhäuser, Boston, MA (2001).

[64] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste* (1892). Reprinted by Dover Publications Inc., New York (1957). Tome III. Invariants intégraux. Solutions périodiques du deuxième genre. Solutions doublement asymptotiques.

[65] M.H. Protter and H.F. Weinberger, *Maximum Principles in Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ (1967).

[66] P.H. Rabinowitz, *Periodic and heteroclinic orbits for a periodic Hamiltonian system*, Ann. Inst. H. Poincaré Anal. Non Linéaire **6** (1989), 331–346.

[67] P.H. Rabinowitz, *Homoclinic and heteroclinic orbits for a class of Hamiltonian systems*, Calc. Var. Partial Differential Equations **1** (1993), 1–36.

[68] P.H. Rabinowitz, *Heteroclinics for a reversible Hamiltonian system. II*, Differential Integral Equations **7** (1994), 1557–1572.

[69] P.H. Rabinowitz, *Homoclinics for an almost periodically forced singular Hamiltonian system*, Topol. Methods Nonlinear Anal. **6** (1995), 49–66.

[70] P.H. Rabinowitz, *On a theorem of Strobel*, Calc. Var. Partial Differential Equations **12** (2001), 399–415.

[71] P.H. Rabinowitz, *A new variational characterization of spatially heteroclinic solutions of a semilinear elliptic PDE*, Discrete Contin. Dyn. Syst. **10** (2004), 507–517.

[72] P.H. Rabinowitz and V. Coti Zelati, *Multichain-type solutions for Hamiltonian systems*, Proceedings of the Conference on Nonlinear Differential Equations, Coral Gables, FL (1999), Electron. J. Differential Equations **5**, Southwest Texas State Univ., San Marcos, TX (2000), 223–235 (electronic).

[73] P.H. Rabinowitz and E. Stredulinsky, *Mixed states for an Allen–Cahn type equation*, Comm. Pure Appl. Math. **56** (2003), 1078–1134.

[74] P.H. Rabinowitz and E. Stredulinsky, *Mixed states for an Allen–Cahn type equation. II*, Calc. Var. Partial Differential Equations **21** (2004), 157–207.

[75] V. Rottschäfer and C.E. Wayne, *Existence and stability of traveling fronts in the extended Fisher–Kolmogorov equation*, J. Differential Equations **176** (2001), 532–560.

[76] P.L. Sachdev, *Nonlinear Ordinary Differential Equations and their Applications*, Monographs and Textbooks in Pure and Applied Mathematics, Vol. 142, M. Dekker, New York (1991).

[77] L. Sanchez, *A note on a nonautonomous O.D.E. related to the Fisher equation*, J. Comput. Appl. Math. **113** (2000), 201–209.

[78] L. Sanchez, *Heteroclinic connections for a class of non-autonomous systems*, Proceedings of the USA–Chile Workshop on Nonlinear Analysis, Viña del Mar-Valparaiso (2000), Electron. J. Differential Equations **6**, Southwest Texas State Univ., San Marcos, TX (2001), 257–266 (electronic).

[79] E. Séré, *Existence of infinitely many homoclinic orbits in Hamiltonian systems*, Math. Z. **209** (1992), 27–42.

[80] E. Serra, *Heteroclinic orbits at infinity for two classes of Hamiltonian systems*, Boll. Un. Mat. Ital. B (7) **8** (1994), 615–639.

[81] S. Smale, *Diffeomorphisms with many periodic points*, Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton Univ. Press, Princeton, NJ (1965) 63–80.

[82] S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817.

[83] D. Smets and J.B. van den Berg, *Homoclinic solutions for Swift–Hohenberg and suspension bridge type equations*, J. Differential Equations **184** (2002), 78–96.

[84] J.B. Swift and P.C. Hohenberg, *Hydrodynamic fluctuations at the convective instability*, Phys. Rev. A **15** (1977), 319–328.

[85] K. Tanaka, *Homoclinic orbits for a singular second order Hamiltonian system*, Ann. Inst. H. Poincaré Anal. Non Linéaire **7** (1990), 427–438.

[86] S. Terracini, *Non-degeneracy and chaotic motions for a class of almost-periodic Lagrangean systems*, Nonlinear Anal. **37** (1999), 337–361.

[87] J.B. van den Berg, *Branches of heteroclinic, homoclinic and periodic solutions in a fourth-order bi-stable system, a numerical study*, Master's thesis, Leiden University (1996).

[88] J.B. van den Berg, *Dynamics and equilibria of fourth order differential equations*, Ph.D. thesis, Leiden University (2000).

[89] J.B. van den Berg, *The phase-plane picture for a class of fourth-order conservative differential equations*, J. Differential Equations **161** (2000), 110–153.

[90] J.B. van den Berg, J. Hulshof and R.C. Vandervorst, *Traveling waves for fourth order parabolic equations*, SIAM J. Math. Anal. **32** (2001), 1342–1374.

[91] W. Walter, *Ordinary Differential Equations*, Graduate Texts in Mathematics, Vol. 182, Springer, New York (1998). Translated from the sixth German (1996) edition by Russell Thompson.

[92] S. Wiggins, *Global bifurcations and chaos*, Applied Mathematical Sciences, Vol. 73, Springer, New York (1988).

CHAPTER 3

A Qualitative Analysis, via Lower and Upper Solutions, of First Order Periodic Evolutionary Equations with Lack of Uniqueness

Colette De Coster

LMPA Liouville, Université du Littoral-Côte d'Opale, 50, rue Ferdinand Buisson, B.P. 699,
F-62228 Calais Cédex, France
E-mail: decoster@lmpa.univ-littoral.fr

Franco Obersnel and Pierpaolo Omari

Dipartimento di Matematica e Informatica, Università degli Studi di Trieste, Via Alfonso Valerio 12/1,
I-34127 Trieste, Italy
E-mails: {obersnel,omari}@units.it

Contents

Preface . . . 205
1. First Order Periodic ODEs . . . 206
  1.1. Introduction . . . 206
  1.2. Preliminaries . . . 209
  1.3. The initial value problem . . . 211
  1.4. The periodic problem . . . 217
  1.5. Structure and dynamics near one-sided isolated periodic solutions . . . 226
  1.6. Stability and detours . . . 229
  1.7. Asymptotic behaviour of bounded solutions . . . 244
2. Second Order Periodic Parabolic PDEs . . . 250
  2.1. Introduction . . . 250
  2.2. Preliminaries . . . 258
  2.3. The initial boundary value problem . . . 265
  2.4. The periodic boundary value problem . . . 284
  2.5. The Order Interval Trichotomy and applications . . . 298
  2.6. Stability matters . . . 310
  2.7. Appendix: The principal eigenvalue . . . 332
References . . . 334

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 3
Edited by A. Cañada, P. Drábek and A. Fonda
© 2006 Elsevier B.V. All rights reserved

This page intentionally left blank

Preface

The method of lower and upper solutions is an elementary but powerful tool in the existence theory of initial and periodic problems for semilinear differential equations for which a maximum principle holds, even in cases where no special structure is assumed on the nonlinearity (see, e.g., [150,118,112,39,41,117,98]). The aim of this work is to show that this method is also quite effective for investigating the qualitative properties of solutions, at the same extent of generality for which the existence theory is developed. Indeed, our main purpose is to work out, under a minimal set of basic assumptions, a qualitative theory for two sample classes of scalar periodic differential equations: the first order ordinary differential equation

$$u' = f(t, u) \tag{1}$$

and the second order parabolic problem

$$\begin{aligned} &\partial_t u + A(x, t, \partial_x)u = f(x, t, u, \nabla_x u) && \text{in } \Omega \times I, \\ &u = 0 && \text{on } \partial\Omega \times I. \end{aligned} \tag{2}$$

Here $I \subseteq \mathbb{R}$ is an interval, $\Omega \subset \mathbb{R}^N$ is a bounded smooth domain, $\partial_t u + A(x, t, \partial_x)u$ is a uniformly parabolic differential operator, with time-periodic coefficients, and the functions f are time-periodic and satisfy suitable Carathéodory conditions. We must stress that under these assumptions neither uniqueness for the initial value problem associated with (1) and (2), nor validity of comparison principles are guaranteed. Accordingly, the qualitative analysis of these equations cannot be performed applying the theory of order preserving dynamical systems as given, for instance, in [60]. In this general setting we will actually ascertain, already for the easy-looking equation (1) and unlike the regular case, the occurrence of very complicated dynamics, so that even the explicit knowledge of a distinguished class of solutions (such as equilibria, or periodic solutions) is generally not sufficient to get a global qualitative portrait.

We are especially interested here in studying, with the aid of lower and upper solutions, the following three basic questions about (1) and (2):

(i) existence of periodic solutions of (1), or (2), and their localization;

(ii) qualitative properties of periodic solutions of (1), or (2), with special reference to their stability or instability;

(iii) asymptotic behaviour of solutions of the initial value problem associated with (1), or (2).

This program is pursued in the first part of this work for (1) and in the second part for (2); however, as it is quite predictable, exhaustive results can be obtained only for (1). Problems (1) and (2) share relevant similarities from the qualitative point of view. In fact Eq. (2) looks like a natural extension of (1), which in turn represents a simple, but nontrivial, paradigm of the patterns that can be possibly displayed by the dynamics of (2). Nevertheless, they exhibit several differences, which appear already evident developing the existence theory for the initial and the periodic problems in the presence of lower and upper solutions. Hence the two equations require separate treatments. We chose to describe here just basic results and to select a small number of examples or applications; further material can be found in [112] for (1) and in [42] for (2).

Although the method of lower and upper solutions applies to a larger class of boundary value problems, we decided to restrict our attention to the above two ones only. In particular, a remarkable topic which is completely left out here concerns the second order periodic ordinary differential equation

$$u'' = f\bigl(t, u, u'\bigr) \quad \text{on } I. \tag{3}$$

In this context a theory of the stability of periodic solutions, exploiting lower and upper solutions, started in [72] and has been recently developed in [116,36,111,109]. This omission is however not only dictated by the necessity of keeping our work within a reasonable length, but it is also motivated by the fact that the study of (3) relies on somewhat different techniques from the ones we use here; further, more regularity is required on f and at the moment the results are less complete.

1. First Order Periodic ODEs

1.1. *Introduction*

In this part we consider the first order scalar ordinary differential equation

$$x' = f(t, x). \tag{1.1}$$

We assume that $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is periodic in t with period $T > 0$ and satisfies the L^1-Carathéodory conditions. Associated with (1.1), we consider the Cauchy problem

$$x' = f(t, x), \quad x(t_0) = x_0, \tag{1.2}$$

where $t_0, x_0 \in \mathbb{R}$ are given, and the periodic problem on $[0, T]$

$$x' = f(t, x), \quad x(0) = x(T). \tag{1.3}$$

Solutions of (1.1), (1.2), (1.3) are intended in the Carathéodory, i.e. $W^{1,1}_{\mathrm{loc}}$-, sense. Since any solution of (1.3) on $[0, T]$ extends to a T-periodic solution of (1.1), it will also be referred to as T-periodic.

Existence, multiplicity, localization and qualitative properties of solutions of (1.3) have been investigated using various approaches (see, e.g., [91,122,147,99,95,148,94,108,28, 142,113] and the included references). In particular, when f is a polynomial in x with coefficients T-periodic in t, the interest of the results that have been obtained stems from their connections with Hilbert Sixteenth Problem about the number of limit cycles of autonomous plane polynomial systems (see, e.g., [87,88,85,135,4,62,63,24]), or from their applications to the study of single species population models in a periodically fluctuating environment (see, e.g., [74,131,31,30,43,129,132,14,96,146,19,97]).

The knowledge of the T-periodic solutions of (1.1) is also the basic step in the realization of the Poincaré program, that is the determination of the qualitative portrait of all solutions of (1.1). If the property of uniqueness of solutions holds for the Cauchy problem (1.2), the behaviour of the solutions of scalar T-periodic differential equations such as (1.1) is, at least conceptually, nearly as simple as the behaviour of autonomous equations. Indeed, in this case the dynamics can be reduced to that of the Poincaré map, which assigns to every initial datum the value of the solution after one period, and, since this is strictly monotone, the orbit structure can be classified in an elementary way (cf. [57, Chapter 4]). In the Carathéodory setting, although generic (cf. [121, Chapter 3]), uniqueness in the future, or in the past, for the solutions of (1.2) is of course not anymore guaranteed. This fact yields evident technical complications also due to the possible failure of comparison principles for the solutions of (1.1). On the other hand, the lack of uniqueness entails, as we shall see, a rich variety of dynamics, such as the possible occurrence of homoclinic, subharmonic and almost periodic solutions. In addition to these theoretical features, that may justify by themselves a close analysis, further motivations for studying (1.1) in the absence of uniqueness come from some models in population dynamics related to those proposed by G.F. Gause in [53, Chapter VI].

The main purpose in this part of our work is to show that a quite satisfactory description of the dynamics of (1.1) can be performed in the Carathéodory setting by a systematic use of lower and upper solutions, i.e. functions satisfying suitable differential inequalities. The lower and upper solutions method is a classical tool for the study of (1.2) and (1.3) and it has already been employed by several authors since the pioneering work of G. Peano [119] (see [120,102,143,101,150,79,2,94,110,38,52,89,28] and the included references). Yet, while this approach has been largely used in the discussion of existence, multiplicity and approximation by monotone iteration, the stability properties of the solutions of (1.3) have been generally detected by other techniques, often based on linearization (see the recent papers [10,11,20,21] and the references therein). Partial exceptions are [95,96,133], where a global qualitative portrait of (1.1) was obtained, but only for a restricted class of functions f satisfying some convexity assumptions with respect to the x-variable, which in turn imply uniqueness of the solutions of (1.2). Of course, these methods do not apply to the general context we wish to discuss here, where Lipschitz or differentiability conditions on f are not required. Our work is therefore in the spirit of [147] and [112], whose declared aim was to pave the way towards a theory of periodic solutions of Eq. (1.1) when no special structure is considered.

This part is organized as follows. In Section 1.2 we formalize some definitions concerning the first order scalar ordinary differential equation (1.1) as well as we settle some notations.

In Section 1.3 we collect some basic facts concerning the solvability in the Carathéodory setting of the Cauchy problem (1.2), when a pair of possibly discontinuous lower and upper solutions is given and no special ordering between them is assumed. Existence and localization of solutions, as well as the structure of the solution set, are discussed in this frame.

In Section 1.4 we initiate the qualitative study of solutions of (1.1). We first discuss a form of monotonicity over periods, named T-monotonicity, which, unlike the case where uniqueness holds for the Cauchy problem (1.2), is not shared by all solutions of (1.1). We prove, however, that T-monotonicity is enjoyed by every solution of (1.1) that is comparable with any possible solution of (1.3). This property represents an essential tool for the description of the asymptotic behaviour of solutions of (1.1). In particular, it allows to obtain an extension of the existence part of the classical Massera Convergence Theorem [91] to the case where uniqueness for (1.2) fails; that is, the existence of a bounded solution x of (1.1) defined on an unbounded interval implies the existence of a T-periodic solution u of (1.1). A different proof of this fact is given in [147], while a discussion of various extensions of the Massera Theorem can be found in [104]. Yet, in the absence of uniqueness for (1.2), x does not generally converge to u: an example where the convergence does not take place is explicitly produced. Afterwards, we prove a technical result which describes the dynamics of solutions stemming from a lower or an upper solution of (1.3). This result plays a central role in the subsequent study of the stability of T-periodic solutions; it resembles the Monotone Convergence Criterion in the theory of monotone maps and may have an independent interest, since it guarantees the convergence to a T-periodic solution of a bounded solution emanating from and lying above a lower solution or emanating from and lying below an upper solution of (1.3). Founded on these results, we prove the main existence and localization theorem for (1.2), under the assumption of the existence of a pair of possibly discontinuous and unordered lower and upper solutions. We also point out by an example how the type of discontinuous lower and upper solutions of (1.3) here considered can be successfully used, in some cases, to get precise information on the localization of branches of T-periodic solutions.

Section 1.5 deals with the description of the dynamics of solutions of (1.1), which lie between two strictly ordered T-periodic solutions, or above the maximum T-periodic solution, or below the minimum one, if they exist. In particular, we show the existence of heteroclinic solutions connecting pairs of ordered T-periodic solutions of (1.1), when there is no further T-periodic solution in between. These results are based on the T-monotonicity of solutions of (1.1), which are comparable with any possible solution of (1.3), and they yield information about the stability and instability of solutions of (1.3), also with reference to the existence of repulsivity and attractivity basins.

Section 1.6, which represents the core of this part, is devoted to the study of the stability properties of the T-periodic solutions of (1.1) by means of lower and upper solutions. We begin with a discussion of the mutual relations of three notions of one-sided stability of a T-periodic solution we believe appropriate to be considered when uniqueness for (1.2) may fail: Lyapunov stability, order stability and weak stability. The second one is standard in the frame of order preserving discrete-time semidynamical systems (cf. [58, 60]). The third one has been used in the context of multivalued semiflows (cf. [130,50]) and is related to the notion of weak invariance (cf. [105,152,140]). If uniqueness in the future, or in the past, holds for (1.2), then Lyapunov stability and order stability are equivalent concepts. Whereas, we show here that order stability and weak stability are always equivalent: thus weak stability yields an alternative and more direct interpretation of order stability, in terms of the dynamics of solutions of (1.1). Using these three notions we can describe in a precise way the stability properties of a T-periodic solution, relating them to the existence of a lower or an upper solution close to it. Hence, when a pair of lower and upper solutions is given, we can discuss the stability or instability of the minimum and the maximum T-periodic solutions v and w wedged between them. Afterwards, we turn to investigate the behaviour of the solutions lying between v and w. If the lower and the upper solutions satisfy the standard ordering condition, i.e. the lower solution is below the upper solution, we find between v and w a weakly stable T-periodic solution; whereas, in the complementary case, Lyapunov unstable T-periodic solutions always exist. We further study the existence and the properties of nondegenerate continua of T-periodic solutions in $C^0([0,T])$. In particular, we see that if two T-periodic solutions are not strictly ordered, they give rise to a nondegenerate continuum of unstable T-periodic solutions. This fact may occur in the presence of a pair of unordered lower and upper solutions of (1.3). On the contrary, the existence of a pair of lower and upper solutions of (1.3) satisfying the standard ordering condition always guarantees the existence of a totally ordered continuum of weakly stable solutions. More generally, we investigate the topological structure of the set of all solutions of (1.3). This set is the union of a family of mutually ordered connected components. In the absence of uniqueness for (1.2) we show that such components may have arbitrarily large (Čech–Lebesgue) dimension. Moreover, when a continuum $\mathcal{K}$ is generated by two T-periodic solutions which are not strictly ordered, the dynamics of the solutions of (1.1), which lie between the minimum and the maximum elements in $\mathcal{K}$, is fairly complicated. Loosely speaking, any discrete, even chaotic, dynamics can be realized (see [114]). In particular, we can find homoclinics, subharmonic solutions of any order and almost periodic solutions which are not periodic of any period. As it is well known [91], this phenomenon cannot take place if the uniqueness property holds for (1.2).

In Section 1.7 we revisit the Massera Convergence Theorem. We have already noticed in Section 1.4 that, even when uniqueness for (1.2) fails, the existence of a solution u of (1.3) is always guaranteed in the presence of a bounded solution x of (1.1) defined on an unbounded interval, in spite of the fact that x does not generally converge to u. In this section we make this statement more precise, providing a complete extension of the Massera Convergence Theorem to the case where the uniqueness assumption for solutions of (1.2) is dropped. Namely, we see that either x converges to u, or the set of all solutions u of (1.3), whose range contains ω-limit points of x, forms a nondegenerate closed connected set in $C^0([0,T])$.

1.2. *Preliminaries*

In this section we formalize some basic definitions concerning the first order scalar ordinary differential equation (1.1). Let $T > 0$ be a fixed number. The following condition is assumed:

(C) $f :]0,T[\times\mathbb{R} \to \mathbb{R}$ satisfies the L^1-Carathéodory conditions, i.e. for every $x \in \mathbb{R}$, $f(\cdot,x)$ is measurable on $]0,T[$; for a.e. $t \in]0,T[$, $f(t,\cdot)$ is continuous on $\mathbb{R}$; for each $\rho > 0$, there exists $\gamma \in L^1(0,T)$ such that $|f(t,x)| \leqslant \gamma(t)$, for a.e. $t \in]0,T[$ and every $x \in [-\rho,\rho]$.

REMARK 1.1. The function f is identified with its T-periodic extension onto $\mathbb{R} \times \mathbb{R}$.

Solutions and lower and upper solutions

DEFINITION 1.1.

- A *solution of* (1.1) *on* $[t_1, t_2]$, with $t_1, t_2 \in \mathbb{R}$ and $t_1 < t_2$, is a function $x \in W^{1,1}(t_1, t_2)$ such that

$$x'(t) = f\big(t, x(t)\big) \quad \text{a.e. on }]t_1, t_2[.$$

- A *solution of* (1.1) *on* J, where J is a noncompact interval having endpoints t_1, t_2, with $-\infty \leqslant t_1 < t_2 \leqslant +\infty$, is a function x such that, for every compact interval $K \subset J$, $u|_K$ is a solution of (1.1) on K.

DEFINITION 1.2. Let $I \subseteq \mathbb{R}$ be an interval having endpoints ω_-, ω_+, with $-\infty \leqslant \omega_- < \omega_+ \leqslant +\infty$. A solution $x : I \to \mathbb{R}$ of (1.1) is said *right-nonextendible* if either $\omega_+ = +\infty$, or $\omega_+ < +\infty$ and there is no solution $\hat{x}$ of (1.1) on $I \cup \{\omega_+\}$ such that $\hat{x}|_I = x$. Similarly, a solution $x : I \to \mathbb{R}$ of (1.1) is said *left-nonextendible* if either $\omega_- = -\infty$, or $\omega_- > -\infty$ and there is no solution $\hat{x}$ of (1.1) on $I \cup \{\omega_-\}$ such that $\hat{x}|_I = x$. A solution $x :]\omega_-, \omega_+[\to \mathbb{R}$ of (1.1) is said *nonextendible* if it is both right-nonextendible and left-nonextendible.

We now introduce a notion of possibly discontinuous lower and upper solutions of (1.1), which have been extensively used in [38,52,86,28,112].

NOTATION 1.3. For a function $x : [t_1, t_2] \to \mathbb{R}$, we write $x|_{]t_1,t_2[} \in W^{1,1}(t_1, t_2)$ if there exists $\widetilde{x} \in W^{1,1}(t_1, t_2) \cap C^0([t_1, t_2])$ such that $\widetilde{x} = x$ on $]t_1, t_2[$.

DEFINITION 1.4.

- A *regular lower solution of* (1.1) *on* $[t_1, t_2]$, with $t_1, t_2 \in \mathbb{R}$ and $t_1 < t_2$, is a function $\alpha : [t_1, t_2] \to \mathbb{R}$, such that $\alpha|_{]t_1,t_2[} \in W^{1,1}(t_1, t_2)$,

$$\alpha'(t) \leqslant f\big(t, \alpha(t)\big) \quad \text{a.e. on } [t_1, t_2],$$

$$\lim_{t \to t_1^+} \alpha(t) \leqslant \alpha(t_1), \qquad \alpha(t_2) \leqslant \lim_{t \to t_2^-} \alpha(t).$$

- A *lower solution of* (1.1) *on* $[t_1, t_2]$ is a function $\alpha : [t_1, t_2] \to \mathbb{R}$ for which there are points $t_1 = \tau_0 < \tau_1 < \cdots < \tau_N = t_2$ such that, for $i = 0, \ldots, N-1$, $\alpha|_{[\tau_i, \tau_{i+1}]}$ is a regular lower solution on $[\tau_i, \tau_{i+1}]$.
- A *lower solution* (respectively a *regular lower solution*) *of* (1.1) *on* J, where J is a noncompact interval having endpoints t_1, t_2 with $-\infty \leqslant t_1 < t_2 \leqslant +\infty$, is a function α such that, for any compact interval $K \subset J$, $\alpha|_K$ is a lower solution (respectively a regular lower solution) of (1.1) on K.
- A *regular upper solution of* (1.1) *on* $[t_1, t_2]$, is a function $\beta : [t_1, t_2] \to \mathbb{R}$, such that $\beta|_{]t_1,t_2[} \in W^{1,1}(t_1, t_2)$,

$$\beta'(t) \geqslant f\big(t, \beta(t)\big) \quad \text{a.e. on } [t_1, t_2],$$

$$\lim_{t \to t_1^+} \beta(t) \geqslant \beta(t_1), \qquad \beta(t_2) \geqslant \lim_{t \to t_2^-} \beta(t).$$

- An *upper solution of* (1.1) *on* $[t_1, t_2]$ is a function $\beta : [t_1, t_2] \to \mathbb{R}$ for which there are points $t_1 = \sigma_0 < \sigma_1 < \cdots < \sigma_N = t_2$ such that, for $i = 0, \ldots, N-1$, $\beta|_{[\sigma_i, \sigma_{i+1}]}$ is a regular upper solution on $[\sigma_i, \sigma_{i+1}]$.
- An *upper solution* (respectively a *regular upper solution*) *of* (1.1) *on* J, where J is a noncompact interval having endpoints t_1, t_2 with $-\infty \leqslant t_1 < t_2 \leqslant +\infty$, is a function β such that, for any compact interval $K \subset J$, $\beta|_K$ is an upper solution (respectively a regular upper solution) of (1.1) on K.

REMARK 1.2. Notice that a regular lower solution α is (absolutely) continuous on $]t_1, t_2[$, but it may be discontinuous at the endpoints of the interval $[t_1, t_2]$. A similar observation holds for a regular upper solution.

Orderings

DEFINITION 1.5. Given functions $x, y : I \to \mathbb{R}$, with $I \subseteq \mathbb{R}$ an interval, we write
- $x \leqslant y$ if $x(t) \leqslant y(t)$ on I;
- $x < y$ if $x \leqslant y$ and $x \neq y$;
- $x \ll y$ if $\inf(y - x)|_K > 0$ for any compact interval $K \subseteq I$.

DEFINITION 1.6. Given functions $x, y : I \to \mathbb{R}$, with $I \subseteq \mathbb{R}$ an interval and $x < y$, we define the *order intervals*

$$\begin{aligned} &[x, y] = \{z : I \to \mathbb{R} \mid x \leqslant z \leqslant y\}, \\ &[x, +\infty[= \{z : I \to \mathbb{R} \mid x \leqslant z\}, \\ &]-\infty, y] = \{z : I \to \mathbb{R} \mid z \leqslant y\}. \end{aligned}$$

DEFINITION 1.7. Let $\mathcal{S}$ be a given set of solutions of (1.1).
- We say that a solution z of (1.1), with $z \in \mathcal{S}$, is a *maximal solution of* (1.1) *in* $\mathcal{S}$ (respectively a *minimal solution of* (1.1) *in* $\mathcal{S}$) if there is no solution x of (1.1), with $x \in \mathcal{S}$, such that $x > z$ (respectively $x < z$).
- We say that a solution z of (1.1), with $z \in \mathcal{S}$, is the *maximum solution of* (1.1) *in* $\mathcal{S}$ (respectively the *minimum solution of* (1.1) *in* $\mathcal{S}$) if every solution x of (1.1), with $x \in \mathcal{S}$, is such that $x \leqslant z$ (respectively $x \geqslant z$).

1.3. *The initial value problem*

In this section we collect some basic facts concerning the solvability of the Cauchy problem (1.2), when a pair of possibly discontinuous lower and upper solutions is given and no special ordering between them is assumed. Existence and localization of solutions, as well as the structure of the solution set, are discussed in this setting.

REMARK 1.3. Since the study of a terminal value problem may be reduced to the study of an initial value problem by reversing time, we shall mainly concentrate on the latter problem. Occasionally we shall need to refer to a terminal value problem; in that case we shall freely use definitions, notations and results about it without explicit references.

Solutions and lower and upper solutions

DEFINITION 1.8. A *solution of the initial value problem* (1.2) on I, with $I \subseteq \mathbb{R}$ an interval containing a right neighbourhood of t_0, is a solution x of (1.1) on I such that $x(t_0) = x_0$.

DEFINITION 1.9.

- A *lower solution* (respectively a *regular lower solution*) of (1.2) on I, with $I \subseteq \mathbb{R}$ an interval containing a right neighbourhood of t_0, is a function $\alpha : I \to \mathbb{R}$, which is a lower solution (respectively a regular lower solution) of (1.1) on I and satisfies $\alpha(t_0) \leqslant x_0$.
- An *upper solution* (respectively a *regular upper solution*) of (1.2) on I, with $I \subseteq \mathbb{R}$ an interval containing a right neighbourhood of t_0, is a function $\beta : I \to \mathbb{R}$, which is an upper solution (respectively a regular upper solution) of (1.1) on I and satisfies $\beta(t_0) \geqslant x_0$.

The following simple observation concerning pairs of unordered lower and upper solutions will be used in the sequel.

PROPOSITION 1.1. *Let α be a lower solution and β be an upper solution of* (1.2) *on* $[t_0, t_1]$ *with $\alpha \nleqslant \beta$. Then there exists $\hat{t} \in [t_0, t_1]$ such that $\alpha(\hat{t}) = \beta(\hat{t})$.*

PROOF. Let us set $\gamma = \beta - \alpha$. According to Definition 1.9, γ satisfies the following condition: there are points $t_0 = \rho_0 < \rho_1 < \cdots < \rho_K = t_1$ such that

- for $i = 0, \ldots, K-1$, $\gamma|_{]\rho_i, \rho_{i+1}[} \in W^{1,1}(\rho_i, \rho_{i+1})$,
- for $i = 1, \ldots, K-1$, $\lim_{t \to \rho_i^-} \gamma(t) \leqslant \gamma(\rho_i) \leqslant \lim_{t \to \rho_i^+} \gamma(t)$,
- $\lim_{t \to t_0^+} \gamma(t) \geqslant \gamma(t_0) \geqslant 0$ and $\gamma(t_1) \geqslant \lim_{t \to t_1^-} \gamma(t)$.

We may assume $\gamma(t_0) > 0$. Since $\alpha \nleqslant \beta$, there is a point $\xi \in]t_0, t_1]$ such that $\gamma(\xi) < 0$. Setting $\hat{t} = \sup\{t \in]t_0, \xi[\mid \gamma(s) > 0 \text{ on } [t_0, t]\}$, the properties of γ imply that $\gamma(\hat{t}) = 0$. □

Existence of solutions and the Hukuhara–Kneser property
We start with a basic existence and localization result for solutions of (1.1).

LEMMA 1.2. *Assume* (C). *Let $\alpha, \beta \in W^{1,1}(t_0, t_1)$ satisfy*

$$\alpha'(t) \leqslant f\big(t, \alpha(t)\big) \quad \textit{and} \quad \beta'(t) \geqslant f\big(t, \beta(t)\big) \quad \textit{a.e. on } [t_0, t_1]. \tag{1.4}$$

- *If $\alpha(t_0) \leqslant \beta(t_0)$, then for every $x_1 \in [\alpha(t_0), \beta(t_0)]$ there exists a solution $x : [t_0, t_1] \to \mathbb{R}$ of the initial value problem*

$$x' = f(t, x), \quad x(t_0) = x_1, \tag{1.5}$$

satisfying

$$\min\{\alpha,\beta\} \leqslant x \leqslant \max\{\alpha,\beta\} \quad \textit{on } [t_0,t_1]. \tag{1.6}$$

- *If* $\alpha(t_1) \geqslant \beta(t_1)$, *then for every* $x_2 \in [\beta(t_1),\alpha(t_1)]$ *there exists a solution* $x:[t_0,t_1] \to \mathbb{R}$ *of the terminal value problem*

$$x' = f(t,x), \quad x(t_1) = x_2, \tag{1.7}$$

  satisfying (1.6).
- *If* $\alpha(t_0) \leqslant \beta(t_0)$ *and* $\alpha(t_1) \geqslant \beta(t_1)$, *then for every* $x_1 \in [\alpha(t_0),\beta(t_0)]$ *and* $x_2 \in [\beta(t_1),\alpha(t_1)]$ *there exists a solution* $x:[t_0,t_1] \to \mathbb{R}$ *of both* (1.5) *and* (1.7) *satisfying* (1.6).

PROOF. We first consider the case where α and β are ordered.

CLAIM 1. *Let* $\alpha,\beta \in W^{1,1}(t_0,t_1)$ *satisfy* (1.4). *Assume that* $\alpha \leqslant \beta$ *on* $[t_0,t_1]$. *Then for each* $x_1 \in [\alpha(t_0),\beta(t_0)]$ *there exists a solution* $x:[t_0,t_1] \to \mathbb{R}$ *of* (1.5) *with* $\alpha \leqslant x \leqslant \beta$ *on* $[t_0,t_1]$.

Set $\gamma(t,x) = \max\{\alpha(t),\min\{x,\beta(t)\}\}$, $\bar{f}(t,x) = f(t,\gamma(t,x))$, and consider the problem

$$x' = \bar{f}(t,x), \quad x(t_0) = x_1. \tag{1.8}$$

Let $S:C^0([t_0,t_1]) \to C^0([t_0,t_1])$ be the Volterra operator defined by $S(x)(t) = x_1 + \int_{t_0}^{t} \bar{f}(s,x(s))\,\mathrm{d}s$. The operator S is continuous and has a relatively compact range. Hence, the existence of a solution x of (1.8) defined on $[t_0,t_1]$ follows from Schauder Theorem. We claim that $\alpha \leqslant x \leqslant \beta$ on $[t_0,t_1]$. Indeed, assuming that there are points s_1, s_2, with $t_0 \leqslant s_1 < s_2 \leqslant t_1$, such that $x(s_1) = \alpha(s_1)$ and $x(s) < \alpha(s)$ on $]s_1,s_2]$, yields $x'(s) - \alpha'(s) \geqslant 0$ a.e. on $]s_1,s_2]$, which leads to a contradiction. Therefore we have $\alpha \leqslant x$. Similarly we obtain $x \leqslant \beta$. Consequently x is a solution of (1.5) as well and the claim is proved.

CLAIM 2. *Let* $\alpha,\beta \in W^{1,1}(t_0,t_1)$ *satisfy* (1.4). *Assume that* $\alpha \geqslant \beta$ *on* $[t_0,t_1]$. *Then for each* $x_2 \in [\beta(t_1),\alpha(t_1)]$ *there exists a solution* $x:[t_0,t_1] \to \mathbb{R}$ *of* (1.7) *satisfying* $\alpha \geqslant x \geqslant \beta$ *on* $[t_0,t_1]$.

Let us consider the problem obtained from (1.7) by reversing time; namely, set $\hat{f}(t,x) = -f(t_0+t_1-t,x)$, $\hat{\alpha}(t) = \beta(t_0+t_1-t)$, $\hat{\beta}(t) = \alpha(t_0+t_1-t)$ and observe that $\hat{\alpha}'(t) \leqslant \hat{f}(t,\hat{\alpha}(t))$, $\hat{\beta}'(t) \geqslant \hat{f}(t,\hat{\beta}(t))$, a.e. on $[t_0,t_1]$, and that $\hat{\alpha}(t_0) \leqslant \hat{\beta}(t_0)$. By Claim 1 there exists a solution $\hat{x}:[t_0,t_1] \to \mathbb{R}$ of

$$x' = \hat{f}(t,x), \quad x(t_0) = x_2,$$

satisfying $\hat{\alpha} \leqslant \hat{x} \leqslant \hat{\beta}$ on $[t_0,t_1]$. The function $x(t) = \hat{x}(t_0+t_1-t)$ is then a solution of (1.7) satisfying $\alpha \geqslant x \geqslant \beta$ on $[t_0,t_1]$.

We now prove the three statements of this lemma.

Case 1: $\alpha(t_0) \leqslant \beta(t_0)$. Pick any $x_1 \in [\alpha(t_0), \beta(t_0)]$. Set $A = \{t \in [t_0, t_1] \mid \alpha(t) \leqslant \beta(t)\}$ and $B = [t_0, t_1] \setminus A$. We can write A as union of countably many disjoint closed intervals (possibly reduced to points) $[a_i, b_i]$, with $a_i \leqslant b_i$ and $a_0 = t_0$, and B as union of countably many disjoint open intervals $]c_j, d_j[$, with $c_j < d_j$. If $a_0 < b_0$, by Claim 1 we can find a solution x_0 of

$$x' = f(t, x), \quad x(t_0) = x_1,$$

satisfying $\alpha(t) \leqslant x_0(t) \leqslant \beta(t)$ on $[t_0, b_0]$ and, for every $i \geqslant 1$ such that $a_i < b_i$, a solution $x_i : [a_i, b_i] \to \mathbb{R}$ of

$$x' = f(t, x), \quad x(a_i) = \alpha(a_i),$$

satisfying $\alpha(t) \leqslant x_i(t) \leqslant \beta(t)$ on $[a_i, b_i]$. Notice that $x_i(b_i) = \alpha(b_i) = \beta(b_i)$ whenever $b_i \neq t_1$. For every j, by Claim 2 we can find a solution $y_j : [c_j, d_j] \to \mathbb{R}$ of

$$x' = f(t, x), \quad x(d_j) = \beta(d_j),$$

satisfying $\beta(t) \leqslant y_j(t) \leqslant \alpha(t)$ on $[c_j, d_j]$. Let us define $x : [t_0, t_1] \to \mathbb{R}$ by $x(t) = x_i(t)$ on $[a_i, b_i]$ and $x(t) = y_j(t)$ on $]c_j, d_j[$. Notice that $x(t_0) = x_1$, $x'(t) = f(t, x)$ a.e. on $[t_0, t_1]$ and (1.6) is satisfied. We only need to verify that x is absolutely continuous. Since x is bounded, by the L^1-Carathéodory conditions, there exists $\gamma \in L^1(t_0, t_1)$ such that $|f(t, x(t))| \leqslant \gamma(t)$ for all $t \in [t_0, t_1]$. Fix $\varepsilon > 0$ and pick points $r_k < s_k$ in $[t_0, t_1]$ with $k = 1, \dots, n$. Then $\sum_{k=1}^{n} |x(s_k) - x(r_k)|$ may be written as a series with terms of the form $|x_i(p_{k_i}) - x_i(q_{k_i})| \leqslant \int_{q_{k_i}}^{p_{k_i}} \gamma(s)\,\mathrm{d}s$ or $|y_j(p_{k_j}) - y_j(q_{k_j})| \leqslant \int_{q_{k_j}}^{p_{k_j}} \gamma(s)\,\mathrm{d}s$, whose sum is smaller than ε whenever $\sum_{k=1}^{n} |s_k - r_k|$ is sufficiently small.

Case 2: $\alpha(t_1) \geqslant \beta(t_1)$. We can argue as in Claim 2 applying our previous conclusion to a reversed time problem.

Case 3: $\alpha(t_0) \leqslant \beta(t_0)$ and $\alpha(t_1) \geqslant \beta(t_1)$. If either $\alpha(t_0) = \beta(t_0)$ or $\alpha(t_1) = \beta(t_1)$ the claim immediately follows from the previous cases. Otherwise, by Proposition 1.1 there exists a point $s_0 \in]t_0, t_1[$ such that $\alpha(s_0) = \beta(s_0)$. The assertion then follows by applying our preceding conclusions to problem (1.5) on $[t_0, s_0]$ and to problem (1.7) on $[s_0, t_1]$. □

We use Lemma 1.2 to recover an existence and localization result established in [28] under a slightly more restrictive notion of lower and upper solutions. We add here some information on the topological structure of the solution set, namely we prove the Hukuhara–Kneser property. In order to show that it is (arcwise) connected, we exploit the following property of compact subsets of $C^0([t_0, t_1])$ which are dense-in-itself with respect to the order.

DEFINITION 1.10. A subset $\mathcal{S}$ of $C^0([t_0, t_1])$ is said *dense-in-itself with respect to the order* if for any $u_1, u_2 \in \mathcal{S}$, with $u_1 < u_2$, there exists $u_3 \in \mathcal{S}$ with $u_1 < u_3 < u_2$.

LEMMA 1.3. *Let $\mathcal{S} \subseteq C^0([t_0, t_1])$ be a compact set which is dense-in-itself with respect to the order. Let $\mathcal{T} \subseteq \mathcal{S}$ be a maximal nondegenerate totally ordered subset of $\mathcal{S}$. Then $\mathcal{T}$ is homeomorphic to a nondegenerate compact interval of $\mathbb{R}$.*

PROOF. We start by verifying that $\mathcal{T}$ is closed in $\mathcal{S}$. Let $(z_n)_n$ be a sequence in $\mathcal{T}$ converging to some function $z \in \mathcal{S}$. Let us show that $z \in \mathcal{T}$, that is, for each $u \in \mathcal{T}$, either $u \leqslant z$ or $u \geqslant z$. Assume by contradiction that there exists $u \in \mathcal{T}$ such that $u \nleqslant z$ and $u \ngeqslant z$, i.e. $\min\{\|(u-z)^+\|_\infty, \|(u-z)^-\|_\infty\} > 0$. Take n such that $\|z_n - z\|_\infty < \min\{\|(u-z)^+\|_\infty, \|(u-z)^-\|_\infty\}$ and suppose, for instance, that $z_n \geqslant u$. We have $(z_n - z)^+ \geqslant (u-z)^+$ and hence $\|(u-z)^+\|_\infty \leqslant \|(z_n - z)^+\|_\infty \leqslant \|z_n - z\|_\infty < \|(u-z)^+\|_\infty$, which is a contradiction. Therefore, we conclude that $z \in \mathcal{T}$.

Moreover, as $\mathcal{T}$ is a maximal totally ordered subset of $\mathcal{S}$, $\mathcal{T}$ is dense-in-itself with respect to the order. Let $\varphi : C^0([t_0, t_1]) \to \mathbb{R}$ be a continuous linear functional such that $\varphi(u) > 0$ for every $u > 0$ (e.g., $\varphi(u) = \int_{t_0}^{t_1} u$). Since $\mathcal{T}$ is compact and $\varphi|_{\mathcal{T}}$ is one-to-one, $\varphi|_{\mathcal{T}}$ is a homeomorphism between $\mathcal{T}$ and $\varphi(\mathcal{T}) \subseteq \mathbb{R}$. Notice that $\varphi(\mathcal{T})$ is also dense-in-itself with respect to the ordering of $\mathbb{R}$ and hence it is a nondegenerate compact interval. □

THEOREM 1.4. *Assume* (C). *Let α be a lower solution and β be an upper solution of* (1.2) *on* $[t_0, t_1]$. *Then there exist the minimum solution v and the maximum solution w in* $[\min\{\alpha, \beta\}, \max\{\alpha, \beta\}]$ *of* (1.2) *on* $[t_0, t_1]$. *Further, the set*

$$\mathcal{K} = \left\{x : [t_0, t_1] \to \mathbb{R} \mid x \text{ is a solution of } (1.2) \text{ with } \min\{\alpha, \beta\} \leqslant x \leqslant \max\{\alpha, \beta\}\right\}$$

is a continuum, i.e. a compact and connected set, in $C^0([t_0, t_1])$.

PROOF. In this proof we use the following notations: if I is an interval, with endpoints a and b, and $u : I \to \mathbb{R}$, then we set $u(a^+) = \lim_{t\to a^+} u(t)$ and $u(b^-) = \lim_{t \to b^-} u(t)$, whenever the limits exist. The proof is divided into three steps.

Step 1. *Existence of a solution.* Let $\tau_0, \tau_1, \ldots, \tau_N$ and $\sigma_0, \sigma_1, \ldots, \sigma_M$ be the points of $[t_0, t_1]$ defining α and β as in Definition 1.9 and relabel them as $\rho_0 < \rho_1 < \cdots < \rho_P$. For each $n = 1, \ldots, P$ denote by α_n, β_n the continuous extensions on $[\rho_{n-1}, \rho_n]$ of $\alpha|_{]\rho_{n-1}, \rho_n[}$ and $\beta|_{]\rho_{n-1}, \rho_n[}$, respectively. We will prove the thesis by recursion on P.

- *There exists a solution y_1 of* (1.2) *such that* $\min\{\alpha, \beta\} \leqslant y_1 \leqslant \max\{\alpha, \beta\}$ *on* $[t_0, \rho_1]$. *Further, if* $\alpha(\rho_1^-) \geqslant \beta(\rho_1^-)$, *then for any x_1 between* $\min\{\alpha(\rho_1^-), \beta(\rho_1)\}$ *and* $\max\{\alpha(\rho_1), \beta(\rho_1^-)\}$, *the solution y_1 can be chosen so that* $y_1(\rho_1) = x_1$.

As $\alpha_1(t_0) = \alpha(t_0^+) \leqslant \alpha(t_0) \leqslant x_0 \leqslant \beta(t_0) \leqslant \beta(t_0^+) = \beta_1(t_0)$, the claim follows by Lemma 1.2.

- *Let* $n \leqslant P - 1$. *Assume there exists a solution y_n of* (1.2) *such that* $\min\{\alpha, \beta\} \leqslant y_n \leqslant \max\{\alpha, \beta\}$ *on* $[t_0, \rho_n]$ *and, in case* $\alpha(\rho_n^-) \geqslant \beta(\rho_n^-)$, *for all x_n between* $\min\{\alpha(\rho_n^-), \beta(\rho_n)\}$ *and* $\max\{\alpha(\rho_n), \beta(\rho_n^-)\}$ *the solution y_n can be chosen so that* $y_n(\rho_n) = x_n$. *Then there exists a solution y_{n+1} of* (1.2) *such that* $\min\{\alpha, \beta\} \leqslant y_{n+1} \leqslant \max\{\alpha, \beta\}$ *on* $[t_0, \rho_{n+1}]$ *and, in case* $\alpha(\rho_{n+1}^-) \geqslant \beta(\rho_{n+1}^-)$, *for all x_{n+1} between* $\min\{\alpha(\rho_{n+1}^-), \beta(\rho_{n+1})\}$ *and* $\max\{\alpha(\rho_{n+1}), \beta(\rho_{n+1}^-)\}$ *the solution y_{n+1} can be chosen so that* $y_{n+1}(\rho_{n+1}) = x_{n+1}$.

To prove this we need to distinguish several cases.

(i) Assume $\alpha(\rho_n^-) \leqslant \beta(\rho_n^-)$. Hence $\alpha(\rho_n) \leqslant \beta(\rho_n)$. Let y_n be a solution as in the assumption. As $\alpha_{n+1}(\rho_n) \leqslant y_n(\rho_n) \leqslant \beta_{n+1}(\rho_n)$, by Lemma 1.2 there exists a solution z of (1.1) such that $z(\rho_n) = y_n(\rho_n)$, $\min\{\alpha_{n+1}, \beta_{n+1}\} \leqslant z \leqslant \max\{\alpha_{n+1}, \beta_{n+1}\}$ on $[\rho_n, \rho_{n+1}]$ and, in case $\alpha_{n+1}(\rho_{n+1}) \geqslant \beta_{n+1}(\rho_{n+1})$, for all $x_{n+1} \in [\beta_{n+1}(\rho_{n+1}), \alpha_{n+1}(\rho_{n+1})]$ such a solution can be chosen so that $z(\rho_{n+1}) = x_{n+1}$. Since in this case the interval $[\beta_{n+1}(\rho_{n+1}), \alpha_{n+1}(\rho_{n+1})]$ is contained in the interval with endpoints $\min\{\alpha(\rho_{n+1}^-), \beta(\rho_{n+1})\}$ and $\max\{\alpha(\rho_{n+1}), \beta(\rho_{n+1}^-)\}$, the function y_{n+1} defined by $y_{n+1}(t) = y_n(t)$ on $[t_0, \rho_n]$ and $y_{n+1}(t) = z(t)$ on $[\rho_n, \rho_{n+1}]$ is as required.

(ii) Assume $\alpha(\rho_n^-) > \beta(\rho_n^-)$ and $\alpha(\rho_n) \leqslant \beta(\rho_n)$. Then $\max\{\alpha(\rho_n), \beta(\rho_n^-)\} \leqslant \min\{\alpha(\rho_n^-), \beta(\rho_n)\}$. Let $x_n \in [\max\{\alpha(\rho_n), \beta(\rho_n^-)\}, \min\{\alpha(\rho_n^-), \beta(\rho_n)\}]$. By assumption there exists a solution y_n of (1.2) such that $\min\{\alpha, \beta\} \leqslant y_n \leqslant \max\{\alpha, \beta\}$ on $[t_0, \rho_n]$ and $y_n(\rho_n) = x_n$. As $\alpha_{n+1}(\rho_n) \leqslant y_n(\rho_n) \leqslant \beta_{n+1}(\rho_n)$, arguing as in (i) we get the existence of the solution y_{n+1} with all the required properties.

(iii) Assume $\alpha(\rho_n) > \beta(\rho_n)$ (hence $\alpha(\rho_n^-) > \beta(\rho_n^-)$) and $\alpha(\rho_n^+) \leqslant \beta(\rho_n^+)$. Let $x_n \in [\alpha(\rho_n^+), \beta(\rho_n^+)] \cap [\beta(\rho_n), \alpha(\rho_n)]$ and argue as in (ii).

(iv) Assume $\alpha(\rho_n^+) > \beta(\rho_n^+)$ (hence $\alpha(\rho_n) > \beta(\rho_n)$ and $\alpha(\rho_n^-) > \beta(\rho_n^-)$) and $\alpha(\rho_{n+1}^-) > \beta(\rho_{n+1}^-)$. For any x_{n+1} between $\min\{\alpha(\rho_{n+1}^-), \beta(\rho_{n+1})\}$ and $\max\{\alpha(\rho_{n+1}), \beta(\rho_{n+1}^-)\}$ we have $\alpha_{n+1}(\rho_{n+1}) \geqslant x_{n+1} \geqslant \beta_{n+1}(\rho_{n+1})$ and hence, by Lemma 1.2, there exists a solution z of (1.1) such that $z(\rho_{n+1}) = x_{n+1}$ and $\min\{\alpha_{n+1}, \beta_{n+1}\} \leqslant z \leqslant \max\{\alpha_{n+1}, \beta_{n+1}\}$ on $[\rho_n, \rho_{n+1}]$. As $\beta(\rho_n) \leqslant \beta(\rho_n^+) \leqslant z(\rho_n) \leqslant \alpha(\rho_n^+) \leqslant \alpha(\rho_n)$, by assumption there exists a solution y_n of (1.2) such that $\min\{\alpha, \beta\} \leqslant y_n \leqslant \max\{\alpha, \beta\}$ on $[t_0, \rho_n]$ and $y_n(\rho_n) = z(\rho_n)$. The function y_{n+1} defined by $y_{n+1}(t) = y_n(t)$ on $[t_0, \rho_n]$ and $y_{n+1}(t) = z(t)$ on $[\rho_n, \rho_{n+1}]$ is as required.

(v) Assume $\alpha(\rho_n^+) > \beta(\rho_n^+)$ (hence $\alpha(\rho_n) > \beta(\rho_n)$ and $\alpha(\rho_n^-) > \beta(\rho_n^-)$) and $\alpha(\rho_{n+1}^-) \leqslant \beta(\rho_{n+1}^-)$. By continuity there is a point $\hat{t} \in]\rho_n, \rho_{n+1}[$ such that $\alpha(\hat{t}) = \beta(\hat{t})$. Arguing as in (iv), by Lemma 1.2 for any x_{n+1} between $\min\{\alpha(\rho_{n+1}^-), \beta(\rho_{n+1})\}$ and $\max\{\alpha(\rho_{n+1}), \beta(\rho_{n+1}^-)\}$ there exists a solution z_2 of (1.1) such that $z_2(\hat{t}) = \alpha(\hat{t})$, $z_2(\rho_{n+1}) = x_{n+1}$ and $\min\{\alpha_{n+1}, \beta_{n+1}\} \leqslant z_2 \leqslant \max\{\alpha_{n+1}, \beta_{n+1}\}$ on $[\hat{t}, \rho_{n+1}]$. Further there exists a solution z_1 of (1.1) such that $z_1(\hat{t}) = \alpha(\hat{t})$ and $\min\{\alpha_{n+1}, \beta_{n+1}\} \leqslant z_1 \leqslant \max\{\alpha_{n+1}, \beta_{n+1}\}$ on $[\rho_n, \hat{t}]$. As $\beta(\rho_n) \leqslant \beta(\rho_n^+) \leqslant z_1(\rho_n) \leqslant \alpha(\rho_n^+) \leqslant \alpha(\rho_n)$, by assumption there exists a solution y_n of (1.2) such that $\min\{\alpha, \beta\} \leqslant y_n \leqslant \max\{\alpha, \beta\}$ on $[t_0, \rho_n]$ and $y_n(\rho_n) = z_1(\rho_n)$. The function y_{n+1} defined by $y_{n+1}(t) = y_n(t)$ on $[t_0, \rho_n]$, $y_{n+1}(t) = z_1(t)$ on $[\rho_n, \hat{t}]$, $y_{n+1}(t) = z_2(t)$ on $[\hat{t}, \rho_{n+1}]$ is as required.

Step 2. Existence of extremal solutions. We have just seen that $\mathcal{K}$ is nonempty. Let us show that $\mathcal{K}$ has a minimum and a maximum. Since the Volterra operator $S : C^0([t_0, t_1]) \to C^0([t_0, t_1])$ associated with (1.2) is completely continuous and the set of its fixed points lying between $\min\{\alpha, \beta\}$ and $\max\{\alpha, \beta\}$ is precisely $\mathcal{K}$, we have that $\mathcal{K}$ is a nonempty compact set. For each $x \in \mathcal{K}$ define the closed set $\mathcal{C}_x = \{z \in \mathcal{K} \mid z \leqslant x\}$. The family $\{\mathcal{C}_x \mid x \in \mathcal{K}\}$ has the finite intersection property since, if $x_1, x_2 \in \mathcal{K}$, then $\min\{x_1, x_2\} \in \mathcal{K}$. By the compactness of $\mathcal{K}$ there exists $v \in \bigcap_{x \in \mathcal{K}} \mathcal{C}_x$; clearly, v is the minimum solution we are looking for. The maximum solution w can be found in a similar way.

Step 3. $\mathcal{K}$ is a continuum. We have only to show that $\mathcal{K}$ is connected. Let us prove that $\mathcal{K}$ is dense-in-itself with respect to the order. Take $x_1, x_2 \in \mathcal{K}$ with $x_1 < x_2$. Let $s_0 \in]t_0, t_1[$ be such that $x_1(s_0) < x_2(s_0)$ and pick a real number $\hat{x}_0$ with $x_1(s_0) < \hat{x}_0 < x_2(s_0)$. By Lemma 1.2 there exists a solution $x_3 : [t_0, t_1] \to \mathbb{R}$ of (1.1), satisfying $x_3(t_0) = x_0$ and $x_3(s_0) = \hat{x}_0$, such that $x_1 \leqslant x_3 \leqslant x_2$; hence $x_3 \in \mathcal{K}$ and $x_1 < x_3 < x_2$. Let us show that $\mathcal{K}$ is arcwise connected. For any $x_1, x_2 \in \mathcal{K}$, with $x_1 \neq x_2$, let $\mathcal{T}_1$ be a maximal totally ordered subset of $\mathcal{K}$ containing x_1 and $\min\{x_1, x_2\}$ and let $\mathcal{T}_2$ be a maximal totally ordered subset of

$\mathcal{K}$ containing x_2 and $\min\{x_1, x_2\}$. By Lemma 1.3, $\mathcal{T}_1$ and $\mathcal{T}_2$ are homeomorphic to compact real intervals. Since $\min\{x_1, x_2\} \in \mathcal{T}_1 \cap \mathcal{T}_2$, we find an arc connecting x_1 to x_2. □

In the sequel we shall also use the following partial extension of Theorem 1.4 to noncompact intervals. Related results can be found in [90].

COROLLARY 1.5. *Assume* (C). *Let α be a lower solution and β be an upper solution of* (1.2) *on* $[t_0, \tau[$, *with* $t_0 < \tau \leqslant +\infty$, *and assume that* $\alpha \leqslant \beta$. *Then there exists the minimum solution v and the maximum solution w in* $[\alpha, \beta]$ *of* (1.2) *on* $[t_0, \tau[$. *Further, the set*

$$\mathcal{K} = \big\{x : [t_0, \tau[\to \mathbb{R} \mid x \textit{ is a solution of } (1.2) \textit{ with } \alpha \leqslant x \leqslant \beta\big\}$$

is a continuum in $C^0([t_0, \tau[)$, *endowed with the topology of uniform convergence on compact intervals.*

PROOF. Let $(\tau_n)_n$, with $\tau_0 = t_0$, be a strictly increasing sequence converging to τ.

Step 1. *Existence of extremal solutions.* In order to prove the existence of the minimum solution v in $[\alpha, \beta]$ of (1.2) on $[t_0, \tau[$, we apply recursively Theorem 1.4 on each interval $[t_0, \tau_n]$, with $n \geqslant 1$ Let us denote by v_n the minimum solution v in $[\alpha, \beta]$ of (1.2) on $[t_0, \tau_n]$. By the minimality of v_n, we have $v_{n+1}|_{[t_0,\tau_n]} \geqslant v_n$. On the other hand, as $\alpha|_{[\tau_n,\tau_{n+1}]}$ and $v_{n+1}|_{[\tau_n,\tau_{n+1}]}$ are, respectively, a lower and an upper solution of the initial value problem $x' = f(t, x)$, $x(\tau_n) = v_n(\tau_n)$, we can continue v_n to a solution $\hat{v}_{n+1}$ of (1.2) on $[t_0, \tau_{n+1}]$ satisfying $\alpha \leqslant \hat{v}_{n+1} \leqslant v_{n+1}$. By the minimality of v_{n+1}, we conclude that $\hat{v}_{n+1} = v_{n+1}$ and hence $v_{n+1}|_{[t_0,\tau_n]} = v_n$. Then we define $v : [t_0, \tau[\to \mathbb{R}$ by setting $v(t) = v_n(t)$ on $[t_0, \tau_n]$. We have that v is the minimum solution in $[\alpha, \beta]$ of (1.2) on $[t_0, \tau[$, because, if x were a solution of (1.2) on $[t_0, \tau[$ with $x \geqslant \alpha$ and $x \ngeqslant v$, then it should follow $x|_{[t_0,\tau_n]} \ngeqslant v_n$ for some n, thus contradicting the minimality of v_n. Similarly, we prove the existence of the maximum solution w in $[\alpha, \beta]$ of (1.2) on $[t_0, \tau[$.

Step 2. *$\mathcal{K}$ is a continuum.* For each $x \in \mathcal{K}$ and $n \geqslant 1$, set $x_n = x|_{[t_0,\tau_n]}$. We denote by $\mathcal{K}_n$ the set of all solutions $x : [t_0, \tau_n] \to \mathbb{R}$ of (1.2) such that $\alpha \leqslant x \leqslant \beta$ on $[t_0, \tau_n]$. By Theorem 1.4, $\mathcal{K}_n$ is a continuum in $C^0([t_0, \tau_n])$. For every $m < n$, let also $\pi^n_m : \mathcal{K}_n \to \mathcal{K}_m$ be the restriction map on $[t_0, \tau_m]$, i.e. $\pi^n_m(x) = x|_{[t_0,\tau_m]}$ for all $x \in \mathcal{K}_n$. Let us define now a function $\chi : \mathcal{K} \to \prod_{n=1}^{+\infty} \mathcal{K}_n$, by setting $\chi(x) = (x_n)_n$. Observe that χ is a homeomorphism of $\mathcal{K}$ into $\prod_{n=1}^{+\infty} \mathcal{K}_n$, when $\prod_{n=1}^{+\infty} \mathcal{K}_n$ is endowed with the Tychonoff product topology, and its range $\chi(\mathcal{K})$ is the set of all sequences $(x_n)_n \in \prod_{n=1}^{+\infty} \mathcal{K}_n$ such that, for all $m < n$, $\pi^n_m(x_n) = x_m$, i.e. $\chi(\mathcal{K})$ is the inverse limit of the sequence $(\mathcal{K}_n)_n$ with bonding maps π^n_m (cf. [47]). As the inverse limit of a sequence of continua is a continuum [47, Theorem 6.1.20] we deduce that $\chi(\mathcal{K})$ is a continuum as well. Since $\mathcal{K}$ is homeomorphic to $\chi(\mathcal{K})$, the conclusion is achieved. □

1.4. *The periodic problem*

In this section we discuss existence and localization of solutions of the periodic problem (1.3), in the presence of a pair of possibly discontinuous and unordered lower and upper solutions. With the aim of obtaining a global portrait of solutions of Eq. (1.1), we also introduce the notion of T-monotonicity. This concept plays a central role in the qualitative study of solutions of (1.1), as it helps to get a thorough classification of their asymptotic behaviour. Therefore we develop some criteria for detecting T-monotonicity. However, it should be stressed that also the study of solutions that are not T-monotone has an interest, since their existence gives rise to T-periodic solutions. All these facts are then used to get an extension of the existence part of the classical Massera Convergence Theorem to the Carathéodory setting.

Solutions and lower and upper solutions

DEFINITION 1.11. A *solution of* (1.3) is a solution u of (1.1) on $\mathbb{R}$ which is T-periodic, i.e. $u(t+T) = u(t)$ on $\mathbb{R}$.

We introduce some notions of possibly discontinuous lower and upper solutions for the periodic problem (1.3).

DEFINITION 1.12.

- A *lower solution of* (1.3) is a T-periodic function $\alpha : \mathbb{R} \to \mathbb{R}$ which is a lower solution of (1.1) on $\mathbb{R}$.
- A *regular lower solution of* (1.3) is a lower solution α of (1.3) such that $\alpha|_{[0,T]}$ is a regular lower solution of (1.1) on $[0, T]$.
- An *upper solution of* (1.3) is a T-periodic function $\beta : \mathbb{R} \to \mathbb{R}$ which is an upper solution of (1.1) on $\mathbb{R}$.
- A *regular upper solution of* (1.3) is an upper solution β of (1.3) such that $\beta|_{[0,T]}$ is a regular upper solution of (1.1) on $[0, T]$.
- A lower solution of (1.3) (respectively an upper solution of (1.3)) is *proper* if it is not a solution of (1.3).
- A proper lower solution α of (1.3) is *strict* if every solution x of (1.3), with $x > \alpha$, is such that $x \gg \alpha$. Similarly, a proper upper solution of (1.3) is *strict* if every solution x of (1.3), with $x < \beta$, is such that $x \ll \beta$.

REMARK 1.4. We notice that even regular lower and upper solutions of (1.3) may be discontinuous at the endpoints of the interval $[0, T]$.

REMARK 1.5. Sometimes we speak of solutions of (1.3) with reference to functions defined on $[t_0, t_0 + T]$ for some $t_0 \in \mathbb{R}$. We also speak of lower and upper solutions of (1.3) with reference to functions defined on $[t_0, t_0 + T[$ for some $t_0 \in \mathbb{R}$. In these cases it is understood that their T-periodic extensions to $\mathbb{R}$ have to be considered.

In the sequel we shall need the following observation, whose proof is a slight modification of that of Proposition 1.1.

PROPOSITION 1.6. *Let α be a lower solution and β be an upper solution of* (1.3), *with $\alpha \not\leqslant \beta$ and $\alpha \not\geqslant \beta$. Then there exists $\hat{t} \in [0, T]$ such that $\alpha(\hat{t}) = \beta(\hat{t})$.*

T-monotonicity

The notion of T-monotonicity plays an important role in the qualitative study of solutions of (1.1). When uniqueness of solutions of (1.2) holds, then any solution is T-monotone and this, in turn, corresponds to the monotonicity of the Poincaré operator. Of course, in the Carathéodory setting this is not anymore true.

DEFINITION 1.13.

- A function $x : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval, is said *T-increasing* (respectively, *T-decreasing*) if $x(t) \leqslant x(t+T)$ (respectively $x(t) \geqslant x(t+T)$) for each $t \in I$ such that $t + T \in I$. A function is said *T-monotone* if it is either T-increasing or T-decreasing.
- A function $x : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval unbounded from above, is said *eventually T-increasing* (respectively, *eventually T-decreasing*) if there exists $t_0 \in I$ such that $x|_{]t_0,+\infty[}$ is T-increasing (respectively, T-decreasing).

REMARK 1.6. Let x be a solution of (1.1) on $[t_0, +\infty[$. If x is T-increasing then the sequence $(\alpha_n)_n$, defined by setting, for each $n \geqslant t_0/T$, $\alpha_n(t) = x(t+nT)$ on $[0, T[$ gives rise to an increasing sequence of regular lower solutions of (1.3). Similarly, any T-decreasing solution of (1.1) on $[t_0, +\infty[$, gives rise to a decreasing sequence $(\beta_n)_n$ of regular upper solutions of (1.3).

The Massera Convergence Theorem for T-monotone solutions

The following simple result may be considered as a first step towards an extension of the classical Massera Convergence Theorem for scalar ordinary differential equations [91] to the case where the right-hand side f of (1.1) is a Carathéodory function.

PROPOSITION 1.7. *Assume* (C). *Let $x :]t_0, +\infty[\to \mathbb{R}$ be a bounded, T-increasing* (*respectively T-decreasing*) *solution of* (1.1). *Then there exists a solution u of* (1.3) *such that $u \geqslant x$* (*respectively $u \leqslant x$*) *on* $]t_0, +\infty[$ *and*

$$\lim_{t\to+\infty} \bigl(x(t) - u(t)\bigr) = 0. \tag{1.9}$$

REMARK 1.7. Symmetrically, we have that, if $x :]-\infty, t_0[\to \mathbb{R}$ is a bounded, T-increasing (respectively T-decreasing) solution of (1.1), then there exists a solution u of (1.3) such that $u \leqslant x$ (respectively $u \geqslant x$) on $]-\infty, t_0[$ and

$$\lim_{t\to-\infty} \bigl(x(t) - u(t)\bigr) = 0.$$

PROOF. Assume, without loss of generality, that $t_0 = n_0 T$, for some $n_0 \in \mathbb{N}$, and suppose that x is T-increasing. Define a sequence $(x_n)_n$ of functions by setting, for every $n \geqslant n_0$ and every $t \in [0, T]$, $x_n(t) = x(t + nT)$. Then x_n is a regular lower solution of (1.3), satisfying

$$x_{n+1}(0) = x_n(T) \tag{1.10}$$

and $x_n \leqslant x_{n+1}$, on $[0, T]$. By assumption, the sequence $(x_n)_n$ is bounded in $C^0([0, T])$ and hence, by monotonicity, it converges pointwise on $[0, T]$ to a function u, which by (1.10) satisfies $u(0) = u(T)$. Moreover, the L^1-Carathéodory conditions imply that $(x_n)_n$ is equicontinuous. Therefore, by Arzelà–Ascoli theorem and monotonicity, $(x_n)_n$ converges uniformly to u. Finally, using the integral representation of solutions of (1.1) and the L^1-Carathéodory conditions again, we conclude that u is a solution and that the convergence takes place in $W^{1,1}(0, T)$.

Let us show now that $\lim_{t\to+\infty}(x(t) - u(t)) = 0$. Indeed, since $(x_n)_n$ converges uniformly to u, given $\varepsilon > 0$ there is $\bar{n}$ such that, for any $n \geqslant \bar{n}$, $|x_n(t) - u(t)| < \varepsilon$ for all $t \in [0, T]$. Hence, if we take $t \geqslant \bar{n}T$, with $t \in [nT, (n + 1)T[$ for some $n \geqslant \bar{n}$, we obtain, by the T-periodicity of u, $|x(t) - u(t)| = |x_n(t - nT) - u(t - nT)| < \varepsilon$. □

A Monotone Convergence Criterion

We now prove a technical result, which will often be used in the sequel. It describes the qualitative behaviour of solutions of (1.1), emanating from and lying above a lower solution of (1.3). Among these we single out a special solution $\widetilde{\alpha}$ which is T-increasing and hence, if bounded, does converge to a T-periodic solution. This statement is reminiscent of the Monotone Convergence Criterion in the theory of monotone maps (cf. [60, Section 5]) and essentially reduces to it when uniqueness of solutions of (1.2) holds.

PROPOSITION 1.8. *Assume* (C).

(i) *Let α be a lower solution of* (1.3) *and let $t_0 \in \mathbb{R}$. Then there exist $\omega \in\,]t_0, +\infty]$ and a T-increasing solution $\widetilde{\alpha} : [t_0, \omega[\to \mathbb{R}$ of* (1.2), *with $x_0 = \alpha(t_0)$, satisfying $\widetilde{\alpha} \geqslant \alpha$ on $[t_0, \omega[$. Further, every solution $x : [t_0, \tau[\to \mathbb{R}$ of* (1.1), *with $x \geqslant \alpha$, is such that $\tau \leqslant \omega$ and $x \geqslant \widetilde{\alpha}$ on $[t_0, \tau[$. Finally, we have that either every right-nonextendible solution $x : [t_0, \tau[\to \mathbb{R}$ of* (1.1), *with $x \geqslant \alpha$, is such that*

$$\limsup_{t\to\tau} x(t) = +\infty$$

and (1.3) *has no solution in $[\alpha, +\infty[$, or there exists the minimum solution v in $[\alpha, +\infty[$ of* (1.3); *in the latter case $\omega = +\infty$, $v \geqslant \widetilde{\alpha}$ and*

$$\lim_{t\to+\infty} \big(\widetilde{\alpha}(t) - v(t)\big) = 0.$$

(ii) *Let β be an upper solution of* (1.3) *and let $t_0 \in \mathbb{R}$. Then there exist $\omega \in\,]t_0, +\infty]$ and a T-decreasing solution $\widetilde{\beta} : [t_0, \omega[\to \mathbb{R}$ of* (1.2), *with $x_0 = \beta(t_0)$, satisfying $\widetilde{\beta} \leqslant \beta$ on $[t_0, \omega[$. Further, every solution $x : [t_0, \tau[\to \mathbb{R}$ of* (1.1), *with $x \leqslant \beta$, is such that $\tau \leqslant \omega$ and $x \leqslant \widetilde{\beta}$ on $[t_0, \tau[$. Finally, we have that either every right-nonextendible solution $x : [t_0, \tau[\to \mathbb{R}$ of* (1.1), *with $x \leqslant \beta$, is such that*

$$\liminf_{t\to\tau} x(t) = -\infty$$

and (1.3) *has no solution in* $]-\infty, \beta]$, *or there exists the maximum solution* w *in* $]-\infty, \beta]$ *of* (1.3); *in the latter case* $\omega = +\infty$, $w \leqslant \widetilde{\beta}$ *and*

$$\lim_{t \to +\infty} \bigl(\widetilde{\beta}(t) - w(t)\bigr) = 0.$$

REMARK 1.8. If $\omega < +\infty$, then $\lim_{t \to \omega} \widetilde{\alpha}(t) = +\infty$. Whereas, it may happen that, if $\omega = +\infty$, $\limsup_{t \to \omega} \widetilde{\alpha}(t) = +\infty$ and $\liminf_{t \to \omega} \widetilde{\alpha}(t) < +\infty$. A simple example is found by taking, in (1.1), $f(t,x) = x^2 \sin t$ if $|t| \leqslant \pi$, $f(t,x) = -x \sin t$ if $\pi < t < 2\pi$, and extending f by 3π-periodicity with respect to the t-variable. A similar remark holds for $\widetilde{\beta}$.

PROOF. We only prove statement (i), as the proof of (ii) is similar.

Let $\omega \in]t_0, +\infty]$ be the supremum of all $\tau > t_0$ such that there exists a solution $x : [t_0, \tau] \to \mathbb{R}$ of (1.2), with $x_0 = \alpha(t_0)$, satisfying $x \geqslant \alpha$ on $[t_0, \tau]$. The existence of such solutions x follows applying Theorem 1.4 to a modified problem. Take a strictly increasing sequence $(\tau_n)_n$ such that $\tau_n \to \omega$ and let $(x_n)_n$ be a corresponding sequence of solutions. Since x_1 is an upper solution and $\alpha|_{[t_0, \tau_1]}$ is a lower solution of (1.2), with $x_0 = \alpha(t_0)$, Theorem 1.4 yields the existence of the minimum solution v_1 in $[\alpha, x_1]$ of (1.2) on $[t_0, \tau_1]$. Actually, v_1 is the minimum solution in $[\alpha, +\infty[$ of (1.2), with $x_0 = \alpha(t_0)$, on $[t_0, \tau_1]$. Since $x_2|_{[\tau_1, \tau_2]}$ is an upper solution and $\alpha|_{[\tau_1, \tau_2]}$ is a lower solution of

$$x' = f(t,x), \quad x(\tau_1) = v_1(\tau_1), \tag{1.11}$$

on $[\tau_1, \tau_2]$, there is the minimum solution v_2 in $[\alpha, x_2]$ of (1.11) on $[\tau_1, \tau_2]$. Actually, v_2 is the minimum solution in $[\alpha, +\infty[$ of (1.11) on $[\tau_1, \tau_2]$. Proceeding in this way, we construct a sequence $(v_n)_n$ such that, for each n, v_n is the minimum solution in $[\alpha, +\infty[$ of

$$x' = f(t,x), \quad x(\tau_{n-1}) = v_{n-1}(\tau_{n-1}),$$

on $[\tau_{n-1}, \tau_n]$. Now, we define a function $\widetilde{\alpha} : [t_0, \omega[\to \mathbb{R}$ by setting

$$\widetilde{\alpha}(t) = v_n(t) \quad \text{if } t \in [\tau_{n-1}, \tau_n[.$$

It is clear that $\widetilde{\alpha}$ is a right-nonextendible solution of (1.2), with $x_0 = \alpha(t_0)$, satisfying $\widetilde{\alpha} \geqslant \alpha$ on $[t_0, \omega[$. Moreover, from the minimality of each v_n we deduce that every solution $x : [t_0, \tau[\to \mathbb{R}$ of (1.2), with $x_0 = \alpha(t_0)$, satisfying $x \geqslant \alpha$, is such that $\tau \leqslant \omega$ and $x \geqslant \widetilde{\alpha}$ on $[t_0, \tau[$. The same holds if $x_0 > \alpha(t_0)$, as $\min\{x, \widetilde{\alpha}\}$ is a solution of (1.2) with $x_0 = \alpha(t_0)$.

If $\omega \leqslant t_0 + T$ the lemma is proved. Suppose next that $t_0 + mT < \omega \leqslant t_0 + (m+1)T$ for some integer $m \geqslant 1$. In this case we may assume that $\tau_k = t_0 + kT$ for $k = 1, \dots, m$, and the T-monotonicity of $\widetilde{\alpha}$ follows from the minimality of the solutions v_n on $[\tau_{n-1}, \tau_n]$. Finally suppose that $\omega = +\infty$. In this case we may assume that $\tau_n = t_0 + nT$ for all n and the T-monotonicity of $\widetilde{\alpha}$ follows as in the previous case. Suppose further that $\limsup_{t \to +\infty} \widetilde{\alpha}(t) < +\infty$. Since $\widetilde{\alpha}$ is T-increasing, Proposition 1.7 implies that there exists a solution v of (1.3) such that $\widetilde{\alpha} \leqslant v$ on $[t_0, +\infty[$ and

$$\lim_{t \to +\infty} \bigl(\widetilde{\alpha}(t) - v(t)\bigr) = 0.$$

Note that v is the minimum solution in $[\alpha, +\infty[$ of (1.3), because the existence of a solution u of (1.3), such that $\alpha \leqslant u < v$, would contradict the minimality properties of $\widetilde{\alpha}$. □

REMARK 1.9. The following version of Proposition 1.8 for a terminal value problem holds.

(j) Let β be an upper solution of (1.3) and let $t_0 \in \mathbb{R}$. Then there exist $\omega \in [-\infty, t_0[$ and a T-decreasing solution $\widetilde{\beta} :]\omega, t_0] \to \mathbb{R}$ of the terminal value problem (1.2), with $x_0 = \beta(t_0)$, satisfying $\widetilde{\beta} \geqslant \beta$ on $]\omega, t_0]$. Further, every solution $x :]\tau, t_0] \to \mathbb{R}$ of (1.1), with $x \geqslant \beta$, is such that $\tau \geqslant \omega$ and $x \geqslant \widetilde{\beta}$ on $]\tau, t_0]$. Finally, we have that either every left-nonextendible solution $x :]\tau, t_0] \to \mathbb{R}$ of (1.1), with $x \geqslant \beta$, is such that

$$\limsup_{t \to \tau} x(t) = +\infty$$

and (1.3) has no solution in $[\beta, +\infty[$, or there exists the minimum solution v in $[\beta, +\infty[$ of (1.3); in the latter case $\omega = -\infty$, $v \geqslant \widetilde{\beta}$ and

$$\lim_{t \to -\infty} \bigl(\widetilde{\beta}(t) - v(t)\bigr) = 0.$$

(jj) Let α be a lower solution of (1.3) and let $t_0 \in \mathbb{R}$. Then there exist $\omega \in [-\infty, t_0[$ and a T-increasing solution $\widetilde{\alpha} :]\omega, t_0] \to \mathbb{R}$ of the terminal value problem (1.2), with $x_0 = \alpha(t_0)$, satisfying $\widetilde{\alpha} \leqslant \alpha$ on $]\omega, t_0]$. Further, every solution $x :]\tau, t_0] \to \mathbb{R}$ of (1.1), with $x \leqslant \alpha$, is such that $\tau \geqslant \omega$ and $x \leqslant \widetilde{\alpha}$ on $]\tau, t_0]$. Finally, we have that either every left-nonextendible solution $x :]\tau, t_0] \to \mathbb{R}$ of (1.1), with $x \leqslant \alpha$, is such that

$$\liminf_{t \to \tau} x(t) = -\infty,$$

and (1.3) has no solution in $]-\infty, \alpha]$, or there exists the maximum solution w in $]-\infty, \alpha]$ of (1.3); in the latter case $\omega = -\infty$, $w \leqslant \widetilde{\alpha}$ and

$$\lim_{t \to -\infty} \bigl(\widetilde{\alpha}(t) - w(t)\bigr) = 0.$$

Existence of solutions

The counterpart of Theorem 1.4 for the periodic problem (1.3) is the following theorem.

THEOREM 1.9. *Assume* (C). *Let α be a lower solution and β be an upper solution of* (1.3). *Then there exist the minimum solution v and the maximum solution w in* $[\min\{\alpha, \beta\}, \max\{\alpha, \beta\}]$ *of* (1.3).

PROOF. We distinguish three cases.

Case 1: $\alpha \leqslant \beta$. Let $\widetilde{\alpha}$ be the T-increasing solution of (1.2), with $x_0 = \alpha(t_0)$, whose existence is guaranteed by Proposition 1.8. Corollary 1.5, together with the minimality property of $\widetilde{\alpha}$, imply that $\widetilde{\alpha}$ is defined on $[t_0, +\infty[$ and $\widetilde{\alpha} \leqslant \beta$. Since $\widetilde{\alpha}$ is bounded, there exists the minimum solution v in $[\alpha, \beta]$ of (1.3). A symmetric argument proves the existence of the maximum solution w in $[\alpha, \beta]$ of (1.3).

Case 2: $\alpha > \beta$. We can reduce this case to the previous one by reversing time, as we did in Claim 2 of Lemma 1.2.

Case 3: $\alpha \nleqslant \beta$ *and* $\beta \nleqslant \alpha$. By Proposition 1.6 there exists $t_0 \in [0, T]$ such that $\alpha(t_0) = \beta(t_0)$. The result follows then immediately by Theorem 1.4 applied to problem (1.2), with $x_0 = \alpha(t_0) = \beta(t_0) = \alpha(t_0 + T) = \beta(t_0 + T)$, on the interval $[t_0, t_0 + T]$. □

We produce a simple example showing how Theorem 1.9 can be effectively applied to detect the existence of branches of T-periodic solutions by an appropriate choice of lower and upper solutions.

EXAMPLE 1.1. Let us consider the equation

$$x' = \sqrt{|x|} + h(t), \tag{1.12}$$

with $h : \mathbb{R} \to \mathbb{R}$ the 1-periodic function defined by $h(t) = -\frac{1}{8}|t|$ on $[-\frac{1}{2}, \frac{1}{2}[$. For any $p \in\,]-2^{-10}, 2^{-10}[$, define $\alpha_p : [-\frac{1}{2}, \frac{1}{2}[\, \to \mathbb{R}$, by setting $\alpha_p(-\frac{1}{2}) = p$ and $\alpha_p(t) = 2^{-4}\operatorname{sgn}(t)t^2$ on $]-\frac{1}{2}, \frac{1}{2}[$, and $\beta_p : [-\frac{1}{2}, \frac{1}{2}[\, \to \mathbb{R}$, by setting $\beta_p(-\frac{1}{2}) = p$ and $\beta_p(t) = -2^{-8}\operatorname{sgn}(t)t^2$ on $]-\frac{1}{2}, \frac{1}{2}[$. Then α_p and β_p are respectively a lower and an upper solution of the 1-periodic problem associated with (1.12). By Theorem 1.9, Eq. (1.12) has, for each $p \in\,]-2^{-10}, 2^{-10}[$, a 1-periodic solution u_p such that $u_p(-\frac{1}{2}) = p$.

A criterion of T-monotonicity

In the next theorem we show that T-monotonicity is a property shared by any solution of (1.1) which is comparable with all T-periodic solutions of (1.1).

THEOREM 1.10. *Assume* (C). *Let* $x :]\omega_-, \omega_+[\, \to \mathbb{R}$ *be a nonextendible solution of* (1.1) *such that, if u is a solution of* (1.3), *then either $x > u$ or $x < u$. Then x is T-monotone.*

PROOF. We assume that $\omega_+ - \omega_- > T$, because otherwise there is nothing to prove. Let us suppose by contradiction that x is not T-monotone, i.e. there are points $t_1, t_2 \in\,]\omega_-, \omega_+[$ such that

$$x(t_1) < x(t_1 + T) \quad \text{and} \quad x(t_2) > x(t_2 + T).$$

We can assume that $t_1 < t_2$, otherwise we set $y(t) = -x(t)$ and replace equation (1.1) with $y' = -f(t, -y)$.

Step 1. Assume $\omega_+ = +\infty$. We distinguish two cases.

• Suppose that there exists $n \in \mathbb{N}^+$ such that $t_2 = t_1 + nT$ and $x(t_1 + kT) \leqslant x(t_1 + (k + 1)T)$ for each $0 \leqslant k \leqslant n - 1$. Let $m \in \mathbb{N}$ be such that $x(t_1 + mT) < x(t_1 + (m + 1)T) = x(t_1 + nT)$. Since $\max\{x(t_1 + mT), x(t_1 + (n + 1)T)\} < x(t_1 + nT)$, we can pick $p \in\,]\max\{x(t_1 + mT), x(t_1 + (n + 1)T)\}, x(t_1 + nT)[$. We define a function $\alpha : [t_1 + mT, t_1 + (m + 1)T[\, \to \mathbb{R}$ by $\alpha(t) = x(t)$ on $]t_1 + mT, t_1 + (m + 1)T[$ and $\alpha(t_1 + mT) = p$, and a function $\beta:[t_1+nT,t_1+(n+1)T[\,\to\mathbb{R}$ by $\beta(t)=x(t)$ on $]t_1+nT,t_1+(n+1)T[$ and $\beta(t_1+nT)=p$. The T-periodic extensions to $\mathbb{R}$ of α and β are respectively a lower and an upper solution of (1.3). Hence Theorem 1.9 yields, for any $p\in\,]\max\{x(t_1+mT),x(t_1+(n+1)T)\},x(t_1+nT)[$, the existence of a solution u of (1.3) such that $\min\{\alpha,\beta\}\leqslant u\leqslant\max\{\alpha,\beta\}$. Since $u(t_1+mT)=p$ and $x(t_1+mT)<p<x(t_1+nT)$, we conclude that $u\nleqslant x$ and $u\ngeqslant x$, which is a contradiction.

- Suppose that there exists $n\in\mathbb{N}$ such that $t_1+nT<t_2<t_1+(n+1)T$ and $x(t_1+kT)\leqslant x(t_1+(k+1)T)$ for each $k\in\mathbb{N}$. Set $\varphi(t)=x(t+T)-x(t)$ on $]\omega_-,+\infty[$. As $\varphi(t_2)<0$ and $\varphi(t_1+(n+1)T)\geqslant 0$, there exists $s_1\in\,]t_2,t_1+(n+1)T]$ such that $\varphi(s_1)=0$, i.e. $x(s_1+T)=x(s_1)$. Let $v:\mathbb{R}\to\mathbb{R}$ be the T-periodic extension to $\mathbb{R}$ of $x|_{[s_1,s_1+T]}$; v is a solution of (1.3). Since $v(t_2)=v(t_2+T)=x(t_2+T)<x(t_2)$ and $v(t_1+nT)=v(t_1+(n+1)T)=x(t_1+(n+1)T)\geqslant x(t_1+nT)$, there exists $s_2\in[t_1+nT,t_2[$ such that $v(s_2)=x(s_2)$. Note that $x(s_2)=x(s_2+T)$. Let $w:\mathbb{R}\to\mathbb{R}$ be the T-periodic extension to $\mathbb{R}$ of $x|_{[s_2,s_2+T]}$; w is a solution of (1.3). Since $v(t_2)=x(t_2+T)<x(t_2)=w(t_2)$, we can pick $p\in\,]v(t_2),w(t_2)[$. Then we define a function $\alpha:[s_2,s_2+T[\,\to\mathbb{R}$, by $\alpha(t)=w(t)$ on $]s_2,t_2[$, $\alpha(t_2)=p$ and $\alpha(t)=v(t)$ on $]t_2,s_2+T[$, and a function $\beta:[s_2,s_2+T[\,\to\mathbb{R}$, by $\beta(t)=v(t)$ on $]s_2,t_2[$, $\beta(t_2)=p$ and $\beta(t)=w(t)$ on $]t_2,s_2+T[$. The T-periodic extensions to $\mathbb{R}$ of α and β are respectively a lower and an upper solution of (1.3). Hence Theorem 1.9 yields, for any $p\in\,]v(t_2),w(t_2)[$, the existence of a solution u of (1.3) such that $\min\{\alpha,\beta\}\leqslant u\leqslant\max\{\alpha,\beta\}$. Since $u(t_2)=p$ and $x(t_2+T)<p<x(t_2)$, we conclude that $u\nleqslant x$ and $u\ngeqslant x$, which is a contradiction.

Step 2. Assume $\omega_-=-\infty$. We set $y(t)=-x(-t)$ and we replace Eq. (1.1) with $y'=f(-t,-y)$. Thus we are reduced to the situation discussed in Step 1.

Step 3. Assume $\omega_-,\omega_+\in\mathbb{R}$. Set $\varphi(t)=x(t+T)-x(t)$ on $]\omega_-,\omega_+-T[$. We essentially distinguish two cases.

- Suppose that

$$\lim_{t\to\omega_-}x(t)=-\infty\quad\text{and}\quad\lim_{t\to\omega_+}x(t)=+\infty;\tag{1.13}$$

the case where $\lim_{t\to\omega_-}x(t)=+\infty$ and $\lim_{t\to\omega_+}x(t)=-\infty$ being treated similarly. Since $\varphi(t_1)>0>\varphi(t_2)$, there exists $t_0\in\,]t_1,t_2[$ such that $\varphi(t_0)=0$, i.e. $x(t_0+T)=x(t_0)$. Let $u:\mathbb{R}\to\mathbb{R}$ be the T-periodic extension to $\mathbb{R}$ of $x|_{[t_0,t_0+T]}$; u is a solution of (1.3). Since (1.13) implies that $u\nleqslant x$ and $u\ngeqslant x$, a contradiction follows.

- Suppose that

$$\lim_{t\to\omega_-}x(t)=\lim_{t\to\omega_+}x(t)=+\infty;\tag{1.14}$$

the case where $\lim_{t\to\omega_-}x(t)=\lim_{t\to\omega_+}x(t)=-\infty$ being treated similarly. Since $\lim_{t\to\omega_-}\varphi(t)=-\infty$ and $\lim_{t\to\omega_+-T}\varphi(t)=+\infty$, there exists $s_1\in\,]\omega_-,\omega_+-T[$ such that $\varphi(s_1)=0$, i.e. $x(s_1+T)=x(s_1)$. Let $w:\mathbb{R}\to\mathbb{R}$ be the T-periodic extension to $\mathbb{R}$ of $x|_{[s_1,s_1+T]}$. Since w is a solution of (1.3) and (1.14) holds, we conclude that $x>w$ and w is the maximum solution of (1.3). Define a function $x_1:\,]\omega_-,\omega_+-T[\,\to\mathbb{R}$ by setting $x_1(t)=x(t)$ on $]\omega_-,s_1]$ and $x_1(t)=x(t+T)$ on $]s_1,\omega_+-T[$. The function x_1 is a solution of (1.1) satisfying $x_1>w$ on $]\omega_-,\omega_+-T[$, $x_1(s_1)=w(s_1)$ and $\lim_{t\to\omega_-}x_1(t)=$

$\lim_{t\to\omega_+-T} x_1(t) = +\infty$. Proceeding recursively we can find $s_n \in]\omega_-, \omega_+ - nT[$ and construct a solution $x_n :]\omega_-, \omega_+ - nT[\to \mathbb{R}$ of (1.1) such that $\omega_+ - nT - \omega_- \leqslant T$, $x_n > w$ on $]\omega_-, \omega_+ - nT[$, $x_n(s_n) = w(s_n)$ and $\lim_{t\to\omega_-} x_n(t) = \lim_{t\to\omega_+-nT} x_n(t) = +\infty$. A contradiction will then be achieved from the following result.

CLAIM. *Assume that there exists the maximum solution w of* (1.3). *Let $x :]\omega_-, \omega_+[\to \mathbb{R}$ be a solution of* (1.1) *such that $\omega_+ - \omega_- \leqslant T$, $x > w$ on $]\omega_-, \omega_+[$ and* $\lim_{t\to\omega_-} x(t) = \lim_{t\to\omega_+} x(t) = +\infty$. *Then $x \gg w$ on $]\omega_-, \omega_+[$.*

Suppose by contradiction that there is $t_0 \in]\omega_-, \omega_+[$ such that $x(t_0) = w(t_0)$ and $x(t) > w(t)$ on $]t_0, \omega_+[$. Denote by $\mathcal{Y}$ the set of all solutions $y : [t_0, \lambda_+(y)[\to \mathbb{R}$ of (1.1), which are nonextendible to the right of $\lambda_+(y)$ and satisfy $y(t_0) = w(t_0)$, $y \geqslant w$ on $[t_0, \lambda_+(y)[$ and $y > w$ on $[t_0, \omega_- + T[\cap [t_0, \lambda_+(y)[$. Note that $\mathcal{Y} \neq \emptyset$, as $x \in \mathcal{Y}$. Set $\lambda_+ = \sup_{y\in\mathcal{Y}} \lambda_+(y)$. Let us show that $\lambda_+ > \omega_- + T$. Otherwise, if $\lambda_+ \leqslant \omega_- + T$, we can pick a solution $z : [\lambda_+ - \delta, \lambda_+ + \delta] \to \mathbb{R}$ of (1.1) such that $z(\lambda_+) > w(\lambda_+)$. Let $y \in \mathcal{Y}$ be such that $\lambda_+(y) > \lambda_+ - \delta$. As $\lim_{t\to\lambda_+(y)} y(t) = +\infty$, we can suppose, possibly for a smaller $\delta > 0$, that $y(t) > z(t)$ on $[\lambda_+ - \delta, \lambda_+(y)[$. Hence z can be continued to a right-nonextendible solution v of (1.1) defined on $[t_0, \lambda_+(v)[$, with $\lambda_+(v) > \lambda_+ + \delta$, and satisfying $v(t_0) = w(t_0)$, $v \geqslant w$ on $[t_0, \lambda_+(v)[$ and $v > w$ on $[t_0, \omega_- + T[\cap [t_0, \lambda_+(v)[$. Since $v \in \mathcal{Y}$ we get a contradiction. Therefore $\lambda_+ > \omega_- + T$. Pick $y \in \mathcal{Y}$ such that $\lambda_+(y) > \omega_- + T$. As $\lim_{t\to(\omega_-+T)^+} x(t - T) = +\infty$, there is $\delta > 0$ such that $w(t) \leqslant y(t) < x(t - T)$ on $]\omega_- + T, \omega_- + T + \delta]$. Hence y can be continued to a solution u of (1.1) defined on $[t_0, t_0 + T]$ and satisfying $w(t) \leqslant u(t) \leqslant x(t - T)$ on $]\omega_- + T, t_0 + T]$. Since u is a solution of (1.3) and $u > w$ on $[t_0, \omega_- + T]$, the maximality of w yields a contradiction. Therefore the claim is proved. □

REMARK 1.10. The following more general conclusion can be achieved. Let $x :]\omega_-, \omega_+[\to \mathbb{R}$ be a nonextendible solution of (1.1) such that $\omega_-, \omega_+ \in \mathbb{R}$, $\omega_+ - \omega_- \leqslant T$ and, for any solution u of (1.3), either $x < u$ or $x > u$. Then two possibilities may occur:

- for all solutions u of (1.3), either $x \ll u$ or $x \gg u$;
- there exist a solution u of (1.3) and $t_0 \in]\omega_-, \omega_+[$ such that $u(t_0) = x(t_0)$ and, e.g., $u < x$. In this case there exists $\bar{t} \in]t_0, t_0 + T[$ such that, for every $\bar{x} > u(\bar{t})$, there is a solution v of (1.3) with $u < v < x$ on $]\omega_-, \omega_+[$ and $v(\bar{t}) = \bar{x}$.

A simple example where the latter possibility occurs follows.

EXAMPLE 1.2. Let us consider Eq. (1.1) taking $f : [-\frac{3}{2}, \frac{3}{2}[\times \mathbb{R} \to \mathbb{R}$ defined by $f(t, x) = 2x^2 \operatorname{sgn}(t)$, if $|x| > 1$, and $f(t, x) = 2\sqrt{|x|} \operatorname{sgn}(t)$, if $|x| \leqslant 1$, and extending f by 3-periodicity with respect to the t-variable. Let $x :]-\frac{3}{2}, \frac{3}{2}[\to \mathbb{R}$ be defined by $x(t) = \frac{1}{3+2t}$ on $]-\frac{3}{2}, -1]$, $x(t) = t^2$ on $]-1, 1]$, $x(t) = \frac{1}{3-2t}$ on $]1, \frac{3}{2}[$. Note that x is a nonextendible solution of (1.1) such that $x > u$ for any solution u of (1.3), 0 is a solution of (1.3), $x(0) = 0$ and, for every $\bar{x} > 0$, there is a solution v of (1.3) with $0 < v < x$ on $] - \frac{3}{2}, \frac{3}{2}[$ and $v(\frac{3}{2}) = \bar{x}$.

The Massera Existence Theorem

When uniqueness of solutions of (1.2) fails, there may exist solutions which are not T-monotone (see Example 1.3). It is useful to notice that existence of a solution that is not T-monotone yields existence of T-periodic solutions. Proposition 1.38 will actually show that infinitely many T-periodic solutions do exist in this case.

COROLLARY 1.11. *Assume* (C). *Let* $x :]t_0, +\infty[\to \mathbb{R}$ *be a solution of* (1.1). *If* x *is not* T*-monotone, then there exists a solution* u *of* (1.3) *such that* $u \nleqslant x$ *and* $u \ngeqslant x$.

PROOF. Since x is not T-monotone, there are points $t_1, t_2 \in]t_0, +\infty[$ such that $x(t_1) < x(t_1 + T)$ and $x(t_2) > x(t_2 + T)$. Hence we can proceed as in the first step of the proof of Theorem 1.10 to get the conclusion. □

We finally observe that combining Corollary 1.11 and Proposition 1.7 immediately yields the following statement.

PROPOSITION 1.12. *Assume* (C). *Suppose that Eq.* (1.1) *has a bounded solution defined on an unbounded interval. Then problem* (1.3) *has at least one solution.*

Proposition 1.12 provides an extension to the Carathéodory setting of the existence part of the Massera Theorem. Yet, it must be observed that, when uniqueness for (1.1) fails, convergence of a bounded solution x to a T-periodic solution u is not generally guaranteed, as the following example will show. A possible extension of the convergence part of the Massera Theorem to the case where the uniqueness assumption is dropped will be discussed in Section 1.7.

EXAMPLE 1.3. We consider Eq. (1.12) again. As described in Example 1.1, for each $p \in]-2^{-10}, 2^{-10}[$, there exists a 1-periodic solution u_p of (1.12) such that $u_p(-\frac{1}{2}) = p$ and $u_p(0) = 0$. Let $p_1 \neq p_2$ and set $x(t) = u_{p_1}(t)$ on $[0, 1[$ and $x(t) = u_{p_2}(t)$ on $[1, 2[$. Since $x(\frac{1}{2}) = p_1 \neq p_2 = x(\frac{3}{2})$, the extension to $\mathbb{R}$ of x by 2-periodicity is a 2-periodic solution of (1.12) which is not 1-periodic; hence it is a bounded solution which does not converge to any 1-periodic solution of (1.12). As we shall see in Example 1.4, x is a subharmonic solution of order 2 of (1.12), i.e. its minimal period is 2.

1.5. *Structure and dynamics near one-sided isolated periodic solutions*

We investigate in this section the dynamics of all solutions of (1.1) in the vicinity of one-sided isolated T-periodic solutions; namely, solutions of (1.1) which lie between a pair of strictly ordered T-periodic solutions, or above the maximum T-periodic solution, or below the minimum T-periodic solution, provided they exist. T-monotonicity plays an important role in this study. Incidentally we notice that the maximum solution of (1.3) always exists, provided the solution set of (1.3) is bounded from above. Similarly the minimum solution of (1.3) always exists, provided the solution set of (1.3) is bounded from below.

Between two ordered T*-periodic solutions*
We start by observing that, given a pair of T-periodic solutions, either they are strictly ordered, or there exist infinitely many T-periodic solutions hitting them.

PROPOSITION 1.13. *Assume* (C). *Let* v, w *be solutions of* (1.3), *with* $v \neq w$. *Then either* $v \ll w$, *or* $v \gg w$, *or for every* $t_0 \in [0, T[$ *and every* $x_0 \in [\min\{v(t_0), w(t_0)\}, \max\{v(t_0), w(t_0)\}]$ *there exists a solution* u *of* (1.3) *with* $u(t_0) = x_0$ *and* $\min\{v, w\} \leqslant u \leqslant \max\{v, w\}$.

PROOF. Assume that there exists $\bar{t} \in [0, T[$ such that $v(\bar{t}) = w(\bar{t})$. Let $t_0 \in [0, T]$ and $x_0 \in [\min\{v(t_0), w(t_0)\}, \max\{v(t_0), w(t_0)\}]$. We may assume that $t_0 \leqslant \bar{t}$, otherwise we replace $\bar{t}$ with $\bar{t} + T$. Let us consider the T-periodic functions α and β defined by $\alpha(t_0) = x_0$, $\alpha(t) = \min\{v(t), w(t)\}$ on $]t_0, \bar{t}]$, $\alpha(t) = \max\{v(t), w(t)\}$ on $]\bar{t}, t_0 + T[$, $\beta(t_0) = x_0$, $\beta(t) = \max\{v(t), w(t)\}$ on $]t_0, \bar{t}]$, $\beta(t) = \min\{v(t), w(t)\}$ on $]\bar{t}, t_0 + T[$. Then α and β are respectively a lower solution and an upper solution of (1.3), hence, by Theorem 1.9, there exists a solution u of (1.3) with $\min\{\alpha, \beta\} \leqslant u \leqslant \max\{\alpha, \beta\}$. In particular $u(t_0) = x_0$ and $\min\{v, w\} \leqslant u \leqslant \max\{v, w\}$. □

We consider here the case where a pair of strictly ordered T-periodic solutions is given, the complementary situation will be discussed in Section 1.6. We prove the existence of heteroclinic solutions connecting a pair of strictly ordered T-periodic solutions, when there is no further T-periodic solution in between.

THEOREM 1.14. *Assume* (C). *Let v, w be solutions of* (1.3) *such that $v \ll w$ and there is no solution u of* (1.3) *satisfying $v < u < w$. Then either any nonextendible solution x of* (1.1), *with $v < x < w$, exists on $\mathbb{R}$, is T-increasing and satisfies*

$$\lim_{t \to -\infty} \bigl(x(t) - v(t)\bigr) = \lim_{t \to +\infty} \bigl(x(t) - w(t)\bigr) = 0,$$

or any nonextendible solution x of (1.1), *with $v < x < w$, exists on $\mathbb{R}$, is T-decreasing and satisfies*

$$\lim_{t \to -\infty} \bigl(x(t) - w(t)\bigr) = \lim_{t \to +\infty} \bigl(x(t) - v(t)\bigr) = 0.$$

PROOF. Pick any nonextendible solution x of (1.1), with $v < x < w$. Clearly, there is a t_0 such that $v(t_0) < x(t_0) < w(t_0)$. Since x exists on $\mathbb{R}$ and satisfies the assumptions of Theorem 1.10, we have that either x is T-increasing, or x is T-decreasing. Assume the former case occurs, the latter one being treated similarly. Then Proposition 1.7 and Remark 1.7 imply that there exist solutions u_1, u_2 of (1.3) satisfying $u_1 < x < u_2$, as x is not a solution of (1.3), and $\lim_{t \to -\infty}(x(t) - u_1(t)) = \lim_{t \to +\infty}(x(t) - u_2(t)) = 0$. Since $v \leqslant u_1 < u_2 \leqslant w$ and there is no solution u of (1.3) satisfying $v < u < w$, we conclude that $u_1 = v$ and $u_2 = w$.

Finally, let us suppose by contradiction that there exist two solutions x_1, x_2 of (1.1), with $v < x_1 < w$ and $v < x_2 < w$ on $\mathbb{R}$, such that $\lim_{t \to -\infty}(x_1(t) - v(t)) = \lim_{t \to +\infty}(x_1(t) - w(t)) = 0$ and $\lim_{t \to -\infty}(x_2(t) - w(t)) = \lim_{t \to +\infty}(x_2(t) - v(t)) = 0$. Set $x_* = \min\{x_1, x_2\}$ and $x^* = \max\{x_1, x_2\}$. Since either $v < x_* < w$, or $v < x^* < w$, we find a solution x of (1.1), with $v < x < w$ in $\mathbb{R}$, such that either $\lim_{t \to -\infty}(x(t) - v(t)) = \lim_{t \to +\infty}(x(t) - v(t)) = 0$, or $\lim_{t \to -\infty}(x(t) - w(t)) = \lim_{t \to +\infty}(x(t) - w(t)) = 0$, thus contradicting the conclusion previously achieved. □

Maximal and minimal T-periodic solutions

We show now that any bounded set of solutions of (1.3) has a maximal and a minimal element.

LEMMA 1.15. *Assume* (C). *Let* $\mathcal{S} \subseteq W^{1,1}(0,T)$ *be a nonempty set of solutions of* (1.3). *Assume that* $\mathcal{S}$ *is uniformly bounded from above, i.e. there exists a constant* M *such that* $\max_{[0,T]} u(t) \leqslant M$ *for all* $u \in \mathcal{S}$. *Then there exists a maximal solution* w *of* (1.3) *in the* C^0*-closure* $\bar{\mathcal{S}}$ *of* $\mathcal{S}$. *Similarly, if* $\mathcal{S}$ *is uniformly bounded from below, then there exists a minimal solution* v *of* (1.3) *in* $\bar{\mathcal{S}}$.

PROOF. We only prove the former case. We show that $(\bar{\mathcal{S}}, \leqslant)$ is inductively ordered. Let $\mathcal{T} = \{u_i \mid i \in I\}$ be a totally ordered subset of $\bar{\mathcal{S}}$. We prove that $\mathcal{T}$ has an upper bound in $\bar{\mathcal{S}}$. Let us set $u(t) = \sup_{i\in I} u_i(t)$ on $[0,T]$. Let $\mathcal{D} = \{t_m \mid m \in \mathbb{N}\}$ be a countable dense subset of $[0,T]$ and define a sequence in $\mathcal{T}$ as follows: for $n=1$, take $u_1 \in \mathcal{T}$ such that $u_1(t_1) \geqslant u(t_1) - 1$, for $n=2$, take $u_2 \in \mathcal{T}$, with $u_2 \geqslant u_1$, such that $u_2(t_2) \geqslant u(t_2) - \frac{1}{2}$, $u_2(t_1) \geqslant u(t_1) - \frac{1}{2}$, and so on. In this way, we construct a sequence $(u_n)_n$ in $\mathcal{T}$, with $u_1 \leqslant u_2 \leqslant \cdots \leqslant u_n \leqslant u_{n+1} \leqslant \cdots$, such that $u_n(t_k) \geqslant u(t_k) - \frac{1}{n}$, for $1 \leqslant k \leqslant n$. It is clear that $(u_n)_n$ converges to u pointwise on $\mathcal{D}$. On the other hand, as $(u_n)_n$ is relatively compact in $C^0([0,T])$, there is a subsequence of $(u_n)_n$ which converges uniformly to some function $\hat{u} \in C^0([0,T])$. Actually, by monotonicity, the whole sequence $(u_n)_n$ converges uniformly to $\hat{u}$. Using the L^1-Carathéodory conditions, we conclude that the convergence takes place in $W^{1,1}(0,T)$, the limit $\hat{u} \in W^{1,1}(0,T)$ and $\hat{u}$ is a solution of (1.3). Moreover, $\hat{u} = u$ on $\mathcal{D}$ and $\hat{u} \leqslant u$ on $[0,T]$. Let us show that $\hat{u} = u$ on $[0,T]$. Indeed, otherwise, one can find a point $t_0 \in [0,T]$ and a function $u_0 \in \mathcal{T}$ such that $\hat{u}(t_0) < u_0(t_0) \leqslant u(t_0)$. The continuity of $\hat{u}$ and u_0 and the density of $\mathcal{D}$ yield a contradiction. This proves that $u \in \bar{\mathcal{S}}$ is an upper bound of $\mathcal{T}$. Since $(\bar{\mathcal{S}}, \leqslant)$ is inductively ordered, Zorn Lemma guarantees the existence of a maximal element $w \in \bar{\mathcal{S}}$. □

By the lattice structure of the solution set of (1.3), we get the following conclusion.

COROLLARY 1.16. *Suppose* (C). *Assume that the set of all solutions of* (1.3) *is uniformly bounded from above. Then there exists the maximum solution of* (1.3). *Assume that the set of all solutions of* (1.3) *is uniformly bounded from below. Then there exists the minimum solution of* (1.3).

Above the maximum, or below the minimum T-periodic solution

We now assume that there exists the maximum solution of (1.3) and describe the qualitative behaviour of all solutions of (1.1) that lie above it. Symmetric conclusions can be established for solutions lying below the minimum T-periodic solution, whenever it exists.

THEOREM 1.17. *Assume* (C). *Suppose that there exists the maximum solution* w *of* (1.3). *Then either any nonextendible solution* x *of* (1.1), *with* $x > w$, *is* T*-increasing or any nonextendible solution* x *of* (1.1), *with* $x > w$, *is* T*-decreasing. Furthermore, for any such* $x :]\omega_-, \omega_+[\to \mathbb{R}$, *either*

$$\omega_-, \omega_+ \in \mathbb{R}, \quad \omega_+ - \omega_- \leqslant T, \quad x \gg w \quad \textit{on }]\omega_-, \omega_+[\quad \textit{and}$$

$$\lim_{t\to\omega_-} x(t) = \lim_{t\to\omega_+} x(t) = +\infty,$$

or

$$\omega_- = -\infty, \quad \lim_{t\to-\infty}\big(x(t) - w(t)\big) = 0 \quad \textit{and} \quad \limsup_{t\to\omega_+} x(t) = +\infty,$$

or

$$\omega_+ = +\infty, \quad \limsup_{t\to\omega_-} x(t) = +\infty \quad \textit{and} \quad \lim_{t\to+\infty}\big(x(t) - w(t)\big) = 0.$$

PROOF. Let $x :]\omega_-, \omega_+[\to \mathbb{R}$ be a nonextendible solution of (1.1) with $x > w$.

Assume first that $\omega_-, \omega_+ \in \mathbb{R}$ and that $\omega_+ - \omega_- \leqslant T$. The Claim stated in the proof of Theorem 1.10 implies that $x \gg w$ on $]\omega_-, \omega_+[$. Then the conclusion follows.

Suppose next that $\omega_+ - \omega_- > T$. By Theorem 1.10 either x is T-increasing, or x is T-decreasing. Assume that the former situation occurs, the latter case being treated similarly. Since x is bounded from below, we have $\omega_- = -\infty$. Let us verify that $\lim_{t\to-\infty}(x(t) - w(t)) = 0$. Remark 1.7 implies that there exists a solution u of (1.3), with $u \leqslant x$, such that $\lim_{t\to-\infty}(x(t) - u(t)) = 0$. Since $u \geqslant w$, the maximality of w yields $u = w$. Of course, if $\omega_+ < +\infty$, then $\lim_{t\to\omega_+} x(t) = +\infty$. Let us show that, if $\omega_+ = +\infty$, then $\limsup_{t\to\omega_+} x(t) = +\infty$. Indeed, otherwise, x would be bounded and Proposition 1.7 would imply the existence of a solution u of (1.3), with $u \geqslant x$, such that $\lim_{t\to+\infty}(x(t) - u(t)) = 0$. Since $u > w$, the maximality of w would yield a contradiction.

Finally, let us suppose, by contradiction, that there exist $x_1 :]-\infty, \omega_+[\to \mathbb{R}$, $x_2 :]\omega_-, +\infty[\to \mathbb{R}$ both solutions of (1.1), with $x_1 > w$ and $x_2 > w$, such that x_1 is T-increasing and x_2 is T-decreasing. Then $\lim_{t\to-\infty}(x_1(t) - w(t)) = 0$ and $\lim_{t\to+\infty}(x_2(t) - w(t)) = 0$. Possibly replacing $x_1(t)$ with $x_1(t - kT)$ for some $k \in \mathbb{N}$, we may assume $\omega_+ > \omega_-$. Set $x_* = \min\{x_1, x_2\}$ and $x^* = \max\{x_1, x_2\}$. Then either $w < x_*$ or $w \not\ll x^*$; in both cases we contradict the conclusions previously achieved. □

1.6. *Stability and detours*

In this section we study the stability properties of T-periodic solutions of (1.1) with the aid of lower and upper solutions. We start by recalling the classical notion of one-sided Lyapunov stability. Yet, in the present context, where uniqueness of the solutions of (1.2) is not guaranteed, such a definition does not generally seem the most appropriate to be considered; indeed, some weaker concept might fit better in order to detect certain residual forms of stability. As an alternative notion to Lyapunov stability we use order stability; this is common in the frame of monotone dynamical systems [93,58,60] and appears suited to our approach based on lower and upper solutions. Lyapunov stability implies order stability, but not vice versa; however, these concepts are equivalent if uniqueness, either in the past or in the future, holds for solutions of (1.2). We further give a characterization of order stability in terms of the asymptotic behaviour of solutions of (1.1) by introducing a third kind of stability, which is named weak stability. This notion was first defined in [130] in the context of multivalued dynamical systems (see also [50]) and is related to that of weak positive invariance considered in [105,152]. We show that weak stability and order stability are always equivalent. Using these concepts, we give a precise description of the stability properties of a T-periodic solution in terms of the existence of a lower or an upper solution close to it. Hence, when a pair of lower and upper solutions is given, we can discuss the stability or instability of the minimum and the maximum T-periodic solutions v and w wedged between them, thus getting a completion of Theorem 1.9. Afterwards, we turn to investigate the behaviour of the solutions lying between v and w. If the lower and the upper solutions are ordered in the standard way, i.e. the lower solution is below the upper solution, we find between v and w a totally ordered continuum of weakly stable T-periodic solutions. Whereas, in the complementary case we see that Lyapunov instability always occurs and, in the absence of uniqueness for (1.2), solutions of (1.1) having very complicated, even chaotic-like, dynamics may exist. In particular, we can find homoclinics, subharmonics of any order and almost periodic solutions, which are not periodic of any fixed period.

One-sided Lyapunov stability

DEFINITION 1.14.

- A solution u of (1.3) is said *Lyapunov stable* (briefly, *$\mathcal{L}$-stable*) *from below* if, for every $\varepsilon > 0$, there is $\delta > 0$ such that, for every $t_0 \in [0, T[$ and every x_0 with $u(t_0) - \delta < x_0 < u(t_0)$, every right-nonextendible solution x of (1.2), with $x \leqslant u$, exists on $[t_0, +\infty[$ and satisfies

$$u(t) - \varepsilon < x(t) \leqslant u(t) \quad \text{on } [t_0, +\infty[. \tag{1.15}$$

- If, further,

$$\lim_{t \to +\infty} \bigl(x(t) - u(t)\bigr) = 0, \tag{1.16}$$

  u is said *$\mathcal{L}$-asymptotically stable from below*.
- *$\mathcal{L}$-stability* and *$\mathcal{L}$-asymptotic stability from above* are defined similarly.
- A solution u of (1.3) is said *$\mathcal{L}$-stable* if, for every $\varepsilon > 0$, there is $\delta > 0$ such that, for every $t_0 \in [0, T[$ and every x_0 with $|x_0 - u(t_0)| < \delta$, every right-nonextendible solution x of (1.2) exists on $[t_0, +\infty[$ and satisfies

$$\bigl|x(t) - u(t)\bigr| < \varepsilon \quad \text{on } [t_0, +\infty[. \tag{1.17}$$

  If, further, (1.16) holds, u is said *$\mathcal{L}$-asymptotically stable*.
- A solution u of (1.3) is said *$\mathcal{L}$-unstable from below* if it is not $\mathcal{L}$-stable from below. *$\mathcal{L}$-instability from above* and *$\mathcal{L}$-instability* are defined similarly.

REMARK 1.11. The notion of $\mathcal{L}$-stability from below given in Definition 1.14 does not require any condition on solutions $x : [t_0, \omega_+[\to \mathbb{R}$ of (1.2), satisfying $u(t_0) - \delta < x_0 < u(t_0)$, but not $x \leqslant u$. A similar remark holds for the $\mathcal{L}$-stability from above.

DEFINITION 1.15. A solution u of (1.3) is said *isolated from below* in $C^0([0, T])$ if there exists $\varepsilon > 0$ such that no solution z of (1.3) exists satisfying $u - \varepsilon < z < u$. Similarly, a solution u of (1.3) is said *isolated from above* in $C^0([0, T])$ if there exists $\varepsilon > 0$ such that no solution z of (1.3) exists satisfying $u < z < u + \varepsilon$.

PROPOSITION 1.18. *Assume* (C). *A solution u of* (1.3) *is $\mathcal{L}$-asymptotically stable from below* (*respectively $\mathcal{L}$-asymptotically stable from above*) *if and only if it is $\mathcal{L}$-stable from below and isolated from below* (*respectively $\mathcal{L}$-stable from above and isolated from above*) *in $C^0([0, T])$.*

PROOF. Assume u is $\mathcal{L}$-asymptotically stable from below, then u is clearly $\mathcal{L}$-stable and isolated from below in $C^0([0, T])$ as a solution of (1.3). Suppose now that u is $\mathcal{L}$-stable from below and isolated from below. Let $\varepsilon > 0$ be such that no solution z of (1.3) exists satisfying $u - \varepsilon < z < u$. By the $\mathcal{L}$-stability from below of u, there is $\delta > 0$ such that every right-nonextendible solution x of (1.2), with $u(t_0) - \delta < x_0 < u(t_0)$ and $x \leqslant u$, exists on $[t_0, +\infty[$ and satisfies $u - \varepsilon/2 < x < u$ on $[t_0, +\infty[$. We claim that any such solution x satisfies $x(t_0) < x(t_0 + T)$. Indeed, otherwise $x|_{[t_0,t_0+T[}$ is a proper upper solution of (1.3). Hence, by Proposition 1.8 there exists a solution v of (1.3) such that $\lim_{t\to+\infty}(x(t) - v(t)) = 0$ and $u - \varepsilon < v < x < u$ on $[t_0, t_0 + T]$, thus contradicting the assumption that u is isolated from below. Hence $x|_{[t_0,t_0+T[}$ is a proper lower solution of (1.3) and, by Proposition 1.8, $\lim_{t\to+\infty}(u(t) - x(t)) = 0$. A similar argument proves the statement about $\mathcal{L}$-stability from above. □

One-sided order stability

DEFINITION 1.16.

- A solution u of (1.3) is said *order stable* (briefly, *$\mathcal{O}$-stable*) (respectively *properly $\mathcal{O}$-stable*, respectively *strictly $\mathcal{O}$-stable*) *from below* if there exists a sequence $(\alpha_n)_n$ of regular lower solutions (respectively proper regular lower solutions, respectively strict regular lower solutions) of (1.3) such that $\alpha_n < u$ for every n and $\alpha_n \to u$ uniformly on $[0, T]$.
- A solution u of (1.3) is said *$\mathcal{O}$-stable* (respectively *properly $\mathcal{O}$-stable*, respectively *strictly $\mathcal{O}$-stable*) *from above* if there exists a sequence $(\beta_n)_n$ of regular upper solutions (respectively proper regular upper solutions, respectively strict regular upper solutions) of (1.3) such that $\beta_n > u$ for every n and $\beta_n \to u$ uniformly on $[0, T]$.
- A solution u of (1.3) is said *$\mathcal{O}$-stable* (respectively *properly $\mathcal{O}$-stable*, respectively *strictly $\mathcal{O}$-stable*) if it is $\mathcal{O}$-stable (respectively properly $\mathcal{O}$-stable, respectively strictly $\mathcal{O}$-stable) from above and from below.

REMARK 1.12. The requirement for the lower and upper solutions to be regular in Definition 1.16 is not essential. Indeed, we see, for instance, that between a proper lower solution α and a solution u of (1.3), with $\alpha < u$, it is always possible to fit in a proper regular lower solution $\bar{\alpha}$. This follows from Proposition 1.20. A similar observation holds for a proper upper solution β.

In order to prove Proposition 1.20 we first need the following result concerning a property of Carathéodory functions. A more general version of it will be proved in Proposition 2.3.

PROPOSITION 1.19. *Assume* (C). *Then, for each* $\rho > 0$, *there exists a* L^1*-Carathéodory function* $h : [0, T] \times [-\rho, \rho] \times [-\rho, \rho] \to \mathbb{R}$ *such that*

(i) *for a.e.* $t \in [0, T]$ *and every* $x \in [-\rho, \rho]$, $h(t, \cdot, x) : [-\rho, \rho] \to \mathbb{R}$ *is strictly increasing*;

(ii) *for a.e.* $t \in [0, T]$ *and every* $y \in [-\rho, \rho]$, $h(t, y, \cdot) : [-\rho, \rho] \to \mathbb{R}$ *is strictly decreasing*;

(iii) *for a.e.* $t \in [0, T]$ *and every* $(x, y) \in [-\rho, \rho] \times [-\rho, \rho]$, $h(t, x, y) = -h(t, y, x)$;

(iv) *for a.e.* $t \in [0, T]$ *and every* $(x, y) \in [-\rho, \rho] \times [-\rho, \rho]$, *with* $x < y$,

$$\left| f(t, y) - f(t, x) \right| < h(t, y, x).$$

PROPOSITION 1.20. *Assume* (C). *Suppose that* u *is a solution of* (1.3). *If* α *is a proper lower solution of* (1.3), *with* $\alpha < u$, *then there exists a proper regular lower solution* $\bar{\alpha}$ *of* (1.3), *satisfying* $\lim_{t\to 0^+} \bar{\alpha}(t) = \bar{\alpha}(0) = \lim_{t\to T^-} \bar{\alpha}(t)$ *and* $\alpha < \bar{\alpha} < u$. *Similarly, if* β *is a proper upper solution of* (1.3), *with* $\beta > u$, *then there exists a proper regular upper solution* $\bar{\beta}$ *of* (1.3), *satisfying* $\lim_{t\to 0^+} \bar{\beta}(t) = \bar{\beta}(0) = \lim_{t\to T^-} \bar{\beta}(t)$ *and* $u < \bar{\beta} < \beta$.

PROOF. Let h be the function associated with f by Proposition 1.19 and corresponding to $\rho = \max\{\|\alpha\|_\infty, \|u\|_\infty\}$. Consider the periodic problem

$$x' = f(t, \alpha) - h(t, x, \alpha), \quad x(0) = x(T).$$

Since α is a proper lower solution and u is a proper upper solution, this problem has a solution $\bar{\alpha}$, satisfying $\alpha < \bar{\alpha} < u$. The properties of h imply that $\bar{\alpha}$ is a proper lower solution of (1.3). $\square$

Lyapunov stability and order stability

We prove now that $\mathcal{L}$-stability implies $\mathcal{O}$-stability. We start with a result, which establishes the $\mathcal{O}$-stability of an isolated solution of (1.3) in the presence of a lower or an upper solution.

PROPOSITION 1.21. *Assume* (C). *Suppose that* z *is a solution of* (1.3).

- *If* α *is a proper lower solution of* (1.3) *such that* $\alpha < z$ *and there is no solution* u *of* (1.3) *satisfying* $\alpha < u < z$, *then there exists a sequence* $(\alpha_n)_n$ *of proper regular lower solutions of* (1.3), *with* $\alpha < \alpha_n < \alpha_{n+1} < z$ *for each* n, *which converges in* $W^{1,1}(0, T)$ *to* z.
- *If* β *is a proper upper solution of* (1.3) *such that* $\beta > z$ *and there is no solution* u *of* (1.3) *satisfying* $z < u < \beta$, *then there exists a sequence* $(\beta_n)_n$ *of proper regular upper solutions of* (1.3), *with* $z < \beta_{n+1} < \beta_n < \beta$ *for each* n, *which converges in* $W^{1,1}(0, T)$ *to* z.

PROOF. We only prove the statement concerning the lower solution. Repeating recursively the argument in the proof of Proposition 1.20, we get a sequence $(\alpha_n)_n$ of proper regular lower solutions of (1.3) such that, for each n, $\alpha < \alpha_n < \alpha_{n+1} < z$ and α_{n+1} is a solution of

$$x' = f(t, \alpha_n) - h(t, x, \alpha_n), \quad x(0) = x(T).$$

Since $(\alpha_n)_n$ is a bounded sequence, we see, arguing as in the proof of Proposition 1.7, that $(\alpha_n)_n$ converges in $W^{1,1}(0, T)$ to a function $\hat{\alpha}$ such that $\hat{\alpha} \leqslant z$ and $\hat{\alpha}(0) = \hat{\alpha}(T)$. Moreover, as $\lim_{n\to+\infty} h(t, \alpha_{n+1}(t), \alpha_n(t)) = 0$ a.e. on $[0, T]$, we see that $\hat{\alpha}$ is a solution of (1.3) and, hence, $\hat{\alpha} = z$. □

PROPOSITION 1.22. *Assume* (C). *If a solution u of* (1.3) *is $\mathcal{L}$-stable from below* (*respectively $\mathcal{L}$-stable from above*, *respectively $\mathcal{L}$-stable*), *then it is $\mathcal{O}$-stable from below* (*respectively $\mathcal{O}$-stable from above*, *respectively $\mathcal{O}$-stable*).

PROOF. Let u be a solution of (1.3) which is $\mathcal{L}$-stable from below. If u is not isolated from below in $C^0([0, T])$ as a solution of (1.3), there is a sequence $(\alpha_n)_n$ of (regular lower) solutions of (1.3) such that $\alpha_n < u$ for every n and $\alpha_n \to u$ uniformly on $[0, T]$, that is, u is $\mathcal{O}$-stable from below. Let us suppose that u is isolated from below in $C^0([0, T])$. Then, arguing as in the proof of Proposition 1.18, we can construct a proper lower solution α of (1.3) such that $\alpha < u$ and there is no solution w of (1.3) with $\alpha < w < u$. By Proposition 1.21, we conclude that u is $\mathcal{O}$-stable from below. □

REMARK 1.13. The converse of Proposition 1.22 does not generally hold. Indeed, there may exist solutions of (1.3) which are $\mathcal{O}$-stable, but $\mathcal{L}$-unstable. A simple example is given by (1.3), with $f(x) = \operatorname{sgn}(x)\sqrt{|x \sin(\frac{1}{x})|}$ if $x \neq 0$ and $f(x) = 0$ if $x = 0$. Here, the equilibrium 0 is obviously $\mathcal{O}$-stable, but it is $\mathcal{L}$-unstable both from below and from above.

Lyapunov stability and order stability in case of uniqueness

We show that $\mathcal{L}$-stability and $\mathcal{O}$-stability are equivalent concepts if we assume uniqueness in the future or in the past for solutions of (1.2). We need two preliminary results: the former concerns the validity of a comparison principle, the latter establishes a form of continuous dependence.

LEMMA 1.23. *Suppose* (C). *Assume that uniqueness in the past holds for solutions of* (1.2). *Let α be a lower solution and β be an upper solution of* (1.1) *on* $[a, b]$ *and suppose that $\alpha(a) < \beta(a)$. Then $\alpha \ll \beta$ on* $[a, b]$.

PROOF. We first show that $\alpha \leqslant \beta$. Notice that α and β are a lower solution and an upper solution for any initial value problem $x' = f(t, x)$, $x(a) = x_0$ with $\alpha(a) \leqslant x_0 \leqslant \beta(a)$. Hence, if we assume, by contradiction, that $\alpha \nleqslant \beta$, then Proposition 1.1 yields the existence of $t_0 \in\,]a, b]$ such that $\alpha(t_0) = \beta(t_0)$. Theorem 1.4 implies then that Eq. (1.1) has solutions x_1, x_2, with $x_1(a) = \alpha(a)$ and $x_2(a) = \beta(a)$, satisfying $\alpha \leqslant x_1 < x_2 \leqslant \beta$ on $[a, t_0]$ and hence $x_1(t_0) = x_2(t_0)$; thus contradicting uniqueness in the past for solutions of (1.2). We just proved that $\alpha < \beta$. Finally, by arguing as above, uniqueness in the past for solutions of (1.2) implies that $\alpha \ll \beta$ on $[a, b]$. □

REMARK 1.14. Assume that uniqueness in the past holds for solutions of (1.2). If α is a lower solution and u is a solution of (1.3), with $\alpha < u$, then $\alpha \ll u$. This implies that, if u is $\mathcal{O}$-stable from below, i.e. if there exists a sequence $(\alpha_n)_n$ of regular lower solutions of (1.3) such that $\alpha_n < u$ for every n and $\alpha_n \to u$ uniformly on $[0, T]$, then there is a subsequence $(\alpha_{n_k})_k$ of $(\alpha_n)_n$ such that $\alpha_{n_k} \ll \alpha_{n_{k+1}} \ll u$ for every k. Further we can see that proper $\mathcal{O}$-stability from below is equivalent to strict $\mathcal{O}$-stability from below.

LEMMA 1.24. *Suppose* (C). *Assume that uniqueness in the future holds for solutions of* (1.2). *Let z be a solution of* (1.1) *on* $[a, b]$. *Then for every* $\varepsilon > 0$ *there is* $\delta > 0$ *such that, for every* $t_0 \in [a, b[$ *and every* x_0, *with* $z(t_0) - \delta < x_0 < z(t_0)$, *any right-nonextendible solution x of* (1.2) *exists* (*at least*) *on* $[t_0, b]$ *and satisfies* $z - \varepsilon < x < z$ *on* $[t_0, b]$.

PROOF. We start with the following basically known statement (cf. [121]), which we prove for the sake of completeness.

CLAIM. *For every* $\varepsilon > 0$ *there is* $\delta > 0$ *such that every right-nonextendible solution x of* (1.1), *with* $z(a) - \delta < x(a) < z(a)$, *exists and satisfies* $z(t) - \varepsilon < x(t) \leqslant z(t)$ *on* $[a, b]$.

Assume by contradiction that there exists $\varepsilon > 0$ and a sequence $(x_n)_n$ of solutions of (1.1) on $[a, b_n]$, with $b_n < b$, such that $x_n(a) \to z(a)$, $z(t) - \varepsilon < x_n(t) < z(t)$ on $[a, b_n[$ and $x_n(b_n) = z(b_n) - \varepsilon$. We can assume that $(x_n(a))_n$ is strictly increasing and hence, by uniqueness in the future, $x_n < x_{n+1}$ on $[a, b_n]$. Therefore, $(b_n)_n$ is increasing, so that $b_n \to \bar{b} \in]a, b]$. Let x be a solution of (1.1), with $x(\bar{b}) = z(\bar{b}) - 2\varepsilon$, defined on $[\bar{b} - \eta, \bar{b}]$, for some $\eta > 0$. By uniqueness in the future, we have $x \leqslant x_n \leqslant z$ on $[\bar{b} - \eta, \bar{b}] \cap [a, b_n]$, for all large n. Hence any such a solution x_n can be continued to $\bar{b}$ and satisfies $z - \varepsilon \leqslant x_n \leqslant z$ on $[a, b_n[$ and $x \leqslant x_n \leqslant z$ on $[b_n, \bar{b}]$. Further, since the sequence $(x_n)_n$ is bounded, by the L^1-Carathéodory conditions, it is also equicontinuous on $[a, \bar{b}]$ and therefore, by monotonicity, it converges to some function y uniformly on $[a, \bar{b}]$. Since the convergence takes place in $W^{1,1}(0, T)$ as well, y is a solution of (1.1), which satisfies $y(a) = z(a)$, as $x_n(a) \to z(a)$. Accordingly, by uniqueness in the future, we conclude that $y = z$ on $[a, \bar{b}]$. On the other hand, as $x_n(b_n) = z(b_n) - \varepsilon \to z(\bar{b}) - \varepsilon$, and, by equicontinuity, $x_n(b_n) \to y(\bar{b})$, we get $y(\bar{b}) = z(\bar{b}) - \varepsilon$, which is a contradiction. Thus, the proof of the claim is finished.

Conclusion. Fix $\varepsilon > 0$ and, according to the Claim, for every $t_0 \in [a, b[$ let $\delta(t_0)$ be the supremum of all $\delta > 0$ such that every right-nonextendible solution x of (1.1), with $z(t_0) - \delta < x(t_0) < z(t_0)$, exists and satisfies $z(t) - \varepsilon < x(t) \leqslant z(t)$ on $[t_0, b]$. This lemma will be proved if we show that $\inf_{t_0 \in [a,b[} \delta(t_0) > 0$. Assume by contradiction that $\inf_{t_0 \in [a,b[} \delta(t_0) = 0$ and let $(t_n)_n$ be a sequence such that $\delta(t_n) \to 0$. We can assume that $(t_n)_n$ is increasing and converges to some $\bar{t} \in]a, b]$. If $(t_n)_n$ is decreasing, the proof is similar. Let $\bar{\delta} > 0$ be the constant associated with $\bar{t}$ by the Claim. We can find a solution $\bar{x}$ of (1.1), with $z(\bar{t}) - \bar{\delta} < \bar{x}(\bar{t}) < z(\bar{t})$, which exists on $[\bar{t} - \eta, b]$, for some $\eta > 0$, and satisfies $z - \varepsilon < \bar{x} < z$ on $[\bar{t} - \eta, b]$ and $\bar{x} \ll z$ on $[\bar{t} - \eta, \bar{t}]$. Then every right-nonextendible solution x of (1.1), with $\bar{x}(\bar{t} - \eta) < x(\bar{t} - \eta) < z(\bar{t} - \eta)$, exists and satisfies $\bar{x} < x < z$ and, hence, $z - \varepsilon < x < z$

on $[\bar{t}-\eta, b]$. This yields a contradiction as we have $\delta(t) \geqslant \min_{[\bar{t}-\eta,\bar{t}]}(z(t)-\bar{x}(t)) > 0$, for every $t \in [\bar{t}-\eta, \bar{t}]$. □

PROPOSITION 1.25. *Assume* (C). *Suppose that uniqueness in the past, or in the future, holds for solutions of* (1.2). *Then a solution u of* (1.3) *is $\mathcal{L}$-stable from below* (*respectively $\mathcal{L}$-stable from above, respectively $\mathcal{L}$-stable*) *if and only if it is $\mathcal{O}$-stable from below* (*respectively $\mathcal{O}$-stable from above, respectively $\mathcal{O}$-stable*).

PROOF. Let u be a solution of (1.3), which is $\mathcal{O}$-stable from below, and let $(\alpha_n)_n$ be a sequence of regular lower solutions of (1.3) such that $\alpha_n < u$ for every n and $\alpha_n \to u$ uniformly on $[0, T]$.

Case 1: *uniqueness in the past holds for* (1.2). By Lemma 1.23 and Remark 1.14, we can suppose that $\alpha_n \ll u$ for each n. Fix $\varepsilon > 0$ and pick n such that $\|\alpha_n - u\|_\infty < \varepsilon$. Set $\delta = \inf_{[0,T]}(u-\alpha_n) > 0$. Then for every $t_0 \in [0, T[$ and x_0 such that $\alpha_n(t_0) \leqslant u(t_0) - \delta < x_0 < u(t_0)$, any right-nonextendible solution x of (1.2) satisfies, according to Lemma 1.23, $\alpha_n \ll x \ll u$ on $[t_0, b]$ for any given $b > t_0$. Hence x exists and satisfies $u - \varepsilon < x < u$ on $[t_0, +\infty[$. This means that u is $\mathcal{L}$-stable from below.

Case 2: *uniqueness in the future holds for* (1.2). The conclusion is an immediate consequence of the following statement.

CLAIM. *Let α be a regular lower solution of* (1.3), *with $\alpha < u$, and set $\eta = \|\alpha - u\|_\infty$. Then there exists $\delta > 0$ such that for every $t_0 \in [0, T[$ every right-nonextendible solution x of* (1.2), *with $u(t_0) - \delta < x_0 < u(t_0)$, exists and satisfies $u - \eta < x < u$ on $[t_0, +\infty[$.*

If $\alpha \ll u$, then the result follows setting $\delta = \inf_{[0,T]}(u-\alpha) > 0$ and using Corollary 1.5. Therefore, suppose $\inf_{[0,T]}(u-\alpha) = 0$. Hence there exist points $t_1 \in [0, T]$ and $t_2 \in]t_1, t_1 + T[$ such that $\alpha(t_1) = u(t_1) - \eta$, $\lim_{t \to t_2^-} \alpha(t) = u(t_2)$ and $u(t) - \eta \leqslant \alpha(t) < u(t)$ on $[t_1, t_2[$. By Lemma 1.24, there is $\delta \in]0, \eta[$ such that, for every $t_0 \in [0, T[$, any right-nonextendible solution x of (1.2), with $u(t_0) - \delta < x_0 < u(t_0)$, exists and satisfies $u(t) - \eta < x(t) \leqslant u(t)$ on $[t_0, t_1 + T]$. Since $\alpha(t_1 + T) = u(t_1 + T) - \eta \leqslant x(t_1 + T) \leqslant u(t_1 + T)$ and α is a lower solution, x exists and satisfies $\alpha(t) \leqslant x(t) \leqslant u(t)$ on $[t_1 + T, t_2 + T[$ and $x(t) = u(t)$ on $[t_2 + T, +\infty[$. □

One-sided weak stability

The following definition of stability is a possible weakening of $\mathcal{L}$-stability suited to deal with cases where uniqueness of the solutions of (1.2) is not guaranteed (see [50]).

DEFINITION 1.17.

- We say that a solution u of (1.3) is *weakly stable* (briefly, *$\mathcal{W}$-stable*) *from below* if, for every $\varepsilon > 0$, there is a uniformly continuous function $\delta : [0, T[\to \mathbb{R}$, with $\delta > 0$, such that, for every $t_0 \in [0, T[$ and every x_0, with $u(t_0) - \delta(t_0) \leqslant x_0 \leqslant u(t_0)$, there exists a solution $x : [t_0, +\infty[\to \mathbb{R}$ of (1.2) which satisfies (1.15).
- If, further, (1.16) holds, we say that u is *$\mathcal{W}$-asymptotically stable from below*.
- *$\mathcal{W}$-stability* and *$\mathcal{W}$-asymptotic stability from above* are defined in a similar way.
- We say that a solution u of (1.3) is *$\mathcal{W}$-stable* if, for every $\varepsilon > 0$, there is a uniformly continuous function $\delta : [0, T[\to \mathbb{R}$, with $\delta > 0$, such that, for every $t_0 \in [0, T[$ and

every x_0, with $|x_0 - u(t_0)| \leqslant \delta(t_0)$, there exists a solution $x : [t_0, +\infty[\to \mathbb{R}$ of (1.2) *which satisfies* (1.17). *If, further,* (1.16) *holds, we say that u is $\mathcal{W}$-asymptotically stable.*

Order stability and weak stability
The notion of $\mathcal{W}$-stability is in any case equivalent to that of $\mathcal{O}$-stability.

PROPOSITION 1.26. *Assume* (C). *Let u be a solution of* (1.3). *Then u is $\mathcal{W}$-stable from below* (*respectively $\mathcal{W}$-stable from above, respectively $\mathcal{W}$-stable*) *if and only if it is $\mathcal{O}$-stable from below* (*respectively $\mathcal{O}$-stable from above, respectively $\mathcal{O}$-stable*).

PROOF. We only discuss stability from below. The proof is divided into two steps.

Step 1. Let u be a solution of (1.3), which is $\mathcal{O}$-stable from below, and let $(\alpha_n)_n$ be a sequence of regular lower solutions of (1.3) such that $\alpha_n < u$ for every n and $\alpha_n \to u$ uniformly on $[0, T]$. Fix $\varepsilon > 0$ and take n such that $\|\alpha_n - u\|_\infty < \varepsilon$. Then set $\delta = u - \alpha_n$ on $[0, T[$. Corollary 1.5 implies that, for every $t_0 \in [0, T]$ and every x_0, with $u(t_0) - \delta(t_0) \leqslant x_0 \leqslant u(t_0)$, there is a solution x which exists and satisfies $u - \varepsilon < x \leqslant u$ on $[t_0, +\infty[$.

Step 2. Let u be a solution of (1.3), which is $\mathcal{W}$-stable from below. If u is not isolated from below as a solution of (1.3), then it is obviously $\mathcal{O}$-stable. Accordingly, suppose that u is isolated from below. Let $\varepsilon > 0$ be such that there is no solution z of (1.3), satisfying $z < u$ and $\|z - u\|_\infty \leqslant \varepsilon$, and let δ be the function associated with ε. Since $\delta > 0$, there exists $t_0 \in [0, T[$ such that $\delta(t_0) > 0$. Fix a strictly increasing sequence $(x_0^{(n)})_n$ such that, for every n, $u(t_0) - \delta(t_0) < x_0^{(n)} < u(t_0)$ and $x_0^{(n)} \to u(t_0)$. Denote by $(x_n)_n$ the corresponding sequence of solutions of (1.2), satisfying, for every n, $x_n(t_0) = x_0^{(n)}$, $u - \varepsilon < x_n \leqslant u$ and $x_n < x_{n+1}$ on $[t_0, +\infty[$. Fix n and observe that $x_n(t_0 + T) > x_n(t_0)$. Indeed, otherwise $x_n|_{[t_0,t_0+T[}$ is a proper upper solution of (1.3) and hence, by Proposition 1.8, there exists a solution z of (1.3) such that $z < x_n$ on $[t_0, +\infty[$, $x_n(t) - z(t) \to 0$, as $t \to +\infty$, and hence $\|z - u\|_\infty \leqslant \varepsilon$, which is a contradiction. Therefore, denoting by α_n the T-periodic extension of $x_n|_{[t_0,t_0+T[}$ to $\mathbb{R}$, we conclude that $(\alpha_n)_n$ is a sequence of proper lower solutions of (1.3), satisfying, for every n, $\alpha_n < \alpha_{n+1} < u$ and $\alpha_n \to u$ uniformly on $[0, T]$. According to Proposition 1.20, we can construct a further sequence $(\bar{\alpha}_n)_n$ of proper regular lower solutions of (1.3) such that, for every n, $\bar{\alpha}_n < u$ and $\bar{\alpha}_n \to u$ uniformly on $[0, T]$. This implies that u is $\mathcal{O}$-stable from below. □

Stability via lower and upper solutions
We now come to the core of this section. In the following results we use lower and upper solutions to give a precise description of the stability properties of the T-periodic solutions of (1.1). In the next theorem we prove the stability of a one-sided isolated solution z of (1.3) in the presence of a lower solution below z, or of an upper solution above z.

THEOREM 1.27. *Assume* (C). *Let z be a solution of* (1.3).

(i) *If α is a proper lower solution of* (1.3) *such that $\alpha < z$ and there is no solution u of* (1.3) *satisfying $\alpha < u < z$, then z is $\mathcal{W}$-asymptotically stable from below.*
(ii) *If β is a proper upper solution of* (1.3) *such that $\beta > z$ and there is no solution u of* (1.3) *satisfying $z < u < \beta$, then z is $\mathcal{W}$-asymptotically stable from above.*

PROOF. We prove only the former statement. Since α is a proper lower solution, we know, by Proposition 1.21, that z is (properly) $\mathcal{O}$-stable from below, hence, by Proposition 1.26, z is $\mathcal{W}$-stable from below. The asymptotic $\mathcal{W}$-stability is then a consequence of Corollary 1.5 and Proposition 1.8. □

Conversely, Theorem 1.30 will show that instability occurs in all the other cases; it is based on the following two propositions.

PROPOSITION 1.28. *Assume* (C). *Let z be a solution of* (1.3).

(i) *If α is a proper lower solution of* (1.3) *such that $\alpha > z$ and there is no solution u of* (1.3) *satisfying $\alpha > u > z$, then z is $\mathcal{L}$-unstable from above.*

(ii) *If β is a proper upper solution of* (1.3) *such that $\beta < z$ and there is no solution u of* (1.3) *satisfying $z > u > \beta$, then z is $\mathcal{L}$-unstable from below.*

PROOF. We prove only the former statement. Let $t_0 \in [0, T]$ be such that $\alpha(t_0) > z(t_0)$. By Remark 1.9 there exists a solution $\widetilde{\alpha} :]-\infty, t_0] \to \mathbb{R}$ of (1.2), with $x_0 = \alpha(t_0)$, satisfying $z \leqslant \widetilde{\alpha} \leqslant \alpha$ on $]-\infty, t_0]$ and such that

$$\lim_{t \to -\infty} \bigl(\widetilde{\alpha}(t) - z(t)\bigr) = 0. \tag{1.18}$$

Set $\varepsilon = \alpha(t_0) - z(t_0) > 0$ and take any $\delta > 0$. By (1.18), there exists a positive integer n such that $z(t_0 - nT) \leqslant \widetilde{\alpha}(t_0 - nT) < z(t_0 - nT) + \delta$. Define, on $]-\infty, t_0 + nT]$, $x(t) = \widetilde{\alpha}(t - nT)$. Then x is a solution of (1.1) with $z(t_0) \leqslant x(t_0) < z(t_0) + \delta$ but $x(t_0 + nT) - z(t_0 + nT) = \varepsilon$. Hence we easily conclude that z is $\mathcal{L}$-unstable from above. □

PROPOSITION 1.29. *Assume* (C). *Let z be a solution of* (1.3).

(i) *If there exist a lower solution α of* (1.3) *and points t_1, t_2 such that $z(t_1) = \alpha(t_1)$ and $z(t_2) < \alpha(t_2)$, then z is $\mathcal{L}$-unstable from above.*

(ii) *If there exist an upper solution β of* (1.3) *and points t_1, t_2 such that $z(t_1) = \beta(t_1)$ and $z(t_2) > \beta(t_2)$, then z is $\mathcal{L}$-unstable from below.*

PROOF. We prove only the former statement. Assume without loss of generality that $t_1 < t_2$. By Proposition 1.8, we know that either every right-nonextendible solution $x : [t_1, \omega_+[\to \mathbb{R}$ of (1.2), with $t_0 = t_1$, $x_0 = \alpha(t_1) = z(t_1)$ and $x \geqslant \alpha$, is such that

$$\limsup_{t \to \omega_+} x(t) = +\infty,$$

or there exist the minimum solution $\widetilde{\alpha} : [t_1, +\infty[\to \mathbb{R}$ in $[\alpha, +\infty[$ of (1.2), with $t_0 = t_1$, $x_0 = \alpha(t_1) = z(t_1)$, and the minimum solution v in $[\alpha, +\infty[$ of (1.3), satisfying $v \geqslant \widetilde{\alpha}$ and

$$\lim_{t \to +\infty} \bigl(\widetilde{\alpha}(t) - v(t)\bigr) = 0.$$

Since there exists t_2 such that $\alpha(t_2) > z(t_2)$, we have $v \neq z$ and hence

$$\limsup_{t\to+\infty}\bigl(\widetilde{\alpha}(t) - z(t)\bigr) > 0.$$

In both cases we easily conclude that z is $\mathcal{L}$-unstable from above. □

THEOREM 1.30. *Assume* (C). *Let z be a solution of* (1.3).

- (i) *If there exists a lower solution α of* (1.3) *such that $z \not\geqslant \alpha$ and, in case $z < \alpha$, there is no solution u of* (1.3) *satisfying $z < u < \alpha$, then z is $\mathcal{L}$-unstable from above.*
- (ii) *If there exists an upper solution β of* (1.3) *such that $z \not\leqslant \beta$ and, in case $z > \beta$, there is no solution u of* (1.3) *satisfying $\beta < u < z$, then z is $\mathcal{L}$-unstable from below.*

PROOF. We prove only the former statement. Two cases may occur: either $\alpha > z$ or, by Proposition 1.6, there exist points t_1, t_2 such that $z(t_1) = \alpha(t_1)$ and $z(t_2) < \alpha(t_2)$. In the former case, the conclusion follows from Proposition 1.28, whereas in the latter it follows from Proposition 1.29. □

A direct consequence of the previous results is the following completion of Theorem 1.9, for what concerns the stability properties of solutions of (1.3) in the presence of a pair of lower and upper solutions.

THEOREM 1.31. *Assume* (C). *Suppose that α is a lower solution and β is an upper solution of* (1.3). *Denote by v and w the minimum and the maximum solutions in* $[\min\{\alpha,\beta\}, \max\{\alpha,\beta\}]$ *of* (1.3).

- (i) *If $\alpha < \beta$, then v is $\mathcal{W}$-asymptotically stable from below, provided α is a proper lower solution, and w is $\mathcal{W}$-asymptotically stable from above, provided β is a proper upper solution.*
- (ii) *If $\alpha > \beta$, then v is $\mathcal{L}$-unstable from below, provided β is a proper upper solution, and w is $\mathcal{L}$-unstable from above, provided α is a proper lower solution.*
- (iii) *If $\alpha \not\leqslant \beta$ and $\beta \not\leqslant \alpha$, then every solution u of* (1.3) *satisfying $v \leqslant u \leqslant w$ is $\mathcal{L}$-unstable from below and from above.*

PROOF. In case (i) the conclusion follows from Theorem 1.27. In case (ii) the conclusion follows from Proposition 1.28. In case (iii), if u is a solution of (1.3) satisfying $\min\{\alpha,\beta\} \leqslant v \leqslant u \leqslant w \leqslant \max\{\alpha,\beta\}$, then according to Proposition 1.6, there exist points t_1 and t_2 such that $\alpha(t_1) = \beta(t_1) = u(t_1)$ and either $\beta(t_2) < u(t_2)$, or $u(t_2) < \alpha(t_2)$. The instability of u then follows from Proposition 1.29. □

REMARK 1.15. In case (i), from Proposition 1.8 we further deduce the following invariance and attractivity result. For every t_0, every solution $x : [t_0, +\infty[\to \mathbb{R}$ of (1.1) is such that: if $\alpha \leqslant x \leqslant \beta$ then $\widetilde{\alpha} \leqslant x \leqslant \widetilde{\beta}$, if $\alpha \leqslant x \leqslant v$ then $\lim_{t\to+\infty}(v(t) - x(t)) = 0$, if $w \leqslant x \leqslant \beta$ then $\lim_{t\to+\infty}(x(t) - w(t)) = 0$.

From Propositions 1.18, 1.25, 1.26 and Remark 1.14 we immediately get the following result.

COROLLARY 1.32. *Assume* (C). *Suppose that α is a proper lower solution, β is a proper upper solution of* (1.3) *and $\alpha < \beta$. Denote by v and w the minimum and the maximum solutions in* $[\alpha, \beta]$ *of* (1.3). *Suppose further that either uniqueness in the past for solutions of* (1.2) *holds, or uniqueness in the future for solutions of* (1.2) *holds and α and β are strict. Then v is $\mathcal{L}$-asymptotically stable from below and w is $\mathcal{L}$-asymptotically stable from above.*

Continua of stable or unstable T-periodic solutions

Theorem 1.31 describes the stability properties of the minimum and the maximum solutions v and w of (1.3) wedged between a given pair of lower and upper solutions. Now, we turn to investigate the behaviour of the solutions between v and w. We start by specifying statement (i) in Theorem 1.31. Namely, we prove the existence of a totally ordered continuum of (two-sided) $\mathcal{W}$-stable T-periodic solutions lying between a pair of lower and upper solutions ordered in the standard way, i.e. the lower solution is below the upper solution.

THEOREM 1.33. *Assume* (C). *Suppose that α is a proper lower solution and β is a proper upper solution of* (1.3), *satisfying $\alpha < \beta$, and denote by v and w respectively the minimum and the maximum solution of* (1.3) *in* $[\alpha, \beta]$. *Then there exists a totally ordered continuum $\mathcal{K}$ in* $C^0([0, T])$, *such that every $u \in \mathcal{K}$ is an $\mathcal{O}$-stable solution of* (1.3) *satisfying $v \leqslant u \leqslant w$; further, $u_1 = \min \mathcal{K}$ is properly $\mathcal{O}$-stable from below and $u_2 = \max \mathcal{K}$ is properly $\mathcal{O}$-stable from above.*

PROOF. Let us denote by $\mathcal{S}_1$ the set of all solutions u of (1.3), with $\alpha < u < \beta$, which are properly $\mathcal{O}$-stable from below. Since, by Proposition 1.21, the minimum solution v is properly $\mathcal{O}$-stable from below, $\mathcal{S}_1$ is non-empty. Notice also that, since $\mathcal{S}_1$ is bounded in $C^0([0, T])$, the L^1-Carathéodory conditions imply that it is also equicontinuous and hence relatively compact in $C^0([0, T])$.

Let us show that the set $(\mathcal{S}_1, \leqslant)$ is inductively ordered. Let $\mathcal{T}$ be a totally ordered subset of $\mathcal{S}_1$. Since $\mathcal{T}$ is uniformly bounded from above, by Lemma 1.15 there exists a maximal solution $\hat{w}$ of (1.3) in the C^0-closure $\bar{\mathcal{T}}$ of $\mathcal{T}$. Actually $\hat{w} \in \mathcal{T}$. Indeed, let $(w_n)_n$ be an increasing sequence in $\mathcal{T}$ converging to $\hat{w}$ in $C^0([0, T])$; since each w_n is properly $\mathcal{O}$-stable from below, by a diagonal argument we conclude that $\hat{w}$ as well is properly $\mathcal{O}$-stable from below. Hence any totally ordered subset of $\mathcal{S}_1$ has an upper bound in $\mathcal{S}_1$. By Zorn Lemma, there exists a maximal element $u_1 \in \mathcal{S}_1$.

Let us denote by $\mathcal{S}_2$ the set of all solutions u of (1.3), with $u_1 \leqslant u < \beta$, which are properly $\mathcal{O}$-stable from above. Since the maximum solution w is properly $\mathcal{O}$-stable from above, $\mathcal{S}_2$ is non-empty. Arguing as above, we see that there exists a minimal solution $u_2 \in \mathcal{S}_2$.

If $u_1 = u_2$, the conclusion is achieved. Therefore, let us suppose that $u_1 < u_2$ and let us denote by $\mathcal{S}_3$ the set of all solutions u of (1.3), with $u_1 \leqslant u \leqslant u_2$. Notice that $\mathcal{S}_3$ is compact in $C^0([0, T])$.

Let us observe that there is no proper lower solution and no proper upper solution of (1.3) between u_1 and u_2. Indeed, if we assume that there exists, for instance, a proper lower solution α^*, with $u_1 < \alpha^* < u_2$, and we denote by z the minimum solution of (1.3), with

$\alpha^* < z \leqslant u_2$, Proposition 1.21 implies that z is properly $\mathcal{O}$-stable from below, thus contradicting the maximality of u_1.

Next, we prove that if $z_1, z_2 \in \mathcal{S}_3$, with $z_1 < z_2$, then there exists a solution z_3 of (1.3) such that $z_1 < z_3 < z_2$. Indeed, if we assume that there is no solution z of (1.3), with $z_1 < z < z_2$, then any nonextendible solution x of (1.1), with $z_1 < x < z_2$, is either a proper lower solution or a proper upper solution, thus contradicting our preceding conclusion. Hence $\mathcal{S}_3$ is dense-in-itself with respect to the order.

Now, let us fix a solution $u_0 \in \mathcal{S}_3$ and denote by $\mathcal{S}(u_0)$ a maximal totally ordered subset of $\mathcal{S}_3$ with $u_0 \in \mathcal{S}(u_0)$. As $u_1, u_2 \in \mathcal{S}(u_0)$, this set is nondegenerate. By Lemma 1.3, $\mathcal{S}(u_0)$ is homeomorphic to a compact interval of $\mathbb{R}$, so it is connected.

Finally, it is clear that every $u \in \mathcal{S}(u_0)$ is $\mathcal{O}$-stable and $u_1 = \min \mathcal{S}(u_0)$ and $u_2 = \max \mathcal{S}(u_0)$ are properly $\mathcal{O}$-stable from below and above, respectively. Therefore, we can set $\mathcal{K} = \mathcal{S}(u_0)$. □

REMARK 1.16. The set $\mathcal{S}_3$, defined in the proof of Theorem 1.33, is obviously connected too. Indeed, for each $u_0 \in \mathcal{S}_3$, the set $\mathcal{S}(u_0)$ is connected. Since $u_1, u_2 \in \bigcap_{u_0 \in \mathcal{S}_3} \mathcal{S}(u_0)$, the set $\mathcal{S}_3 = \bigcup_{u_0 \in \mathcal{S}_3} \mathcal{S}(u_0)$ is connected.

REMARK 1.17. We notice that if α is a lower solution and β is an upper solution of (1.3), with $\alpha > \beta$, then one cannot generally guarantee the existence in between of a solution of (1.3) which is $\mathcal{L}$-unstable both from below and from above. A simple example is given by (1.3), with $f(x) = x|\sin(1/x)|$ for $x \neq 0$ and $f(x) = 0$ for $x = 0$. Here, $\beta = -1$ is a strict upper solution and $\alpha = 1$ is a strict lower solution; in between there is a countable set of equilibria, none of which is simultaneously $\mathcal{L}$-unstable from below and from above. Of course, according to Theorem 1.31, there are equilibria which are $\mathcal{L}$-unstable either from below or from above.

Finally, we consider the case where α and β are not comparable. A counterpart of Theorem 1.33, which specifies statement (iii) in Theorem 1.31, is based on the following preliminary result.

LEMMA 1.34. *Assume* (C). *Let v, w be solutions of* (1.3), *such that $v \neq w$ and $v(t_0) = w(t_0)$ for some t_0. Then the set*

$$\mathcal{K} = \bigl\{u : [t_0, t_0 + T] \to \mathbb{R} \mid u \text{ is a solution of (1.1) with } \min\{v, w\} \leqslant u \leqslant \max\{v, w\}\bigr\}$$

is a nondegenerate continuum in $C^0([t_0, t_0 + T])$ and every $u \in \mathcal{K}$, extended by T-periodicity onto $\mathbb{R}$, is a $\mathcal{L}$-unstable solution of (1.3).

PROOF. Every solution u of (1.1), with $\min\{v, w\} \leqslant u \leqslant \max\{v, w\}$ on $[t_0, t_0 + T]$, extended by T-periodicity onto $\mathbb{R}$, is a solution of (1.3). Since $\min\{v, w\}$ and $\max\{v, w\}$ are solutions of (1.2), with $x_0 = v(t_0) = w(t_0)$, Theorem 1.4 guarantees that the set $\mathcal{K}$ is a continuum in $C^0([t_0, t_0 + T])$. The instability finally follows from Proposition 1.29. □

THEOREM 1.35. *Assume* (C). *Suppose that* α *is a lower solution and* β *is an upper solution of* (1.3), *with* $\alpha \nleqslant \beta$ *and* $\alpha \ngeqslant \beta$. *Then the set*

$$\mathcal{K} = \{u \mid u \text{ is a solution of } (1.3) \text{ with } \min\{\alpha, \beta\} \leqslant u \leqslant \max\{\alpha, \beta\}\}$$

is a continuum in $C^0([0, T])$ *and every* $u \in \mathcal{K}$ *is* $\mathcal{L}$*-unstable.*

Complicated dynamics

In the frame of Lemma 1.34, the existence of a nondegenerate continuum entails fairly complicated dynamics for the solutions of (1.1) between $\min\{v, w\}$ and $\max\{v, w\}$. In particular, there exist homoclinics, subharmonics of any order and almost periodic solutions, which are not periodic of any fixed period; this last assertion will now be proved.

THEOREM 1.36. *Assume* (C). *Let* v, w *be solutions of* (1.3) *such that* $v \neq w$ *and* $v(t_0) = w(t_0)$ *for some* t_0. *Then there exists a* (*Čech–Lebesgue*) *infinite dimensional set* $\mathcal{X} \subseteq L^\infty(\mathbb{R})$ *such that every* $x \in \mathcal{X}$ *is a* (*Bohr*) *almost periodic solution of* (1.1), *with* $\min\{v, w\} < x < \max\{v, w\}$, *which is not periodic of any fixed period.*

PROOF. Possibly replacing v with $\min\{v, w\}$ and w with $\max\{v, w\}$, we can assume $v < w$ and $v(t_0) = w(t_0)$ for some t_0. It is not even restrictive to suppose $t_0 = 0$. Let $\mathcal{K} \subseteq C^0([0, T])$ be the set of all solutions x of (1.3) with $v \leqslant x \leqslant w$. By Lemma 1.34, $\mathcal{K}$ is a nondegenerate continuum and it is dense-in-itself with respect to the order (see also Proposition 1.13).

Step 1. *Existence of almost periodic solutions.* Without loss of generality we may assume $T = 1$. Fix a sequence $(u_n)_n$ of solutions of (1.3) such that $v \leqslant u_n \leqslant w$ and

$$0 < \|u_{n+\nu} - u_n\|_\infty < 2^{-n} \tag{1.19}$$

for every $n \in \mathbb{N}$ and $\nu \in \mathbb{N}^+$. Hence $(u_n)_n$ uniformly converges to a solution u_∞ of (1.3), with $v \leqslant u_\infty \leqslant w$. Let us define a function $x : \mathbb{R} \to \mathbb{R}$ by setting $x(t) = u_\infty(t)$ on $[0, 1[$ and, for each $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, $x(t) = u_n(t)$ on $[(2k + 1)2^n, (2k + 1)2^n + 1[$. The function x is well-defined, since every $m \in \mathbb{Z} \setminus \{0\}$ can be represented in a unique way as $m = (2k + 1)2^n$ for some $k \in \mathbb{Z}$ and $n \in \mathbb{N}$. Moreover, since $u_n(0) = u_\infty(0)$ for each n, x is a solution of (1.1).

We want to prove that x is almost periodic and it is not periodic of any fixed period. Let us recall that a continuous function $x : \mathbb{R} \to \mathbb{R}$ is almost periodic if, for every $\varepsilon > 0$, there exists $l > 0$ such that, for every t_0, there is $\tau \in [t_0, t_0 + l]$ with $|x(t + \tau) - x(t)| < \varepsilon$ on $\mathbb{R}$.

Fix $\varepsilon > 0$ and pick $n \in \mathbb{N}^+$ such that

$$2^{-n} < \varepsilon \tag{1.20}$$

and

$$\|u_j - u_\infty\|_\infty < \varepsilon \tag{1.21}$$

for all $j \geqslant n$. Set $l = 2^{n+2}$ and take $t_0 \in \mathbb{R}$. Let $k \in \mathbb{Z}$ be such that $(2k-1)2^n < t_0 \leqslant (2k+1)2^n$. Accordingly, set $\tau = (2k+1)2^n$ and note that $\tau \in [t_0, t_0 + 2^{n+1}[\subseteq [t_0, t_0 + l]$. We want to prove that, for every t, $|x(t+\tau) - x(t)| < \varepsilon$.

Let $t \in [0, 1[$ and hence $t + \tau \in [(2k+1)2^n, (2k+1)2^n + 1[$. We have, by (1.21),

$$\left|x(t+\tau) - x(t)\right| = \left|u_n(t) - u_\infty(t)\right| < \varepsilon.$$

Let $t \in [(2j+1)2^p, (2j+1)2^p + 1[$, for some $j \in \mathbb{Z}$ and $p \in \mathbb{N}$, and hence $t + \tau \in [(2j+1)2^p + (2k+1)2^n, (2j+1)2^p + (2k+1)2^n + 1[$. In order to estimate $x(t+\tau) - x(t)$, we distinguish three cases.

- Let $p < n$. Then there exists $r \in \mathbb{N}$ such that $p + 1 + r = n$ and $(2j+1)2^p + (2k+1)2^n = (2j+1)2^p + (2k+1)2^{p+1+r} = (2(j + (2k+1)2^r) + 1)2^p$. Hence we have $x(t+\tau) = u_p(t) = x(t)$.
- Let $p > n$. Then there exists $r \in \mathbb{N}$ such that $n + 1 + r = p$ and $(2j+1)2^p + (2k+1)2^n = (2j+1)2^{n+r+1} + (2k+1)2^n = (2(k + (2j+1)2^r) + 1)2^n$. Hence, by (1.19) and (1.20), we have $x(t+\tau) = u_n(t)$ and

$$\left|x(t+\tau) - x(t)\right| = \left|u_n(t) - u_p(t)\right| < 2^{-n} < \varepsilon.$$

- Let $p = n$. Then we have $(2j+1)2^p + (2k+1)2^n = (j+k+1)2^{p+1}$. If $j+k+1 = 0$, then $t + \tau \in [0, 1[$ and, by (1.20) and (1.21),

$$\left|x(t+\tau) - x(t)\right| = \left|u_\infty(t) - u_p(t)\right| = \left|u_\infty(t) - u_n(t)\right| < \varepsilon.$$

If $j + k + 1 \neq 0$, then there exist $r \in \mathbb{N}$ and $h \in \mathbb{Z}$ such that $j + k + 1 = (2h+1)2^r$; hence $(j+k+1)2^{p+1} = (2h+1)2^{p+r+1}$. Accordingly, by (1.19) and (1.20), we get

$$\left|x(t+\tau) - x(t)\right| = \left|u_{p+r+1}(t) - u_p(t)\right| = \left|u_{n+r+1}(t) - u_n(t)\right| < 2^{-n} < \varepsilon.$$

Finally, we prove that x is not T-periodic for any $T > 0$. We first recall that, for each $n \in \mathbb{N}$, $x = u_n$ on $[2^n, 2^n + 1]$ and $x = u_{n+1}$ on $[2^{n+1}, 2^{n+1} + 1]$. Since $u_n \neq u_{n+1}$, x is not T-periodic for any $T \in \mathbb{Q}^+$. Assume by contradiction that x is T-periodic for some $T \in \mathbb{R}^+ \setminus \mathbb{Q}$. For each $n \in \mathbb{Z}$, we have $x(n) = x(0)$, by construction, and $x(n + mT) = x(0)$ for every $m \in \mathbb{Z}$, by T-periodicity. Using the density in $\mathbb{R}$ of the set $\{n + mT \mid n, m \in \mathbb{Z}\}$ and the continuity of x, we conclude that x is constant, which is a contradiction.

Step 2. *Existence of an infinite dimensional set of almost periodic solutions.* Fix a sequence $(z_n)_n$ of solutions of (1.3) satisfying $v < z_{n+1} < z_n < w$ and $\|z_n - z_{n+\nu}\|_\infty < 2^{-n}$ for every $n, \nu \in \mathbb{N}$. For each n, let $\mathcal{S}_n$ be the set of all solutions u of (1.3) such that $z_{2n+1} \leqslant u \leqslant z_{2n}$. By Lemma 1.34, $\mathcal{S}_n \subseteq C^0([0, T])$ is compact and dense-in-itself with respect to the order. Hence Lemma 1.3 implies that there exists a continuum $\mathcal{T}_n \subseteq \mathcal{S}_n$ which is homeomorphic to $[0, 1]$. For any sequence $(u_n)_n \in \prod_{n=0}^{+\infty} \mathcal{T}_n$ denote by $\Phi((u_n)_n)$ the almost periodic solution x of (1.1) constructed in Step 1 and set $\mathcal{X} = \Phi\big(\prod_{n=0}^{+\infty} \mathcal{T}_n\big)$. Fix a sequence $(\hat{u}_n)_n \in \prod_{n=0}^{+\infty} \mathcal{T}_n$ and let N be any positive integer. The map $\Phi_N : \prod_{n=0}^{N-1} \mathcal{T}_n \to L^\infty(\mathbb{R})$ defined by setting $\Phi_N(u_0, u_1, \ldots, u_{N-1}) = \Phi(u_0, u_1, \ldots, u_{N-1}, \hat{u}_N, \hat{u}_{N+1}, \ldots)$ is one-to-one and continuous, once $\prod_{n=0}^{N-1} \mathcal{T}_n$ is endowed with the product topology. Actually, since

$\prod_{n=0}^{N-1} \mathcal{T}_n$ is compact, Φ_N is a homeomorphism between $\prod_{n=0}^{N-1} \mathcal{T}_n$ and $\Phi_N(\prod_{n=0}^{N-1} \mathcal{T}_n) = \mathcal{X}_N$. Since $\mathcal{X}_N \subseteq \mathcal{X}$ and $\dim \mathcal{X}_N = N$, we conclude that $\dim \mathcal{X} = +\infty$ (see [47, Chapter 7]). □

EXAMPLE 1.4. We consider once more equation (1.12). Let $\alpha : [-\frac{1}{2}, \frac{1}{2}[\to \mathbb{R}$ be defined by $\alpha(t) = (\frac{\sqrt{2}-1}{4} t)^2$ on $[-\frac{1}{2}, 0]$, and $\alpha(t) = (\frac{1}{4}t)^2$ on $]0, \frac{1}{2}[$. Then α is a solution of (1.12) and a proper lower solution of the associated 1-periodic problem. Note that $\beta = 0$ is a proper upper solution of the 1-periodic problem. By Theorem 1.9, there exists a 1-periodic solution w of (1.12) such that $0 < w < \alpha$ and $w(0) = 0$. Observe that, by symmetry, if x is a solution of (1.12), then $y(t) = -x(-t)$ is also a solution of (1.12). Hence the function $v(t) = -w(-t)$ is a 1-periodic solution of (1.12) such that $v < 0$ and $v(0) = 0$. Applying Lemma 1.34 we see that the set $\mathcal{K}$ of all solutions $u : [0, 1] \to \mathbb{R}$ of (1.12), with $v \leqslant u \leqslant w$, is a nondegenerate continuum in $C^0([0, 1])$ and every $u \in \mathcal{K}$ is a $\mathcal{L}$-unstable 1-periodic solution of (1.12). In addition, we can see that $v(t) < 0 < w(t)$ on $]0, 1[$ and the continuum $\mathcal{K}$ is totally ordered. Indeed, by a result in [29, Corollary 3.5], uniqueness for the Cauchy problem associated with Eq. (1.12) is guaranteed for any t_0, x_0 with $(t_0, x_0) \neq (0, 0)$. Hence, if $u_1, u_2 \in \mathcal{K}$, we have $u_1(t) = u_2(t)$ if and only if $t \in \mathbb{Z}$; the 1-periodicity of u_1 and u_2 yields either $u_1 < u_2$ or $u_2 < u_1$. By Theorem 1.36 there exists an infinite dimensional set $\mathcal{X} \subseteq L^\infty(\mathbb{R})$ such that every $x \in \mathcal{X}$ is an almost periodic solution of (1.12), with $v < x < w$, and x is not periodic of any fixed period. Furthermore, there exist periodic solutions x of (1.12), with $v < x < w$, that are subharmonics of any order. To verify this let $p_1, p_2, \dots, p_n$ be two by two distinct numbers such that $v(\frac{1}{2}) \leqslant p_i \leqslant w(\frac{1}{2})$ for $i = 1, \dots, n$. Let $u_i \in \mathcal{K}$ be such that $u_i(\frac{1}{2}) = p_i$. Define $x(t) = u_i(t)$ on $[i - 1, i[$, for $i = 1, \dots, n$, and extend x by n-periodicity to $\mathbb{R}$. Then x is an n-periodic solution of (1.12). Clearly, x is not k-periodic for any positive integer $k < n$. Suppose by contradiction that x is T-periodic, for some non-integer $T < n$. Then $x(T) = x(0) = 0$ and $x'(T) = x'(0) = 0$, while $x'(T) = \sqrt{|x(T)|} + h(T) = h(T) < 0$. Finally, we observe that homoclinic solutions of (1.12) can be trivially constructed as well.

REMARK 1.18. The solution set of (1.3) is the union of a family of mutually ordered connected components in $C^0([0, T])$. Yet, unlike the case where the property of uniqueness holds for (1.2) and hence nondegenerate continua are necessarily one-dimensional, such components may have arbitrarily large dimension (see [114]). We have seen that the dynamics of the trajectories of (1.1), which lie in the plane region filled by the graphs of the T-periodic solutions in a given connected component, may be quite complicated. Therefore any attempt to get a global qualitative portrait of solutions of (1.1) appears at least awkward, if solutions having graph in such regions are included in the analysis. Instead, it seems more meaningful to study only the dynamics of the solutions whose graph lies outside. In the light of Theorem 1.10, the qualitative behaviour of these solutions closely resembles the case where uniqueness holds. Then one could be naturally lead to look at the connected components of the set of solutions of (1.3) as single objects and to discuss the stability of each of them, by checking the stability from below of its minimum element and the stability from above of its maximum element, whenever they exist. Of course the consideration of a multivalued Poincaré operator would play a role here.

1.7. *Asymptotic behaviour of bounded solutions*

We already noticed in Section 1.4 that the existence of a solution u of (1.3) is always guaranteed in the presence of a bounded solution x of (1.1) defined on an unbounded interval. In this section we make this statement more precise, providing a complete extension of the Massera Convergence Theorem to the case where the uniqueness assumption for solutions of (1.2) is dropped. Namely, we see that either x converges to u, or the set of all solutions u of (1.3), whose range contains ω-limit points of x, forms a nondegenerate closed connected set in $C^0([0,T])$.

Lack of T-monotonicity and continua of T-periodic solutions

We start with a technical lemma saying that if a solution x of (1.1) and a solution v of (1.3) attain the same values at two points $a < b$, then there is a nondegenerate closed connected set of solutions u of (1.3) whose graphs completely fill the graph of x on $[a,b]$.

LEMMA 1.37. *Assume* (C). *Let v be a solution of* (1.3) *and $x:[a,b]\to\mathbb{R}$ a solution of* (1.1). *Suppose that $x(a)=v(a)$, $x(b)=v(b)$ and $x(t)>v(t)$ on $]a,b[$. Then the set $\mathcal{C}$ of all solutions $u\geqslant v$ of* (1.3) *such that there exists $t_0\in[a,b]$ with $u(t_0)=x(t_0)$ is a nondegenerate closed connected subset of $C^0([0,T])$. Furthermore, for each $t_0\in[a,b]$ there exists $u\in\mathcal{C}$ with $u(t_0)=x(t_0)$.*

PROOF. The set $\mathcal{C}$ is clearly closed and non-empty.

Step 1. Let us prove that $\mathcal{C}$ is dense-in-itself with respect to the order. Indeed, let $u_1,u_2\in\mathcal{C}$ be such that $u_1<u_2$. If $u_1(t_0)=u_2(t_0)$ for some t_0, then Lemma 1.34 yields the existence of a solution u of (1.3) with $u_1<u<u_2$. Clearly, $u\in\mathcal{C}$. If $u_1\ll u_2$ and no solution u of (1.3) exists with $u_1<u<u_2$, we set $\hat{x}=x$ on $[a,b]$ and $\hat{x}=v$ on $\mathbb{R}\setminus[a,b]$, $y=\min\{\max\{\hat{x},u_1\},u_2\}$. Then y is a nonextendible solution of (1.1) such that $u_1<y<u_2$ and $\lim_{t\to-\infty}(y(t)-u_1(t))=\lim_{t\to+\infty}(y(t)-u_1(t))=0$, contradicting the conclusions of Theorem 1.14. Hence there is a solution u of (1.3) such that $u_1<u<u_2$. Clearly, $u\in\mathcal{C}$.

Step 2. Let us prove that $\mathcal{C}$ is not a singleton. We distinguish two cases.

- Let $b-a\leqslant T$. Consider the function $w:[a,a+T]\to\mathbb{R}$ defined by $w(t)=x(t)$ on $[a,b]$ and $w(t)=v(t)$ on $[b,a+T]$, and extend w by T-periodicity onto $\mathbb{R}$. Then both v and w belong to $\mathcal{C}$.
- Let $b-a>T$. We show that v is not isolated from above in the set of all solutions of (1.3). Indeed, if v were the maximum solution of (1.3), then the function $\hat{x}:\mathbb{R}\to\mathbb{R}$ defined by $\hat{x}(t)=x(t)$ on $[a,b]$ and $\hat{x}(t)=v(t)$ on $\mathbb{R}\setminus[a,b]$ would be a solution of (1.1) contradicting the conclusions of Theorem 1.17. Whereas, if w were a solution of (1.3) with $v\ll w$ and such that there is no solution u of (1.3) with $v<u<w$, then the function $\hat{x}:\mathbb{R}\to\mathbb{R}$ defined by $\hat{x}(t)=\min\{x(t),w(t)\}$ on $[a,b]$ and $\hat{x}(t)=v(t)$ on $\mathbb{R}\setminus[a,b]$ would be a solution of (1.1) contradicting the conclusions of Lemma 1.14. Hence there exists a sequence $(u_n)_n$ of solutions of (1.3), with $u_n>v$ for every n, which converges uniformly to v. Then $u_n\in\mathcal{C}$ for all large n.

Step 3. Let us prove that $\mathcal{C}$ is connected. Pick $w\in\mathcal{C}$, with $w>v$. Set $\mathcal{K}=\{u\mid u$ is a solution of (1.3) with $v\leqslant u\leqslant w\}$. Notice that $\mathcal{K}\subseteq\mathcal{C}$, $\mathcal{K}$ is compact and, by Step 1, dense-in-itself with respect to the order. Let $\mathcal{T}\subseteq\mathcal{K}$ be a maximal totally ordered subset of $\mathcal{K}$ and notice that both v and w belong to $\mathcal{T}$. By Lemma 1.3, $\mathcal{T}$ is homeomorphic to a compact interval of $\mathbb{R}$. This shows that any pair $w_1, w_2 \in \mathcal{C}$ can be connected by a continuum (in fact an arc) containing v and contained in $\mathcal{C}$.

Step 4. Let us prove that for each $t_0 \in [a,b]$ there exists $u \in \mathcal{C}$ with $u(t_0) = x(t_0)$. We distinguish two cases again.

- If $b - a \leqslant T$, the conclusion follows arguing as in Step 2.
- If $b - a > T$, we prove that for each $t_0 \in [a,b]$ there exists $u \in \mathcal{C}$ with $u(t_0) \geqslant x(t_0)$.

As $\mathcal{C}$ is connected the conclusion will follow. Accordingly, let us suppose by contradiction there is a point $t_0 \in [a,b]$ such that $u(t_0) < x(t_0)$ for all $u \in \mathcal{C}$. Assume first that $\mathcal{C}$ is bounded. Then $\mathcal{C}$ is a continuum. Let $\mathcal{T}$ be a nondegenerate maximal totally ordered subset of $\mathcal{C}$. Since $\mathcal{C}$ is dense-in-itself, by Lemma 1.3, $\mathcal{T}$ is homeomorphic to a compact interval. Set $w = \max \mathcal{T}$. There are points a_1, b_1 such that $a \leqslant a_1 < t_0 < b_1 \leqslant b$, $x(a_1) = w(a_1)$, $x(b_1) = w(b_1)$ and $x(t) > w(t)$ on $]a_1, b_1[$. By Step 2 there exist a solution u of (1.3), with $u > w$, and a point $t_1 \in [a_1, b_1]$ such that $u(t_1) = x(t_1)$. Hence we conclude that $u \in \mathcal{C}$ and $u > w$, thus contradicting the maximality of w. Assume now that $\mathcal{C}$ is unbounded from above and let $u \in \mathcal{C}$ be such that $\max u > \max x$. Since $u(t_0) < x(t_0)$, there are points a_1, b_1, with $a_1 < t_0 < b_1 < a_1 + T$, such that $x(a_1) = u(a_1)$, $x(b_1) = u(b_1)$ and $x(t) > u(t)$ on $]a_1, b_1[$. Since $b_1 - a_1 < T$, a contradiction is obtained, arguing as in Step 2. □

Next, we specify Corollary 1.11 by showing that existence of solutions of (1.1) which are not T-monotone yields existence of nondegenerate continua of solutions of (1.3).

PROPOSITION 1.38. *Assume* (C). *Let* $x :]\omega_-, \omega_+[\to \mathbb{R}$ *be a nonextendible solution of* (1.1) *and suppose that* x *is not* T*-monotone. Then there exists a nondegenerate continuum* $\mathcal{K}$ *of solutions of* (1.3) *such that, for every* $u \in \mathcal{K}$, *there is* $t_0 \in]\omega_-, \omega_+[$ *with* $u(t_0) = x(t_0)$.

PROOF. We proceed in two steps.

Step 1. Assume either $\omega_- = -\infty$, or $\omega_+ = +\infty$. We know from the first two steps in the proof of Theorem 1.10 that there are points t_1, t_2, p_1, p_2, with $p_1 < p_2$, such that, for every $p \in]p_1, p_2[$, there are a lower solution α and an upper solution β of (1.3) such that $\alpha(t_1) = \beta(t_1) = x(t_1)$ and $\alpha(t_2) = \beta(t_2) = p$. Then Theorem 1.9 and Lemma 1.34 yield the existence of a nondegenerate continuum $\mathcal{K}$ satisfying the required conditions.

Step 2. Assume $\omega_+, \omega_- \in \mathbb{R}$, with $\omega_+ - \omega_- > T$. We distinguish two cases.

- Suppose that

$$\lim_{t \to \omega_-} x(t) = \lim_{t \to \omega_+} x(t) = -\infty;$$

the case where $\lim_{t \to \omega_-} x(t) = \lim_{t \to \omega_+} x(t) = +\infty$ being treated similarly. Since x is not T-monotone, Theorem 1.10 implies the existence of a solution v of (1.3) such that $x(t_0) > v(t_0)$ for some $t_0 \in]\omega_-, \omega_+[$. Hence we can find $a, b \in]\omega_-, \omega_+[$, with $a < b$, such that $x(a) = v(a)$, $x(b) = v(b)$ and $x(t) > v(t)$ on $]a, b[$. Lemma 1.37 yields the existence of a nondegenerate continuum $\mathcal{K}$ having the desired properties.

- Suppose that

$$\lim_{t \to \omega_-} x(t) = -\infty \quad \text{and} \quad \lim_{t \to \omega_+} x(t) = +\infty;$$

the case where $\lim_{t\to\omega_-} x(t) = +\infty$ and $\lim_{t\to\omega_+} x(t) = -\infty$ being treated similarly. Set $\varphi(t) = x(t+T) - x(t)$ on $]\omega_-, \omega_+ - T[$. We have $\lim_{t\to\omega_-} \varphi(t) = \lim_{t\to\omega_+ - T} \varphi(t) = +\infty$ and, as x is not T-monotone, $\varphi(t_0) = x(t_0 + T) - x(t_0) < 0$ for some $t_0 \in\,]\omega_-, \omega_+ - T[$. Define

$$t_1 = \min\bigl\{t \in\,]\omega_-, \omega_+ - T[\; \mid \varphi(t) = 0\bigr\},$$

$$t_2 = \max\bigl\{t \in\,]\omega_-, \omega_+ - T[\; \mid \varphi(t) = 0\bigr\}$$

and set

$$u_1 = x|_{[t_1, t_1+T]}, \quad u_2 = x|_{[t_2, t_2+T]}.$$

The T-periodic extensions to $\mathbb{R}$ of u_1, u_2 are solutions of (1.3).

If $u_1 = u_2$, we set $v = u_1 = u_2$. Observe that either $x(t_0) > v(t_0)$, or $x(t_0 + T) < v(t_0 + T)$. Suppose that $x(t_0) > v(t_0)$; the other case being treated similarly. Since $t_0 \in\,]t_1, t_2[$, we have that $\hat{x} = \max\{v, x|_{[t_1,t_2]}\}$ is a solution of (1.1) satisfying $\hat{x} > v$ on $[t_1, t_2]$, $\hat{x}(t_1) = v(t_1)$ and $\hat{x}(t_2) = v(t_2)$. Lemma 1.37 yields the existence of a nondegenerate continuum $\mathcal{K}$ satisfying the required conditions.

If $u_1 \neq u_2$ and $u_1(\bar{t}) = u_2(\bar{t})$ for some $\bar{t}$, then Lemma 1.34 implies the existence of a nondegenerate continuum $\mathcal{K}$ having the desired properties.

If $u_1 \ll u_2$, then $t_2 - t_1 > T$. The function $\hat{x} : \mathbb{R} \to \mathbb{R}$ defined by $\hat{x}(t) = u_1(t)$ on $]-\infty, t_1[$, $\hat{x}(t) = x(t)$ on $[t_1, t_2]$ and $\hat{x}(t) = u_2(t)$ on $]t_2, +\infty[$ is a solution of (1.1) satisfying $\hat{x}(t_1 + nT) = u_2(t_1 + nT) = u_2(t_1) > u_1(t_1) = \hat{x}(t_1)$, for all large $n \in \mathbb{N}$. Accordingly, $\hat{x}$ is not T-decreasing. On the other hand, $\hat{x}$ is not T-increasing. Otherwise $\varphi(t) \geqslant 0$ on $]\omega_-, \omega_+ - T[$ and hence x should be T-increasing as well, a contradiction. Therefore, $\hat{x}$ is not T-monotone and, by Step 1, we can find a nondegenerate continuum $\mathcal{K}$ satisfying the required conditions.

If $u_1 \gg u_2$, then there exist $a, b \in\,]\omega_-, \omega_+[$ such that $x(a) = u_2(a)$, $x(b) = u_2(b)$ and $x(t) > u_2(t)$ on $]a, b[$. Lemma 1.37 implies the existence of a nondegenerate continuum $\mathcal{K}$ having the desired properties. □

The Massera Convergence Theorem revisited

We are now in position of proving the announced extension of the Massera Convergence Theorem.

THEOREM 1.39. *Assume* (C). *Suppose that* $x :]t_0, +\infty[\to \mathbb{R}$ *is a bounded solution of* (1.1). *Then either*

- *there exists a solution u of* (1.3) *such that*

$$\lim_{t\to+\infty} \bigl|x(t) - u(t)\bigr| = 0, \tag{1.22}$$

or

- *the set* $\mathcal{C}$ *of all solutions* u *of* (1.3) *such that*

$$\liminf_{t\to+\infty}\left|x(t)-u(t)\right|=0 \tag{1.23}$$

is a nondegenerate closed connected subset of $C^0([0,T])$.

REMARK 1.19. It may happen that there exists a solution u of (1.3) such that x converges to u, nevertheless the set $\mathcal{C}$ is nondegenerate and even unbounded (see Example 1.2).

PROOF. Let us suppose that there is no solution u of (1.3) such that (1.22) holds. Then, by Proposition 1.7, x is not eventually T-monotone. We start by showing that $\mathcal{C}$ is not empty. Next we prove that $\mathcal{C}$ has at least two elements. Finally we show that $\mathcal{C}$ is closed and connected.

Step 1. *$\mathcal{C}$ is nonempty.* Let us denote by $\mathcal{S}_x$ the set of all solutions u of (1.3) for which there are unbounded sequences of real numbers $(a_n)_n$ and $(b_n)_n$ such that $u(a_n)<x(a_n)$ and $u(b_n)>x(b_n)$ for all n. Notice that $\mathcal{S}_x\subseteq\mathcal{C}$.

We claim that $\mathcal{S}_x$ is not empty. Assume by contradiction that $\mathcal{S}_x=\emptyset$. Since x is not eventually T-monotone, by Corollary 1.11 we can construct an increasing unbounded sequence $(a_n)_n$ of real numbers and a sequence $(u_n)_n$ of solutions of (1.3) such that either

$$u_n(a_n)<x(a_n)\quad\text{and}\quad u_n\geqslant x\quad\text{on } [a_{n+1},+\infty[,$$

or

$$u_n(a_n)>x(a_n)\quad\text{and}\quad u_n\leqslant x\quad\text{on } [a_{n+1},+\infty[,$$

for all n. We shall suppose that the former eventuality occurs. A similar argument can be used in the latter case. For each n, we set $\tilde{u}_n=\min\{u_1,u_2,\dots,u_n\}$. Notice that $\tilde{u}_n$ is a solution of (1.3), $\tilde{u}_n(a_n)<x(a_n)$ and $\tilde{u}_n\geqslant x$ on $[a_{n+1},+\infty[$. Since $\tilde{u}_n\geqslant\tilde{u}_{n+1}\geqslant\inf\{x(t)\mid t\in[t_0,+\infty[\}$ on $[0,T]$, the sequence $(\tilde{u}_n)_n$ converges uniformly to a solution u of (1.3), which satisfies $u(a_n)\leqslant\tilde{u}_n(a_n)<x(a_n)$ for all n.

Now, either $u\leqslant x$ on some interval $[b,+\infty[$ and hence $u\leqslant x\leqslant\tilde{u}_n$ on $[a_{n+1},+\infty[$ for all large n, or there is an increasing unbounded sequence of real numbers $(b_n)_n$ such that $u(b_n)>x(b_n)$. In the former case we conclude that (1.22) holds. In the latter case we conclude that $u\in\mathcal{S}_x$. In both cases we get a contradiction.

Step 2. *$\mathcal{C}$ is not a singleton.* Pick $u_0\in\mathcal{C}$. Since

$$0=\liminf_{t\to+\infty}\left|x(t)-u_0(t)\right|<\limsup_{t\to+\infty}\left|x(t)-u_0(t)\right|,$$

we can find a positive number ε and an increasing unbounded sequence of real numbers $(a_n)_n$ such that either $x(a_n)-u_0(a_n)>\varepsilon$ for all n or $u_0(a_n)-x(a_n)>\varepsilon$ for all n. We shall suppose that the former eventuality occurs. A similar argument can be used in the latter case.

Let us show that u_0 cannot be isolated from above in the set of all solutions u of (1.3). We proceed by contradiction. First assume that u_0 is the maximum solution of (1.3).

Let $y = \max\{u_0, x\}$. By Theorem 1.17 we should have either $\lim_{t\to+\infty} y(t) = +\infty$ or $\lim_{t\to+\infty}(y(t) - u_0(t)) = 0$; in both cases we get a contradiction. Next notice that, if there are a solution w of (1.3) and a point $\hat{t}$ such that $u_0 < w$ and $u_0(\hat{t}) = w(\hat{t})$, then, by Lemma 1.34, u_0 is not isolated from above. Finally suppose that there exists a solution w of (1.3) such that $u_0 \ll w$ and no solution u of (1.3) exists with $u_0 < u < w$. Set $y = \min\{\max\{x, u_0\}, w\}$. By Lemma 1.14 we should have either $\lim_{t\to+\infty}(y(t) - u_0(t)) = 0$ or $\lim_{t\to+\infty}(y(t) - w(t)) = 0$; in both cases we get again a contradiction.

Since u_0 is not isolated from above, we can pick a solution u of (1.3) with $u > u_0$ and $\|u - u_0\|_\infty < \varepsilon$. Then we have $u(a_n) < u_0(a_n) + \varepsilon < x(a_n)$ for all n. Since $\liminf_{t\to+\infty} |x(t) - u_0(t)| = 0$, we conclude that $u \in \mathcal{C}$ and therefore $\mathcal{C}$ has at least two elements.

Step 3. *$\mathcal{C}$ is closed and connected.* Since $\mathcal{C}$ is obviously closed, we only need to prove that it is connected. We start showing that $\mathcal{C}$ is dense-in-itself with respect to the order. Notice that if u is a solution of (1.3) and $u_1 < u < u_2$, then $u \in \mathcal{C}$. Assume first that $u_1(\hat{t}) = u_2(\hat{t})$ for some $\hat{t}$. Then, by Lemma 1.34, there exists a nondegenerate continuum of solutions u of (1.3) with $u_1 \leqslant u \leqslant u_2$. Assume next that $u_1 \ll u_2$ and that no solution u of (1.3) exists with $u_1 < u < u_2$. Set $y = \min\{\max\{x, u_1\}, u_2\}$. Then y is a nonextendible solution of (1.1) such that $u_1 < y < u_2$. By Lemma 1.14 we should have either $\lim_{t\to+\infty}(u_2(t) - y(t)) = 0$ or $\lim_{t\to+\infty}(y(t) - u_1(t)) = 0$, contradicting the fact that both u_1 and u_2 belong to $\mathcal{C}$.

Now, let $u_1, u_2 \in \mathcal{C}$ and assume $u_1 < u_2$. Let $\mathcal{T}$ be a nondegenerate maximal totally ordered subset of $\mathcal{C} \cap [u_1, u_2]$. By Lemma 1.3, $\mathcal{T}$ is an arc connecting u_1 and u_2. Accordingly, in order to prove that $\mathcal{C}$ is (arcwise) connected, it is enough to show that for every $u_1, u_2 \in \mathcal{C}$ there exists $u \in \mathcal{C}$ which is order-comparable with both u_1 and u_2. To achieve this aim, we set $v = \min\{u_1, u_2\}$ and $w = \max\{u_1, u_2\}$. We claim that either $v \in \mathcal{C}$ or $w \in \mathcal{C}$. Suppose $v \notin \mathcal{C}$. Then there exists $\varepsilon > 0$ and $\hat{t} \in \mathbb{R}$ such that $x(t) - v(t) > \varepsilon$ for all $t > \hat{t}$. Let us verify that $w \in \mathcal{C}$. Since $u_1 \in \mathcal{C}$, there exists an increasing unbounded sequence of real numbers $(a_n)_n$ with $|x(a_n) - u_1(a_n)| < \varepsilon$ for every n. If $\hat{n}$ is such that $a_{\hat{n}} > \hat{t}$, then $x(a_n) - v(a_n) > \varepsilon > |x(a_n) - u_1(a_n)|$ for all $n > \hat{n}$. Hence we deduce that $v(a_n) \neq u_1(a_n)$; therefore $u_1(a_n) = w(a_n)$ and $|x(a_n) - w(a_n)| < \varepsilon$ for all $n > \hat{n}$. ☐

The following statement conversely shows that each ω-limit point of a bounded solution defined on an unbounded interval must be in the range of a T-periodic solution.

PROPOSITION 1.40. *Assume* (C). *Suppose that* $x : [t_0, +\infty[\to \mathbb{R}$ *is a bounded solution of* (1.1). *Then, for each* $\hat{x}$, *with* $\liminf_{t\to+\infty} x(t) \leqslant \hat{x} \leqslant \limsup_{t\to+\infty} x(t)$, *and for each sequence* $(t_n)_n$, *with* $\lim_{n\to+\infty} t_n = +\infty$ *and* $\lim_{n\to+\infty} x(t_n) = \hat{x}$, *there exists a solution* u *of* (1.3) *such that* $\liminf_{n\to+\infty} |x(t_n) - u(t_n)| = 0$.

PROOF. If there exists a solution u of (1.3) such that (1.22) holds, the conclusion trivially follows. Hence we may assume that this is not the case. Let $\mathcal{C}$ be the set of all solutions u of (1.3) such that (1.23) holds. By Theorem 1.39, $\mathcal{C}$ is a nondegenerate closed connected set.

Step 1. Suppose that, for each n, $t_n = \hat{t} + k_n T$ with $\hat{t} \in [t_0, t_0 + T]$ and $k_n \in \mathbb{N}^+$. Let us show that there exists a function $u \in \mathcal{C}$ such that $u(\hat{t}) = \hat{x}$. We start by assuming, by contradiction, that $u(\hat{t}) < \hat{x}$ for all $u \in \mathcal{C}$.

Assume that $\mathcal{C}$ is bounded from above. Then, by Lemma 1.15, there exists a maximal element w of $\mathcal{C}$. We first observe that there is no $\bar{t} > t_0$ such that $x > w$ on $[\bar{t}, +\infty[$. Suppose not. Since $0 = \liminf_{t\to+\infty} |x(t) - w(t)| < \limsup_{t\to+\infty} |x(t) - w(t)|$, by Theorem 1.17 there exists a solution u of (1.3) such that $w < u$, and, by Lemma 1.14, w is not isolated from above in the set of all solutions u of (1.3). Hence there exists a sequence $(w_k)_k$ of solutions of (1.3), converging uniformly to w and such that $w_k > w$ for all k. Since $\limsup_{t\to+\infty} |x(t) - w(t)| > 0$, there exists k such that $w_k \in \mathcal{C}$, thus yielding a contradiction to the maximality of w. Nevertheless, since $w(\hat{t}) < \hat{x} = \lim_{n\to+\infty} x(t_n)$, there are $\varepsilon > 0$ and $\bar{n}$ such that $w(\hat{t}) = w(t_n) < \hat{x} - \varepsilon < x(t_n)$ for all $n \geqslant \bar{n}$. Hence there exist a sequence of positive integers $(n_j)_j$, with $n_0 \geqslant \bar{n}$, and sequences of real numbers $(a_j)_j$ and $(b_j)_j$, with $a_j < t_{n_j} < b_j$ such that $w(a_j) = x(a_j)$, $w(b_j) = x(b_j)$ and $x(t) > w(t)$ on $]a_j, b_j[$ for all j. By Lemma 1.37, there exists a solution u_0 of (1.3) such that $u_0 \geqslant w$ and $u_0(t_{n_0}) = \hat{x} - \varepsilon$. Notice that, for all j, we both have $u_0(t_{n_j}) = \hat{x} - \varepsilon < x(t_{n_j})$ and $x(a_j) = w(a_j) \leqslant u_0(a_j)$. Hence $u_0 \in \mathcal{C}$, a contradiction with the maximality of w.

Assume now that $\mathcal{C}$ is unbounded from above. Then there exists $w \in \mathcal{C}$ such that $\max w > \sup x$. Let $a \in]\hat{t} - T, \hat{t}[$ be such that $w(a) = \max w$. Since $w(\hat{t}) < \hat{x} = \lim_{n\to+\infty} x(t_n)$, there are an integer $\bar{n}$ and an increasing sequence $(p_n)_n$ of real numbers such that $\lim_{n\to+\infty} p_n = \hat{x}$ and, for all $n \geqslant \bar{n}$, $w(\hat{t}) = w(t_n) < p_n < \min\{x(t_n), \hat{x}\}$. For each $n \geqslant \bar{n}$, let $w_n \colon [a + k_n T, a + (k_n + 1)T] \to \mathbb{R}$ be the function defined by $w_n(t) = \max\{w(t), x(t)\}$. Notice that $w_n(a + k_n T) = w(a + k_n T) = w(a)$, $w_n(a + (k_n + 1)T) = w(a + (k_n + 1)T) = w(a)$ and $w_n > w$ on $[a + k_n T, a + (k_n + 1)T]$. Since $w(t_n) < p_n < w_n(t_n)$, by Lemma 1.37, there exists a solution v_n of (1.3) such that $v_n \geqslant w$ and $v_n(t_n) = p_n$. Let $u_n = \min\{v_n, w_n\}$. Then, for each $n \geqslant \bar{n}$, $u_n(a) = w(a)$ and $u_n(\hat{t}) = p_n$. Hence, for all $n \geqslant \bar{n}$, both $u_{\bar{n}}(a + k_n T) = \max w > x(a + k_n T)$ and $u_{\bar{n}}(t_n) = p_{\bar{n}} \leqslant p_n < x(t_n)$, and therefore $u_{\bar{n}} \in \mathcal{C}$. Furthermore, the sequence $(u_n)_n$ is bounded and hence there is a subsequence $(u_{n_j})_j$ converging uniformly to a function $u \in \mathcal{C}$. Since $u_{n_j}(\hat{t}) = p_{n_j}$, we get $u(\hat{t}) = \hat{x}$, a contradiction.

In a similar way we can show that the assumption that $u(\hat{t}) > \hat{x}$ for all $u \in \mathcal{C}$ leads to a contradiction. Hence either there exists $u \in \mathcal{C}$ such that $u(\hat{t}) = \hat{x}$, or there exist $u_1, u_2 \in \mathcal{C}$ with $u_1(\hat{t}) \leqslant \hat{x} \leqslant u_2(\hat{t})$. In the latter case the connectedness of $\mathcal{C}$ yields anyhow the existence of a function $u \in \mathcal{C}$ such that $u(\hat{t}) = \hat{x}$.

Step 2. Let us show that we can always reduce ourselves to the situation considered in Step 1. Namely, for each n let k_n be the integer such that $t_n \in [t_0 + k_n T, t_0 + (k_n + 1)T[$ and set $\hat{t}_n = t_n - k_n T$. Let $\hat{t} \in [t_0, t_0 + T]$ be a cluster point of the sequence $(\hat{t}_n)_n$ and pick a subsequence $(\hat{t}_{n_j})_j$ converging to $\hat{t}$. Set $s_j = \hat{t} + k_{n_j} T$. Notice that $\lim_{j\to+\infty} |t_{n_j} - s_j| = \lim_{j\to+\infty} |\hat{t}_{n_j} - \hat{t}| = 0$. Since the function x is uniformly continuous on $[t_0, +\infty[$, we have $\lim_{j\to+\infty} |x(t_{n_j}) - x(s_j)| = 0$ and therefore $\lim_{j\to+\infty} x(s_j) = \hat{x}$. By Step 1 there exists a solution u of (1.3) such that $\liminf_{j\to+\infty} |x(s_j) - u(s_j)| = 0$. Since $|x(t_{n_j}) - u(t_{n_j})| \leqslant |x(t_{n_j}) - x(s_j)| + |x(s_j) - u(s_j)| + |u(s_j) - u(t_{n_j})|$, $\lim_{j\to+\infty} |x(t_{n_j}) - x(s_j)| = 0$ and $\lim_{j\to+\infty} |u(s_j) - u(t_{n_j})| = \lim_{j\to+\infty} |u(\hat{t}) - u(\hat{t}_{n_j})| = 0$, we conclude that $\liminf_{n\to+\infty} |x(t_n) - u(t_n)| = 0$. □

REMARK 1.20. In the light of Proposition 1.38, Proposition 1.7 and Theorem 1.39, one might guess that, if $x :]t_0, +\infty[\to \mathbb{R}$ is a bounded solution of (1.1), which is not eventually T-monotone, then the set $\mathcal{C}$ of all solutions u of (1.3), such that $\liminf_{t\to+\infty} |u(t) - x(t)| =$

0, is always a nondegenerate connected set. Yet, as the next example shows, this is not actually true.

EXAMPLE 1.5. We construct a function f, which is 1-periodic with respect to the t-variable, such that equation (1.1) has a bounded solution x that is not eventually 1-monotone, but the set $\mathcal{C}$ of all solutions u of (1.3), such that $\liminf_{t\to+\infty} |u(t) - x(t)| = 0$, is a singleton $\mathcal{C} = \{u\}$ and actually $\lim_{t\to+\infty}(u(t) - x(t)) = 0$. Our construction is based on a modification of Example 1.4. Let $h : \mathbb{R} \to \mathbb{R}$ be the 1-periodic function defined by $h(t) = -\frac{1}{8}|t|$ on $[-\frac{1}{2}, \frac{1}{2}[$. Let us define f by setting, for every t, $f(t, x) = \sqrt{|x|} + h(t)$, if $x \leqslant -\frac{1}{8}$, $f(t, x) = \sqrt{|x - 2 + \frac{1}{2^{n-1}}|} + \frac{1}{2^n}h(t)$, if $n \geqslant 0$ and $x \in [2 - \frac{17}{2^{3+n}}, 2 - \frac{15}{2^{3+n}}]$, $f(t, x) = 0$, if $x \geqslant 2$. Then we extend f onto $\mathbb{R}^2$ by linear interpolation with respect to x. We can see that Eq. (1.1) has a sequence $(\mathcal{K}_n)_n$ of continua of 1-periodic solutions, which approximate the equilibrium $w \equiv 2$. Each continuum $\mathcal{K}_n$ is a "squeezing" of the continuum $\mathcal{K}$ produced in Example 1.4. Denote by v_n and w_n respectively the minimum and the maximum of $\mathcal{K}_n$. One can verify that for each n, between w_n and v_{n+1}, all solutions are heteroclinics that leave w_n and reach v_{n+1} in a finite time. Indeed, for any $m_n \in \mathbb{Z}$ we can find a solution y_n of (1.1) which pulls ahead of w_n at m_n and reaches the value $2 - \frac{15}{2^{3+n}}$ in a finite time. Since f is positive there, y_n enters a region where f is defined linearly. There y_n is still increasing until it reaches the value $2 - \frac{17}{2^{4+n}}$, where f is still positive. Then y_n enters a region where all solutions reach v_{n+1} in a finite time.

Let us now outline a possible construction of the solution x of (1.1) having the required properties. Start taking x on [0, 1] to be any $x_1 \in \mathcal{K}_1$. Then there is a solution y_1 of (1.1), which pulls ahead of w_1 at $m_1 = 1$ and reaches v_2 in a finite time. On this interval x is y_1. Afterwards, we take x to be for a while an almost periodic solution lying between v_2 and w_2, careful enough to destroy 1-monotonicity. Then we pull ahead of w_2 at some m_2 and we define x to be a function of the form of y_2. Proceeding in this way we get our desired solution x.

A careful reading of our construction actually shows that x can be defined so that it is not T-monotone for any fixed $T > 0$.

2. Second Order Periodic Parabolic PDEs

2.1. *Introduction*

In this part we consider the parabolic boundary value problem

$$\begin{cases} \partial_t u + A(x, t, \partial_x)u = f(x, t, u, \nabla_x u) & \text{in } \Omega \times I, \\ u = 0 & \text{on } \partial\Omega \times I. \end{cases} \tag{2.1}$$

Here Ω is a bounded smooth domain in $\mathbb{R}^N$, with $N \geqslant 1$, I is a real interval, $\partial_t + A(x, t, \partial_x)$ is a linear second order uniformly parabolic operator. The coefficients of A are T-periodic in t, $T > 0$ being a given period. The function $f : \Omega \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is T-periodic in t and satisfies the L^p-Carathéodory conditions, for some $p > N + 2$, as well as a Nagumo condition, that is, $f(x, t, s, \xi)$ grows at most quadratically with respect to ξ. According to these assumptions, it is natural to consider solutions of (2.1) intended in the strong, i.e. $W_p^{2,1}$-, sense. In this work for simplicity we only deal with homogeneous Dirichlet data in (2.1), however more general boundary conditions, such as Robin and mixed ones, could be considered as well (see [42]).

Nonlinear diffusion equations such as (2.1) typically occur in the mathematical description of various biological, chemical, physical and economical phenomena and they have particular interest in population ecology models where the data periodically depend on time due to seasonal fluctuations (see, e.g., [49,58,155,25,115,103]).

We are mainly interested here in studying existence, localization and qualitative properties of the T-periodic solutions of (2.1), assuming that a lower solution α and an upper solution β of the T-periodic problem associated with (2.1) are given.

If we suppose that the L^p-Carathéodory function f satisfies further regularity assumptions, e.g., f is locally Lipschitz continuous with respect to s and ξ, then the initial value problem associated with (2.1) is well-posed, i.e. it has a unique solution which depends continuously on the given data; moreover, the regularity of f, in conjunction with the parabolic maximum principle, yields the validity of comparison principles of Nagumo-Westphal type (see [151,106]). In this frame the Poincaré operator associated with (2.1), which assigns to every initial datum the value of the solution after one period, is well-defined and strictly order preserving; so that the above mentioned questions can be tackled by the theory of order preserving discrete-time semidynamical systems as in [58,60]. Our perspective in this work is different. Here we assume that f satisfies only the L^p-Carathéodory conditions: in this case evidently such theory does not apply, due to the failure both of uniqueness for the initial value problem and of comparison principles. Nevertheless, we shall see that a careful direct analysis of the partial differential equation (2.1) allows to reinterpret, or to adapt to a more general setting, or to extend the validity of several results which are known within the frame of order preserving dynamical systems, always keeping minimal regularity assumptions.

The study of the existence of periodic solutions of certain special forms of (2.1) seems to have been initiated in the late thirties and continued in the forties by D.H. Karimov [64–67]. In the early fifties, G. Prodi [125–127] introduced lower and upper solutions and a Nagumo-type condition for studying the solvability of the periodic problem associated with (2.1). The stability of periodic solutions was also discussed by G. Prodi in [124], using the concept of Lyapunov stability as extended to parabolic equations by R. Bellmann [17]. In the fifties and the sixties, further progress was obtained, concerning both existence and stability (see, e.g., [107,137,75,23,69,48,12,138,70,78]). These authors approached the problem by various methods and techniques, but still only special cases of (2.1) were considered. A general answer to the question of the existence of a T-periodic solution, when the lower and the upper solutions satisfy

$$\alpha(x,t) \leqslant \beta(x,t) \quad \text{in } \Omega \times \left]0,T\right[, \tag{2.2}$$

was eventually given by Ju.S. Kolesov [72,73] in the late sixties. Namely, under additional regularity assumptions on the coefficients and the nonlinearity, he proved the existence of a classical T-periodic solution u of (2.1) such that

$$\alpha(x,t) \leqslant u(x,t) \leqslant \beta(x,t) \quad \text{in } \Omega \times \left]0,T\right[. \tag{2.3}$$

This result was improved by H. Amann [6,8] in the seventies. In particular, he proved some basic estimates in connection with the above mentioned Nagumo condition, established the existence of a minimum and of a maximum T-periodic solution between α and β and also allowed Robin boundary conditions in (2.1).

The method of Ju.S. Kolesov and H. Amann consists in finding periodic solutions of (2.1) as fixed points of the Poincaré operator associated with (2.1), whose existence is guaranteed by the smoothness assumptions on f. Further, it was observed apparently for the first time in [8] that the Poincaré operator has the important property of being strongly order preserving with respect to the order induced by the positive cone of $C_0^1(\overline{\Omega})$. In [8] and other related works, this property follows from the validity of a strong comparison principle of the Nagumo–Westphal type for (2.1), which is in turn a consequence of the linear parabolic maximum principle and, again, of the smoothness of f.

These ideas are useful for solving both theoretically and numerically the periodic problem associated with (2.1), but they also revealed to be quite fruitful for investigating the stability properties of the periodic solutions of (2.1). Indeed, strongly influenced by some works of H. Matano [92,93] who successfully integrated monotonicity methods with dynamical systems techniques for the study of autonomous parabolic initial value problems (see also M.W. Hirsch [59] and H.L. Smith [136]), E.N. Dancer and P. Hess [33] were able to provide, remaining in the same frame of [8], an answer to the question of the stability of periodic solutions of (2.1). Namely, they proved that, if α and β are strict lower and upper solutions, for which (2.2) holds, then there exists at least one T-periodic solution u satisfying (2.3) which is stable. Although preliminary results in this direction can be found, e.g., in [71,72,134,92], the novelty of the approach of E.N. Dancer and P. Hess lies in the introduction of a strongly order preserving discrete-time semidynamical system, which allows to detect many qualitative properties of the initial value problems associated with (2.1). In [33] this discrete-time semidynamical system is determined by the solution operator corresponding to the T-periodic problem related to (2.1), although the smoothness assumptions imposed on (2.1) would allow to consider as well the discrete-time semidynamical system induced by the Poincaré operator. In their successive paper [34], as well as in the monograph by P. Hess [58], the approach based on the Poincaré operator was developed in an abstract setting and systematically employed to describe the dynamics of (2.1). Thus, several stability, instability and stabilization results for the T-periodic solutions of (2.1) were obtained. In particular, the existence of heteroclinic orbits, connecting periodic solutions, was established, extending previous results, such as those in [93]. In [32] E.N. Dancer carried on these studies further, investigating the relations between fixed point index, one-sided stability and existence of lower and upper solutions close to a periodic solution, as well as the possibility of weakening the order preserving requirements on the semiflow. Further progress in the theory of discrete-time order preserving semidynamical systems has been obtained in the last ten years; we refer to the very recent works [60,123,61] for updated surveys on this theory, as well as on its applications to parabolic boundary value problems of the form of (2.1).

We notice that all the above cited papers deal with classical solutions of (2.1); whereas, the existence and localization of weak T-periodic solutions in the presence of lower and upper solutions were discussed in [44], even for more general quasilinear equations. In this frame, the existence of a minimum and a maximum solution was established in [26] and their one-sided stability in [46]. To the best of our knowledge, strong solutions have not been explicitly considered in the literature in connection with lower and upper solutions, although they naturally occur in the Carathéodory frame.

So far we have considered situations dealing with pairs of lower and upper solutions satisfying (2.2). Now, we turn to discuss the complementary case where there exist a lower solution α and an upper solution β for which condition (2.2) fails, that is,

$$\alpha(x_0, t_0) > \beta(x_0, t_0) \quad \text{for some } (x_0, t_0) \in \Omega \times \,]0, T[. \tag{2.4}$$

There are concrete motivations to consider this situation. Indeed, condition (2.2) turns out to be restrictive in some situations. This was remarked by J. Kazdan and F. Warner [68] for elliptic equations, but their observation immediately extends to periodic parabolic problems. Let us denote by λ_1 the principal eigenvalue of the linear T-periodic problem associated with (2.1). Assume for example that the function f in (2.1) is independent of $\nabla_x u$. If f satisfies

$$\operatorname*{ess\,sup}_{\Omega\times]0,T[\times\mathbb{R}} \partial_s f(x, t, s) < \lambda_1, \tag{2.5}$$

then one can construct lower and upper solutions α, β satisfying (2.2). On the other hand, if f satisfies

$$\operatorname*{ess\,inf}_{\Omega\times]0,T[\times\mathbb{R}} \partial_s f(x, t, s) > \lambda_1 \tag{2.6}$$

then, for any pair of lower and upper solutions α, β satisfying (2.2), one easily verifies that $\alpha = \beta$ and they must already be solutions. Hence any result involving such lower and upper solutions is of no use here. Yet, in this situation lower and upper solutions satisfying (2.4) occur quite naturally. Indeed, one can show that condition (2.6) implies the existence of pairs of lower and upper solutions α, β such that

$$\alpha(x, t) > \beta(x, t) \quad \text{in } \Omega \times \,]0, T[. \tag{2.7}$$

Moreover, the principle of linearized instability suggests that, if a T-periodic solution of (2.1) existed in the presence of lower and upper solutions satisfying (2.4), it should be presumably unstable, unlike what has been established when (2.2) holds. This was conjectured in the frame of elliptic boundary value problems by D.H. Sattinger [134], who was the first to pose these questions in the early seventies. Afterwards, it was pointed out in [7] that the mere existence of a lower solution and an upper solution for which (2.4) holds is generally not sufficient to guarantee the solvability. Adapted to our setting, the example in [7] is of the following type:

$$\begin{cases} \partial_t u - \Delta u = \lambda_m u + \varphi_m & \text{in } \Omega \times \,]0, T[, \\ u = 0 & \text{on } \partial\Omega \times [0, T], \\ u(\cdot, 0) = u(\cdot, T) & \text{in } \Omega, \end{cases} \tag{2.8}$$

where λ_m is an eigenvalue of $-\Delta$ on $H_0^1(\Omega)$ greater than λ_1 and φ_m is a corresponding nonzero eigenfunction. Of course, problem (2.8) has no solution, although one can construct a lower solution α and an upper solution β satisfying (2.4). Accordingly, one is led to conclude that, to achieve solvability, some further conditions must be assumed, in order to prevent the interference of f with the higher part of the spectrum. This can be expressed in various ways. In [9], where the first important contribution was given in the frame of elliptic equations, H. Amann, A. Ambrosetti and G. Mancini proved that the existence of a lower solution α and of an upper solution β, not necessarily satisfying any ordering condition, implies solvability, provided f is a bounded perturbation of $\lambda_1 s$. Further results, mainly in the direction of weakening this last condition, as well as of obtaining information on the localization of the solution, were later proved, always for elliptic equations, in [54,55,40]. In [41] the existence result of [40] was extended to the periodic problem for (2.1) and, apparently for the first time, the study of the stability properties was faced in this context. Indeed, assuming the existence of strict lower and upper solutions satisfying (2.4) and some growth restrictions on f, it was proved in [41] the existence of a solution u of (2.1), which is unstable and satisfies, for some $(x_1, t_1), (x_2, t_2) \in \Omega \times]0, T[$,

$$u(x_1, t_1) < \alpha(x_1, t_1) \quad \text{and} \quad u(x_2, t_2) > \beta(x_2, t_2). \tag{2.9}$$

We stress that this result is the direct counterpart of the one in [33] when condition (2.2) is replaced by (2.4), because it yields the existence of a solution, information about its localization in terms of α and β and its instability. We further notice that in [41] the analysis was entirely performed in the frame of strong solutions of (2.1) without requiring, unlike [33] and the other quoted results, any further regularity on f in addition to the Carathéodory conditions.

It should be observed that in several cases the principles of linearized stability, or the construction of Lyapunov functionals, can be successfully used to detect the stability of periodic solutions (see, e.g., [37,154,123]). On the other hand, the use of lower and upper solutions for discussing the stability properties of a periodic solution appears appropriate in some critical cases where such methods fail, in particular when the periodic solution under consideration is degenerate, or the nonlinearity f does not satisfy the necessary differentiability requirements, and obviously in all situations where the existence of a Lyapunov functional is not known. This remark also provides a strong motivation for studying the applicability of the lower and upper solution method to detect stability, or instability, and to investigate other qualitative properties, assuming the least of regularity on f. In this work we pursue this program.

The remainder of this introduction is devoted to describe with some detail the plan of this part of our work.

In Section 2.2 we collect some definitions and preliminary statements, having a basic or technical nature, which will be extensively used in the sequel. Most of these are variations or extensions of known ones. Nevertheless, we state them in a precise form, providing even proofs, whenever we are not able to supply adequate references. We introduce in particular a rather general notion of lower and upper solutions for problem (2.1), which appear appropriate in the frame of strong solutions. Namely, a lower solution is defined, locally in time, as the pointwise maximum of a finite number of $W_p^{2,1}$-lower solutions. Similarly an upper solution is defined, locally in time, as the pointwise minimum of a finite number of $W_p^{2,1}$-upper solutions. In this way we allow "angles" in the space variables and "jumps" in the time variable. This level of generality is motivated by the fact that "angles" and "jumps" naturally occur in practical and theoretical situations even in the frame of smooth solutions: typically when considering minima or maxima of solutions, which are not solutions due to the lack of lattice structure, or studying solutions of the initial value problem lying between lower and upper solutions of the periodic problem. After reviewing a parabolic maximum principle for strong solutions, some notions of orderings, order norms and other topics, we further prove in Proposition 2.3 an elementary fact, which, roughly speaking, says that any Carathéodory function satisfies a local regularity condition, which is the natural generalization of the Lipschitz or Hölder continuity and will allow us to compare problem (2.1) with other ones possessing a certain extent of monotonicity. This simple observation has also some interesting connections with the use of a monotone iterative scheme, namely the Chaplyghin method (see [100,13,8]), for the construction of the minimum and the maximum solution of (2.1), lying between a lower and an upper solution satisfying (2.2). However, for brevity, we do not discuss this topic here and we refer to [42].

In Section 2.3 we describe some basic facts concerning the initial value problem associated with (2.1). Namely, we discuss existence, localization and structure of the set of strong solutions of (2.1) in the presence of a pair of possibly discontinuous lower and upper solutions satisfying (2.2). A crucial step of this section is the Nagumo-type result stated in Proposition 2.5 (a different version suited for the periodic problem will be proved later in Proposition 2.10): these statements basically show that suitable $W_p^{2-2/p}$-, or L^∞-, bounds on the initial values force $W_p^{2,1}$-bounds on the solutions of (2.1), provided f satisfies a Nagumo condition. We reproduce in both cases their rather delicate proofs in detail, even if these follow the argument in [6], because some specifications are needed in order to extend their validity to our setting. Then we prove the main results of this section: Theorem 2.6, dealing with compact cylinders, and Corollary 2.7, dealing with noncompact cylinders. They extend to strong solutions and possibly discontinuous lower and upper solutions some existence and localization results for the initial value problem associated with (2.1) which are well known in the frame of classical solutions. We also describe the topological structure of the solution set, proving the Hukuhara–Kneser property. Since our analysis is performed just assuming that f satisfies the Carathéodory conditions, we get an improvement of previous statements obtained in [76], for f Hölder continuous, and in [81], for f continuous. Our proof uses an abstract theorem in [77,153], but differs from the standard ones in the approximation method, as we adapt here an idea introduced in [139] dealing with ordinary differential equations. As a byproduct we get, by a simple proof, a basically known result concerning local existence, continuation and ultimate behaviour of solutions of (2.1).

Section 2.4 is devoted to some existence and localization statements for the periodic problem associated with (2.1), when a pair α, β of possibly discontinuous lower and upper solutions is given. After some preliminaries, mainly devoted to state the above mentioned Nagumo-type result for the periodic problem, we introduce in the frame of (2.1) the notion of T-monotonicity. This property is a simple tool to prove convergence of a bounded solution defined in $\overline{\Omega} \times [t_0, +\infty[$ to a T-periodic solution. We use it to get an extension to the present context of the Monotone Convergence Criterion known in the frame of order preserving discrete-time semidynamical systems (see, e.g., [60, Section 5]). The version of this criterion we give in Proposition 2.13 says that any bounded solution of (2.1) emanating from and lying above a lower solution of the T-periodic problem associated with (2.1), or respectively emanating from and lying below an upper solution, is T-monotone and hence converges to a T-periodic solution. We point out that, in space dimension $N \geqslant 2$, convergence to a T-periodic solution is not the behaviour of all bounded solutions of (2.1), due to the possible occurrence of subharmonic solutions (see, e.g., [141,35] and [123, Section 3]). In our frame additional difficulties come from the lack of regularity of f, so that even generic convergence results fail (see [123, Section 3]). We then give an application of this result to the existence of strong periodic solutions of the linear problem associated with (2.1). Although this topic was recently considered in [84], we show here that a simple approach, based on a monotone iterative scheme, yields a constructive existence proof. This technique is very much in the spirit of this work, since the iteration here introduced is used several times elsewhere in the paper. Additional statements concerning the principal eigenvalue λ_1 of the linear problem associated with (2.1), as well as some Fredholm-type results, are produced in Appendix.

After these premises, which however have an independent interest, we prove the main theorems of this section: the existence and localization of solutions of the periodic problem associated with (2.1) when $\alpha \leqslant \beta$, or when $\alpha \nleqslant \beta$. Theorem 2.15 yields the existence of a minimum and a maximum T-periodic solution of (2.1) between two possibly discontinuous lower and upper solutions α, β satisfying the condition $\alpha \leqslant \beta$. This extends to strong solutions previous results obtained, e.g., in [73,8,6,33,58] for classical solutions. The proof of Theorem 2.15 makes use of a quite standard argument, substantially similar to that used for proving Theorem 2.6. Nevertheless, as its counterpart for the initial value problem, the modification and truncation method we use are new in this frame, since some specific devices are required in order to deal with nonregular lower and upper solutions, as well as with Carathéodory gradient dependent nonlinearities, for handling which some care is required. We also show the existence of a minimum and of a maximum solution using a neat argument based on compactness. For later use, relations with the topological degree are also established, when the lower and upper solutions are strict.

Afterwards we discuss the existence of T-periodic solutions of (2.1) in the presence of generalized lower and upper solutions α, β such that $\alpha \nleqslant \beta$. As we already pointed out, in this frame the sole existence of a pair of lower and upper solutions, even satisfying $\alpha > \beta$, does not generally guarantee the existence of a solution. Here we prove in Theorem 2.17 a result which is related to the classical Amann–Kolesov Three Solutions Theorem [73,5,7] and requires the existence of a further pair of lower and upper solutions α_1, β_1, with $\alpha_1 \leqslant \beta_1$, bracketing α, β. Theorem 2.17 then yields the existence of maximal and minimal T-periodic solutions of (2.1) lying between α_1, β_1 and belonging to the $C^{1,0}$-closure of the set of functions satisfying (2.9). The solvability is established evaluating, by means of the additivity and excision properties, the topological degree of a solution operator associated with (2.1), whose fixed points are precisely the periodic solutions of (2.1). The existence of minimal and maximal solutions is proved using some a-priori bounds in connection with Zorn Lemma. For related results we refer to [41,42], where alternative conditions are used and several applications are indicated.

The aim of Section 2.5 is to extend to our general context the Order Interval Trichotomy, first proved in [34] for order preserving discrete-time semidynamical systems. We stress once more that our results are obtained without requiring any regularity on the function f besides the Carathéodory conditions. Nevertheless with the aid of Proposition 2.3, which plays a crucial role in this context, we can make deformations of our original problem to other problems displaying a certain amount of monotonicity, from which we eventually infer the information needed to prove the existence of some type of connections. Namely, if we fix an ordered pair of T-periodic solutions, say u_1, u_2, then we prove the existence either of a further T-periodic solution between u_1 and u_2, or of a double sequence of lower, or upper, solutions connecting u_1 and u_2. The last assertion expresses a form of stability, or instability, of u_1 and u_2, which is referred to as order stability. Once this version of the Order Interval Trichotomy is established, we apply it to the study of the dynamics near one-sided isolated T-periodic solutions of (2.1). Namely we prove the existence of T-monotone heteroclinic solutions connecting a pair of comparable T-periodic solutions and we study the qualitative behaviour of certain solutions of (2.1) lying above a maximal, or below a minimal, T-periodic solution. Our statements extend previous results in [34,58, 32] (see also [60,123]).

In Section 2.6 we study the stability properties of the T-periodic solutions of (2.1) with the aid of lower and upper solutions. We start by recalling the definition of one-sided Lyapunov–Bellmann stability. Instead of using the more usual C^1-, or L^∞-, norm for measuring the distance between solutions, we use the order norm induced by a fixed function belonging to the interior of the positive cone in $C_0^1(\overline{\Omega})$. This seems natural in order to fit better with the features of (2.1) and with our approach. Since we have to take care of the fact that uniqueness for the initial value problem and validity of comparison principles are not assumed, the notion of Lyapunov stability may be not the most appropriate to be considered here; indeed, some weaker concept might be more suited to detect certain residual forms of stability. As an alternative definition to Lyapunov stability we consider that of order stability; this is commonly used in the frame of order preserving semidynamical systems [93,33,58,60] and appears suited to our approach based on lower and upper solutions. We perform a rather detailed examination of the relationship occurring between the notions of Lyapunov and order stability. In particular we discuss the meaning of these notions of stability, when some versions of the comparison principle hold for solutions of (2.1), implying or not uniqueness for the initial value problem. Explicit conditions on the nonlinearity f, entailing their validity, are also exhibited. We show in particular that Lyapunov stability implies order stability, whereas the converse implication is not generally true; however, these concepts are equivalent if a comparison principle holds. Using these notions, we give a precise description of the stability properties of a T-periodic solution in terms of the existence of a lower or an upper solution close to it. Hence, when a pair of lower and upper solutions α, β, with $\alpha \leqslant \beta$, is given, we discuss the stability of the minimum and the maximum T-periodic solutions v and w lying in between, thus getting a completion of Theorem 2.15. Namely, as a consequence of Proposition 2.13 we prove, in Theorem 2.33, an invariance property of the set of T-periodic solutions which lie between a lower and an upper solution satisfying (2.2) and, in particular, a form of relative attractivity of the minimum and the maximum solutions. Afterwards, we discuss the structure of the set of T-periodic solutions lying between v and w, as well as the dynamics occurring between them. In particular, we prove the existence of a totally ordered continuum of order stable solutions. Assuming the validity of some comparison principle, we obtain refinements and strengthenings of the results we have established in a more general context. In particular we can recover several known results, about stability, asymptotic stability, instability of periodic solutions, using an approach which singles out the essential hypotheses. Finally, we describe with some detail the instability properties of maximal T-periodic solutions which are not above a strict lower solution, or not below a strict upper solution. As a consequence, in the case where $\alpha \not\leqslant \beta$, we get an instability result which completes the conclusions of Theorem 2.17. The last subsection is devoted to a discussion of some relations between stability properties and multiplicity of T-periodic solutions.

2.2. *Preliminaries*

In this section we state some basic facts concerning the parabolic problem

$$\begin{cases} \partial_t u + A(x,t,\partial_x)u = f(x,t,u,\nabla_x u) & \text{in } \Omega \times I, \\ u = 0 & \text{on } \partial\Omega \times I, \end{cases} \tag{2.10}$$

where $I \subseteq \mathbb{R}$ is an interval.

Basic assumptions

We suppose that

(D) $\Omega \subset \mathbb{R}^N$ is a bounded domain, having a boundary $\partial\Omega$ of class C^2, and $T > 0$ is a fixed number.

Hereafter, we set $Q_T = \Omega \times]0,T[$ and $\Sigma_T = \partial\Omega \times [0,T]$. The differential operator has the form

$$A(x,t,\partial_x)u = -\sum_{i,j=1}^{N} \partial_{x_i}\big(a_{ij}(x,t)\partial_{x_j}u\big) + \sum_{i=1}^{N} a_i(x,t)\partial_{x_i}u + a_0(x,t)u.$$

We assume that

(A) $a_{ij} \in C^{1,0}(\overline{Q}_T)$, $a_{ij}(x,t) = a_{ji}(x,t)$ in $\overline{Q}_T$, $a_{ij}(x,0) = a_{ij}(x,T)$ in $\overline{\Omega}$, for $i,j = 1,\dots,N$, and there exists $\eta > 0$ such that, for all $(x,t) \in \overline{Q}_T$ and $\xi \in \mathbb{R}^N$,

$$\sum_{i,j=1}^{N} a_{ij}(x,t)\xi_i\xi_j \geqslant \eta|\xi|^2;$$

$a_i \in L^\infty(Q_T)$ for $i = 0,\dots,N$; $\operatorname{ess\,inf}_{Q_T} a_0 > 0$.

We further suppose that

(C) $f : Q_T \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ satisfies the L^p-Carathéodory conditions for some $p > N+2$, i.e. for every $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$, $f(\cdot,\cdot,s,\xi)$ is measurable on Q_T; for a.e. $(x,t) \in Q_T$, $f(x,t,\cdot,\cdot)$ is continuous on $\mathbb{R} \times \mathbb{R}^N$; for each $\rho > 0$, there exists $\gamma \in L^p(Q_T)$ such that $|f(x,t,s,\xi)| \leqslant \gamma(x,t)$, for a.e. $(x,t) \in Q_T$ and every $(s,\xi) \in [-\rho,\rho] \times [-\rho,\rho]^N$;

and

(N) for every $M > 0$, there exist $h \in L^p(Q_T)$, with $p > N+2$, and $K > 0$ such that for a.e. $(x,t) \in Q_T$, every $s \in [-M, M]$ and every $\xi \in \mathbb{R}^N$,

$$\left|f(x,t,s,\xi)\right| \leqslant h(x,t) + K|\xi|^2. \tag{2.11}$$

REMARK 2.1. All functions considered in (A) and (C) are identified with their T-periodic extensions onto $\Omega \times \mathbb{R}$ and $\Omega \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N$, respectively.

Solutions and lower and upper solutions

We notice that, assuming (D) and $p > N+2$, the Sobolev space $W_p^{2,1}(\Omega \times]t_1, t_2[)$, with $t_1, t_2 \in \mathbb{R}$ and $t_1 < t_2$, is (compactly) embedded into $C^{1,0}(\overline{\Omega} \times [t_1, t_2])$ (see [80, Chapter II, Section 3]). Accordingly, the following definitions make sense.

DEFINITION 2.1.

- A *solution of* (2.10) *in* $\overline{\Omega} \times [t_1, t_2]$, with $t_1, t_2 \in \mathbb{R}$ and $t_1 < t_2$, is a function $u \in W_p^{2,1}(\Omega \times]t_1, t_2[)$ such that

$$\begin{aligned} &\partial_t u + A(x,t,\partial_x)u = f(x,t,u,\nabla_x u) && \text{a.e. in } \Omega \times]t_1, t_2[, \\ &u(x,t) = 0 && \text{on } \partial\Omega \times [t_1, t_2]. \end{aligned}$$

- A *solution of* (2.10) *in* $\overline{\Omega} \times J$, where J is a noncompact interval having endpoints t_1, t_2 with $-\infty \leqslant t_1 < t_2 \leqslant +\infty$, is a function u such that, for every compact interval $K \subset J$, $u|_{\overline{\Omega} \times K}$ is a solution of (2.10) in $\overline{\Omega} \times K$.

REMARK 2.2. In the literature this kind of solutions of (2.10) are usually referred to as *strong* solutions.

DEFINITION 2.2. A solution $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$ of (2.10) is said *nonextendible* (*to the right of* ω), if either $\omega = +\infty$, or $\omega < +\infty$ and there is no solution $\hat{u}$ of (2.10) in $\overline{\Omega} \times [t_0, \omega]$ such that $\hat{u}|_{\overline{\Omega} \times [t_0, \omega[} = u$.

We now introduce two notions of lower and upper solutions for (2.10) of increasing generality, allowing to the utmost "angles" in the space variables and "jumps" in the time variable.

NOTATION 2.3. For a function $u : \overline{\Omega} \times [t_1, t_2] \to \mathbb{R}$, we write $u|_{\overline{\Omega} \times]t_1, t_2[} \in W_p^{2,1}(\Omega \times]t_1, t_2[)$ if there exists $\widetilde{u} \in W_p^{2,1}(\Omega \times]t_1, t_2[) \cap C^{1,0}(\overline{\Omega} \times [t_1, t_2])$ such that $\widetilde{u} = u$ on $\overline{\Omega} \times]t_1, t_2[$.

DEFINITION 2.4.

- A *regular lower solution of* (2.10) *in* $\overline{\Omega} \times [t_1, t_2]$, with $t_1, t_2 \in \mathbb{R}$ and $t_1 < t_2$, is a function $\alpha : \bar{\Omega} \times [t_1, t_2] \to \mathbb{R}$, such that $\alpha|_{\overline{\Omega} \times]t_1, t_2[} \in W_p^{2,1}(\Omega \times]t_1, t_2[)$,

$$\begin{aligned} &\partial_t \alpha + A(x,t,\partial_x)\alpha \leqslant f(x,t,\alpha,\nabla_x \alpha) && \text{a.e. in } \Omega \times]t_1, t_2[, \\ &\alpha(x,t) \leqslant 0 && \text{on } \partial\Omega \times [t_1, t_2], \end{aligned}$$

and, for every $x \in \overline{\Omega}$,

$$\alpha(x,t_1) \geqslant \lim_{t \to t_1^+} \alpha(x,t), \qquad \alpha(x,t_2) \leqslant \lim_{t \to t_2^-} \alpha(x,t).$$

- A *lower solution of* (2.10) *in* $\overline{\Omega} \times [t_1,t_2]$ is a function $\alpha : \overline{\Omega} \times [t_1,t_2] \to \mathbb{R}$ for which there are points $t_1 = \sigma_0 < \sigma_1 < \cdots < \sigma_h < \sigma_{h+1} < \cdots < \sigma_k = t_2$, such that, for each $h \in \{0,\dots,k-1\}$, $\alpha|_{\bar{\Omega}\times[\sigma_h,\sigma_{h+1}]} = \max_{1 \leqslant i \leqslant m} \alpha_i(x,t)$ in $\overline{\Omega} \times [\sigma_h,\sigma_{h+1}]$, where, for each $i = 1,\dots,m$, α_i is a regular lower solution of (2.10) in $\overline{\Omega} \times [\sigma_h,\sigma_{h+1}]$.
- A *lower solution* (respectively a *regular lower solution*) *of* (2.10) *in* $\overline{\Omega} \times J$, where J is a noncompact interval having endpoints t_1, t_2 with $-\infty \leqslant t_1 < t_2 \leqslant +\infty$, is a function α such that, for every compact interval $K \subset J$, $\alpha|_{\overline{\Omega}\times K}$ is a lower solution (respectively a regular lower solution) of (2.10) in $\overline{\Omega} \times K$.
- A *regular upper solution of* (2.10) *in* $\overline{\Omega} \times [t_1,t_2]$, with $t_1, t_2 \in \mathbb{R}$ and $t_1 < t_2$, is a function $\beta : \overline{\Omega} \times [t_1,t_2] \to \mathbb{R}$ such that $\beta|_{\overline{\Omega}\times]t_1,t_2[} \in W_p^{2,1}(\Omega \times]t_1,t_2[)$,

$$\begin{array}{ll} \partial_t\beta + A(x,t,\partial_x)\beta \geqslant f(x,t,\beta,\nabla_x\beta) & \text{a.e. in } \Omega \times]t_1,t_2[, \\ \beta(x,t) \geqslant 0 & \text{on } \partial\Omega \times [t_1,t_2], \end{array}$$

and, for every $x \in \bar{\Omega}$,

$$\beta(x,t_1) \leqslant \lim_{t \to t_1^+} \beta(x,t), \qquad \beta(x,t_2) \geqslant \lim_{t \to t_2^-} \beta(x,t).$$

- An *upper solution of* (2.10) *in* $\overline{\Omega} \times [t_1,t_2]$ is a function $\beta : \overline{\Omega} \times [t_1,t_2] \to \mathbb{R}$ for which there are points $t_1 = \rho_0 < \rho_1 < \cdots < \rho_h < \rho_{h+1} < \cdots < \rho_l = t_2$, such that, for each $h \in \{0,\dots,l-1\}$, $\beta|_{\overline{\Omega}\times[\rho_h,\rho_{h+1}]} = \min_{1 \leqslant j \leqslant n} \beta_j(x,t)$ in $\overline{\Omega} \times [\rho_h,\rho_{h+1}]$, where, for each $j = 1,\dots,n$, β_j is a regular upper solution of (2.10) in $\bar{\Omega} \times [\rho_h,\rho_{h+1}]$.
- An *upper solution* (respectively a *regular upper solution*) *of* (2.10) *in* $\overline{\Omega} \times J$, where J is a noncompact interval having endpoints t_1, t_2 such that $-\infty \leqslant t_1 < t_2 \leqslant +\infty$, is a function β such that, for every compact interval $K \subset J$, $\beta|_{\overline{\Omega}\times K}$ is an upper solution (respectively a regular upper solution) of (2.10) in $\overline{\Omega} \times K$.

REMARK 2.3. We notice that a regular lower solution α is continuous in $\bar{\Omega} \times]t_1,t_2[$, but it may be discontinuous with respect to t at the endpoints of the interval $[t_1,t_2]$. A similar observation holds for a regular upper solution β.

REMARK 2.4. Whenever a lower solution and an upper solution are simultaneously considered, we can assume, without loss of generality, that the sequences $(\sigma_h)_h$ and $(\rho_h)_h$ coincide.

Orderings

DEFINITION 2.5. Assume (D). Given functions $u, v : \bar{\Omega} \times I \to \mathbb{R}$, with $I \subseteq \mathbb{R}$ an interval, we write

- $u \leqslant v$ if $u(x,t) \leqslant v(x,t)$ in $\bar{\Omega} \times I$;
- $u < v$ if $u \leqslant v$ and $u \neq v$;
- $u \ll v$ if there exists a function $w \in C^{1,0}(\overline{\Omega} \times I)$, with $w(x,t) > 0$ for every $(x,t) \in \Omega \times I$, $w(x,t) = 0$ on $\partial\Omega \times I$ and $\frac{\partial w}{\partial \nu}(x,t) < 0$ on $\partial\Omega \times I$, where $\nu = \nu(x) \in \mathbb{R}^N$ is the unit outer normal to Ω at $x \in \partial\Omega$, such that $u + w \leqslant v$.

DEFINITION 2.6. Given functions $v, w : \overline{\Omega} \times I \to \mathbb{R}$, with $I \subseteq \mathbb{R}$ an interval and $v < w$, we define the *order intervals*

$$[v, w] = \{u : \overline{\Omega} \times I \to \mathbb{R} \mid v \leqslant u \leqslant w\},$$

$$[v, +\infty[= \{u : \overline{\Omega} \times I \to \mathbb{R} \mid v \leqslant u\},$$

$$]-\infty, w] = \{u : \overline{\Omega} \times I \to \mathbb{R} \mid u \leqslant w\}.$$

All these notations, with the same meaning, will be also used for functions defined on $\partial\Omega \times I$, or $\overline{\Omega}$, or $\partial\Omega$.

DEFINITION 2.7. Let $\mathcal{S}$ be a given set of solutions of (2.10).

- We say that a solution z of (2.10), with $z \in \mathcal{S}$, is a *maximal solution of* (2.10) *in* $\mathcal{S}$ (respectively a *minimal solution of* (2.10) *in* $\mathcal{S}$) if there is no solution u of (2.10), with $u \in \mathcal{S}$, such that $u > z$ (respectively $u < z$).
- We say that a solution z of (2.10), with $z \in \mathcal{S}$, is the *maximum solution of* (2.10) *in* $\mathcal{S}$ (respectively the *minimum solution of* (2.10) *in* $\mathcal{S}$) if every solution u of (2.10), with $u \in \mathcal{S}$, is such that $u \leqslant z$ (respectively $u \geqslant z$).

The following elementary statement is related to the properties of the positive cone in $C^{1,0}(\bar{\Omega} \times [t_1, t_2])$ and can be proved repeating almost verbatim the argument of [55, Lemma 3.1].

PROPOSITION 2.1. *Assume* (D) *and let* $t_1, t_2 \in \mathbb{R}$, *with* $t_1 \leqslant t_2$, *be given. Let* $u \in C^{1,0}(\bar{\Omega} \times [t_1, t_2])$ *be such that* $u = 0$ *on* $\partial\Omega \times [t_1, t_2]$ *and* $u \gg 0$. *Then the following conclusions hold*:

(i) *for every* $c \in]0, 1[$, *there exists* $\delta > 0$ *such that, for any* $v \in C^{1,0}(\bar{\Omega} \times [t_1, t_2])$ *with* $v \geqslant 0$ *on* $\partial\Omega \times [t_1, t_2]$, *if* $\|u - v\|_{C^{1,0}(\bar{\Omega} \times [t_1, t_2])} < \delta$ *then* $v \geqslant cu$ *in* $\bar{\Omega} \times [t_1, t_2]$;

(ii) *there exists* $\delta > 0$ *such that, for any* $v \in C^{1,0}(\bar{\Omega} \times [t_1, t_2])$ *with* $v \leqslant 0$ *on* $\partial\Omega \times [t_1, t_2]$, *if* $\|v\|_{C^{1,0}(\bar{\Omega} \times [t_1, t_2])} < \delta$ *then* $v \leqslant u$ *in* $\bar{\Omega} \times [t_1, t_2]$.

REMARK 2.5. Conclusion (ii) of Proposition 2.1 is still valid if $u \geqslant 0$ on $\partial\Omega \times [t_1, t_2]$.

Order norm

Following [7] we introduce the notion of order norm.

DEFINITION 2.8. Assume (D) and let $t_1, t_2 \in \mathbb{R}$, with $t_1 \leqslant t_2$, be given. Fix a function $e \in C^1(\bar{\Omega})$ such that $e(x) > 0$ in Ω, $e(x) = 0$ and $\partial_\nu e(x) < 0$ on $\partial\Omega$, where $\nu = \nu(x)$ is the unit outer normal to Ω at $x \in \partial\Omega$. For any function $u : \overline{\Omega} \times [t_1, t_2] \to \mathbb{R}$, satisfying $u = 0$ on $\partial\Omega \times [t_1, t_2]$, we set

$$\|u\|_e = \inf\bigl\{\lambda > 0 \mid |u| \leqslant \lambda e \text{ in } \overline{\Omega} \times [t_1, t_2]\bigr\} \ (\in [0, +\infty]).$$

REMARK 2.6. There are constants $c_1, c_2 > 0$ such that, for every $u \in C^{1,0}(\overline{\Omega} \times [t_1, t_2])$ satisfying $u = 0$ on $\partial\Omega \times [t_1, t_2]$, $c_1 \|u\|_{L^\infty(\Omega \times]t_1, t_2[)} \leqslant \|u\|_e \leqslant c_2 \|u\|_{C^{1,0}(\overline{\Omega} \times [t_1, t_2])}$.

In the remainder of this section we collect some basic results that will be systematically used in the sequel.

Parabolic maximum principle

The following is a version of the parabolic strong maximum principle, expressed both in the interior and in the boundary forms (see [37] and also [45,83,128,145]).

PROPOSITION 2.2. *Assume* (D) *and* (A). *Let* $t_1, t_2 \in \mathbb{R}$, *with* $t_1 < t_2$, *be fixed and let* $q \in L^\infty(\Omega \times]t_1, t_2[)$ *be such that*

$$a_0(x, t) + q(x, t) \geqslant 0 \quad \textit{a.e. in } \Omega \times]t_1, t_2[.$$

Assume that $u \in W_p^{2,1}(\Omega \times]t_1, t_2[)$, *with* $p > N + 2$, *satisfy*

$$\partial_t u + A(x, t, \partial_x)u + q(x, t)u \geqslant 0 \quad \textit{a.e. in } \Omega \times]t_1, t_2[.$$

Moreover, set $m = \min_{\overline{\Omega} \times [t_1, t_2]} u$ *and suppose that* $m \leqslant 0$. *Then the following conclusions hold*:

(i) *if* m *is attained at* $(x_0, t_0) \in \Omega \times]t_1, t_2]$, *then* $u = m$ *in* $\bar{\Omega} \times [t_1, t_0]$;

(ii) *if* m *is attained at* $(x_0, t_0) \in \partial\Omega \times]t_1, t_2]$ *and* u *is not constant in* $\Omega \times]t_1, t_0]$, *then* $\partial_\mu u(x_0, t_0) < 0$, *where* $\mu \in \mathbb{R}^N$ *is such that* $\mu \cdot \nu > 0$, $\nu = \nu(x_0)$ *being the unit outer normal to* Ω *at* $x_0 \in \partial\Omega$.

REMARK 2.7. In the special case where $m = 0$, Proposition 2.2 remains valid without the assumption $a_0(x, t) + q(x, t) \geqslant 0$ a.e. in $\Omega \times]t_1, t_2[$. Indeed, in that case, as $u \geqslant 0$, we have a.e. in $\Omega \times]t_1, t_2[$

$$\partial_t u + A(x, t, \partial_x)u - a_0(x, t)u + (a_0 + q)^+(x, t)u \geqslant (a_0 + q)^-(x, t)u \geqslant 0$$

and we can apply the previous result.

A generalized Hölder-type continuity

The following result points out a property, shared by all Carathéodory functions, which can be seen as a generalization of the usual Hölder continuity. In spite of its quite elementary character, it will turn out crucial in the sequel.

PROPOSITION 2.3. *Assume that, for some* $p \geqslant 1$, $f : \Omega \times]t_1, t_2[\times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ *satisfies the* L^p*-Carathéodory conditions. Then for each* $\rho > 0$, *there exists a* L^p*-Carathéodory function* $h : \Omega \times]t_1, t_2[\times [-\rho, \rho] \times [-\rho, \rho] \times \mathbb{R}^N \to \mathbb{R}$ *such that*

- (i) *for a.e.* $(x, t) \in \Omega \times]t_1, t_2[$ *and every* $(r, \xi) \in [-\rho, \rho] \times \mathbb{R}^N$, $h(x, t, \cdot, r, \xi) : [-\rho, \rho] \to \mathbb{R}$ *is strictly increasing*;
- (ii) *for a.e.* $(x, t) \in \Omega \times]t_1, t_2[$ *and every* $(s, \xi) \in [-\rho, \rho] \times \mathbb{R}^N$, $h(x, t, s, \cdot, \xi) : [-\rho, \rho] \to \mathbb{R}$ *is strictly decreasing*;
- (iii) *for a.e.* $(x, t) \in \Omega \times]t_1, t_2[$ *and every* $(r, s, \xi) \in [-\rho, \rho] \times [-\rho, \rho] \times \mathbb{R}^N$, $h(x, t, s, r, \xi) = -h(x, t, r, s, \xi)$;
- (iv) *for a.e.* $(x, t) \in \Omega \times]t_1, t_2[$ *and every* $(r, s, \xi) \in [-\rho, \rho] \times [-\rho, \rho] \times \mathbb{R}^N$, *with* $r < s$,

$$\left| f(x, t, s, \xi) - f(x, t, r, \xi) \right| < h(x, t, s, r, \xi);$$

- (v) *if, moreover, there exist* $K > 0$ *and* $k \in L^p(\Omega \times]t_1, t_2[)$ *such that for a.e.* $(x, t) \in \Omega \times]t_1, t_2[$, *every* $s \in [-\rho, \rho]$ *and every* $\xi \in \mathbb{R}^N$,

$$\left| f(x, t, s, \xi) \right| \leqslant k(x, t) + K|\xi|^2,$$

then, for a.e. $(x, t) \in \Omega \times]t_1, t_2[$, *every* $(r, s) \in [-\rho, \rho] \times [-\rho, \rho]$ *and every* $\xi \in \mathbb{R}^N$,

$$\left| h(x, t, s, r, \xi) \right| \leqslant 2\left(k(x, t) + \rho + K|\xi|^2 \right).$$

PROOF. The function h will be constructed by using the modulus of continuity ω of f. We start by showing that ω is a continuous function.

Step 1. *Let* $g : [a, b] \to \mathbb{R}$ *be a continuous function and let* $\omega : [0, +\infty[\to [0, +\infty[$ *be the modulus of continuity of* g, *defined by setting, for every* $\delta \geqslant 0$,

$$\omega(\delta) = \max_{|x-y| \leqslant \delta} \left| g(x) - g(y) \right|.$$

Then ω *is a continuous function on* $[0, +\infty[$. Clearly, ω is increasing on $[0, +\infty[$ and is continuous at 0 with $\omega(0) = 0$. Let us prove that ω is continuous at any point $\delta_0 > 0$. Fix such a $\delta_0 > 0$. By monotonicity, we have $\lim_{\delta \to \delta_0^-} \omega(\delta) \leqslant \omega(\delta_0) \leqslant \lim_{\delta \to \delta_0^+} \omega(\delta)$. To show that $\lim_{\delta \to \delta_0^+} \omega(\delta) \leqslant \omega(\delta_0)$, pick a decreasing sequence $(\delta_n)_n$ such that $\delta_n \to \delta_0$. For each n, let $x_n, y_n \in [a, b]$ be such that $|x_n - y_n| \leqslant \delta_n$ and $|g(x_n) - g(y_n)| = \omega(\delta_n)$. Possibly passing to subsequences, we may assume $x_n \to \bar{x}$ and $y_n \to \bar{y}$, so that

$$\omega(\delta_n) = \left| g(x_n) - g(y_n) \right| \to \left| g(\bar{x}) - g(\bar{y}) \right| \leqslant \omega(\delta_0)$$

and therefore $\lim_{\delta \to \delta_0^+} \omega(\delta) = \lim_{n \to +\infty} \omega(\delta_n) \leqslant \omega(\delta_0)$. To show that $\lim_{\delta \to \delta_0^-} \omega(\delta) \geqslant \omega(\delta_0)$, let $x_0, y_0 \in [a, b]$ be such that $|x_0 - y_0| \leqslant \delta_0$ and $|g(x_0) - g(y_0)| = \omega(\delta_0)$. Pick an increasing sequence $(\delta_n)_n$ such that $\delta_n \to \delta_0$. By convexity, there exist sequences $(x_n)_n$ and $(y_n)_n$ such that $x_n \to x_0$, $y_n \to y_0$ and $|x_n - y_n| \leqslant \delta_n$, for each n. Hence it follows that

$$\omega(\delta_0) = \bigl|g(x_0) - g(y_0)\bigr| = \lim_{n\to+\infty} \bigl|g(x_n) - g(y_n)\bigr| \leqslant \lim_{n\to+\infty} \omega(\delta_n) = \lim_{\delta\to\delta_0^-} \omega(\delta).$$

Step 2. *Let* $g : [a, b] \times \mathbb{R}^N \to \mathbb{R}$ *be a continuous function and let* $\omega : [0, +\infty[\times \mathbb{R}^N \to [0, +\infty[$ *be the modulus of continuity of* g *with respect to the first variable, defined by setting, for each* $\delta \geqslant 0$ *and* $\xi \in \mathbb{R}^N$,

$$\omega(\delta, \xi) = \max_{|x-y|\leqslant\delta} \bigl|g(x, \xi) - g(y, \xi)\bigr|.$$

Then ω *is a continuous function in* $[0, +\infty[\times \mathbb{R}^N$. Assume, by contradiction, that ω is not continuous at $(\delta_0, \xi_0) \in [0, +\infty[\times \mathbb{R}^N$. Then there exist $\varepsilon > 0$ and, for each $\eta > 0$, a point $(\delta_\eta, \xi_\eta) \in [0, +\infty[\times \mathbb{R}^N$ such that $|\delta_0 - \delta_\eta| + |\xi_0 - \xi_\eta| < \eta$ and $|\omega(\delta_0, \xi_0) - \omega(\delta_\eta, \xi_\eta)| > \varepsilon$. By Step 1, there is $\eta_1 > 0$ such that, if $|\delta_0 - \delta| < \eta_1$, then

$$\bigl|\omega(\delta_0, \xi_0) - \omega(\delta, \xi_0)\bigr| < \varepsilon/2. \tag{2.12}$$

By the continuity of g, there is $\eta_2 > 0$ such that, if $|\xi_0 - \xi| < \eta_2$, then

$$\bigl|g(x, \xi) - g(x, \xi_0)\bigr| < \varepsilon/6$$

for all $x \in [a, b]$. Pick $\eta = \min\{\eta_1, \eta_2\}$ and the corresponding point (δ_η, ξ_η) such that $|\delta_0 - \delta_\eta| + |\xi_0 - \xi_\eta| < \eta$, $|\omega(\delta_0, \xi_0) - \omega(\delta_\eta, \xi_\eta)| > \varepsilon$ and, using (2.12),

$$\bigl|\omega(\delta_\eta, \xi_0) - \omega(\delta_\eta, \xi_\eta)\bigr| > \varepsilon/2.$$

Assume that $\omega(\delta_\eta, \xi_0) > \omega(\delta_\eta, \xi_\eta) + \varepsilon/2$, the other case being similar. Take $x_0, y_0 \in [a, b]$ such that $|x_0 - y_0| \leqslant \delta_\eta$ and

$$\omega(\delta_\eta, \xi_0) = \bigl|g(x_0, \xi_0) - g(y_0, \xi_0)\bigr|.$$

Take further $x_\eta, y_\eta \in [a, b]$ such that $|x_\eta - y_\eta| \leqslant \delta_\eta$ and

$$\omega(\delta_\eta, \xi_\eta) = \bigl|g(x_\eta, \xi_\eta) - g(y_\eta, \xi_\eta)\bigr|.$$

Then we get the contradiction

$$\begin{aligned}
\omega(\delta_\eta, \xi_\eta) &\geqslant \bigl|g(x_0, \xi_\eta) - g(y_0, \xi_\eta)\bigr| \\
&\geqslant \bigl|g(x_0, \xi_0) - g(y_0, \xi_0)\bigr| - \bigl|g(x_0, \xi_0) - g(x_0, \xi_\eta)\bigr| \\
&\quad - \bigl|g(y_0, \xi_0) - g(y_0, \xi_\eta)\bigr| \\
&> \omega(\delta_\eta, \xi_0) - \varepsilon/3 > \omega(\delta_\eta, \xi_\eta) + \varepsilon/6.
\end{aligned}$$

Step 3. *Definition of* h. Fix $\rho > 0$ and set, for a.e. $(x,t) \in \Omega \times]t_1, t_2[$, every $\delta \geqslant 0$ and every $\xi \in \mathbb{R}^N$,

$$\omega(x,t,\delta,\xi) = \max\bigl\{\bigl|f(x,t,r,\xi) - f(x,t,s,\xi)\bigr| \mid |r| \leqslant \rho, |s| \leqslant \rho, |r-s| \leqslant \delta\bigr\}.$$

We have, for a.e. $(x,t) \in \Omega \times]t_1, t_2[$, every $r, s \in [-\rho, \rho]$ with $r \leqslant s$ and $\xi \in \mathbb{R}^N$,

$$\bigl|f(x,t,r,\xi) - f(x,t,s,\xi)\bigr| \leqslant \omega(x,t,s-r,\xi).$$

Let us define for a.e. $(x,t) \in \Omega \times]t_1, t_2[$ and every $r, s \in [-\rho, \rho]$, $\xi \in \mathbb{R}^N$,

$$h(x,t,s,r,\xi) = \operatorname{sgn}(s-r)\bigl(\omega\bigl(x,t,|s-r|,\xi\bigr) + |s-r|\bigr).$$

Then the following conclusions hold:

- for a.e. $(x,t) \in \Omega \times]t_1, t_2[$, $h(x,t,\cdot,\cdot,\cdot)$ is continuous;
- for every $(r,s,\xi) \in [-\rho,\rho] \times [-\rho,\rho] \times \mathbb{R}^N$, $h(\cdot,\cdot,s,r,\xi)$ is measurable (this is due to the fact that $\omega(x,t,\delta,\xi) = \sup\{|f(x,t,r,\xi) - f(x,t,s,\xi)| \mid |r-s| \leqslant \delta$, with $r, s \in \mathbb{Q}\}$);
- for each $\sigma > 0$, there exists $\gamma \in L^p(\Omega \times]t_1, t_2[)$ such that, for a.e. $(x,t) \in \Omega \times]t_1, t_2[$ and every $(r,s,\xi) \in [-\rho,\rho] \times [-\rho,\rho] \times [-\sigma,\sigma]^N$, $|h(x,t,s,r,\xi)| \leqslant \gamma(x,t)$;
- for a.e. $(x,t) \in \Omega \times]t_1, t_2[$ and every $(r,\xi) \in [-\rho,\rho] \times \mathbb{R}^N$, $h(x,t,\cdot,r,\xi) : [-\rho,\rho] \to \mathbb{R}$ is strictly increasing;
- for a.e. $(x,t) \in \Omega \times]t_1, t_2[$ and every $(s,\xi) \in [-\rho,\rho] \times \mathbb{R}^N$, $h(x,t,s,\cdot,\xi) : [-\rho,\rho] \to \mathbb{R}$ is strictly decreasing;
- for a.e. $(x,t) \in \Omega \times]t_1, t_2[$ and every $(r,s,\xi) \in [-\rho,\rho] \times [-\rho,\rho] \times \mathbb{R}^N$, $h(x,t,s,r,\xi) = -h(x,t,r,s,\xi)$;
- for a.e. $(x,t) \in \Omega \times]t_1, t_2[$ and every $(r,s,\xi) \in [-\rho,\rho] \times [-\rho,\rho] \times \mathbb{R}^N$, with $r < s$,

  $$\bigl|f(x,t,s,\xi) - f(x,t,r,\xi)\bigr| < h(x,t,s,r,\xi);$$

- if, moreover, there exist $K > 0$ and $k \in L^p(\Omega \times]t_1, t_2[)$ such that for a.e. $(x,t) \in \Omega \times]t_1, t_2[$, every $s \in [-\rho,\rho]$ and every $\xi \in \mathbb{R}^N$,

  $$\bigl|f(x,t,s,\xi)\bigr| \leqslant k(x,t) + K|\xi|^2,$$

  then, for a.e. $(x,t) \in \Omega \times]t_1, t_2[$, every $(r,s) \in [-\rho,\rho] \times [-\rho,\rho]$ and every $\xi \in \mathbb{R}^N$,

  $$\bigl|h(x,t,s,r,\xi)\bigr| \leqslant 2\bigl(k(x,t) + \rho + K|\xi|^2\bigr). \qquad \Box$$

2.3. *The initial boundary value problem*

Let us consider the parabolic initial value problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f(x,t,u,\nabla_x u) & \text{in } \Omega \times I, \\ u = 0 & \text{on } \partial\Omega \times I, \\ u(\cdot, t_0) = u_0 & \text{in } \Omega, \end{array} \tag{2.13}$$

where $I \subseteq \mathbb{R}$ is an interval containing a right neighbourhood of t_0 and $u_0 \in W_p^{2-2/p}(\Omega)$ is such that $u_0 = 0$ on $\partial\Omega$.

In this section we collect some basic facts concerning problem (2.13). Namely, we discuss existence, localization and structure of the set of strong solutions of (2.13) in the presence of a pair of possibly nonregular lower and upper solutions.

Solutions and lower and upper solutions

We notice that, assuming (D) and $p > N + 2$, $W_p^{2-2/p}(\Omega)$ is (compactly) embedded in $C^1(\overline{\Omega})$ (see [1, Chapter VII]). Therefore the following definitions make sense.

DEFINITION 2.9. A *solution of* (2.13) in $\overline{\Omega} \times I$, with $I \subseteq \mathbb{R}$ an interval containing a right neighbourhood of t_0, is a solution u of (2.10) in $\overline{\Omega} \times I$ such that

$$u(x, t_0) = u_0(x) \quad \text{in } \overline{\Omega}.$$

The following notion of lower and upper solutions for the initial value problem (2.13) is used.

DEFINITION 2.10.

- A *lower solution* (respectively a *regular lower solution*) *of* (2.13) *in* $\overline{\Omega} \times I$, with $I \subseteq \mathbb{R}$ an interval containing a right neighbourhood of t_0, is a function α which is a lower solution (respectively a regular lower solution) of (2.10) in $\overline{\Omega} \times I$ and satisfies

  $$\alpha(x, t_0) \leqslant u_0(x) \quad \text{in } \Omega.$$

- An *upper solution* (respectively a *regular upper solution*) *of* (2.13) *in* $\overline{\Omega} \times I$, with $I \subseteq \mathbb{R}$ an interval containing a right neighbourhood of t_0, is a function β which is an upper solution (respectively a regular upper solution) of (2.10) in $\overline{\Omega} \times I$ and satisfies

  $$\beta(x, t_0) \geqslant u_0(x) \quad \text{in } \Omega.$$

The linear initial value problem

The next statement, which follows from [80, Chapter IV, Section 9 and Chapter VII, Section 10], regards existence, uniqueness and regularity of solutions of the linear initial value problem

$$\begin{array}{ll} \partial_t u + A(x, t, \partial_x)u = h(x, t) & \text{in } \Omega \times]t_0, \tau[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, \tau], \\ u(\cdot, t_0) = u_0 & \text{in } \Omega. \end{array} \tag{2.14}$$

PROPOSITION 2.4. *Assume* (D) *and* (A). *Fix* $p > N + 2$ *and let* $t_0 \in \mathbb{R}$ *and* $\tau \in]t_0, +\infty[$ *be given. Then for every* $h \in L^p(\Omega \times]t_0, \tau[)$ *and every* $u_0 \in W_p^{2-2/p}(\Omega)$ *with* $u_0 = 0$ *on* $\partial\Omega$, *problem* (2.14) *has a unique solution* $u \in W_p^{2,1}(\Omega \times]t_0, \tau[)$. *Moreover, there exists an increasing function* $C(\tau)$, *which is independent of* h *and* u_0, *such that*

$$\|u\|_{W_p^{2,1}(\Omega \times]t_0, \tau[)} \leqslant C(\tau)\big(\|h\|_{L^p(\Omega \times]t_0, \tau[)} + \|u_0\|_{W_p^{2-2/p}(\Omega)}\big).$$

A Nagumo-type result

The following result will be systematically used in the sequel.

PROPOSITION 2.5. *Assume* (D) *and* (A). *Let* $p > N + 2$, $t_0 \in \mathbb{R}$, $\omega \in \,]t_0, +\infty[$, $h \in L^p(\Omega \times \,]t_0, \omega[)$, $K > 0$ *and* $R > 0$ *be given. Then there exists a constant* $C > 0$ *such that for each* $u : \bar{\Omega} \times [t_0, \omega[\,\to \mathbb{R}$, *with* $u \in W_p^{2,1}(\Omega \times \,]t_0, \tau[)$ *for every* $\tau < \omega$, *satisfying*

$$\begin{array}{ll} \left|\partial_t u + A(x, t, \partial_x)u\right| \leqslant h(x, t) + K|\nabla_x u|^2 & \textit{a.e. in } \Omega \times \,]t_0, \omega[, \\ u = 0 & \textit{on } \partial\Omega \times [t_0, \omega[\end{array}$$

and

$$\left\|u(\cdot, t_0)\right\|_{W_p^{2-2/p}(\Omega)} \leqslant R,$$

we have $u \in W_p^{2,1}(\Omega \times \,]t_0, \omega[)$ *and*

$$\|u\|_{W_p^{2,1}(\Omega \times]t_0, \omega[)} \leqslant C.$$

PROOF. We start with a preliminary result, whose proof closely follows an argument in [6, Lemma 2.1] (see also [149]). We present here a detailed proof in order to verify that the constant γ_0, appearing below, is independent of $\tau \in \,]\delta, \omega[$, whenever $\delta \in \,]t_0, \omega[$ is fixed. Hereafter, for any $\sigma \in \,]t_0, \omega]$, we set $Q_\sigma = \Omega \times \,]t_0, \sigma[$ and $\Sigma_\sigma = \partial\Omega \times [t_0, \sigma[$.

CLAIM. *Fix any constant* $\delta \in \,]t_0, \omega[$. *Let a function* $g \in L^\infty(Q_\omega)$ *and* $u_0 \in W_p^{2-2/p}(\Omega)$ *with* $u_0 = 0$ *on* $\partial\Omega$ *be given. Then, for every* $\tau \in \,]\delta, \omega[$, *the initial value problem*

$$\begin{array}{ll} \partial_t u + A(x, t, \partial_x)u = g(x, t)\big(1 + |\nabla_x u|^2\big) & \text{in } Q_\tau, \\ u = 0 & \text{on } \Sigma_\tau, \\ u(\cdot, t_0) = u_0 & \text{in } \Omega \end{array} \tag{2.15}$$

has a unique solution u, which satisfies

$$\|u\|_{W_p^{2,1}(Q_\tau)} \leqslant \gamma_0\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}\big),$$

where γ_0 *depends in an increasing way on the indicated quantities, but is independent of* $\tau \in \,]\delta, \omega[$.

Let us consider, for $\tau \in \,]t_0, \omega[$ and $\lambda \in [0, 1]$, the initial value problem

$$\begin{array}{ll} \partial_t u + A(x, t, \partial_x)u = g(x, t)\big(\lambda + |\nabla_x u|^2\big) & \text{in } Q_\tau, \\ u = 0 & \text{on } \Sigma_\tau, \\ u(\cdot, t_0) = \lambda u_0 & \text{in } \Omega. \end{array} \tag{2.16}$$

Step 1. *For every* $\tau \in \,]t_0, \omega[$ *and* $\lambda \in [0, 1]$, (2.16) *has at most one solution*. Indeed, if u, v are solutions of (2.16), then $w = u - v$ satisfies

$$\begin{array}{ll} \partial_t w + A(x,t,\partial_x)w - g\nabla_x(u+v)\cdot\nabla_x w = 0 & \text{in } Q_\tau, \\ w = 0 & \text{on } \Sigma_\tau, \\ w(\cdot,t_0) = 0 & \text{in } \Omega \end{array}$$

and hence $w = 0$, as it easily follows from Proposition 2.2.

For each $\tau \in]t_0, \omega[$, let us denote by Λ_τ the set of all $\lambda \in [0,1]$ such that (2.16) has a solution.

Step 2. There exists a constant $\varepsilon = \varepsilon(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}) > 0$ *such that, for every* $\tau \in]t_0, \omega[$ *and* $\lambda_1, \lambda_2 \in \Lambda_\tau$, *with* $|\lambda_1 - \lambda_2| \leqslant \varepsilon$, *the corresponding solutions* u_1, u_2 *of* (2.16) *satisfy*

$$\|u_1 - u_2\|_{W_p^{2,1}(Q_\tau)} \leqslant \gamma_1\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, \|u_1\|_{C^{1,0}(\overline{Q}_\tau)}\big), \tag{2.17}$$

where γ_1 *depends in an increasing way on the indicated quantities.* Set $w = u_1 - u_2$ and observe that w satisfies

$$\begin{array}{ll} \partial_t w + A(x,t,\partial_x)w = g\nabla_x(u_1+u_2)\cdot\nabla_x w + (\lambda_1-\lambda_2)g & \text{in } Q_\tau, \\ w = 0 & \text{on } \Sigma_\tau, \\ w(\cdot,t_0) = (\lambda_1-\lambda_2)u_0 & \text{in } \Omega. \end{array} \tag{2.18}$$

Let $\eta = \min\{1, \operatorname{ess\,inf}_{\Omega\times\mathbb{R}} a_0\} > 0$ and define

$$M = \eta^{-1}|\lambda_1 - \lambda_2| \max\big\{\|g\|_{L^\infty(Q_\tau)}, \|u_0\|_{L^\infty(\Omega)}\big\}.$$

The function $M - w$ satisfies

$$\begin{array}{ll} \multicolumn{2}{l}{\partial_t(M-w) + A(x,t,\partial_x)(M-w) - g\nabla_x(u_1+u_2)\cdot\nabla_x(M-w)} \\ \quad = a_0 M - (\lambda_1-\lambda_2)g \geqslant 0 & \text{in } Q_\tau, \\ M - w \geqslant 0 & \text{on } \Sigma_\tau, \\ M - w(\cdot,t_0) = M - (\lambda_1-\lambda_2)u_0 \geqslant 0 & \text{in } \Omega. \end{array}$$

Proposition 2.2 then implies that $M - w \geqslant 0$ in Q_τ. Similarly, one proves that $M + w \geqslant 0$ in Q_τ and hence

$$\|w\|_{L^\infty(Q_\tau)} \leqslant \eta^{-1}|\lambda_1 - \lambda_2| \max\big\{\|g\|_{L^\infty(Q_\tau)}, \|u_0\|_{L^\infty(\Omega)}\big\}. \tag{2.19}$$

Next, using the inequality

$$\big|\nabla_x(u_1+u_2)\cdot\nabla_x w\big| \leqslant |\nabla_x u_1|^2 + 2|\nabla_x w|^2,$$

we get from (2.18)

$$\big\|\partial_t w + A(x,t,\partial_x)w\big\|_{L^p(Q_\tau)}$$

$$\leqslant 2\|g\|_{L^\infty(Q_\tau)}\big\||\nabla_x w|^2\big\|_{L^p(Q_\tau)} \\ + \big(\mathrm{meas}(Q_\tau)\big)^{1/p}\|g\|_{L^\infty(Q_\tau)}\big[\|u_1\|^2_{C^{1,0}(\overline{Q}_\tau)} + 1\big]. \tag{2.20}$$

Now, by Hölder inequality and Fubini theorem, we have

$$\begin{aligned}\big\||\nabla_x w|^2\big\|_{L^p(Q_\tau)} &\leqslant \sum_{i=1}^N \big\||\partial_{x_i} w|^2\big\|_{L^p(Q_\tau)} \\ &= \sum_{i=1}^N \bigg(\int_0^\tau \bigg(\int_\Omega \big|\partial_{x_i} w(x,t)\big|^{2p}\,\mathrm{d}x\bigg)\,\mathrm{d}t\bigg)^{1/p} \\ &= \sum_{i=1}^N \bigg(\int_0^\tau \big(\big\|\partial_{x_i} w(\cdot,t)\big\|^2_{L^{2p}(\Omega)}\big)^p\,\mathrm{d}t\bigg)^{1/p}\end{aligned}$$

and hence, by an interpolation inequality [51, Theorem 1.10.1],

$$\begin{aligned}\big\||\nabla_x w|^2\big\|_{L^p(Q_\tau)} &\leqslant \sum_{i=1}^N \bigg(\int_0^\tau \big(\big\|\partial_{x_i} w(\cdot,t)\big\|^2_{L^{2p}(\Omega)}\big)^p\,\mathrm{d}t\bigg)^{1/p} \\ &\leqslant \sum_{i=1}^N \bigg(\int_0^\tau C_1^p \big\|w(\cdot,t)\big\|^p_{W^2_p(\Omega)}\big\|w(\cdot,t)\big\|^p_{L^\infty(\Omega)}\,\mathrm{d}t\bigg)^{1/p} \\ &\leqslant NC_1\|w\|_{L^\infty(Q_\tau)}\bigg(\int_0^\tau \|w(\cdot,t)\|^p_{W^2_p(\Omega)}\,\mathrm{d}t\bigg)^{1/p} \\ &\leqslant NC_1\|w\|_{L^\infty(Q_\tau)}\|w\|_{W^{2,1}_p(Q_\tau)},\end{aligned} \tag{2.21}$$

where C_1 is a constant depending only on Ω and p. Combining (2.21) with (2.19) yields

$$\big\||\nabla_x w|^2\big\|_{L^p(Q_\tau)} \leqslant \frac{NC_1}{\eta}|\lambda_1 - \lambda_2|\max\big\{\|g\|_{L^\infty(Q_\tau)}, \|u_0\|_{L^\infty(\Omega)}\big\}\|w\|_{W^{2,1}_p(Q_\tau)} \tag{2.22}$$

and hence, inserting (2.22) into (2.20),

$$\begin{aligned}&\big\|\partial_t w + A(x,t,\partial_x)w\big\|_{L^p(Q_\tau)} \\ &\quad\leqslant \frac{2NC_1}{\eta}\|g\|_{L^\infty(Q_\tau)}|\lambda_1 - \lambda_2|\max\big\{\|g\|_{L^\infty(Q_\tau)}, \|u_0\|_{L^\infty(\Omega)}\big\}\|w\|_{W^{2,1}_p(Q_\tau)} \\ &\qquad + \big(\mathrm{meas}(Q_\tau)\big)^{1/p}\|g\|_{L^\infty(Q_\tau)}\big(\|u_1\|^2_{C^{1,0}(\overline{Q}_\tau)} + 1\big).\end{aligned} \tag{2.23}$$

On the other hand, Proposition 2.4 implies that

$$\|w\|_{W_p^{2,1}(Q_\tau)} \leqslant C_2(\tau)\big(\big\|\partial_t w + A(x,t,\partial_x)w\big\|_{L^p(Q_\tau)} + |\lambda_1 - \lambda_2|\|u_0\|_{W_p^{2-2/p}(\Omega)}\big), \tag{2.24}$$

where, for every $\tau \in]t_0, \omega[$, $C_2(\tau) \leqslant C_2(\omega) < +\infty$ as $\omega < +\infty$. Therefore, if we set

$$\begin{aligned} &C_3\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{L^\infty(\Omega)}\big) \\ &\quad = C_2(\omega)2\eta^{-1}NC_1\|g\|_{L^\infty(Q_\omega)}\max\big\{\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{L^\infty(\Omega)}\big\} \end{aligned} \tag{2.25}$$

and

$$\begin{aligned} &C_4\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, \|u_1\|_{C^{1,0}(\overline{Q}_\tau)}\big) \\ &\quad = C_2(\omega)\|u_0\|_{W_p^{2-2/p}(\Omega)} \\ &\qquad + C_2(\omega)\big(\text{meas}(Q_\omega)\big)^{1/p}\|g\|_{L^\infty(Q_\omega)}\big(\|u_1\|^2_{C^{1,0}(\overline{Q}_\tau)} + 1\big), \end{aligned} \tag{2.26}$$

we get

$$\begin{aligned} \|w\|_{W_p^{2,1}(Q_\tau)} \leqslant{}& C_3\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{L^\infty(\Omega)}\big)|\lambda_1 - \lambda_2|\|w\|_{W_p^{2,1}(Q_\tau)} \\ &+ C_4\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, \|u_1\|_{C^{1,0}(\overline{Q}_\tau)}\big). \end{aligned}$$

Accordingly, letting

$$\varepsilon = \varepsilon\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{L^\infty(\Omega)}\big) = \big[2C_3\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{L^\infty(\Omega)}\big)\big]^{-1},$$

we conclude that

$$\|w\|_{W_p^{2,1}(Q_\tau)} \leqslant 2C_4\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, \|u_1\|_{C^{1,0}(\overline{Q}_\tau)}\big),$$

provided $|\lambda_1 - \lambda_2| \leqslant \varepsilon$. Hence (2.17) follows setting

$$\begin{aligned} &\gamma_1\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, \|u_1\|_{C^{1,0}(\overline{Q}_\tau)}\big) \\ &\quad = 2C_4\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, \|u_1\|_{C^{1,0}(\overline{Q}_\tau)}\big), \end{aligned}$$

where γ_1 depends in an increasing way on the indicated quantities.

Step 3. *For every* $\tau \in]t_0, \omega[$ *and* $\lambda \in [0, \varepsilon]$, *the initial value problem* (2.16) *has a solution* u_λ, *which is unique by Step* 1 *and satisfies*

$$\|u_\lambda\|_{W_p^{2,1}(Q_\tau)} \leqslant \gamma_1\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, 0\big).$$

For each $\tau \in]t_0, \omega[$, $0 \in \Lambda_\tau$ and for every $\lambda \in [0, \varepsilon]$, any possible solution u_λ of (2.16) satisfies, by letting $u_1 = 0$ and $u_2 = u_\lambda$ in (2.17),

$$\|u_\lambda\|_{C^{1,0}(\overline{Q}_\tau)} < R(\tau) = C_5(\tau)\gamma_1\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, 0\big) + 1,$$

where $C_5(\tau)$ is the embedding constant of $W_p^{2,1}(Q_\tau)$ into $C^{1,0}(\overline{Q}_\tau)$. Therefore, the homotopy invariance of the degree implies that

$$\deg\big(I - S_{\tau,\lambda}, B\big(0, R(\tau)\big)\big) = \deg\big(I - S_{\tau,0}, B\big(0, R(\tau)\big)\big),$$

where $S_{\tau,\lambda} : C^{1,0}(\overline{Q}_\tau) \to C^{1,0}(\overline{Q}_\tau)$ is the solution operator associated with (2.16) and $B(0, R(\tau))$ is the open ball in $C^{1,0}(\overline{Q}_\tau)$ of center 0 and radius $R(\tau)$. As for $\mu \in [0, 1]$, $u = \mu S_{\tau,0} u$ has only the trivial solution, we have also

$$\begin{aligned}\deg\big(I - S_{\tau,0}, B\big(0, R(\tau)\big)\big) &= \deg\big(I - \mu S_{\tau,0}, B\big(0, R(\tau)\big)\big)\\ &= \deg\big(I, B\big(0, R(\tau)\big)\big) = 1,\end{aligned}$$

and hence

$$\deg\big(I - S_{\tau,\lambda}, B\big(0, R(\tau)\big)\big) = 1.$$

The existence property of the degree yields the conclusion.

Conclusion. Now, arguing as in Step 3, letting in (2.17) $u_1 = u_\varepsilon$ and $u_2 = u_\lambda$ with $\lambda \in [\varepsilon, \min\{2\varepsilon, 1\}]$, we get $[0, \min\{2\varepsilon, 1\}] \subseteq \Lambda_\tau$ and

$$\|u_{2\varepsilon}\|_{W_p^{2,1}(Q_\tau)} \leqslant 2\gamma_1\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, \|u_\varepsilon\|_{C^{1,0}(\overline{Q}_\tau)}\big),$$

where

$$\|u_\varepsilon\|_{C^{1,0}(\overline{Q}_\tau)} \leqslant C_5(\delta, \omega)\gamma_1\big(\|g\|_{L^\infty(Q_\omega)}, \|u_0\|_{W_p^{2-2/p}(\Omega)}, 0\big),$$

$C_5(\delta, \omega)$ denoting a constant such that $C_5(\delta, \omega) > C_5(\tau)$ for every $\tau \in [\delta, \omega]$ (cf. [80, Lemma II.3.3]). Finally, iterating this process a finite number of times, the assertion of the claim follows.

We are now in position to conclude the proof of this proposition. Let $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$ be a function such that $u \in W_p^{2,1}(Q_\tau)$ for every $\tau \in]t_0, \omega[$, and

$$\begin{aligned}&\big|\partial_t u + A(x, t, \partial_x)u\big| \leqslant h(x, t) + K|\nabla_x u|^2 && \text{a.e. in } Q_\omega,\\ &u = 0 && \text{on } \Sigma_\omega,\\ &u(\cdot, t_0) = u_0 && \text{in } \Omega,\end{aligned}$$

where $u_0 \in W_p^{2-2/p}(\Omega)$ is such that $u_0 = 0$ on $\partial\Omega$ and

$$\|u_0\|_{W_p^{2-2/p}(\Omega)} \leqslant R.$$

Define $k = \frac{\partial_t u + A(x,t,\partial_x)u}{1+h+K|\nabla_x u|^2} \in L^\infty(Q_\omega)$. Let $v \in W_p^{2,1}(Q_\omega)$ be the solution of

$$\begin{array}{ll} \partial_t v + A(x,t,\partial_x)v = k(x,t)h(x,t) & \text{in } Q_\omega, \\ v = 0 & \text{on } \Sigma_\omega, \\ v(\cdot, t_0) = 0 & \text{in } \Omega, \end{array}$$

which satisfies, by Proposition 2.4,

$$\|v\|_{W_p^{2,1}(Q_\omega)} \leqslant C_2(\omega)\|h\|_{L^p(Q_\omega)} \tag{2.27}$$

and hence, by the embedding of $W_p^{2,1}(Q_\omega)$ into $C^{1,0}(\overline{Q}_\omega)$,

$$\|v\|_{C^{1,0}(\overline{Q}_\omega)} \leqslant C_6\|h\|_{L^p(Q_\omega)}. \tag{2.28}$$

Set $w = u - v$ and $g = k\frac{1+K|\nabla_x u|^2}{1+|\nabla_x w|^2}$. The function w satisfies

$$\partial_t w + A(x,t,\partial_x)w = k(x,t)\big(1 + K|\nabla_x u|^2\big) = g(x,t)\big(1 + |\nabla_x w|^2\big)$$

a.e. in Q_ω and therefore, for every $\tau \in]0, \omega[$, it is a solution of

$$\begin{array}{ll} \partial_t w + A(x,t,\partial_x)w = g(x,t)\big(1 + |\nabla_x w|^2\big) & \text{in } Q_\tau, \\ w = 0 & \text{on } \Sigma_\tau, \\ w(\cdot, t_0) = u_0 & \text{in } \Omega. \end{array}$$

Since, by (2.28),

$$\|\nabla_x u\|^2_{L^\infty(Q_\omega)} \leqslant 2\|\nabla_x w\|^2_{L^\infty(Q_\omega)} + 2\big(C_6\|h\|_{L^p(Q_\omega)}\big)^2,$$

it follows that $g \in L^\infty(Q_\omega)$ and

$$\|g\|_{L^\infty(Q_\omega)} \leqslant \max\{1, 2K\} + 2K\big(C_6\|h\|_{L^p(Q_\omega)}\big)^2.$$

Therefore, if we fix $\delta \in]t_0, \omega[$, the above claim implies that, for every $\tau \in]\delta, \omega[$,

$$\|w\|_{W_p^{2,1}(Q_\tau)} \leqslant \gamma_0\Big(\max\{1, 2K\} + 2K\big(C_6\|h\|_{L^p(Q_\omega)}\big)^2, R\Big),$$

which, in turn, yields

$$\|w\|_{C^{1,0}(\overline{Q}_\tau)} \leqslant C_5(\delta, \omega)\gamma_0\Big(\max\{1, 2K\} + 2K\big(C_6\|h\|_{L^p(Q_\omega)}\big)^2, R\Big)$$

and, in particular,

$$\|\nabla_x w\|_{L^\infty(Q_\omega)} \leqslant C_5(\delta, \omega)\gamma_0\Big(\max\{1, 2K\} + 2K\big(C_6\|h\|_{L^p(Q_\omega)}\big)^2, R\Big).$$

Accordingly, denoting by $z \in W_p^{2,1}(Q_\omega)$ the unique solution of

$$\begin{array}{ll} \partial_t z + A(x,t,\partial_x)z = g(x,t)\big(1+|\nabla_x w|^2\big) & \text{in } Q_\omega, \\ z = 0 & \text{on } \Sigma_\omega, \\ z(\cdot,t_0) = u_0 & \text{in } \Omega, \end{array}$$

we have $w = z$ in $\overline{\Omega} \times [t_0,\omega[$ and hence $w \in W_p^{2,1}(Q_\omega)$. Finally, using (2.27), we conclude that $u \in W_p^{2,1}(Q_\omega)$ and

$$\|u\|_{W_p^{2,1}(Q_\omega)} \leqslant \gamma_0\big(\max\{1,2K\} + 2K\big(C_6\|h\|_{L^p(Q_\omega)}\big)^2, R\big) \\ + C_2(\omega)\|h\|_{L^p(Q_\omega)}. \qquad \square$$

Existence of solutions and the Hukuhara–Kneser property

In this section we extend to strong solutions and nonregular lower and upper solutions some existence and localization results for the initial value problem (2.13) which are well known in the frame of classical solutions. We also describe the topological structure of the solution set, proving the Hukuhara–Kneser property. Since our analysis is performed just assuming that f satisfies the Carathéodory conditions, we get an improvement of results obtained in [76], for f Hölder continuous, and in [81], for f continuous. Our proof, which differs from the standard ones in the approximation method, adapts an idea introduced in [139] dealing with ordinary differential equations.

We first deal with the initial value problem on a compact interval.

THEOREM 2.6. *Assume* (D), (A), (C), (N) *and let* $u_0 \in W_p^{2-2/p}(\Omega)$ *be such that* $u_0 = 0$ *on* $\partial\Omega$. *Let* $t_0, t_1 \in \mathbb{R}$ *be such that* $t_0 < t_1$ *and suppose that* α *is a lower solution and* β *is an upper solution of* (2.13) *in* $\overline{\Omega} \times [t_0,t_1]$, *satisfying* $\alpha \leqslant \beta$. *Then there exist the minimum solution* v *and the maximum solution* w *in* $[\alpha,\beta]$ *of problem* (2.13) *in* $\overline{\Omega} \times [t_0,t_1]$. *Further, the set*

$$\mathcal{K} = \big\{u : \overline{\Omega} \times [t_0,t_1] \to \mathbb{R} \mid u \text{ is a solution of (2.13) with } \alpha \leqslant u \leqslant \beta \\ \text{in } \overline{\Omega} \times [t_0,t_1]\big\}$$

is a continuum, i.e. a compact and connected set, in $C^{1,0}(\overline{\Omega} \times [t_0,t_1])$.

PROOF. The proof is divided into three parts.

Part 1. *Existence of a solution* u *of* (2.13) *in* $\overline{\Omega} \times [t_0,t_1]$, *with* $\alpha \leqslant u \leqslant \beta$. We can assume that the sequences $(\sigma_h)_{0 \leqslant h \leqslant k}$ and $(\rho_h)_{0 \leqslant h \leqslant l}$ coincide. We also notice that, by our definition of lower and upper solutions and the condition $\alpha \leqslant \beta$, there exists a constant $M > 0$ such that $-M \leqslant \alpha(x,t) \leqslant \beta(x,t) \leqslant M$ in $\overline{\Omega} \times [t_0,t_1]$. Hence, by condition (N) and Proposition 2.5, there exists $R > 0$ such that, for every function f satisfying (2.11) and every solution u of (2.13) in $\overline{\Omega} \times [t_0,t_1]$ with $\alpha \leqslant u \leqslant \beta$, we have

$$\|u\|_{C^{1,0}(\overline{\Omega}\times[t_0,t_1])} < R. \tag{2.29}$$

For $h \in \{0, \ldots, k-1\}$, let $\alpha = \max_{1 \leqslant i \leqslant m_h} \alpha_i^{(h)}$ and $\beta = \min_{1 \leqslant j \leqslant n_h} \beta_j^{(h)}$ in $\overline{\Omega} \times]\sigma_h, \sigma_{h+1}[$, with $\alpha_i^{(h)}, \beta_j^{(h)} \in W_p^{2,1}(\Omega \times]\sigma_h, \sigma_{h+1}[) \cap C^{1,0}(\overline{\Omega} \times [\sigma_h, \sigma_{h+1}])$. Take $\overline{R}$ such that

$$\begin{aligned}\overline{R} > R + \max_{0 \leqslant h \leqslant k-1} \Big\{ & \max_{1 \leqslant i \leqslant m_h} \big\| \alpha_i^{(h)} \big\|_{C^{1,0}(\overline{\Omega} \times [\sigma_h, \sigma_{h+1}])}, \\ & \max_{1 \leqslant j \leqslant n_h} \big\| \beta_j^{(h)} \big\|_{C^{1,0}(\overline{\Omega} \times [\sigma_h, \sigma_{h+1}])} \Big\}\end{aligned}$$

and set

$$\bar{f}(x, t, s, \xi) = \begin{cases} f(x, t, s, \xi) & \text{if } |\xi| \leqslant \overline{R}, \\ f\big(x, t, s, \overline{R} \frac{\xi}{|\xi|}\big) & \text{if } |\xi| > \overline{R}, \end{cases}$$

for a.e. $(x, t) \in \Omega \times \mathbb{R}$ and every $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$. For any given $h \in \{0, \ldots, k-1\}$, we define the functions

$$g_i^{(h)}(x, t, s, \xi) = \begin{cases} \bar{f}\big(x, t, \alpha_i^{(h)}(x, t), \xi\big) + \omega_{1i}^{(h)}\big(x, t, \alpha_i^{(h)}(x, t) - s\big) \\ \qquad\qquad\qquad \text{if } s < \alpha_i^{(h)}(x, t), \\ \bar{f}(x, t, s, \xi) \quad \text{if } s \geqslant \alpha_i^{(h)}(x, t), \end{cases}$$

$$h_j^{(h)}(x, t, s, \xi) = \begin{cases} \bar{f}\big(x, t, \beta_j^{(h)}(x, t), \xi\big) - \omega_{2j}^{(h)}\big(x, t, s - \beta_j^{(h)}(x, t)\big) \\ \qquad\qquad\qquad \text{if } s > \beta_j^{(h)}(x, t), \\ \bar{f}(x, t, s, \xi) \quad \text{if } s \leqslant \beta_j^{(h)}(x, t), \end{cases}$$

where

$$\begin{aligned} \omega_{1i}^{(h)}(x, t, \delta) = \max_{|\xi| \leqslant \delta} \big| & \bar{f}\big(x, t, \alpha_i^{(h)}(x, t), \nabla_x \alpha_i^{(h)}(x, t) + \xi\big) \\ & - \bar{f}\big(x, t, \alpha_i^{(h)}(x, t), \nabla_x \alpha_i^{(h)}(x, t)\big)\big|, \\ \omega_{2j}^{(h)}(x, t, \delta) = \max_{|\xi| \leqslant \delta} \big| & \bar{f}\big(x, t, \beta_j^{(h)}(x, t), \nabla_x \beta_j^{(h)}(x, t) + \xi\big) \\ & - \bar{f}\big(x, t, \beta_j^{(h)}(x, t), \nabla_x \beta_j^{(h)}(x, t)\big)\big|, \end{aligned}$$

for $i \in \{1, \ldots, m_h\}$ and $j \in \{1, \ldots, n_h\}$, and

$$F^{(h)}(x, t, s, \xi) = \begin{cases} \max\limits_{1 \leqslant i \leqslant m_h} g_i^{(h)}(x, t, s, \xi) & \text{if } s \leqslant \alpha(x, t), \\ \bar{f}(x, t, s, \xi) & \text{if } \alpha(x, t) < s < \beta(x, t), \\ \min\limits_{1 \leqslant j \leqslant n_h} h_j^{(h)}(x, t, s, \xi) & \text{if } s \geqslant \beta(x, t), \end{cases} \tag{2.30}$$

for a.e. $(x,t) \in \Omega \times]\sigma_h, \sigma_{h+1}[$ and every $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$. Then we consider the modified problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = F(x,t,u,\nabla_x u) & \text{in } \Omega \times]t_0, t_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, t_1], \\ u(\cdot, t_0) = u_0 & \text{in } \Omega, \end{array} \tag{2.31}$$

where $F : \Omega \times]t_0, t_1[\times \mathbb{R} \times \mathbb{R}^N$ is a L^p-Carathéodory function such that, for each $h \in \{0, \ldots, k-1\}$,

$$F(x,t,s,\xi) = F^{(h)}(x,t,s,\xi),$$

for a.e. $(x,t) \in \Omega \times]\sigma_h, \sigma_{h+1}[$ and every $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$. Notice that there exists $\gamma \in L^p(\Omega \times]t_0, t_1[)$ such that

$$\left|F(x,t,s,\xi)\right| \leqslant \gamma(x,t) \tag{2.32}$$

for a.e. $(x,t) \in \Omega \times]t_0, t_1[$ and every $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$.

Step 1. *Every solution u of* (2.31) *satisfies $\alpha \leqslant u \leqslant \beta$.* Assume there is a solution u of (2.31) such that $\inf_{\Omega \times]t_0, t_1[}(u - \alpha) < 0$. By our definition of a lower solution of (2.31), we can find $h \in \{0, \ldots, k-1\}$ such that $\inf_{\Omega \times]\sigma_h, \sigma_{h+1}[}(u - \alpha) < 0$ and $u(x, \sigma_h) \geqslant \lim_{t \to \sigma_h^+} \alpha(x,t)$ in Ω. Since $\alpha = \max_{1 \leqslant i \leqslant m_h} \alpha_i^{(h)}$, there is $i \in \{1, \ldots, m_h\}$ such that, setting $v = u - \alpha_i^{(h)} \in W_p^{2,1}(\Omega \times]\sigma_h, \sigma_{h+1}[)$, we have $\min_{\overline{\Omega} \times [\sigma_h, \sigma_{h+1}]} v < 0$, $v(x, \sigma_h) \geqslant 0$ in Ω and $v(x,t) \geqslant 0$ on $\partial\Omega \times [\sigma_h, \sigma_{h+1}]$. Hence there exists $(\bar{x}, \bar{t}) \in \Omega \times]\sigma_h, \sigma_{h+1}]$ such that $\min_{\overline{\Omega} \times [\sigma_h, \sigma_{h+1}]} v = v(\bar{x}, \bar{t})$. As $\bar{x} \in \Omega$, we have $\nabla_x v(\bar{x}, \bar{t}) = 0$ and there is an open ball $B \subseteq \Omega$, with $\bar{x} \in B$, and a point $t_1 \in]\sigma_h, \bar{t}[$ such that, a.e. in $B \times]t_1, \bar{t}]$, $|\nabla_x v(x,t)| \leqslant |v(x,t)|$, $v(x,t) < 0$ and

$$\begin{aligned} &\partial_t v + A(x,t,\partial_x)v \\ &\quad \geqslant F^{(h)}\big(x,t,u(x,t),\nabla_x u(x,t)\big) - f\big(x,t,\alpha_i^{(h)}(x,t),\nabla_x \alpha_i^{(h)}(x,t)\big) \\ &\quad \geqslant \bar{f}\big(x,t,\alpha_i^{(h)}(x,t),\nabla_x u(x,t)\big) + \omega_{1i}^{(h)}\big(x,t,\alpha_i^{(h)}(x,t) - u(x,t)\big) \\ &\qquad - \bar{f}\big(x,t,\alpha_i^{(h)}(x,t),\nabla_x \alpha_i^{(h)}(x,t)\big) \\ &\quad \geqslant -\omega_{1i}^{(h)}\big(x,t,\left|\nabla_x v(x,t)\right|\big) + \omega_{1i}^{(h)}\big(x,t,\left|v(x,t)\right|\big) \\ &\quad \geqslant 0, \end{aligned}$$

because $\omega_{1i}^{(h)}(x,t,\cdot)$ is increasing. Hence Proposition 2.2 implies that $v(x,t) = v(\bar{x}, \bar{t}) = \min_{\overline{\Omega} \times [\sigma_h, \sigma_{h+1}]} v$ in $B \times]t_1, \bar{t}]$ and, as $\operatorname{ess\,inf} a_0 > 0$,

$$\partial_t v + A(x,t,\partial_x)v = a_0 v < 0,$$

a.e. in $B \times]t_1, \bar{t}]$, thus contradicting the previous inequality. Therefore, we conclude that $u \geqslant \alpha$. Similarly, one proves that $u \leqslant \beta$.

Step 2. *Every solution u of* (2.31) *is a solution of* (2.13). In Step 1, we proved that every solution u of (2.31) satisfies $\alpha \leqslant u \leqslant \beta$ and hence it is a solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = \bar{f}(x,t,u,\nabla_x u) & \text{in } \Omega \times]t_0,t_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0,t_1], \\ u(\cdot,t_0) = u_0 & \text{in } \Omega. \end{array}$$

As $\bar{f}$ satisfies (2.11), we have $\|u\|_{C^{1,0}(\overline{\Omega}\times[t_0,t_1])} < \overline{R}$ and hence u is a solution of (2.13).

Step 3. *Problem* (2.31) *has at least one solution*. Let us consider the solution operator $S : C^{1,0}(\overline{\Omega} \times [t_0,t_1]) \to C^{1,0}(\overline{\Omega} \times [t_0,t_1])$ associated with (2.31), which sends any function $u \in C^{1,0}(\overline{\Omega} \times [t_0,t_1])$ onto the unique solution $v \in W_p^{2,1}(\overline{\Omega} \times [t_0,t_1])$ of

$$\begin{array}{ll} \partial_t v + A(x,t,\partial_x)v = F(x,t,u,\nabla_x u) & \text{in } \Omega \times]t_0,t_1[, \\ v = 0 & \text{on } \partial\Omega \times [t_0,t_1], \\ v(\cdot,t_0) = u_0 & \text{in } \Omega. \end{array}$$

The operator S is continuous, has a relatively compact range and its fixed points are the solutions of (2.31). In particular, there exists a constant $\widetilde{R} > 0$ such that

$$\|Su\|_{C^{1,0}(\overline{\Omega}\times[t_0,t_1])} < \widetilde{R}$$

for every $u \in C^{1,0}(\overline{\Omega} \times [t_0,t_1])$. Then standard results of degree theory imply that

$$\deg\bigl(I - S, B\bigl(0,\widetilde{R}\bigr)\bigr) = 1,$$

where I is the identity operator in $C^{1,0}(\overline{\Omega} \times [t_0,t_1])$ and $B(0,\widetilde{R})$ is the open ball of center 0 and radius $\widetilde{R}$ in $C^{1,0}(\overline{\Omega} \times [t_0,t_1])$. Therefore S has a fixed point. By the conclusions of Steps 1 and 2, we get the existence of a solution u of (2.13) in $\overline{\Omega} \times [t_0,t_1]$ satisfying $\alpha \leqslant u \leqslant \beta$.

Part 2. *Existence of extremal solutions*. We know, from Part 1, that the solutions u of (2.13) in $\overline{\Omega} \times [t_0,t_1]$, with $\alpha \leqslant u \leqslant \beta$, are precisely the fixed points of the solution operator S associated with (2.31), i.e.

$$\mathcal{K} = \bigl\{u \in C^{1,0}\bigl(\overline{\Omega} \times [t_0,t_1]\bigr) \mid u = Su\bigr\},$$

and $\mathcal{K}$ is a nonempty compact subset of $C^{1,0}(\overline{\Omega} \times [t_0,t_1])$. Next, for each $u \in \mathcal{K}$, define the closed set $\mathcal{C}_u = \{z \in \mathcal{K} \mid z \leqslant u\}$. The family $\{\mathcal{C}_u \mid u \in \mathcal{K}\}$ has the finite intersection property, as it follows from Part 1 observing that if $u_1, u_2 \in \mathcal{K}$, then $\min\{u_1,u_2\}$ is an upper solution of (2.31) with $\alpha \leqslant \min\{u_1,u_2\}$. By the compactness of $\mathcal{K}$ there exists $v \in \bigcap_{u\in\mathcal{K}} \mathcal{C}_u$; clearly, v is the minimum solution in $[\alpha,\beta]$ of (2.13) in $\overline{\Omega} \times [t_0,t_1]$. The maximum solution w can be found in a similar way.

Part 3. *$\mathcal{K}$ is a continuum.* We know from Part 2 that $\mathcal{K}$ is compact in $C^{1,0}(\bar{\Omega} \times [t_0, t_1])$; we apply [77, Theorem 48.2] or [153, Theorem 13.E] to show that $\mathcal{K}$ is connected too. We already proved in Part 1 that there is $R > 0$ such that

$$\deg\bigl(I - S, B(0, R)\bigr) = 1,$$

where $S : C^{1,0}(\overline{\Omega} \times [t_0, t_1]) \to C^{1,0}(\overline{\Omega} \times [t_0, t_1])$ is the solution operator associated with (2.31). Next, for each $n \geqslant 1$, set $\tau_n = \frac{t_1 - t_0}{n}$ and denote by $S_n : C^{1,0}(\overline{\Omega} \times [t_0, t_1]) \to C^{1,0}(\overline{\Omega} \times [t_0, t_1])$ the operator which sends any function $v \in C^{1,0}(\overline{\Omega} \times [t_0, t_1])$ onto the function $w \in C^{1,0}(\overline{\Omega} \times [t_0, t_1])$ defined by

$$w(x,t) = \begin{cases} u_0^*(x,t) & \text{in } \overline{\Omega} \times [t_0, t_0 + \tau_n], \\ u(x,t) & \text{in } \overline{\Omega} \times]t_0 + \tau_n, t_1], \end{cases}$$

where u_0^* is the solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = 0 & \text{in } \Omega \times]t_0, t_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, t_1], \\ u(\cdot, t_0) = u_0 & \text{in } \Omega \end{array}$$

and u is the solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u & \\ \quad = F\bigl(x, t - \tau_n, v(x, t - \tau_n), \nabla_x v(x, t - \tau_n)\bigr) & \text{in } \Omega \times]t_0 + \tau_n, t_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0 + \tau_n, t_1], \\ u(\cdot, t_0 + \tau_n) = u_0^*(\cdot, t_0 + \tau_n) & \text{in } \Omega. \end{array} \tag{2.33}$$

Each operator S_n is continuous and has a relatively compact range. Further, $I - S_n$ is one-to-one. Indeed, let $v_1, v_2 \in C^{1,0}(\overline{\Omega} \times [t_0, t_1])$ be such that $(I - S_n)v_1 = (I - S_n)v_2$. This means that $v_1(x,t) = v_2(x,t)$ in $\overline{\Omega} \times [t_0, t_0 + \tau_n]$ and $v_1(x,t) - u_1(x,t) = v_2(x,t) - u_2(x,t)$ in $\bar{\Omega} \times [t_0 + \tau_n, t_1]$, where, for $i = 1, 2$, u_i is the solution of (2.33) corresponding to $v = v_i$. Hence we get by recursion that $u_1(x,t) = u_2(x,t)$ and $v_1(x,t) = v_2(x,t)$ in $\bar{\Omega} \times [t_0 + (k-1)\tau_n, t_0 + k\tau_n]$, for $k = 2, \ldots, n$, and then $u_1(x,t) = u_2(x,t)$ and $v_1(x,t) = v_2(x,t)$ in $\overline{\Omega} \times [t_0 + \tau_n, t_1]$. This yields $v_1 = v_2$.

At last we prove that the sequence $(S_n)_n$ converges to S uniformly in $C^{1,0}(\overline{\Omega} \times [t_0, t_1])$. Assume by contradiction this is false, i.e. there exist a constant $\varepsilon > 0$ and a sequence $(v_k)_k$ in $C^{1,0}(\overline{\Omega} \times [t_0, t_1])$ such that

$$\|S_{n_k} v_k - S v_k\|_{C^{1,0}(\overline{\Omega} \times [t_0, t_1])} \geqslant \varepsilon. \tag{2.34}$$

For each k, define

$$g_k(x,t) = F\bigl(x, t, v_k(x,t), \nabla_x v_k(x,t)\bigr) \quad \text{a.e. in } \Omega \times]t_0, t_1[,$$

and

$$\ell_k(x,t) = \begin{cases} 0 \quad \text{a.e. in } \Omega \times]t_0, t_0 + \tau_{n_k}[, \\ F\big(x, t-\tau_{n_k}, v_k(x, t-\tau_{n_k}), \nabla_x v_k(x, t-\tau_{n_k})\big) \\ \qquad \text{a.e. in } \Omega \times]t_0 + \tau_{n_k}, t_1[. \end{cases}$$

Then $Sv_k = w_k$ is the solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = g_k(x,t) & \text{in } \Omega \times]t_0, t_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, t_1], \\ u(\cdot, t_0) = u_0 & \text{in } \Omega \end{array}$$

and $S_{n_k} v_k = z_k$ is the solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = \ell_k(x,t) & \text{in } \Omega \times]t_0, t_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, t_1], \\ u(\cdot, t_0) = u_0 & \text{in } \Omega. \end{array}$$

By condition (2.32), we can suppose, possibly passing to subsequences, that $(g_k)_k$ converges weakly in $L^p(\Omega \times]t_0, t_1[)$ to g and $(w_k)_k$ converges weakly in $W_p^{2,1}(\Omega \times]t_0, t_1[)$ to the solution w of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = g(x,t) & \text{in } \Omega \times]t_0, t_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, t_1], \\ u(\cdot, t_0) = u_0 & \text{in } \Omega. \end{array}$$

Similarly, we can suppose that $(\ell_k)_k$ converges weakly in $L^p(\Omega \times]t_0, t_1[)$ to ℓ and $(z_k)_k$ converges weakly in $W_p^{2,1}(\Omega \times]t_0, t_1[)$ to the solution z of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = \ell(x,t) & \text{in } \Omega \times]t_0, t_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, t_1], \\ u(\cdot, t_0) = u_0 & \text{in } \Omega. \end{array}$$

We shall prove that $g = \ell$. Hence the sequences $(w_k)_k$ and $(z_k)_k$ will have the same limit in $C^{1,0}(\overline{\Omega} \times [t_0, t_1])$, thus contradicting (2.34). For any $k \geqslant 1$ and $\psi \in L^{\frac{p}{p-1}}(\Omega \times]t_0, t_1[)$, let us compute

$$\begin{aligned} &\int_{\Omega \times]t_0, t_1[} \ell_k \psi \\ &\quad = \int_\Omega \left(\int_{t_0+\tau_{n_k}}^{t_1} F\big(x, t-\tau_{n_k}, v_k(x, t-\tau_{n_k}), \nabla_x v_k(x, t-\tau_{n_k})\big) \psi(x,t)\, \mathrm{d}t \right) \mathrm{d}x \\ &\quad = \int_\Omega \left(\int_{t_0}^{t_1-\tau_{n_k}} F\big(x, s, v_k(x,s), \nabla_x v_k(x,s)\big) \psi(x, s+\tau_{n_k})\, \mathrm{d}s \right) \mathrm{d}x \\ &\quad = \int_{\Omega \times]t_0, t_1[} g_k \psi_k, \end{aligned}$$

where we set

$$\psi_k(x,t) = \begin{cases} \psi(x, t+\tau_{n_k}) & \text{a.e. in } \Omega \times \,]t_0, t_1 - \tau_{n_k}[, \\ 0 & \text{a.e. in } \Omega \times [t_1 - \tau_{n_k}, t_1[. \end{cases}$$

Since $(g_k)_k$ is bounded in $L^p(\Omega \times \,]t_0, t_1[)$ and

$$\int_{\Omega\times]t_0,t_1[} \ell_k \psi = \int_{\Omega\times]t_0,t_1[} g_k \psi + \int_{\Omega\times]t_0,t_1[} g_k(\psi_k - \psi),$$

for proving that $g = \ell$ it is enough to show that $\|\psi_k - \psi\|_{L^{\frac{p}{p-1}}(\Omega\times]t_0,t_1[)} \to 0$, as $k \to +\infty$. Let us denote by $\widetilde{\psi}$ the zero extension of ψ outside $\Omega \times \,]t_0, t_1[$ and define

$$\widetilde{\psi}_k(x,t) = \widetilde{\psi}(x, t + \tau_{n_k}) \quad \text{a.e. in } \mathbb{R}^{N+1}.$$

By [22, Lemma IV.4] $(\widetilde{\psi}_k)_k$ converges to $\widetilde{\psi}$ in $L^{\frac{p}{p-1}}(\mathbb{R}^{N+1})$. Hence we obtain

$$\begin{aligned}
&\|\psi_k - \psi\|^{\frac{p}{p-1}}_{L^{\frac{p}{p-1}}(\Omega\times]t_0,t_1[)} \\
&\quad = \int_\Omega \left(\int_{t_0}^{t_1-\tau_{n_k}} \left|\psi_k(x,t) - \psi(x,t)\right|^{\frac{p}{p-1}} \mathrm{d}t \right) \mathrm{d}x \\
&\qquad + \int_\Omega \left(\int_{t_1-\tau_{n_k}}^{t_1} \left|\psi(x,t)\right|^{\frac{p}{p-1}} \mathrm{d}t \right) \mathrm{d}x \\
&\quad = \int_\Omega \left(\int_{\mathbb{R}} \left|\widetilde{\psi}_k(x,t) - \widetilde{\psi}(x,t)\right|^{\frac{p}{p-1}} \mathrm{d}t \right) \mathrm{d}x \\
&\qquad - \int_\Omega \left(\int_{t_0-\tau_{n_k}}^{t_0} \left|\widetilde{\psi}_k(x,t)\right|^{\frac{p}{p-1}} \mathrm{d}t \right) \mathrm{d}x \\
&\quad = \left\|\widetilde{\psi}_k - \widetilde{\psi}\right\|^{\frac{p}{p-1}}_{L^{\frac{p}{p-1}}(\mathbb{R}^{N+1})} - \int_\Omega \left(\int_{t_0}^{\tau_{n_k}+t_0} \left|\widetilde{\psi}(x,t)\right|^{\frac{p}{p-1}} \mathrm{d}t \right) \mathrm{d}x \to 0,
\end{aligned}$$

as $k \to +\infty$. This concludes the proof. □

REMARK 2.8. Condition (N) can be obviously replaced in Theorem 2.6 by

(N′) there exist $h \in L^p(\Omega \times \,]t_0, t_1[)$, with $p > N + 2$, and $K > 0$ such that, for a.e. $(x,t) \in \Omega \times \,]t_0, t_1[$, every $s \in [\alpha(x,t), \beta(x,t)]$ and every $\xi \in \mathbb{R}^N$,

$$\left|f(x,t,s,\xi)\right| \leqslant h(x,t) + K|\xi|^2.$$

According to Proposition 2.5, these Nagumo-type conditions prevent the formation of singularities of the space gradient of locally bounded solutions of (2.13).

Now we extend Theorem 2.6 to noncompact intervals.

COROLLARY 2.7. *Assume* (D), (A), (C), (N) *and let* $u_0 \in W_p^{2-2/p}(\Omega)$ *be such that* $u_0 = 0$ *on* $\partial\Omega$. *Let* $t_0 \in \mathbb{R}$ *and* $\tau \in \,]t_0, +\infty]$. *Suppose that* α *is a lower solution and* β *is an upper solution of* (2.13) *in* $\overline{\Omega} \times [t_0, \tau[$ *satisfying* $\alpha \leqslant \beta$. *Then there exist the minimum solution* v *and the maximum solution* w *in* $[\alpha, \beta]$ *of* (2.13) *in* $\overline{\Omega} \times [t_0, \tau[$. *Further, the set*

$$\mathcal{K} = \big\{u : \overline{\Omega} \times [t_0, \tau[\, \to \mathbb{R} \mid u \text{ is a solution of (2.13) with } \alpha \leqslant u \leqslant \beta\big\}$$

is a continuum in $C^{1,0}(\overline{\Omega} \times [t_0, \tau[)$, *endowed with the topology of* $C^{1,0}$*-convergence on compact subsets of* $\overline{\Omega} \times [t_0, \tau[$.

PROOF. Let $(\tau_n)_n$, with $\tau_0 = t_0$, be a strictly increasing sequence converging to τ.

Step 1. *Existence of extremal solutions.* In order to prove the existence of the minimum solution v in $[\alpha, \beta]$ of (2.13) in $\overline{\Omega} \times [t_0, \tau[$, we apply recursively Theorem 2.6 to (2.13) in $\overline{\Omega} \times [t_0, \tau_n]$ for each $n \geqslant 1$. Let us denote by v_n the minimum solution v in $[\alpha, \beta]$ of (2.13) in $\bar{\Omega} \times [t_0, \tau_n]$. By the minimality of v_n, we have $v_{n+1}|_{\overline{\Omega}\times[t_0,\tau_n]} \geqslant v_n$. On the other hand, $v_n(\cdot, \tau_n) \in W_p^{2-2/p}(\Omega)$, $v_n(\cdot, \tau_n) = 0$ on $\partial\Omega$ and the functions $\alpha|_{\overline{\Omega}\times[\tau_n,\tau_{n+1}]}$ and $v_{n+1}|_{\bar{\Omega}\times[\tau_n,\tau_{n+1}]}$ are, respectively, a lower and an upper solution of the initial value problem

$$\begin{array}{ll}
\partial_t u + A(x, t, \partial_x)u = f(x, t, u, \nabla_x u) & \text{in } \Omega \times \,]\tau_n, \tau_{n+1}[, \\
u = 0 & \text{on } \partial\Omega \times [\tau_n, \tau_{n+1}], \\
u(\cdot, \tau_n) = v_n(\cdot, \tau_n) & \text{in } \Omega.
\end{array}$$

Hence this problem has a solution v_n^* satisfying $\alpha|_{\bar{\Omega}\times[\tau_n,\tau_{n+1}]} \leqslant v_n^* \leqslant v_{n+1}|_{\bar{\Omega}\times[\tau_n,\tau_{n+1}]}$. Then we define $\hat{v}_{n+1} : \bar{\Omega} \times [t_0, \tau_{n+1}] \to \mathbb{R}$ by setting

$$\hat{v}_{n+1}(x, t) = \begin{cases} v_n(x, t) & \text{in } \overline{\Omega} \times [t_0, \tau_n[, \\ v_n^*(x, t) & \text{in } \overline{\Omega} \times [\tau_n, \tau_{n+1}]. \end{cases}$$

Let us prove that $\hat{v}_{n+1} \in W_p^{2,1}(\Omega \times \,]t_0, \tau_{n+1}[)$. Let $w \in W_p^{2,1}(\Omega \times \,]t_0, \tau_{n+1}[)$ be the unique solution of the linear initial value problem

$$\begin{array}{ll}
\partial_t u + A(x, t, \partial_x)u = f(x, t, \hat{v}_{n+1}, \nabla_x \hat{v}_{n+1}) & \text{in } \Omega \times \,]t_0, \tau_{n+1}[, \\
u = 0 & \text{on } \partial\Omega \times [t_0, \tau_{n+1}], \\
u(\cdot, t_0) = u_0 & \text{in } \Omega.
\end{array}$$

Since both v_n and $w|_{\overline{\Omega}\times[t_0,\tau_n]}$ are solutions of the linear initial value problem

$$\begin{array}{ll}
\partial_t u + A(x, t, \partial_x)u = f(x, t, v_n, \nabla_x v_n) & \text{in } \Omega \times \,]t_0, \tau_n[, \\
u = 0 & \text{on } \partial\Omega \times [t_0, \tau_n], \\
u(\cdot, t_0) = u_0 & \text{in } \Omega,
\end{array}$$

by uniqueness, we get $v_n = w$ in $\overline{\Omega} \times [t_0, \tau_n]$. Further, as $w(\cdot, \tau_n) = v_n(\cdot, \tau_n)$ in Ω, v_n^* and $w|_{\overline{\Omega}\times[\tau_n,\tau_{n+1}]}$ are both solutions of the linear initial value problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f(x,t,v_n^*,\nabla_x v_n^*) & \text{in } \Omega \times]\tau_n, \tau_{n+1}[, \\ u = 0 & \text{on } \partial\Omega \times [\tau_n, \tau_{n+1}], \\ u(\cdot, \tau_n) = v_n(\cdot, \tau_n) & \text{in } \Omega. \end{array}$$

Hence, by uniqueness, we have $v_n^* = w$ in $\overline{\Omega} \times [\tau_n, \tau_{n+1}]$ and we conclude that $\hat{v}_{n+1} = w$ in $\overline{\Omega} \times [t_0, \tau_{n+1}]$. Therefore $\hat{v}_{n+1}$ is a solution of (2.13) in $\overline{\Omega} \times [t_0, \tau_{n+1}]$ satisfying $\alpha \leqslant \hat{v}_{n+1} \leqslant v_{n+1}$. By the minimality of v_{n+1}, we conclude that $\hat{v}_{n+1} = v_{n+1}$ and hence $v_{n+1}|_{\overline{\Omega}\times[t_0,\tau_n]} = v_n$. Then we define $v : \overline{\Omega} \times [t_0, \tau[\to \mathbb{R}$ by setting $v(x,t) = v_n(x,t)$ in $\overline{\Omega} \times [t_0, \tau_n]$. We have that v is the minimum solution in $[\alpha, \beta]$ of (2.13) in $\overline{\Omega} \times [t_0, \tau[$, because, if u were a solution of (2.13) in $\overline{\Omega} \times [t_0, \tau[$ with $u \geqslant \alpha$ and $u \ngeqslant v$, then it should follow $u|_{\overline{\Omega}\times[t_0,\tau_n]} \ngeqslant v_n$ for some n, thus contradicting the minimality of v_n. Similarly, we prove the existence of the maximum solution w in $[\alpha, \beta]$ of (2.13) in $\overline{\Omega} \times [t_0, \tau[$.

Step 2. *$\mathcal{K}$ is a continuum.* We denote by $\mathcal{K}_n$ the set of all solutions $u : \overline{\Omega} \times [t_0, \tau_n] \to \mathbb{R}$ of (2.13) such that $\alpha \leqslant u \leqslant \beta$ on $\overline{\Omega} \times [t_0, \tau_n]$. By Theorem 2.6, $\mathcal{K}_n$ is a continuum in $C^{1,0}(\overline{\Omega} \times [t_0, \tau_n])$. For every $m < n$, let also $\pi_m^n : \mathcal{K}_n \to \mathcal{K}_m$ be the restriction map on $\overline{\Omega} \times [t_0, \tau_m]$, i.e. $\pi_m^n(u) = u|_{\overline{\Omega}\times[t_0,\tau_m]}$ for all $u \in \mathcal{K}_n$. For each $u \in \mathcal{K}$, let us set now $u_n = u|_{\overline{\Omega}\times[t_0,\tau_n]}$ and define a function $\chi : \mathcal{K} \to \prod_{n=1}^{+\infty} \mathcal{K}_n$, by $\chi(u) = (u_n)_n$. Observe that χ is a homeomorphism of $\mathcal{K}$ into $\prod_{n=1}^{+\infty} \mathcal{K}_n$, when $\prod_{n=1}^{+\infty} \mathcal{K}_n$ is endowed with the Tychonoff product topology, and its range $\chi(\mathcal{K})$ is the set of all sequences $(u_n)_n \in \prod_{n=1}^{+\infty} \mathcal{K}_n$ such that, for all $m < n$, $\pi_m^n(u_n) = u_m$, i.e. $\chi(\mathcal{K})$ is the inverse limit of the sequence $(\mathcal{K}_n)_n$ with bonding maps π_m^n (see [47]). As the inverse limit of a sequence of continua is a continuum [47, Theorem 6.1.20] we deduce that $\chi(\mathcal{K})$ is a continuum as well. Since $\mathcal{K}$ is homeomorphic to $\chi(\mathcal{K})$, the conclusion is achieved. □

Continuation and ultimate behaviour of local solutions

The following two results concern existence, continuation and ultimate behaviour of local solutions of problem (2.13); although they are basically known, we provide here simple proofs based on Proposition 2.5 and Theorem 2.6.

LEMMA 2.8. *Assume* (D), (A), (C), (N) *and let* $u_0 \in W_p^{2-2/p}(\Omega)$ *be such that* $u_0 = 0$ *on* $\partial\Omega$. *Fix* $t_0 \in \mathbb{R}$ *and* $t_1 \in]t_0, +\infty]$. *If* α *is a lower solution of* (2.13) *in* $\Omega \times]t_0, t_1[$, *then there exist* $\omega \in]t_0, t_1]$ *and a solution* v *of* (2.13) *in* $\overline{\Omega} \times [t_0, \omega[$, *with* $v \geqslant \alpha$, *such that every upper solution* β *of* (2.13) *in* $\overline{\Omega} \times [t_0, \sigma]$, *with* $\beta \geqslant \alpha$, *satisfies* $\sigma \leqslant \omega$ *and* $\beta \geqslant v$ *in* $\overline{\Omega} \times [t_0, \sigma]$. *Similarly, if* β *is an upper solution of* (2.13) *in* $\Omega \times]t_0, t_1[$, *then there exist* $\omega \in]t_0, t_1]$ *and a solution* w *of* (2.13) *in* $\overline{\Omega} \times [t_0, \omega[$, *with* $w \leqslant \beta$, *such that every lower solution* α *of* (2.13) *in* $\overline{\Omega} \times [t_0, \sigma]$, *with* $\alpha \leqslant \beta$, *satisfies* $\sigma \leqslant \omega$ *and* $\alpha \leqslant w$ *in* $\overline{\Omega} \times [t_0, \sigma]$.

PROOF. The proof is carried out through two steps.

Step 1. *Existence of a local solution $u \geqslant \alpha$.* Let $\alpha = \max_{1\leqslant i\leqslant m} \alpha_i$ in $\overline{\Omega} \times [t_0, \sigma_1]$, where $\sigma_1 > t_0$ comes from the definition of a lower solution. Define, for each $i \in \{1, \dots, m\}$,

$$g_i(x,t,s,\xi) = \begin{cases} f\big(x,t,\alpha_i(x,t),\xi\big) + \omega_i\big(x,t,\alpha_i(x,t) - s\big) & \text{if } s < \alpha_i(x,t), \\ f(x,t,s,\xi) & \text{if } s \geqslant \alpha_i(x,t), \end{cases}$$

where

$$\begin{aligned} &\omega_i(x,t,\delta) \\ &\quad = \max_{|\xi|\leqslant\delta} \big| f\big(x,t,\alpha_i(x,t), \nabla_x\alpha_i(x,t) + \xi\big) - f\big(x,t,\alpha_i(x,t), \nabla_x\alpha_i(x,t)\big)\big| \end{aligned}$$

for a.e. $(x,t) \in \Omega \times]t_0, \sigma_1[$ and every $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$. Fix a constant

$$R > \max\big\{\|\alpha\|_{L^\infty(\Omega\times]t_0,\sigma_1[)}, \|u_0\|_{L^\infty(\Omega)}\big\}$$

and consider the modified problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = F(x,t,u,\nabla_x u) & \text{in } \Omega \times]t_0,\sigma_1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0,\sigma_1], \\ u(\cdot,t_0) = u_0 & \text{in } \Omega, \end{array} \tag{2.35}$$

where

$$F(x,t,s,\xi) = \begin{cases} \max\limits_{1\leqslant i\leqslant m} g_i(x,t,s,\xi) & \text{if } s \leqslant \alpha(x,t), \\ f(x,t,s,\xi) & \text{if } \alpha(x,t) < s < R, \\ \big(R + 1 - |s|\big) f(x,t,s,\xi) & \text{if } R \leqslant s \leqslant R+1, \\ 0 & \text{if } s > R+1, \end{cases}$$

for a.e. $(x,t) \in \Omega \times]t_0,\sigma_1[$ and every $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$. Problem (2.35) admits any large positive constant as an upper solution. Therefore, by Theorem 2.6, it has at least one solution, which gives rise to a local solution of (2.13) defined in $\overline{\Omega} \times [t_0, \tau]$, for some $\tau > t_0$.

Step 2. *Existence of ω and v.* Let ω be the supremum of all $\sigma \in]t_0, t_1[$ such that problem (2.13) has a solution u in $\overline{\Omega} \times [t_0, \sigma]$ with $u \geqslant \alpha$. Let $(\sigma_n)_n$ be a strictly increasing sequence converging to ω and $(u_n)_n$ be the corresponding sequence of solutions of (2.13) defined in $\overline{\Omega} \times [t_0, \sigma_n]$ with $u_n \geqslant \alpha$. By Theorem 2.6, for each n, there exists the minimum solution v_n in $[\alpha, u_n]$ of (2.13) in $\overline{\Omega} \times [t_0, \sigma_n]$. As in the proof of Corollary 2.7, we see that the function $v : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$, defined by setting $v(x,t) = v_n(x,t)$ in $\overline{\Omega} \times [t_0, \sigma_n]$, is the minimum solution in $[\alpha, +\infty[$ of (2.13) in $\overline{\Omega} \times [t_0, \omega[$. Finally, we notice that every upper solution β of (2.13) in $\overline{\Omega} \times [t_0, \sigma]$, with $\beta \geqslant \alpha$, is such that $\sigma \leqslant \omega$ and $\beta \geqslant v$ in $\overline{\Omega} \times [t_0, \sigma]$. Indeed, there is a solution z of (2.13) in $\overline{\Omega} \times [t_0, \sigma]$, satisfying $\alpha \leqslant z \leqslant \beta$ and hence, by construction of v, $z \geqslant v$ in $\overline{\Omega} \times [t_0, \sigma]$. □

PROPOSITION 2.9. *Assume* (D), (A), (C), (N) *and let $u_0 \in W_p^{2-2/p}(\Omega)$ be such that $u_0 = 0$ on $\partial\Omega$. Then problem* (2.13) *has at least one solution $u : \overline{\Omega} \times [t_0, \tau] \to \mathbb{R}$, for some $\tau > t_0$,* *and every such a solution can be continued to a nonextendible solution. Further every nonextendible solution* $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$, *with* $\omega < +\infty$, *satisfies*

$$\limsup_{t \to \omega} \|u(\cdot, t)\|_{L^\infty(\Omega)} = +\infty.$$

PROOF. We proceed through three steps.

Step 1. *Existence of local solutions*. We fix a constant $R > \|u_0\|_{L^\infty(\Omega)}$ and we set, for a.e. $(x, t) \in \Omega \times \mathbb{R}$ and every $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$,

$$F(x,t,s,\xi) = \begin{cases} f(x,t,s,\xi) & \text{if } |s| < R, \\ (R + 1 - |s|) f(x,t,s,\xi) & \text{if } R \leqslant |s| \leqslant R + 1, \\ 0 & \text{if } |s| > R + 1. \end{cases}$$

The corresponding initial value problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x) u = F(x,t,u,\nabla_x u) & \text{in } \Omega \times]t_0, t_0 + 1[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, t_0 + 1], \\ u(\cdot, t_0) = u_0 & \text{in } \Omega, \end{array}$$

admits any large negative constant as a lower solution and any large positive constant as an upper solution. Therefore, by Theorem 2.6, it has a solution, which gives rise to a local solution of (2.13) defined on $\overline{\Omega} \times [t_0, \tau]$, for some $\tau > 0$.

Step 2. *Existence of nonextendible solutions*. Let u be a solution of (2.13) in $\overline{\Omega} \times [t_0, \tau]$, with $\tau \in]t_0, +\infty[$. Consider the set $\mathcal{V}$ of all functions $v : \overline{\Omega} \times [t_0, \rho[\to \mathbb{R}$, for some $\rho > \tau$, such that v is a solution of (2.13) in $\overline{\Omega} \times [t_0, \sigma]$, for every $\sigma < \rho$, and $v|_{\overline{\Omega} \times [t_0, \tau]} = u$. This set $\mathcal{V}$ is non-empty, since $u(\cdot, \tau) \in W_p^{2-2/p}(\Omega)$ and $u(\cdot, \tau) = 0$ on $\partial\Omega$ and therefore u can be locally continued to the right of τ. If $v_1, v_2 \in \mathcal{V}$, we set $v_1 \prec v_2$ whenever $\rho_1 \leqslant \rho_2$ and $v_2|_{\overline{\Omega} \times [t_0, \rho_1]} = v_1$. It is easily verified that $\mathcal{V}$ is inductively ordered. Hence Zorn Lemma yields the existence of maximal elements in $\mathcal{V}$, i.e. the existence of nonextendible solutions of (2.13).

Step 3. *Behaviour of nonextendible solutions*. Let $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$, with $\omega < +\infty$, be a nonextendible solution of (2.13) and assume that

$$\limsup_{t \to \omega} \|u(\cdot, t)\|_{L^\infty(\Omega)} < +\infty,$$

i.e. there is a constant $M > 0$ such that $|u(x, t)| \leqslant M$ in $\overline{\Omega} \times [t_0, \omega[$. By the Nagumo condition (N), there exist $h \in L^p(\Omega \times]t_0, \omega[)$ and $K > 0$ such that

$$|f(x,t,u,\nabla_x u)| \leqslant h(x,t) + K |\nabla_x u|^2$$

for a.e. $(x, t) \in \Omega \times]t_0, \omega[$. Proposition 2.5 implies that $u \in W_p^{2,1}(\Omega \times]t_0, \omega[)$. Hence, in particular, $u(\cdot, \omega) \in W_p^{2-2/p}(\Omega)$ and $u(\cdot, \omega) = 0$ on $\partial\Omega$. Accordingly, u can be continued to the right of ω, thus contradicting the nonextendibility of u. This allows to conclude that

$$\limsup_{t \to \omega} \|u(\cdot, t)\|_{L^\infty(\Omega)} = +\infty.$$

□

REMARK 2.9. Without the Nagumo condition (N), the conclusions of Proposition 2.9 remain true, but now

$$\limsup_{t\to\omega}\|u(\cdot,t)\|_{C^1(\overline{\Omega})}=+\infty.$$

To prove the first step in that case, we define F as in the above proof and

$$\overline{F}(x,t,s,\xi)=\max\bigl\{-m(x,t),\min\bigl\{F(x,t,s,\xi),m(x,t)\bigr\}\bigr\}$$

where $m\in L^p(\Omega\times\,]t_0,t_0+1[)$ satisfies $m(x,t)\geqslant|f(x,t,u,\xi)|+1$, for a.e. $(x,t)\in\Omega\times\,]t_0,t_0+1[$ and every $(u,\xi)\in\mathbb{R}\times\mathbb{R}^N$ with $|u_0(x)-u|\leqslant 1$ and $|\nabla_x u_0(x)-\xi|\leqslant 1$. We then consider the modified problem

$$\begin{array}{ll}\partial_t u+A(x,t,\partial_x)u=\overline{F}(x,t,u,\nabla_x u) & \text{in } \Omega\times\,]t_0,t_0+1[,\\ u=0 & \text{on } \partial\Omega\times[t_0,t_0+1],\\ u(\cdot,t_0)=u_0 & \text{in } \Omega\end{array}$$

and obtain the existence of a local solution as above. The remainder of the proof is similar.

2.4. *The periodic boundary value problem*

Let us consider the periodic problem

$$\begin{array}{ll}\partial_t u+A(x,t,\partial_x)u=f(x,t,u,\nabla_x u) & \text{in } Q_T,\\ u=0 & \text{on } \Sigma_T,\\ u(\cdot,0)=u(\cdot,T) & \text{in } \Omega,\end{array}\tag{2.36}$$

where, we recall, Q_T and Σ_T stand for $\Omega\times\,]0,T[$ and $\partial\Omega\times[0,T]$ respectively. When dealing with problem (2.36) it is convenient to introduce the following space of functions, satisfying the Dirichlet-periodic boundary conditions,

$$C_B^{1,0}\bigl(\overline{Q}_T\bigr)=\bigl\{u\in C^{1,0}\bigl(\overline{Q}_T\bigr)\mid u=0\text{ on }\Sigma_T\text{ and }u(\cdot,0)=u(\cdot,T)\text{ in }\Omega\bigr\},$$

endowed with the $C^{1,0}$-norm. We notice that, for $u\in C_B^{1,0}(\overline{Q}_T)$, we have $u\gg 0$ if and only if $u(x,t)>0$ in $\Omega\times[0,T]$ and $\partial_\nu u(x,t)<0$ on $\partial\Omega\times[0,T]$, where $\nu=\nu(x)$ is the unit outer normal to Ω at $x\in\partial\Omega$.

In this section we discuss existence and localization of solutions of the periodic problem (2.36) in the presence of a pair α, β of possibly discontinuous lower and upper solutions. We further introduce the notion of T-monotonicity and we use it to extend to the present context the Monotone Convergence Criterion known in the frame of order preserving discrete-time semidynamical systems (see, e.g., [60, Section 5]). We then give a simple application of this result to the study of a linear periodic parabolic problem. After these preliminaries we prove the main theorems of this section: existence and localization of solutions of (2.36) when $\alpha\leqslant\beta$, or when $\alpha\nleqslant\beta$. It is worthy to notice again that for problem

(2.36) the mere existence of a pair of lower and upper solutions, even satisfying $\alpha > \beta$, does not generally guarantee the existence of a solution.

Solutions and lower and upper solutions

DEFINITION 2.11. A *solution of* (2.36) is a solution u of (2.10) in $\overline{\Omega} \times \mathbb{R}$ which is T-periodic with respect to the second variable, i.e. $u(x, t + T) = u(x, t)$ in $\overline{\Omega} \times \mathbb{R}$.

DEFINITION 2.12.

- A *lower solution of* (2.36) is a function $\alpha : \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$, T-periodic with respect to the second variable, which is a lower solution of (2.10) in $\overline{\Omega} \times \mathbb{R}$.
- A *regular lower solution of* (2.36) is a lower solution α of (2.36) such that $\alpha|_{\overline{\Omega}\times[0,T]}$ is a regular lower solution of (2.10) in $\overline{\Omega} \times [0, T]$.
- An *upper solution of* (2.36) is a function $\beta : \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$, T-periodic with respect to the second variable, which is an upper solution of (2.10) in $\overline{\Omega} \times \mathbb{R}$.
- A *regular upper solution of* (2.36) is an upper solution β of (2.36) such that $\beta|_{\overline{\Omega}\times[0,T]}$ is a regular upper solution of (2.10) in $\overline{\Omega} \times [0, T]$.
- A lower solution of (2.36) (respectively an upper solution of (2.36)) is *proper* if it is not a solution of (2.36).
- A proper lower solution α of (2.36) is *strict* if every solution u of (2.36), with $u > \alpha$, is such that $u \gg \alpha$. Similarly, a proper upper solution β of (2.36) is *strict* if every solution u of (2.36), with $u < \beta$, is such $u \ll \beta$.

REMARK 2.10. We notice that even regular lower and upper solutions of (2.36) may be discontinuous with respect to t at the endpoints of the interval $[0, T]$.

REMARK 2.11. Sometimes we speak of solutions of (2.36) with reference to functions defined on $\overline{\Omega} \times [t_0, t_0 + T]$, and of lower and upper solutions of (2.36) with reference to functions defined on $\overline{\Omega} \times [t_0, t_0 + T[$ for some $t_0 \in \mathbb{R}$. In these cases it is understood that their T-periodic extensions to $\overline{\Omega} \times \mathbb{R}$ have to be considered.

REMARK 2.12. If α is a lower solution of (2.36) and $u \in C^{1,0}_B(\overline{Q}_T)$ is such that $u > \alpha$, then there exist an open ball $B \subseteq \Omega$ and points $t_1, t_2 \in [0, T]$, with $t_1 < t_2$, such that $u(x, t) > \alpha(x, t)$ in $B \times]t_1, t_2[$. If α is a lower solution of (2.36) and $u \in C^{1,0}_B(\overline{Q}_T)$ is such that $u \not\geq \alpha$, then there exist an open ball $B \subseteq \Omega$ and points $t_1, t_2 \in [0, T]$, with $t_1 < t_2$, such that $u(x, t) < \alpha(x, t)$ in $B \times]t_1, t_2[$. Similar conclusions hold for an upper solution β and a function $u \in C^{1,0}_B(\overline{Q}_T)$, such that $u < \beta$, or $u \not\leq \beta$.

A further Nagumo-type result

In the frame of problem (2.36) we need a Nagumo-type result that is a variant of Proposition 2.5, where the $W^{2-2/p}_p$-bound on the initial condition is replaced by a L^∞-bound on the solution itself.

PROPOSITION 2.10. *Assume* (D) *and* (A). *Let* $p > N+2$, $h \in L^p(\Omega \times]-T,T[)$, $K > 0$ *and* $R > 0$ *be given. Then there exists a constant* $C > 0$ *such that, for each* $u \in W_p^{2,1}(\Omega \times]-T,T[)$ *satisfying*

$$\begin{array}{ll} \left|\partial_t u + A(x,t,\partial_x)u\right| \leqslant h(x,t) + K|\nabla_x u|^2 & \textit{a.e. in } \Omega \times]-T,T[, \\ u = 0 & \textit{on } \partial\Omega \times [-T,T], \end{array}$$

and

$$\|u\|_{L^\infty(\Omega\times]-T,T[)} \leqslant R,$$

we have

$$\|u\|_{W_p^{2,1}(\Omega\times]0,T[)} \leqslant C.$$

PROOF. As in the proof of Proposition 2.5, it is enough to prove the following result.

CLAIM. *Let* $g \in L^\infty(\Omega \times]-T,T[)$ *and* $R > 0$ *be given. Then there is a constant* $C > 0$ *such that every* $u \in W_p^{2,1}(\Omega \times]-T,T[)$, *for which*

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = g(x,t)\big(1 + |\nabla_x u|^2\big) & \textit{in } \Omega \times]-T,T[, \\ u = 0 & \textit{on } \partial\Omega \times [-T,T], \end{array} \tag{2.37}$$

and

$$\|u\|_{L^\infty(\Omega\times]-T,T[)} \leqslant R \tag{2.38}$$

hold, satisfies

$$\|u\|_{W_p^{2,1}(\Omega\times]-T,T[)} \leqslant C.$$

Fix a function u satisfying (2.37) and (2.38), and consider, for each $\lambda \in [0,1]$, the initial value problem

$$\begin{array}{ll} \partial_t u_\lambda + A(x,t,\partial_x)u_\lambda = g(x,t)\big(\lambda + |\nabla_x u_\lambda|^2\big) & \text{in } \Omega \times]-T,T[, \\ u_\lambda = 0 & \text{on } \partial\Omega \times [-T,T], \\ u_\lambda(\cdot,-T) = \lambda u(\cdot,-T) & \text{in } \Omega. \end{array} \tag{2.39}$$

As in the proof of Proposition 2.5, we see that (2.39) has at most one solution. Let $\eta = \operatorname{ess\,inf}_{\Omega\times]-T,T[} a_0 > 0$ and $M = \max\{\eta^{-1}\|g\|_{L^\infty(\Omega\times]-T,T[)}, R\}$. Since $-M$ is a lower solution and M is an upper solution of (2.39), Theorem 2.6 implies that, for every $\lambda \in [0,1]$, (2.39) has a unique solution u_λ.

Moreover, as in Proposition 2.5, we can prove that, for every $\lambda_1, \lambda_2 \in [0,1]$, the corresponding solutions $u_{\lambda_1}, u_{\lambda_2}$ of (2.39) satisfy

$$\|u_{\lambda_1} - u_{\lambda_2}\|_{L^\infty(\Omega\times]-T,T[)} \leqslant C_2|\lambda_1 - \lambda_2|. \tag{2.40}$$

Next, we set in $\Omega \times [-T, T]$,

$$v_\lambda(x,t) = \chi(t)u_\lambda(x,t),$$

where $\chi \in C^\infty(\mathbb{R}, [0,1])$ is such that

$$\chi(t) = \begin{cases} 0 & \text{if } t \leqslant -T/2, \\ 1 & \text{if } t \geqslant 0. \end{cases}$$

Each v_λ satisfies

$$\begin{array}{ll} \partial_t v_\lambda + A(x,t,\partial_x)v_\lambda = \chi(t)g(x,t)\big(\lambda + |\nabla_x u_\lambda|^2\big) + \chi'(t)u_\lambda & \text{in } \Omega \times]-T, T[, \\ v_\lambda = 0 & \text{on } \partial\Omega \times [-T, T], \\ v_\lambda(\cdot, -T) = 0 & \text{in } \Omega. \end{array} \tag{2.41}$$

For each $\lambda_1, \lambda_2 \in [0,1]$, set $w = v_{\lambda_1} - v_{\lambda_2} = \chi(u_{\lambda_1} - u_{\lambda_2})$, which satisfies

$$\begin{array}{ll} \partial_t w + A(x,t,\partial_x)w = \chi g(\lambda_1 - \lambda_2) + g\nabla_x(u_{\lambda_1} + u_{\lambda_2}) \cdot \nabla_x w + \chi'(u_{\lambda_1} - u_{\lambda_2}) & \\ \qquad\qquad\qquad \text{in } \Omega \times]-T, T[, & \\ w = 0 \qquad\qquad \text{on } \partial\Omega \times [-T, T], & \\ w(\cdot, -T) = 0 \quad \text{in } \Omega. & \end{array} \tag{2.42}$$

and, by (2.40),

$$\|w\|_{L^\infty(\Omega\times]-T,T[)} \leqslant C_2|\lambda_1 - \lambda_2|.$$

Since

$$\begin{aligned} \big|\nabla_x(u_{\lambda_1} + u_{\lambda_2}) \cdot \nabla_x w\big| &= \chi\big|\nabla_x(u_{\lambda_1} + u_{\lambda_2}) \cdot \nabla_x(u_{\lambda_1} - u_{\lambda_2})\big| \\ &\leqslant \chi|\nabla_x u_{\lambda_1}|^2 + 2\chi\big|\nabla_x(u_{\lambda_1} - u_{\lambda_2})\big|^2, \end{aligned}$$

we get

$$\begin{aligned} &\big\|\partial_t w + A(x,t,\partial_x)w\big\|_{L^p(\Omega\times]-T,T[)} \\ &\quad \leqslant C_4\big(\big\|\chi|\nabla_x u_{\lambda_1}|^2\big\|_{L^p(\Omega\times]-T,T[)} + \big\|\chi|\nabla_x(u_{\lambda_1} - u_{\lambda_2})|^2\big\|_{L^p(\Omega\times]-T,T[)}\big) + C_3. \end{aligned} \tag{2.43}$$

As in the proof of Proposition 2.5, we also have, by (2.40),

$$\begin{aligned}&\big\|\chi\,|\nabla_x(u_{\lambda_1}-u_{\lambda_2})|^2\big\|_{L^p(\Omega\times]-T,T[)}\\&\quad\leqslant C_5\|u_{\lambda_1}-u_{\lambda_2}\|_{L^\infty(\Omega\times]-T,T[)}\big\|\chi(u_{\lambda_1}-u_{\lambda_2})\big\|_{W_p^{2,1}(\Omega\times]-T,T[)}\\&\quad\leqslant C_2C_5|\lambda_1-\lambda_2|\|w\|_{W_p^{2,1}(\Omega\times]-T,T[)},\end{aligned}\tag{2.44}$$

and in particular,

$$\big\|\chi|\nabla_x u_{\lambda_1}|^2\big\|_{L^p(\Omega\times]-T,T[)}\leqslant C_2C_5\lambda_1\|v_{\lambda_1}\|_{W_p^{2,1}(\Omega\times]-T,T[)}.\tag{2.45}$$

Using (2.43), (2.44), (2.45) and Proposition 2.4, we get

$$\begin{aligned}&\|w\|_{W_p^{2,1}(\Omega\times]-T,T[)}\\&\quad\leqslant C_6+C_7\|v_{\lambda_1}\|_{W_p^{2,1}(\Omega\times]-T,T[)}+C_8|\lambda_1-\lambda_2|\|w\|_{W_p^{2,1}(\Omega\times]-T,T[)}.\end{aligned}$$

This also yields

$$\|w\|_{W_p^{2,1}(\Omega\times]-T,T[)}\leqslant 2C_6+2C_7\|v_{\lambda_1}\|_{W_p^{2,1}(\Omega\times]-T,T[)}$$

provided that $|\lambda_1-\lambda_2|\leqslant\varepsilon=(2C_8)^{-1}$. Hence, in particular, we have, for each $\lambda\in[0,\varepsilon]$,

$$\|v_\lambda\|_{W_p^{2,1}(\Omega\times]-T,T[)}\leqslant 2C_6.$$

Finally, iterating this process a finite number of times, we conclude that there is a constant C_9 such that

$$\|v_1\|_{W_p^{2,1}(\Omega\times]-T,T[)}\leqslant C_9$$

and hence

$$\|u\|_{W_p^{2,1}(\Omega\times]0,T[)}\leqslant C_9.$$ □

T-monotonicity and asymptotic behaviour of bounded solutions

DEFINITION 2.13. A function $u:\overline{\Omega}\times I\to\mathbb{R}$, where $I\subseteq\mathbb{R}$ is an interval, is said *T-increasing* (respectively *T-decreasing*) if $u(x,t)\leqslant u(x,t+T)$ (respectively $u(x,t+T)\geqslant u(x,t)$) for every $x\in\overline{\Omega}$ and each $t\in I$ such that $t+T\in I$. A function is said *T-monotone* if it is either T-increasing or T-decreasing.

REMARK 2.13. Let u be a solution of (2.10) on $\overline{\Omega}\times[t_0,+\infty[$. If u is T-increasing then the sequence $(\alpha_n)_n$, defined by setting, for each $n\geqslant t_0/T$, $\alpha_n(x,t)=u(x,t+nT)$ in $\overline{\Omega}\times[0,T[$, gives rise to an increasing sequence of regular lower solutions of (2.36). Similarly, any T-decreasing solution of (2.10) in $\overline{\Omega}\times[t_0,+\infty[$, gives rise to a decreasing sequence $(\beta_n)_n$ of regular upper solutions of (2.36).

The T-monotonicity property is a basic tool to prove convergence of a bounded solution u defined on $\overline{\Omega} \times [t_0, +\infty[$ to a T-periodic solution v. We start with a simple preliminary result where the T-monotonicity of u is replaced by a weaker request: the monotonicity of the sequence $(u(\cdot, nT))_n$, with $n \geqslant t_0/T$. However, in this case only the existence of a T-periodic solution v lying in the ω-limit set of u is guaranteed, not the convergence of u to v.

LEMMA 2.11. *Assume* (D), (A), (C) *and* (N). *Let u be a solution of* (2.10) *in $\overline{\Omega} \times [t_0, +\infty[$, which is bounded in $\overline{\Omega} \times [t_0, +\infty[$ and such that either $u(\cdot, nT) \leqslant u(\cdot, (n+1)T)$ in $\overline{\Omega}$ for every $n \geqslant t_0/T$, or $u(\cdot, nT) \geqslant u(\cdot, (n+1)T)$ in $\overline{\Omega}$ for every $n \geqslant t_0/T$. Then problem* (2.36) *has at least one solution.*

PROOF. Assume that $u(\cdot, nT) \leqslant u(\cdot, (n+1)T)$ in $\overline{\Omega}$ for every $n \geqslant t_0/T$. Define a sequence $(u_n)_n$ by setting $u_n(x,t) = u(x, t + nT)$ in $\overline{Q}_T$. By Proposition 2.10, we know that the sequence $(u_n)_n$ is bounded in $W^{2,1}_p(Q_T)$. Hence there exists a subsequence $(u_{n_k})_k$ which converges to a solution v of (2.10) in $\overline{Q}_T$ weakly in $W^{2,1}_p(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$. Moreover, as the sequences $(u_n(\cdot, 0))_n$ and $(u_n(\cdot, T))_n$ satisfy $u_n(\cdot, 0) \leqslant u_n(\cdot, T) = u_{n+1}(\cdot, 0)$, we have $u_{n_k}(\cdot, 0) \leqslant u_{n_k}(\cdot, T) \leqslant u_{n_{k+1}}(\cdot, 0)$ and, passing to the limit, $v(\cdot, 0) \leqslant v(\cdot, T) \leqslant v(\cdot, 0)$. Therefore v is a solution of (2.36). □

PROPOSITION 2.12. *Assume* (D), (A), (C) *and* (N). *Let u be a solution of* (2.10) *in $\overline{\Omega} \times [t_0, +\infty[$, which is bounded in $\bar{\Omega} \times [t_0, +\infty[$ and T-monotone. Then there exists a solution v of* (2.36) *such that*

$$\lim_{t \to +\infty} \left\| u(\cdot, t) - v(\cdot, t) \right\|_{C^1(\overline{\Omega})} = 0.$$

PROOF. Assume that u is T-increasing. Define a sequence $(u_n)_n$ by setting, for every $n \geqslant t_0/T$, $u_n(x,t) = u(x, t + nT)$ in $\overline{Q}_T$. From the proof of Lemma 2.11, we know that a subsequence of $(u_n)_n$ converges to a solution v of (2.36) weakly in $W^{2,1}_p(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$. The monotonicity of $(u_n)_n$ then implies that the whole sequence $(u_n)_n$ converges to v. It remains to prove that

$$\lim_{t \to +\infty} \left\| u(\cdot, t) - v(\cdot, t) \right\|_{C^1(\overline{\Omega})} = 0.$$

Indeed, since $(u_n)_n$ converges in $C^{1,0}(\overline{Q}_T)$ to v, given $\varepsilon > 0$ there is $\bar{n}$ such that, for any $n \geqslant \bar{n}$,

$$\left\| u_n(\cdot, t) - v(\cdot, t) \right\|_{C^1(\bar{\Omega})} < \varepsilon, \quad \text{for all } t \in [0, T].$$

Hence, if we take $t \geqslant \bar{n}T$, with $t \in [nT, (n+1)T[$ for some $n \geqslant \bar{n}$, we obtain

$$\left\| u(\cdot, t) - v(\cdot, t) \right\|_{C^1(\overline{\Omega})} = \left\| u_n(\cdot, t - nT) - v(\cdot, t - nT) \right\|_{C^1(\overline{\Omega})} < \varepsilon,$$

by the periodicity of v. □

A Monotone Convergence Criterion

We now prove that any bounded solution of (2.10) emanating from and lying above a lower solution of (2.36), or respectively emanating from and lying below an upper solution of (2.36), is T-monotone and hence converges to a T-periodic solution. We point out that, in space dimension $N \geqslant 2$, convergence to a T-periodic solution is not the behaviour of all bounded solutions of (2.10), due to the possible occurrence of (linearly stable) subharmonic solutions (see, e.g., [141,35] and [123, Section 3]). In our frame additional difficulties come from the lack of regularity of f, so that even generic convergence results, possibly to subharmonic solutions of (2.10), are not available (see [123, Section 3]). For a discussion of the regular case in space dimension $N = 1$ we refer to [56].

PROPOSITION 2.13. *Assume* (D), (A), (C) *and* (N).

(i) *Let* α *be a lower solution of* (2.36) *such that, for some* t_0, $\alpha(\cdot, t_0) \in W_p^{2-2/p}(\Omega)$ *and* $\alpha(\cdot, t_0) = 0$ *on* $\partial\Omega$. *Then there exist* $\omega \in\,]t_0, +\infty]$ *and a* T*-increasing solution* $\widetilde{\alpha} : \overline{\Omega} \times [t_0, \omega[\, \to \mathbb{R}$ *of* (2.13), *with* $u_0 = \alpha(\cdot, t_0)$, *satisfying* $\widetilde{\alpha} \geqslant \alpha$ *in* $\overline{\Omega} \times [t_0, \omega[$. *Further, every solution* $u : \overline{\Omega} \times [t_0, \sigma[\, \to \mathbb{R}$ *of* (2.10), *with* $u \geqslant \alpha$, *is such that* $\sigma \leqslant \omega$ *and* $u \geqslant \widetilde{\alpha}$ *in* $\overline{\Omega} \times [t_0, \sigma[$. *Finally, we have that either every nonextendible solution* $u : \overline{\Omega} \times [t_0, \sigma[\, \to \mathbb{R}$ *of* (2.10), *with* $u \geqslant \alpha$, *is such that*

$$\limsup_{t \to \sigma} \Bigl(\max_{\overline{\Omega}} u(\cdot, t) \Bigr) = +\infty$$

and (2.36) *has no solution in* $[\alpha, +\infty[$, *or there exists the minimum solution* v *in* $[\alpha, +\infty[$ *of* (2.36); *in the latter case* $\omega = +\infty$, $v \geqslant \widetilde{\alpha}$ *and*

$$\lim_{t \to +\infty} \bigl\| \widetilde{\alpha}(\cdot, t) - v(\cdot, t) \bigr\|_{C^1(\overline{\Omega})} = 0.$$

(ii) *Let* β *be an upper solution of* (2.36) *such that, for some* t_0, $\beta(\cdot, t_0) \in W_p^{2-2/p}(\Omega)$ *and* $\beta(\cdot, t_0) = 0$ *on* $\partial\Omega$. *Then there exist* $\omega \in\,]t_0, +\infty]$ *and a* T*-decreasing solution* $\widetilde{\beta} : \overline{\Omega} \times [t_0, \omega[\, \to \mathbb{R}$ *of* (2.13), *with* $u_0 = \beta(\cdot, t_0)$, *satisfying* $\widetilde{\beta} \leqslant \beta$ *in* $\overline{\Omega} \times [t_0, \omega[$. *Further, every solution* $u : \overline{\Omega} \times [t_0, \sigma[\, \to \mathbb{R}$ *of* (2.10), *with* $u \leqslant \beta$, *is such that* $\sigma \leqslant \omega$ *and* $u \leqslant \widetilde{\beta}$ *in* $\overline{\Omega} \times [t_0, \sigma[$. *Finally, we have that either every nonextendible solution* $u : \overline{\Omega} \times [t_0, \sigma[\, \to \mathbb{R}$ *of* (2.10), *with* $u \leqslant \beta$, *is such that*

$$\liminf_{t \to \sigma} \Bigl(\min_{\overline{\Omega}} u(\cdot, t) \Bigr) = -\infty$$

and (2.36) *has no solution in* $]-\infty, \beta]$, *or there exists the maximum solution* w *in* $]-\infty, \beta]$ *of* (2.36); *in the latter case* $\omega = +\infty$, $w \leqslant \widetilde{\beta}$ *and*

$$\lim_{t \to +\infty} \bigl\| \widetilde{\beta}(\cdot, t) - w(\cdot, t) \bigr\|_{C^1(\overline{\Omega})} = 0.$$

PROOF. We only prove statement (i), as the proof of (ii) is similar.

The existence of $\widetilde{\alpha}$ follows from Lemma 2.8. Namely, there exist $\omega \in\,]t_0, +\infty]$ and a solution $\widetilde{\alpha} : \overline{\Omega} \times [t_0, \omega[\, \to \mathbb{R}$ of (2.13), with $u_0 = \alpha(\cdot, t_0)$, satisfying $\widetilde{\alpha} \geqslant \alpha$ in $\overline{\Omega} \times [t_0, \omega[$;

further, every solution $u:\overline{\Omega} \times [t_0, \sigma] \to \mathbb{R}$ of (2.10), with $u \geqslant \alpha$, is such that $\sigma \leqslant \omega$ and $u \geqslant \widetilde{\alpha}$ in $\overline{\Omega} \times [t_0, \sigma]$.

Let us prove that $\widetilde{\alpha}$ is T-increasing. We can assume that $\omega > t_0 + T$ as otherwise there is nothing to prove. Notice first that $\widetilde{\alpha}(\cdot, t_0 + T) \geqslant \alpha(\cdot, t_0 + T) = \alpha(\cdot, t_0) = \widetilde{\alpha}(\cdot, t_0)$ in $\overline{\Omega}$. Assume, by contradiction, there exist $x_1 \in \Omega$ and $t_1 \geqslant t_0$ such that

$$\widetilde{\alpha}(x_1, t_1 + T) < \widetilde{\alpha}(x_1, t_1).$$

If $\omega < +\infty$, we set $\beta(x, t) = \min\{\widetilde{\alpha}(x, t + T), \widetilde{\alpha}(x, t)\}$ in $\overline{\Omega} \times [t_0, \omega - T[$. Then β is an upper solution of (2.13), with $u_0 = \alpha(\cdot, t_0)$, in $\overline{\Omega} \times [t_0, \omega - T[$, satisfying $\beta \geqslant \alpha$. Hence, by Theorem 2.6, there exists a solution w of (2.13), with $u_0 = \alpha(\cdot, t_0)$, in $\overline{\Omega} \times [t_0, \omega - T[$, satisfying $\alpha \leqslant w \leqslant \beta$. If $\omega = +\infty$, we set $\beta(x, t) = \min\{\widetilde{\alpha}(x, t + T), \widetilde{\alpha}(x, t)\}$ in $\overline{\Omega} \times [t_0, +\infty[$. Then β is an upper solution of (2.13), with $u_0 = \alpha(\cdot, t_0)$, in $\overline{\Omega} \times [t_0, +\infty[$, satisfying $\beta \geqslant \alpha$. Hence, by Corollary 2.7, there exists a solution w of (2.13), with $u_0 = \alpha(\cdot, t_0)$, in $\overline{\Omega} \times [t_0, +\infty[$, satisfying $\alpha \leqslant w \leqslant \beta$. In both cases, since $\beta < \widetilde{\alpha}$, we get $w < \widetilde{\alpha}$, thus contradicting the definition of $\widetilde{\alpha}$.

By Proposition 2.9, either every nonextendible solution $u:\overline{\Omega} \times [0, \sigma[\to \mathbb{R}$ of (2.10), with $u \geqslant \alpha$, is such that

$$\limsup_{t \to \sigma} \Big(\max_{\overline{\Omega}} u(\cdot, t) \Big) = +\infty$$

and in particular (2.36) has no solution in $[\alpha, +\infty[$, or $\omega = +\infty$ and $\widetilde{\alpha}$ is bounded in $\overline{\Omega} \times [t_0, +\infty[$. In the latter case, as $\tilde{\alpha}$ is T-increasing, by Proposition 2.12, there exists a solution $v \geqslant \widetilde{\alpha}$ of (2.36) such that

$$\lim_{t \to +\infty} \left\| \widetilde{\alpha}(\cdot, t) - v(\cdot, t) \right\|_{C^1(\overline{\Omega})} = 0.$$

Let us prove that v is the minimum solution of (2.36) in $[\alpha, +\infty[$. Otherwise, if $\bar{v}$ were a solution of (2.36) with $\bar{v} \geqslant \alpha$ and $\bar{v} \not\geqslant v$, then $\min\{v, \bar{v}\}$ should be an upper solution of (2.13), with $u_0 = \alpha(\cdot, t_0)$, in $\overline{\Omega} \times [t_0, +\infty[$, such that $\min\{v, \bar{v}\} \geqslant \alpha$ and $\min\{v, \bar{v}\} \not\geqslant \widetilde{\alpha}$. By Corollary 2.7, there should exist a solution z of (2.13), with $u_0 = \alpha(\cdot, t_0)$, in $\overline{\Omega} \times [t_0, +\infty[$, such that $\alpha \leqslant z \leqslant \min\{v, \bar{v}\}$, thus contradicting the definition of $\widetilde{\alpha}$. □

A linear periodic problem

As a simple application of Proposition 2.13, we get a result that concerns existence, uniqueness and regularity of the solution of the linear periodic problem

$$\begin{aligned} &\partial_t u + A(x, t, \partial_x) u + \sigma u = h(x, t) && \text{in } Q_T, \\ &u = 0 && \text{on } \Sigma_T, \\ &u(\cdot, 0) = u(\cdot, T) && \text{in } \Omega, \end{aligned} \tag{2.46}$$

where $\sigma \geqslant 0$ is fixed and $h \in L^p(Q_T)$ for some $p > N + 2$. Although proofs can be found in [84, Lemma 4.1 and Corollary 5.6], we use here an alternative approach which is more in the spirit of this work and will yield, as a byproduct, the global asymptotic stability of the solution (see Proposition 2.35). We associate with problem (2.46) the linear operator

$$L : W^{2,1}_{p,B}(Q_T) \to L^p(Q_T), \tag{2.47}$$

defined by

$$Lu = \partial_t u - \sum_{i,j=1}^{N} \partial_{x_i}\big(a_{ij}(x,t)\partial_{x_j} u\big) + \sum_{i=1}^{N} a_i(x,t)\partial_{x_i} u + a_0(x,t)u, \tag{2.48}$$

where

$$\begin{aligned} W^{2,1}_{p,B}(Q_T) &= W^{2,1}_p(Q_T) \cap C^{1,0}_B\big(\overline{Q}_T\big) \\ &= \big\{u \in W^{2,1}_p(Q_T) \mid u = 0 \text{ on } \Sigma_T \text{ and } u(\cdot,0) = u(\cdot,T) \text{ in } \Omega\big\}. \end{aligned}$$

Notice that $W^{2,1}_{p,B}(Q_T)$ is a Banach subspace of $W^{2,1}_p(Q_T)$ and L is a bounded operator.

PROPOSITION 2.14. *Assume* (D) *and* (A). *Let* $p > N + 2$ *be fixed. Then, for every* $\sigma \geqslant 0$ *and for any given* $h \in L^p(Q_T)$, *problem* (2.46) *has a unique solution* $u \in W^{2,1}_p(Q_T)$. *Moreover, there exists a constant* C, *independent of* h, *such that*

$$\|u\|_{W^{2,1}_p(Q_T)} \leqslant C\|h\|_{L^p(Q_T)}.$$

PROOF. The proof consists of four steps.

Step 1. *Problem* (2.46) *has at most one solution.* The conclusion immediately follows showing that any function $u \in W^{2,1}_p(Q_T)$, satisfying

$$\begin{aligned} &\partial_t u + A(x,t,\partial_x)u + \sigma u \geqslant 0 && \text{in } Q_T, \\ &u \geqslant 0 && \text{on } \Sigma_T, \\ &u(\cdot,0) = u(\cdot,T) && \text{in } \Omega, \end{aligned}$$

is such that $u \geqslant 0$. Indeed, assume by contradiction that $\min_{\overline{Q}_T} u < 0$. Since $u(\cdot,0) = u(\cdot,T)$ in Ω and $u \geqslant 0$ on Σ_T, there exists $(x_0,t_0) \in \Omega \times\,]0,T]$ such that $u(x_0,t_0) = \min_{\overline{Q}_T} u$. Then Proposition 2.2 implies that u is constant in $\bar{\Omega} \times [0,t_0]$. This yields a contradiction, as then

$$\partial_t u + A(x,t,\partial_x)u + \sigma u = (a_0 + \sigma)u < 0, \quad \text{a.e. in } \Omega \times [0,t_0].$$

Step 2. *Problem* (2.46) *has a solution for any given* $h \geqslant 0$. The function $\alpha = 0$ is obviously a lower solution of (2.46) with $\alpha(\cdot,0) \in W^{2-2/p}_p(\Omega)$ and $\alpha(\cdot,0) = 0$ on $\partial\Omega$. Let us build an upper solution of (2.46). Let z be the unique solution of

$$\begin{aligned} &\partial_t u + A(x,t,\partial_x)u + \sigma u = h(x,t) && \text{in } Q_T, \\ &u = 0 && \text{on } \Sigma_T, \\ &u(\cdot,0) = 0 && \text{in } \Omega. \end{aligned}$$

Define a function $\beta \in W_p^{2,1}(Q_T)$, by setting

$$\beta(x,t) = z(x,t) + at + b \quad \text{in } \overline{\Omega} \times [0, T[,$$

where $a \leqslant 0$ and $b \geqslant 0$ are constants chosen in such a way that

$$aT + \left\| z(\cdot, T) \right\|_{L^\infty(\Omega)} \leqslant 0 \quad \text{and} \quad a + \Bigl(\sigma + \operatorname*{ess\,inf}_{Q_T} a_0\Bigr)(aT + b) \geqslant 0.$$

Accordingly, β satisfies

$$\begin{aligned} &\partial_t \beta + A(x,t,\partial_x)\beta + \sigma\beta \geqslant h(x,t) && \text{in } Q_T, \\ &\beta \geqslant 0 && \text{on } \Sigma_T, \\ &\beta(\cdot, 0) \geqslant \lim_{t \to T^-} \beta(\cdot, t) && \text{in } \Omega \end{aligned}$$

and hence β is an upper solution of (2.46). We apply Corollary 2.7 and Proposition 2.13 to deduce the existence of a T-increasing solution $\widetilde{\alpha} : \overline{\Omega} \times [0, +\infty[\to \mathbb{R}$ of

$$\begin{aligned} &\partial_t u + A(x,t,\partial_x)u + \sigma u = h(x,t) && \text{in } \Omega \times]0, +\infty[, \\ &u = 0 && \text{on } \partial\Omega \times [0, +\infty[, \\ &u(\cdot, 0) = 0 && \text{in } \Omega \end{aligned}$$

and of a solution v of (2.46) in $[0, \beta]$, which by Step 1 is the unique solution of (2.46). Further, we have

$$\lim_{t \to +\infty} \left\| \widetilde{\alpha}(\cdot, t) - v(\cdot, t) \right\|_{C^1(\bar{\Omega})} = 0.$$

Step 3. *Problem* (2.46) *has a solution for any given h*. Let us define the functions $h^+ = \max\{h, 0\}$, $h^- = h^+ - h$ and consider the periodic problems

$$\begin{aligned} &\partial_t u + A(x,t,\partial_x)u + \sigma u = h^+(x,t) && \text{in } Q_T, \\ &u = 0 && \text{on } \Sigma_T, \\ &u(\cdot, 0) = u(\cdot, T) && \text{in } \Omega \end{aligned}$$

and

$$\begin{aligned} &\partial_t u + A(x,t,\partial_x)u + \sigma u = h^-(x,t) && \text{in } Q_T, \\ &u = 0 && \text{on } \Sigma_T, \\ &u(\cdot, 0) = u(\cdot, T) && \text{in } \Omega. \end{aligned}$$

Denote by u_+ and u_- the corresponding solutions, constructed in Step 2. By linearity, we conclude that $u = u_+ - u_-$ is the solution of (2.46).

Step 4. *Continuous dependence.* Since the linear operator $L + \sigma I : W^{2,1}_{p,B}(Q_T) \to L^p(Q_T)$ is continuous, one-to-one and onto and its domain $W^{2,1}_{p,B}(Q_T)$ is a Banach space, the Open Mapping Theorem yields the estimate

$$\|u\|_{W^{2,1}_p(Q_T)} \leqslant C\|h\|_{L^p(Q_T)}. \qquad \square$$

Existence in case $\alpha \leqslant \beta$

The following result yields the existence of extremal solutions of (2.36) in the presence of lower and upper solutions α, β satisfying the condition $\alpha \leqslant \beta$. This extends to strong solutions previous results obtained, e.g., in [73,8,6,33,58] for classical solutions. For later use we also evaluate the degree of the fixed point operator associated with (2.36), when the lower and the upper solutions are strict. To this end we denote by $S : C^{1,0}_B(\overline{Q}_T) \to C^{1,0}_B(\overline{Q}_T)$ the operator, which according to Proposition 2.14 sends any function $u \in C^{1,0}_B(\overline{Q}_T)$ onto the unique solution $v \in W^{2,1}_p(Q_T)$ of

$$\begin{array}{ll} \partial_t v + A(x,t,\partial_x)v = f(x,t,u,\nabla_x u) & \text{in } Q_T, \\ v = 0 & \text{on } \Sigma_T, \\ v(\cdot,0) = v(\cdot,T) & \text{in } \Omega. \end{array} \tag{2.49}$$

THEOREM 2.15. *Assume* (D), (A), (C) *and* (N). *Suppose that α is a lower solution and β is an upper solution of* (2.36) *satisfying*

$$\alpha \leqslant \beta.$$

Then there exist the minimum solution v and the maximum solution w of (2.36) *in* $[\alpha, \beta]$. *Further, if α and β are strict, then the operator S defined by* (2.49) *has no fixed point on the boundary of*

$$\mathcal{U} = \left\{u \in C^{1,0}_B\left(\overline{Q}_T\right) \mid \alpha \ll u \ll \beta\right\},$$

the set of fixed points of S in $\mathcal{U}$ is bounded in $C^{1,0}_B(\overline{Q}_T)$ and

$$\deg\bigl(I - S, \mathcal{U} \cap B(0,R)\bigr) = 1,$$

where I is the identity operator in $C^{1,0}_B(\overline{Q}_T)$ and $B(0,R)$ is the open ball of center 0 *and radius R in $C^{1,0}_B(\overline{Q}_T)$, with R so large that all fixed points of S in $\mathcal{U}$ belong to $B(0,R)$.*

PROOF. The proof is divided into three parts.

Part 1. *Existence of a solution u of* (2.36) *with $\alpha \leqslant u \leqslant \beta$.* According to Remark 2.4, we can assume that the sequences $(\sigma_h)_{0\leqslant h\leqslant k}$ and $(\rho_h)_{0\leqslant h\leqslant l}$ coincide. We also notice that, by our definition of lower and upper solutions of (2.36) and the condition $\alpha \leqslant \beta$, there exists $M > 0$ such that $-M \leqslant \alpha(x,t) \leqslant \beta(x,t) \leqslant M$ in $\overline{Q}_T$. Hence, by condition (N) and

Proposition 2.10, there exists $R > 0$ such that, for every function f satisfying (2.11) and every solution u of (2.36) with $\alpha \leqslant u \leqslant \beta$, we have

$$\|u\|_{C^{1,0}(\overline{Q}_T)} < R. \tag{2.50}$$

Exactly as in the first part of the proof of Theorem 2.6, we define a function $F : Q_T \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ with $\Omega \times]t_0, t_1[$ replaced by $\Omega \times]0, T[$. Then we consider the modified problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = F(x,t,u,\nabla_x u) & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega. \end{array} \tag{2.51}$$

Step 1. *Every solution u of* (2.51) *satisfies $\alpha \leqslant u \leqslant \beta$.* This can be proved as in the corresponding step of the proof of Theorem 2.6. We just observe that, if u is a solution of (2.51) such that $\inf_{Q_T}(u - \alpha) < 0$, then by our definition of a lower solution of (2.51) either there exists $h \in \{0, \dots, k-1\}$ such that $\inf_{\Omega\times]\sigma_h,\sigma_{h+1}[}(u - \alpha) < 0$ and, for all $x \in \Omega$, $\lim_{t\to\sigma_h^+}(u - \alpha)(x,t) > \inf_{\Omega\times]\sigma_h,\sigma_{h+1}[}(u - \alpha)$, or $\inf_{\Omega\times]\sigma_h,\sigma_{h+1}[}(u - \alpha) = \inf_{\Omega\times]\sigma_{h-1},\sigma_h[}(u - \alpha)$ for every $h \in \{1, \dots, k-1\}$. In any case we can find $h \in \{0, \dots, k-1\}$ and $i \in \{1, \dots, m_h\}$ such that, setting $v = u - \alpha_i^{(h)} \in W_p^{2,1}(\Omega \times]\sigma_h, \sigma_{h+1}[)$, we have $v(x,t) \geqslant 0$ on $\partial\Omega \times [\sigma_h, \sigma_{h+1}]$ and there exists $(\bar{x}, \bar{t}) \in \Omega \times]\sigma_h, \sigma_{h+1}]$ such that $\min_{\overline{\Omega}\times[\sigma_h,\sigma_{h+1}]} v = v(\bar{x}, \bar{t}) < 0$.

Step 2. *Every solution u of* (2.51) is *a solution of* (2.36). Again this can be proved as in the corresponding step of the proof of Theorem 2.6.

Step 3. *Problem* (2.51) *has at least one solution.* Let us consider the solution operator $\overline{S} : C_B^{1,0}(\overline{Q}_T) \to C_B^{1,0}(\overline{Q}_T)$ associated with (2.51), which sends any function $u \in C_B^{1,0}(\overline{Q}_T)$ onto the unique solution $v \in W_p^{2,1}(Q_T)$ of

$$\begin{array}{ll} \partial_t v + A(x,t,\partial_x)v = F(x,t,u,\nabla_x u) & \text{in } Q_T, \\ v = 0 & \text{on } \Sigma_T, \\ v(\cdot,0) = v(\cdot,T) & \text{in } \Omega. \end{array}$$

The operator $\overline{S}$ is continuous, has a relatively compact range and its fixed points are the solutions of (2.51). Hence there exists a constant $\overline{R} > 0$, that we can suppose larger than R, such that

$$\|\overline{S}u\|_{C^{1,0}(\overline{Q}_T)} < \overline{R},$$

for every $u \in C_B^{1,0}(\overline{Q}_T)$, and

$$\deg\bigl(I - \overline{S}, B\bigl(0, \overline{R}\bigr)\bigr) = 1, \tag{2.52}$$

where I is the identity operator in $C_B^{1,0}(\overline{Q}_T)$ and $B(0, \overline{R})$ is the open ball of center 0 and radius $\overline{R}$ in $C_B^{1,0}(\overline{Q}_T)$. Therefore $\overline{S}$ has a fixed point. By the conclusions of Steps 1 and 2, we get the existence of a solution u of (2.13) satisfying $\alpha \leqslant u \leqslant \beta$.

Part 2. *Extremal solutions.* Again this can be proved as in the corresponding part of the proof of Theorem 2.6.

Part 3. *Degree computation.* Now, let us assume that α and β are strict lower and upper solutions respectively. Since there exists a solution u of (2.36), which satisfies $\alpha \leqslant u \leqslant \beta$, and every such a solution satisfies $\alpha \ll u \ll \beta$, it follows that $\alpha \ll \beta$. Hence $\mathcal{U}$ is a non-empty open set in $C_B^{1,0}(\overline{Q}_T)$ and there is no fixed point either of S or of $\overline{S}$ on its boundary $\partial\mathcal{U}$. Moreover, by (2.50), the sets of fixed points of S and $\overline{S}$ coincide on $\mathcal{U} \cap B(0, R)$ and we have

$$\deg\bigl(I - S, \mathcal{U} \cap B(0, R)\bigr) = \deg\bigl(I - \overline{S}, \mathcal{U} \cap B(0, R)\bigr).$$

Furthermore, by the excision property of the degree, we get from (2.50) and (2.52)

$$\deg\bigl(I - \overline{S}, B(0, R)\bigr) = 1.$$

Finally, since all fixed points of $\overline{S}$ are in $\mathcal{U} \cap B(0, R)$, we conclude

$$\begin{aligned}\deg\bigl(I - S, \mathcal{U} \cap B(0, R)\bigr) &= \deg\bigl(I - \overline{S}, \mathcal{U} \cap B(0, R)\bigr)\\ &= \deg\bigl(I - \overline{S}, B(0, R)\bigr) = 1.\end{aligned}$$

□

REMARK 2.14. We assumed condition (N) in Theorem 2.15 in order to unify the presentation with the rest of this work. However, here it would be enough to assume the existence of $K > 0$ and $h \in L^p(Q_T)$, with $p > N + 2$, such that, for a.e. $(x, t) \in Q_T$, every $s \in [\alpha(x, t), \beta(x, t)]$ and every $\xi \in \mathbb{R}^N$,

$$\bigl|f(x, t, s, \xi)\bigr| \leqslant h(x, t) + K|\xi|^2.$$

Existence in case $\alpha \nleqslant \beta$

We now discuss the existence of solutions of (2.36) in the presence of lower and upper solutions α, β such that $\alpha \nleqslant \beta$. As we already pointed out, in the frame of problem (2.36) the sole existence of a pair of lower and upper solutions, even satisfying $\alpha > \beta$, does not generally guarantee the existence of a solution. At the light of the example presented in the introduction of this part (cf. (2.8)), we prove a result which is related to the classical Amann–Kolesov Three Solutions Theorem [73,5,7] and requires the existence of a further pair of lower and upper solutions α_1, β_1 such that $\alpha_1 \leqslant \beta_1$ and $\alpha, \beta \in [\alpha_1, \beta_1]$. Theorem 2.17 yields the existence of maximal and minimal solutions of (2.36) in $[\alpha_1, \beta_1]$ belonging to the $C^{1,0}$-closure of the set

$$\mathcal{V} = \bigl\{u \in C_B^{1,0}\bigl(\overline{Q}_T\bigr) \mid u \ngeqslant \alpha \text{ and } u \nleqslant \beta\bigr\}. \tag{2.53}$$

The proof of Theorem 2.17 makes use of the following statement concerning the existence of extremal solutions of (2.36).

LEMMA 2.16. *Assume* (D), (A), (C) *and* (N). *Let* $\mathcal{S}$ *be a nonempty set of solutions of* (2.36). *Suppose that* $\mathcal{S}$ *is uniformly bounded from above, i.e. there is a constant* M *such that* $\max_{Q_T} u \leqslant M$ *for all* $u \in \mathcal{S}$. *Then there exists a maximal solution* w *of* (2.36) *in the* $C^{1,0}$*-closure* $\bar{\mathcal{S}}$ *of* $\mathcal{S}$. *Similarly, if* $\mathcal{S}$ *is uniformly bounded from below, then there exists a minimal solution* v *of* (2.36) *in the* $C^{1,0}$*-closure* $\bar{\mathcal{S}}$ *of* $\mathcal{S}$.

PROOF. We can repeat step by step the proof of Lemma 1.15, with just the following specifications. The set $\mathcal{D} = \{(x_m, t_m) \mid m \in \mathbb{N}\}$ is a countable dense subset of Q_T. The sequence $(u_n)_n$, constructed as in Lemma 1.15, is bounded in $L^\infty(Q_T)$ and, by Proposition 2.10, it is bounded in $W_p^{2,1}(Q_T)$. Hence there is a subsequence converging weakly in $W_p^{2,1}(Q_T)$ to some function $\hat{u} \in W_p^{2,1}(Q_T)$, which is a solution of (2.36). The conclusion then follows as in Lemma 1.15. □

THEOREM 2.17. *Assume* (D), (A), (C) *and* (N). *Suppose that* α *is a lower solution and* β *is an upper solution of* (2.36) *satisfying*

$$\alpha \not\leqslant \beta.$$

Moreover, assume that there exist a lower solution α_1 *and an upper solution* β_1 *of* (2.36) *such that* $\alpha_1 \leqslant \beta_1$ *and* $\alpha, \beta \in [\alpha_1, \beta_1]$. *Then problem* (2.36) *has at least one minimal solution* v *and at least one maximal solution* w *in* $\overline{\mathcal{V}} \cap [\alpha_1, \beta_1]$ *with* $\mathcal{V}$ *defined by* (2.53). *Further, if* α, β *and* α_1, β_1 *are strict, then the solution operator* S *corresponding to* (2.49) *has no fixed point on the boundary of* $\mathcal{V} \cap [\alpha_1, \beta_1]$, *the set of fixed points of* S *in* $\mathcal{V} \cap [\alpha_1, \beta_1]$ *is bounded in* $C_B^{1,0}(\overline{Q}_T)$ *and*

$$\deg\bigl(I - S, \mathcal{V} \cap \mathcal{U} \cap B(0, R)\bigr) = -1,$$

where I *is the identity operator in* $C_B^{1,0}(\overline{Q}_T)$, $\mathcal{U} = \{u \in C_B^{1,0}(\overline{Q}_T) \mid \alpha_1 \ll u \ll \beta_1\}$ *and* $B(0, R)$ *is the open ball of center* 0 *and radius* R *in* $C_B^{1,0}(\overline{Q}_T)$, *with* R *so large that all fixed points of* S *in* $\mathcal{V} \cap \mathcal{U}$ *belongs to* $B(0, R)$.

PROOF. In the course of this proof we denote the given lower and upper solutions α, β by α_0, β_0, respectively. Let us consider problem (2.51), where the function F is defined as in the first part of the proof of Theorem 2.15, with α and β replaced by α_1 and β_1 respectively. In case α_1 or β_1 are not strict for (2.36), and hence for (2.51), we further replace α_1 with $\alpha_1 - 1$ and β_1 with $\beta_1 + 1$. Hence we can suppose that α_1 and β_1 are strict lower and upper solutions of (2.51) such that $\alpha_1 \ll \alpha_0 \ll \beta_1$ and $\alpha_1 \ll \beta_0 \ll \beta_1$. Let us set, for $(i, j) \in \{0, 1\}^2 \setminus \{(0, 0)\}$,

$$\mathcal{U}_{(i,j)} = \bigl\{u \in C_B^{1,0}\bigl(\overline{Q}_T\bigr) \mid \alpha_i \ll u \ll \beta_j\bigr\}.$$

In particular, we have $\mathcal{U}_{(1,1)} = \mathcal{U}$. Notice that $\mathcal{U}_{(i,j)}$ are nonempty open sets in $C_B^{1,0}(\overline{Q}_T)$, since $\alpha_i \ll \beta_j$ for $(i, j) \in \{0, 1\}^2 \setminus \{(0, 0)\}$. Moreover, we have that

$$\mathcal{V} \cap \mathcal{U}_{(1,1)} = \mathcal{U}_{(1,1)} \setminus \overline{(\mathcal{U}_{(0,1)} \cup \mathcal{U}_{(1,0)})}.$$

Let S be the solution operator corresponding to (2.49) and observe that, by Theorem 2.15, all fixed points $u \in \overline{\mathcal{V} \cap \mathcal{U}_{(1,1)}}$ of S satisfy

$$\|u\|_{C^{1,0}(\overline{Q}_T)} < R,$$

for some constant $R > 0$. Then, either there is a solution $u \in \partial\mathcal{V}$ of (2.36) and the existence part of the theorem is proved, or there is no solution $u \in \partial\mathcal{V}$ of (2.36), which means that α_0 and β_0 are strict. In that case, by the additivity-excision property of the degree and Theorem 2.15, we get

$$\begin{aligned}1 &= \deg\bigl(I - S, \mathcal{U}_{(1,1)} \cap B(0,R)\bigr)\\ &= \deg\bigl(I - S, \mathcal{U}_{(0,1)} \cap B(0,R)\bigr) + \deg\bigl(I - S, \mathcal{U}_{(1,0)} \cap B(0,R)\bigr)\\ &\quad + \deg\bigl(I - S, \mathcal{V} \cap \mathcal{U}_{(1,1)} \cap B(0,R)\bigr)\\ &= 2 + \deg\bigl(I - S, \mathcal{V} \cap \mathcal{U}_{(1,1)} \cap B(0,R)\bigr)\end{aligned}$$

and hence

$$\deg\bigl(I - S, \mathcal{V} \cap \mathcal{U}_{(1,1)} \cap B(0,R)\bigr) = -1,$$

where $B(0,R)$ is the open ball of center 0 and radius R in $C^{1,0}_B(\overline{Q}_T)$. This argument yields, in any case, the existence of a solution $u \in \overline{\mathcal{V}} \cap [\alpha_1, \beta_1]$ of (2.36).

Finally, the existence of minimal and maximal solutions follows from Lemma 2.16 with $\mathcal{S} = \mathcal{V} \cap [\alpha_1, \beta_1]$. □

2.5. *The Order Interval Trichotomy and applications*

The aim of this section is to extend to our general context the Order Interval Trichotomy, first proved in [34] for order preserving discrete-time semidynamical systems. We stress once more that our results are obtained without requiring any regularity on the function f besides the Carathéodory conditions. Therefore no continuous or discrete semiflow can be naturally associated with (2.10). Moreover, even if a semiflow were defined by (2.13), there is no apparent reason for which it should be order preserving. Nevertheless with the aid of Proposition 2.3, which plays a crucial role in this context, we can deform our original problem to other problems possessing a certain amount of monotonicity, from which we infer the information needed to prove the existence of some kind of connections.

Once a version of the Order Interval Trichotomy is established, we apply it to the study of the dynamics near one-sided isolated solutions of (2.36), namely we prove the existence of heteroclinic solutions connecting a pair of comparable solutions of (2.36) and we study the qualitative behaviour of solutions of (2.10) lying above a maximal, or below a minimal, solution of (2.36).

Connecting T-periodic solutions by lower or upper solutions

The basic result in this context is the following statement.

PROPOSITION 2.18. *Assume* (D), (A), (C) *and* (N). *Suppose that* u_1, u_2 *are solutions of* (2.36) *such that* $u_1 < u_2$ *and there is no solution* u *of* (2.36) *with* $u_1 < u < u_2$. *Then either*

- *there exists a sequence* $(\alpha_n)_n$ *of proper regular lower solutions of* (2.36) *such that, for each* n, $\alpha_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ *and* $u_1 < \alpha_n < u_2$, *which converges weakly in* $W^{2,1}_p(Q_T)$ *and strongly in* $C^{1,0}(\overline{Q}_T)$ *to* u_1 *as* $n \to +\infty$, *or*
- *there exists a sequence* $(\beta_n)_n$ *of proper regular upper solutions of* (2.36) *such that, for each* n, $\beta_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ *and* $u_1 < \beta_n < u_2$, *which converges weakly in* $W^{2,1}_p(Q_T)$ *and strongly in* $C^{1,0}(\overline{Q}_T)$ *to* u_2 *as* $n \to +\infty$.

PROOF. Let us set, for every $(x,t) \in \overline{Q}_T$ and $s \in \mathbb{R}$,

$$\gamma(x,t,s) = \max\bigl\{u_1(x,t), \min\bigl\{s, u_2(x,t)\bigr\}\bigr\}.$$

Clearly, $\gamma : \overline{Q}_T \times \mathbb{R} \to \mathbb{R}$ is continuous and, for each $(x,t) \in \overline{Q}_T$, $\gamma(x,t,\cdot) : \mathbb{R} \to \mathbb{R}$ is increasing. Moreover, for $i = 1, 2$, let us set, for a.e. $(x,t) \in Q_T$ and every $\varepsilon > 0$,

$$\begin{aligned}&\omega_i(x,t,\varepsilon)\\ &\quad = \max_{|\xi| \leqslant \varepsilon} \bigl| f\bigl(x,t,u_i(x,t), \nabla_x u_i(x,t) + \xi\bigr) - f\bigl(x,t,u_i(x,t), \nabla_x u_i(x,t)\bigr)\bigr|.\end{aligned}$$

Then define, for a.e. $(x,t) \in Q_T$, every $s \in \mathbb{R}$ and $\xi \in \mathbb{R}^N$,

$$\omega(x,t,s) = \begin{cases} \omega_1\bigl(x,t,u_1(x,t) - s\bigr) & \text{if } s < u_1(x,t),\\ 0 & \text{if } u_1(x,t) \leqslant s \leqslant u_2(x,t),\\ -\omega_2\bigl(x,t,s - u_2(x,t)\bigr) & \text{if } s > u_2(x,t) \end{cases}$$

and

$$\bar{f}(x,t,s,\xi) = f\bigl(x,t,\gamma(x,t,s),\xi\bigr) + \omega(x,t,s).$$

Clearly, $\bar{f} : Q_T \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ satisfies the Carathéodory condition (C) and the Nagumo condition (N). Let us consider the modified problem

$$\begin{aligned} &\partial_t u + A(x,t,\partial_x)u = \bar{f}(x,t,u,\nabla_x u) && \text{in } Q_T,\\ &u = 0 && \text{on } \Sigma_T,\\ &u(\cdot,0) = u(\cdot,T) && \text{in } \Omega. \end{aligned} \tag{2.54}$$

Further, take $\rho = \max\{\|u_1\|_{L^\infty(Q_T)}, \|u_2\|_{L^\infty(Q_T)}\}$ and let h be the function associated with f and ρ whose existence is guaranteed by Proposition 2.3. Consider, for each $\mu \in [0,1]$, the following problems

$$\begin{aligned} &\partial_t u + A(x,t,\partial_x)u = \mu \bar{f}(x,t,u,\nabla_x u) + (1-\mu)\bigl[\bar{f}(x,t,u_1,\nabla_x u)\\ &\qquad + h(x,t,u_1,\gamma(x,t,u),\nabla_x u) + \omega(x,t,u)\bigr] && \text{in } Q_T,\\ &u = 0 && \text{on } \Sigma_T,\\ &u(\cdot,0) = u(\cdot,T) && \text{in } \Omega \end{aligned} \tag{2.55}$$

and

$$\begin{array}{ll}\partial_t u + A(x,t,\partial_x)u = \mu \bar{f}(x,t,u,\nabla_x u) + (1-\mu)\big[\bar{f}(x,t,u_2,\nabla_x u) & \\ \quad +h(x,t,u_2,\gamma(x,t,u),\nabla_x u) + \omega(x,t,u)\big] & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega. \end{array} \tag{2.56}$$

It is clear that, if $\mu = 1$, (2.55) and (2.56) reduce to (2.54). Note that the right hand side of the equations in (2.55) and (2.56) satisfy the Carathéodory condition (C) and the Nagumo condition (N), uniformly with respect to $\mu \in [0,1]$. Proposition 2.10 implies that there is a constant $C > 0$ such that, if u is a solution of (2.55), or (2.56), for any $\mu \in [0,1]$, satisfying $u_1 \leqslant u \leqslant u_2$, then

$$\|u\|_{W_p^{2,1}(Q_T)} \leqslant C. \tag{2.57}$$

CLAIM 1. *For any* $\mu \in [0,1]$, *every solution* u *of* (2.55), *or* (2.56), *satisfies* $u_1 \leqslant u \leqslant u_2$ *and, hence,* (2.57). *In particular,* u_1 *and* u_2 *are the only solutions of* (2.54).

Let u be a solution of (2.55) and prove that $u \geqslant u_1$. The argument is similar to the proof of Theorem 2.6 observing that, if $v = u - u_1$ is such that $v(x,t) < 0$ and $|\nabla_x v(x,t)| \leqslant |v(x,t)|$ on $B \times]t_1, t_0]$, for some ball $B \subseteq \Omega$, then on this set

$$\begin{aligned} \partial_t v + A(x,t,\partial_x)v &= \mu\big[\bar{f}(x,t,u,\nabla_x u) - f(x,t,u_1,\nabla_x u_1)\big] \\ &\quad + (1-\mu)\big[\bar{f}(x,t,u_1,\nabla_x u) - f(x,t,u_1,\nabla_x u_1) \\ &\quad + \omega(x,t,u) + h(x,t,u_1,\gamma(x,t,u),\nabla_x u)\big] \\ &= f(x,t,u_1,\nabla_x u) + \omega_1\big(x,t,\big|v(x,t)\big|\big) - f(x,t,u_1,\nabla_x u_1) \\ &\geqslant f(x,t,u_1,\nabla_x u) - f(x,t,u_1,\nabla_x u_1 + \nabla_x v) \\ &\quad + f(x,t,u_1,\nabla_x u_1) - f(x,t,u_1,\nabla_x u_1) \\ &= 0. \end{aligned}$$

Now, let us prove that $u \leqslant u_2$. Again, we argue as in the proof of Theorem 2.6 observing that, if $v = u_2 - u$ is such that $v(x,t) < 0$ and $|\nabla_x v(x,t)| \leqslant |v(x,t)|$ in $B \times]t_1, t_0]$, for some ball $B \subseteq \Omega$, then on this set

$$\begin{aligned} &\partial_t v + A(x,t,\partial_x)v \\ &\quad = \mu\big[f(x,t,u_2,\nabla_x u_2) - \bar{f}(x,t,u,\nabla_x u)\big] + (1-\mu)\big[f(x,t,u_2,\nabla_x u_2) \\ &\qquad - \bar{f}(x,t,u_1,\nabla_x u) - \omega(x,t,u) - h(x,t,u_1,\gamma(x,t,u),\nabla_x u)\big] \\ &\quad = \mu\big[f(x,t,u_2,\nabla_x u_2) - f(x,t,u_2,\nabla_x u) + \omega_2\big(x,t,\big|v(x,t)\big|\big)\big] \\ &\qquad + (1-\mu)\big[f(x,t,u_2,\nabla_x u_2) - f(x,t,u_1,\nabla_x u) + \omega_2\big(x,t,\big|v(x,t)\big|\big) \\ &\qquad - h(x,t,u_1,u_2,\nabla_x u)\big] \end{aligned}$$

$$
\begin{aligned}
&\geqslant \mu\big[f(x,t,u_2,\nabla_x u_2) - f(x,t,u_2,\nabla_x u) - f(x,t,u_2,\nabla_x u_2) \\
&\quad + f(x,t,u_2,\nabla_x u_2 - \nabla_x v)\big] \\
&\quad + (1-\mu)\big[f(x,t,u_2,\nabla_x u_2) - f(x,t,u_1,\nabla_x u) \\
&\quad - f(x,t,u_2,\nabla_x u_2) + f(x,t,u_2,\nabla_x u_2 - \nabla_x v) + h(x,t,u_2,u_1,\nabla_x u)\big] \\
&\geqslant (1-\mu)\big[-f(x,t,u_1,\nabla_x u) + f(x,t,u_2,\nabla_x u) + h(x,t,u_2,u_1,\nabla_x u)\big] \\
&\geqslant 0.
\end{aligned}
$$

In a completely similar way, it can be seen that the same conclusion holds for (2.56). Hence Claim 1 is proved.

CLAIM 2. *For any* $\mu \in [0,1]$, *every solution of* (2.55) *is a lower solution of* (2.36) *and every solution of* (2.56) *is an upper solution of* (2.36).

We verify only the former statement, since the latter can be proved similarly. Indeed, if u is a solution of (2.55), we have, as $u_1 \leqslant u \leqslant u_2$,

$$
\begin{aligned}
&\partial_t u + A(x,t,\partial_x)u \\
&\quad = \mu \bar f(x,t,u,\nabla_x u) \\
&\qquad + (1-\mu)\big[\bar f(x,t,u_1,\nabla_x u) + h(x,t,u_1,\gamma(x,t,u),\nabla_x u) + \omega(x,t,u)\big] \\
&\quad = \mu f(x,t,u,\nabla_x u) + (1-\mu)\big[f(x,t,u_1,\nabla_x u) + h(x,t,u_1,u,\nabla_x u)\big] \\
&\quad \leqslant \mu f(x,t,u,\nabla_x u) + (1-\mu) f(x,t,u,\nabla_x u) = f(x,t,u,\nabla_x u).
\end{aligned}
$$

Thus, Claim 2 is proved.

Now, let us associate with problems (2.55) and (2.56) the corresponding solution operators $S_{1,\mu}, S_{2,\mu} : C_B^{1,0}(\overline{Q}_T) \to C_B^{1,0}(\overline{Q}_T)$, with $\mu \in [0,1]$, whose fixed points are precisely the solutions of (2.55) and (2.56), respectively. Note that $S_{1,1} = S_{2,1}$ is the solution operator corresponding to (2.54).

CLAIM 3. *For every* $\delta > 0$, $u_1 - \delta$ *and* $u_1 + \delta$ *are, respectively, a lower solution and an upper solution of* (2.55), *with* $\mu = 0$.

As $\operatorname{ess\,inf}_{Q_T} a_0 > 0$, this can be deduced from the fact that

$$
\begin{aligned}
&\bar f\big(x,t,u_1,\nabla_x(u_1-\delta)\big) + h\big(x,t,u_1,\gamma\big(x,t,(u_1-\delta)\big),\nabla_x(u_1-\delta)\big) \\
&\quad = f(x,t,u_1,\nabla_x u_1), \\
&\bar f\big(x,t,u_1,\nabla_x(u_1+\delta)\big) + h\big(x,t,u_1,\gamma\big(x,t,(u_1+\delta)\big),\nabla_x(u_1+\delta)\big) \\
&\quad = f(x,t,u_1,\nabla_x u_1) + h\big(x,t,u_1,\gamma\big(x,t,(u_1+\delta)\big),\nabla_x u_1\big) \\
&\quad \leqslant f(x,t,u_1,\nabla_x u_1)
\end{aligned}
$$

and

$$
\omega(x,t,u_1-\delta) \geqslant 0 \geqslant \omega(x,t,u_1+\delta).
$$

Thus, Claim 3 is proved.

Moreover, $u_1 - \delta$ is a strict lower solution. Indeed, if u is a solution of (2.55) with $\mu = 0$, satisfying $u \geqslant u_1 - \delta$, then $u \geqslant u_1$ and hence $u \gg u_1 - \delta$. On the other hand, for what concerns $u_1 + \delta$, either there is $\delta_0 > 0$ such that, for every $\delta \in \left]0, \delta_0\right[$, $u_1 + \delta$ is a strict upper solution, or for every $\delta_0 > 0$ there are $\delta \in \left]0, \delta_0\right[$ and a solution u_δ of (2.55) with $\mu = 0$, such that $u_1 \leqslant u_\delta \leqslant \min\{u_2, u_1 + \delta\}$ and $u_\delta(x_0, t_0) = u_1(x_0, t_0) + \delta$ for some $(x_0, t_0) \in \overline{Q}_T$. Hence, in the latter case, we have $\|u_1 - u_\delta\|_{L^\infty(Q_T)} = \delta$ and therefore $u_\delta \to u_1$ in $L^\infty(Q_T)$, as $\delta \to 0$. By (2.57) there is a constant $C > 0$ such that $\|u_\delta\|_{W_p^{2,1}(Q_T)} \leqslant C$, for every $\delta \in \left]0, \delta_0\right[$, and therefore $u_\delta \to u_1$ in $C^{1,0}(\overline{Q}_T)$, as $\delta \to 0$. Recalling that each u_δ is a lower solution of (2.36), the conclusion of this Proposition is achieved.

Therefore, let us suppose that there is $\delta_0 > 0$ such that, for every $\delta \in \left]0, \delta_0\right[$, $u_1 + \delta$ is a strict upper solution of (2.55). Accordingly Theorem 2.15 guarantees that, for some $R > 0$ and every $\delta \in \left]0, \delta_0\right[$,

$$\deg\big(I - S_{1,0}, \mathcal{U}_1^\delta \cap B(0, R)\big) = 1, \tag{2.58}$$

where $\mathcal{U}_1^\delta = \{u \in C_B^{1,0}(\overline{Q}_T) \mid u_1 - \delta \ll u \ll u_1 + \delta\}$.

In a similar way one proves the following result.

CLAIM 4. *For every* $\delta > 0$, $u_2 - \delta$ *and* $u_2 + \delta$ *are, respectively, a lower solution and an upper solution of* (2.56), *with* $\mu = 0$.

Moreover, $u_2 + \delta$ is a strict upper solution. On the other hand, for what concerns $u_2 - \delta$, as in Claim 3, we can suppose that there is $\delta_0 > 0$ such that, for every $\delta \in \left]0, \delta_0\right[$, $u_2 - \delta$ is a strict lower solution of (2.56) as otherwise the conclusion is achieved. Accordingly Theorem 2.15 guarantees that, for some $R > 0$ and for every $\delta \in \left]0, \delta_0\right[$,

$$\deg\big(I - S_{2,0}, \mathcal{U}_2^\delta \cap B(0, R)\big) = 1, \tag{2.59}$$

where $\mathcal{U}_2^\delta = \{u \in C_B^{1,0}(\overline{Q}_T) \mid u_2 - \delta \ll u \ll u_2 + \delta\}$.

CLAIM 5. *The functions* $u_1 - 1$ *and* $u_2 + 1$ *are, respectively, a strict lower and a strict upper solution of* (2.54), *i.e. of* (2.55) *and* (2.56) *with* $\mu = 1$.

We have

$$\begin{aligned}
&\partial_t(u_1 - 1) + A(x, t, \partial_x)(u_1 - 1) = \partial_t u_1 + A(x, t, \partial_x)u_1 - a_0 \\
&\quad = f(x, t, u_1, \nabla_x u_1) - a_0 \\
&\quad \leqslant f\big(x, t, \gamma\big(x, t, (u_1 - 1)\big), \nabla_x(u_1 - 1)\big) + \omega_1(x, t, 1) - a_0 \\
&\quad \leqslant \bar{f}\big(x, t, u_1 - 1, \nabla_x(u_1 - 1)\big)
\end{aligned}$$

and

$$\begin{aligned}
&\partial_t(u_2 + 1) + A(x, t, \partial_x)(u_2 + 1) = \partial_t u_2 + A(x, t, \partial_x)u_2 + a_0 \\
&\quad = f(x, t, u_2, \nabla_x u_2) + a_0 \\
&\quad \geqslant f\big(x, t, \gamma\big(x, t, (u_2 + 1)\big), \nabla_x(u_2 + 1)\big) - \omega_2(x, t, 1) + a_0 \\
&\quad \geqslant \bar{f}\big(x, t, u_2 + 1, \nabla_x(u_2 + 1)\big).
\end{aligned}$$

Moreover, $u_1 - 1$ and $u_2 + 1$ are strict, since u_1 and u_2 are the only solutions of (2.54). Thus, Claim 5 is proved.

Accordingly Theorem 2.15 guarantees that, for some $R > 0$,

$$\deg\big(I - S_{1,1}, \mathcal{U} \cap B(0,R)\big) = \deg\big(I - S_{2,1}, \mathcal{U} \cap B(0,R)\big) = 1, \tag{2.60}$$

where $\mathcal{U} = \{u \in C_B^{1,0}(\overline{Q}_T) \mid u_1 - 1 \ll u \ll u_2 + 1\}$. Using again the fact that u_1 and u_2 are the only solutions of (2.55) with $\mu = 1$, we conclude, by the excision and additivity properties of the degree, that

$$\begin{aligned}\deg\big(I - S_{1,1}, \mathcal{U} \cap B(0,R)\big) &= \deg\big(I - S_{1,1}, \big(\mathcal{U}_1^\delta \cup \mathcal{U}_2^\delta\big) \cap B(0,R)\big)\\ &= \deg\big(I - S_{1,1}, \mathcal{U}_1^\delta \cap B(0,R)\big)\\ &\quad + \deg\big(I - S_{1,1}, \mathcal{U}_2^\delta \cap B(0,R)\big)\end{aligned} \tag{2.61}$$

for every $\delta \in\,]0, \min\{1, \delta_0, \|u_1 - u_2\|_\infty\}[$.

Now, let us assume that for every $\delta_0 > 0$ there exists $\delta \in\,]0, \delta_0[$ such that, for every $\mu \in [0,1]$, (2.55) has no solution on $\partial\mathcal{U}_1^\delta$ and (2.56) has no solution on $\partial\mathcal{U}_2^\delta$. The homotopy property of the degree then implies, by (2.58) and (2.59),

$$\deg\big(I - S_{1,1}, \mathcal{U}_1^\delta \cap B(0,R)\big) = \deg\big(I - S_{1,0}, \mathcal{U}_1^\delta \cap B(0,R)\big) = 1$$

and

$$\deg\big(I - S_{2,1}, \mathcal{U}_2^\delta \cap B(0,R)\big) = \deg\big(I - S_{2,0}, \mathcal{U}_2^\delta \cap B(0,R)\big) = 1. \tag{2.62}$$

Finally, combining relations (2.60)–(2.62) and using the fact that $S_{1,1} = S_{2,1}$, we obtain,

$$\begin{aligned}2 &= \deg\big(I - S_{1,0}, \mathcal{U}_1^\delta \cap B(0,R)\big) + \deg\big(I - S_{2,0}, \mathcal{U}_2^\delta \cap B(0,R)\big)\\ &= \deg\big(I - S_{1,1}, \mathcal{U}_1^\delta \cap B(0,R)\big) + \deg\big(I - S_{2,1}, \mathcal{U}_2^\delta \cap B(0,R)\big)\\ &= \deg\big(I - S_{1,1}, \mathcal{U}_1^\delta \cap B(0,R)\big) + \deg\big(I - S_{1,1}, \mathcal{U}_2^\delta \cap B(0,R)\big)\\ &= \deg\big(I - S_{1,1}, \mathcal{U} \cap B(0,R)\big) = 1,\end{aligned}$$

which is a contradiction. Hence we can conclude that there is $\delta_0 > 0$ such that for every $\delta \in\,]0, \delta_0[$ either there is a solution u_δ of (2.55), for some $\mu \in [0,1]$, with $u_\delta \in \partial\mathcal{U}_1^\delta$, and hence $u_1 \leqslant u_\delta \leqslant \min\{u_2, u_1 + \delta\}$ and $\|u_1 - u_\delta\|_{L^\infty(Q_T)} = \delta$, or there is a solution u_δ of (2.56), for some $\mu \in [0,1]$, with $u_\delta \in \partial\mathcal{U}_2^\delta$, and hence $\max\{u_1, u_2 - \delta\} \leqslant u_\delta \leqslant u_2$ and $\|u_2 - u_\delta\|_{L^\infty(Q_T)} = \delta$. Condition (2.57) and recalling that solutions u of (2.55), with $u \neq u_1$, are proper lower solutions of (2.36) and solutions u of (2.56), with $u \neq u_2$, are proper upper solutions of (2.36) yield the conclusion. □

REMARK 2.15. Condition (N) can be obviously replaced in Proposition 2.18 by assuming that there exist $h \in L^p(Q_T)$ and $K > 0$ such that for a.e. $(x,t) \in Q_T$, every $s \in [u_1(x,t), u_2(x,t)]$ and every $\xi \in \mathbb{R}^N$,

$$\left| f(x,t,s,\xi) \right| \leqslant h(x,t) + K|\xi|^2.$$

REMARK 2.16. From the proof of Proposition 2.18 it follows that, if u_1 is isolated in $C^{1,0}(\overline{Q}_T)$ as a fixed point of $I - S_{1,1}$ and $\operatorname{ind}(I - S_{1,1}, u_1) \neq 1$, then there exists a sequence $(\alpha_n)_n$ of proper lower solutions of (2.36) such that, for each n, $\alpha_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $u_1 < \alpha_n < u_2$, which converges weakly in $W^{2,1}_p(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$ to u_1 as $n \to +\infty$. A similar result holds for u_2.

In order to prove the existence of connecting sequences of lower and upper solutions, the following two results are needed.

LEMMA 2.19. *Assume* (D), (A), (C) *and* (N) *and let z be a solution of* (2.36).

(i) *If α is a proper lower solution of* (2.36) *such that $\alpha < z$, then there exists a proper regular lower solution $\bar{\alpha}$ of* (2.36), *satisfying $\bar{\alpha}|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $\alpha < \bar{\alpha} < z$.*

(ii) *If β is a proper upper solution of* (2.36) *such that $\beta > z$, then there exists a proper regular upper solution $\bar{\beta}$ of* (2.36), *satisfying $\bar{\beta}|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $z < \bar{\beta} < \beta$.*

PROOF. We only prove the former statement. Let h be the function associated with f by Proposition 2.3 and corresponding to $\rho = \max\{\|\alpha\|_{L^\infty(Q_T)}, \|z\|_{L^\infty(Q_T)}\}$. Consider the periodic problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f(x,t,\alpha,\nabla_x u) - h(x,t,u,\alpha,\nabla_x u) & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega. \end{array} \tag{2.63}$$

The right-hand side of the equation satisfies the Carathéodory condition (C) and the Nagumo condition (N). Since α is a proper lower solution and z is a proper upper solution of (2.63), with $\alpha < z$, by Theorem 2.15, this problem has a solution $\bar{\alpha}$, satisfying $\alpha < \bar{\alpha} < z$. The properties of h imply that $\bar{\alpha}$ is a proper lower solution of (2.36). □

PROPOSITION 2.20. *Assume* (D), (A), (C) *and* (N) *and let z be a solution of* (2.36).

(i) *Let α be a proper lower solution of* (2.36) *such that $\alpha < z$ and there is no solution u of* (2.36) *with $\alpha < u < z$. Then there exists a sequence $(\alpha_n)_n$ of proper regular lower solutions of* (2.36), *such that, for each $n \geqslant 1$, $\alpha_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $\alpha < \alpha_n < \alpha_{n+1} < z$, which converges weakly in $W^{2,1}_p(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$ to z.*

(ii) *Let β be a proper upper solution of* (2.36) *such that $z < \beta$ and there is no solution u of* (2.36) *with $z < u < \beta$. Then there exists a sequence $(\beta_n)_n$ of proper regular upper solutions of* (2.36), *such that, for each $n \geqslant 1$, $\beta_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $z < \beta_{n+1} < \beta_n < \beta$, which converges weakly in $W^{2,1}_p(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$ to z.*

PROOF. We only prove the former statement. Repeating recursively the argument of the proof of Lemma 2.19, we get a sequence of proper regular lower solutions $(\alpha_n)_n$ such that $\alpha_0 = \alpha$ and, for each $n \geqslant 1$, $\alpha_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$, $\alpha < \alpha_{n-1} < \alpha_n < z$ and α_n is a solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f(x,t,\alpha_{n-1},\nabla_x u) - h(x,t,u,\alpha_{n-1},\nabla_x u) & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega. \end{array} \tag{2.64}$$

Since the right-hand side of the equation in (2.64) satisfies conditions (C) and (N), as in the proof of Proposition 2.12, we see that the sequence $(\alpha_n)_n$ converges weakly in $W^{2,1}_p(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$ to a solution of (2.36), which must be z. □

As a direct consequence of Proposition 2.18 and Proposition 2.20 we are finally able to state our general version of the Order Interval Trichotomy, which extends a result in [34] (see also [58,60]).

THEOREM 2.21. *Assume* (D), (A), (C) *and* (N). *Suppose that* u_1, u_2 *are solutions of* (2.36) *such that* $u_1 < u_2$. *Then either*

- *there is a solution* u *of* (2.36) *with* $u_1 < u < u_2$,

or

- *there exists a double sequence* $(\alpha_m)_{m\in\mathbb{Z}}$ *of proper regular lower solutions of* (2.36), *such that, for each* m, $\alpha_m|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ *and* $u_1 < \alpha_m < u_2$, *which converges weakly in* $W^{2,1}_p(Q_T)$ *and strongly in* $C^{1,0}(\overline{Q}_T)$ *to* u_2 *as* $m \to -\infty$ *and to* u_1 *as* $m \to +\infty$,

or

- *there exists a double sequence* $(\beta_m)_{m\in\mathbb{Z}}$ *of proper regular upper solutions of* (2.36), *such that, for each* m, $\beta_m|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ *and* $u_1 < \beta_m < u_2$, *which converges weakly in* $W^{2,1}_p(Q_T)$ *and strongly in* $C^{1,0}(\overline{Q}_T)$ *to* u_1 *as* $m \to -\infty$ *and to* u_2 *as* $m \to +\infty$.

REMARK 2.17. In the frame of Theorem 2.21, if $u_1 \ll u_2$ and there is no solution u of (2.36) with $u_1 < u < u_2$, then the remaining two alternatives are mutually exclusive.

Heteroclinic solutions

We now interpret Theorem 2.21 in terms of the dynamics of solutions of (2.10) by showing that T-monotone solutions connecting two ordered T-periodic solutions do exist. Our statement extends previous results in [34,58,60]. We point out that the proof of the existence of heteroclinic solutions is more involved here than in the case where uniqueness for the initial value problem and validity of comparison principles are assumed.

Theorem 2.22. *Assume* (D), (A), (C) *and* (N). *Suppose that* u_1, u_2 *are solutions of* (2.36) *such that there is no solution* u *of* (2.36) *with* $u_1 < u < u_2$. *Then there exists a* T*-monotone solution* $z : \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$ *of* (2.10) *such that either*

$$\lim_{t \to -\infty} \left\| z(\cdot, t) - u_1(\cdot, t) \right\|_{C^1(\overline{\Omega})} = 0 \quad \textit{and} \quad \lim_{t \to +\infty} \left\| z(\cdot, t) - u_2(\cdot, t) \right\|_{C^1(\overline{\Omega})} = 0, \tag{2.65}$$

or

$$\lim_{t \to -\infty} \left\| z(\cdot, t) - u_2(\cdot, t) \right\|_{C^1(\overline{\Omega})} = 0 \quad \textit{and} \quad \lim_{t \to +\infty} \left\| z(\cdot, t) - u_1(\cdot, t) \right\|_{C^1(\bar{\Omega})} = 0. \tag{2.66}$$

Proof. According to Proposition 2.18, we assume, for example, the existence of a sequence $(\alpha_n)_n$ of proper regular lower solutions of (2.36) such that, for each n, $\alpha_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $u_1 < \alpha_n < u_2$, which converges weakly in $W^{2,1}_p(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$ to u_1 as $n \to +\infty$.

For each n, we also consider the solution $\widetilde{\alpha}_n : \overline{\Omega} \times [0, +\infty[\to \mathbb{R}$ of (2.10), with $\widetilde{\alpha}_n(\cdot, 0) = \alpha_n(\cdot, 0)$, $u_1 < \widetilde{\alpha}_n < u_2$ in $\overline{\Omega} \times [0, +\infty[$ and $\|\widetilde{\alpha}_n(\cdot, t) - u_2(\cdot, t)\|_{C^1(\overline{\Omega})} \to 0$ as $t \to +\infty$, whose existence is guaranteed by Proposition 2.13.

We are going to select a subsequence of $(\widetilde{\alpha}_n)_n$, we shall still denote for simplicity by $(\widetilde{\alpha}_n)_n$, as follows. Let us define $\delta = \frac{1}{2}\|u_1 - u_2\|_{L^\infty(Q_T)}$. We first choose $\widetilde{\alpha}_1$ such that there is $m_1 \geqslant 1$ for which $\|\widetilde{\alpha}_1(\cdot, kT) - u_1(\cdot, 0)\|_{L^\infty(\Omega)} < \delta$, if $k \in \{0, \ldots, m_1 - 1\}$, and $\|\widetilde{\alpha}_1(\cdot, m_1 T) - u_1(\cdot, 0)\|_{L^\infty(\Omega)} \geqslant \delta$. Next, we pick $\widetilde{\alpha}_2$ such that there is $m_2 \geqslant m_1 + 1$ for which $\|\widetilde{\alpha}_2(\cdot, kT) - u_1(\cdot, 0)\|_{L^\infty(\Omega)} < \delta$, if $k \in \{0, \ldots, m_2 - 1\}$, and $\|\widetilde{\alpha}_2(\cdot, m_2 T) - u_1(\cdot, 0)\|_{L^\infty(\Omega)} \geqslant \delta$. If we cannot find such a function $\widetilde{\alpha}_2$, then there exist $\ell \in \{1, \ldots, m_1\}$ and a subsequence $(\widetilde{\alpha}_{n_k})_k$ such that $\|\widetilde{\alpha}_{n_k}(\cdot, \ell T) - u_1(\cdot, 0)\|_{L^\infty(\Omega)} \geqslant \delta$. Since $(\widetilde{\alpha}_{n_k})_k$ is bounded in $L^\infty(\Omega \times]0, \ell T[)$, $(\widetilde{\alpha}_{n_k}(\cdot, 0))_k$ is bounded in $W^{2-2/p}_p(\Omega)$ and $\|\widetilde{\alpha}_{n_k}(\cdot, 0) - u_1(\cdot, 0)\|_{L^\infty(\Omega)} \to 0$, by Proposition 2.5 there is a subsequence of $(\widetilde{\alpha}_{n_k})_k$ converging weakly in $W^{2,1}_p(\Omega \times]0, \ell T[)$ to a solution v of

$$\begin{array}{ll} \partial_t u + A(x, t, \partial_x) u = f(x, t, u, \nabla_x u) & \text{in } \Omega \times]0, \ell T[, \\ u = 0 & \text{on } \partial\Omega \times [0, \ell T], \end{array}$$

satisfying $v(\cdot, 0) = u_1(\cdot, 0)$ and $\|v(\cdot, \ell T) - u_1(\cdot, 0)\|_{L^\infty(\Omega)} \geqslant \delta$. Let $i \in \{0, \ldots, \ell - 1\}$ be such that $v(\cdot, (i+1)T) > v(\cdot, iT) = u_1(\cdot, 0)$. Then the function

$$\alpha(x, t) = v(x, t + iT) \quad \text{in } \overline{\Omega} \times [0, T[,$$

is a proper lower solution of (2.36), with $u_1 < \alpha < u_2$ and $\alpha(\cdot, 0) = u_1(\cdot, 0)$. Proposition 2.13 yields the existence of a T-increasing solution $\widetilde{\alpha}$ of (2.13), with $t_0 = 0$ and $u_0 = u_1(\cdot, 0)$, such that

$$\lim_{t \to +\infty} \left\| \widetilde{\alpha}(\cdot, t) - u_2(\cdot, t) \right\|_{C^1(\bar{\Omega})} \to 0.$$

Accordingly there exists a T-increasing solution $z:\overline{\Omega}\times\mathbb{R}\to\mathbb{R}$ of (2.10) satisfying (2.65). Then, proceeding by induction, once $\widetilde{\alpha}_{n-1}$ has been defined, either we construct just as above, a T-increasing solution $z:\overline{\Omega}\times\mathbb{R}\to\mathbb{R}$ of (2.10) satisfying (2.65), or we find $\widetilde{\alpha}_n$ for which there is $m_n\geqslant m_{n-1}+1$ such that $\|\widetilde{\alpha}_n(\cdot,kT)-u_1(\cdot,0)\|_{L^\infty(\Omega)}<\delta$, if $k\in\{0,\dots,m_n-1\}$, and $\|\widetilde{\alpha}_n(\cdot,m_nT)-u_1(\cdot,0)\|_{L^\infty(\Omega)}\geqslant\delta$. Suppose that the latter eventuality always occurs and set, for each n,

$$v_n(x,t)=\widetilde{\alpha}_n(x,t+m_nT)\quad\text{in }\overline{\Omega}\times[-m_nT,0]$$

and, for $k\in\{1,\dots,m_n\}$,

$$v_n^{(k)}(x,t)=v_n(x,t)\quad\text{in }\overline{\Omega}\times\bigl[-kT,-(k-1)T\bigr].$$

Fix $k\geqslant 1$ and observe that the sequence $(v_n^{(k)})_n$, with n such that $m_n\geqslant k+1$, is bounded in $L^\infty(\Omega\times]-(k+1)T,-(k-1)T[)$. Hence, by Proposition 2.10, there is a subsequence $(v^{(k)}_{n^{(k)}_j})_j$ converging weakly in $W^{2,1}_p(\Omega\times]-kT,-(k-1)T[)$ to a solution w_k of

$$\begin{array}{ll}\partial_t u+A(x,t,\partial_x)u=f(x,t,u,\nabla_x u) & \text{in }\Omega\times\bigl]-kT,-(k-1)T\bigr[,\\ u=0 & \text{on }\partial\Omega\times\bigl[-kT,-(k-1)T\bigr].\end{array}$$

Since we can choose the sequence $(n_j^{(k+1)})_j$ to be a subsequence of $(n_j^{(k)})_j$, the sequence $(w_k)_k$ is such that, for each $k\geqslant 1$, $w_{k+1}(\cdot,-kT)=w_k(\cdot,-kT)$ and $\|w_1(\cdot,0)-u_1(\cdot,0)\|_{L^\infty(\Omega)}\geqslant\delta$. Moreover, as each $\widetilde{\alpha}_n$ is T-increasing, we have $w_k(\cdot,-kT)\leqslant w_k(\cdot,-(k-1)T)$. Setting for each $k\geqslant 1$

$$z(x,t)=w_k(x,t)\quad\text{in }\overline{\Omega}\times\bigl[-kT,-(k-1)T\bigr],$$

we get a solution of

$$\begin{array}{ll}\partial_t u+A(x,t,\partial_x)u=f(x,t,u,\nabla_x u) & \text{in }\Omega\times]-\infty,0[,\\ u=0 & \text{on }\partial\Omega\times]-\infty,0],\end{array}$$

which satisfies

$$z(\cdot,-kT)\leqslant z\bigl(\cdot,-(k-1)T\bigr),\tag{2.67}$$

$\|z(\cdot,-kT)-u_1(\cdot,0)\|_{L^\infty(\Omega)}\leqslant\delta$, for any $k\geqslant 1$, and $\|z(\cdot,0)-u_1(\cdot,0)\|_{L^\infty(\Omega)}\geqslant\delta$.

Let us define, for each $k\geqslant 1$,

$$z_k(x,t)=z\bigl(x,t-(k-1)T\bigr)\quad\text{in }\overline{\Omega}\times[-T,0].$$

As z is bounded in $\overline{\Omega} \times \left]-\infty, 0\right]$ and satisfies (2.67), arguing as in the proof of Lemma 2.11, we prove that the sequence $(z_k)_k$ has a subsequence $(z_{k_j})_j$ converging weakly in $W_p^{2,1}(\Omega \times \left]-T, 0\right[)$ and strongly in $C^{1,0}(\overline{\Omega} \times [-T, 0])$ to a solution ζ of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f(x,t,u,\nabla_x u) & \text{in } \Omega \times \left]-T, 0\right[, \\ u = 0 & \text{on } \partial\Omega \times [-T, 0], \\ u(\cdot, 0) = u(\cdot, -T) & \text{in } \Omega, \end{array}$$

that is a solution of (2.36). As $\|\zeta(\cdot,0) - u_1(\cdot,0)\|_{L^\infty(\Omega)} \leqslant \delta$ and $u_1 \leqslant \zeta \leqslant u_2$, we have $\zeta = u_1$. Since this argument applies to show that any subsequence of $(z_k)_k$ has a subsequence which converges weakly in $W_p^{2,1}(\Omega \times \left]-T, 0\right[)$ to u_1, we can conclude that

$$\lim_{t \to -\infty} \left\| z(\cdot,t) - u_1(\cdot,t) \right\|_{C^1(\overline{\Omega})} = 0.$$

At last, as $z(\cdot,-T) \leqslant z(\cdot,0)$, $z(\cdot,0) > u_1(\cdot,0)$ and $\|z(\cdot,-T) - u_1(\cdot,-T)\|_{L^\infty(\Omega)} \leqslant \delta$, we see that $z|_{\overline{\Omega} \times [-T,0[}$ gives rise to a proper lower solution of (2.36) and, by Proposition 2.13, it can be extended to a T-increasing solution of (2.10) defined in $\overline{\Omega} \times \mathbb{R}$ and satisfying (2.65). □

REMARK 2.18. In the frame of Theorem 2.22, if $u_1 \ll u_2$, then the two alternatives are mutually exclusive. Further, if (2.65) holds, then for every $t_0 \in \mathbb{R}$ and every $u_0 \in W_p^{2-2/p}(\Omega)$, with $u_0 = 0$ on $\partial\Omega$ and $u_1(\cdot,t_0) \ll u_0 \leqslant u_2(\cdot,t_0)$ in Ω, there is a solution $u : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ of (2.13) such that

$$\lim_{t \to +\infty} \left\| u(\cdot,t) - u_2(\cdot,t) \right\|_e = 0.$$

Similar conclusions are valid if (2.66) holds.

Above a maximal, or below a minimal, T-periodic solution

We now apply Theorem 2.22 to discuss the behaviour of solutions of (2.10) lying above a maximal solution of (2.36). Of course, symmetric conclusions can be established for solutions lying below a minimal solution of (2.36).

THEOREM 2.23. *Assume* (D), (A), (C) *and* (N). *Suppose that u_2 is a maximal solution of* (2.36). *Then either there exist $\omega_+ \in \left]-\infty, +\infty\right]$ and a T-increasing solution $z : \overline{\Omega} \times \left]-\infty, \omega_+\right[\to \mathbb{R}$ of* (2.10) *such that $z > u_2$ and*

$$\lim_{t \to -\infty} \left\| z(\cdot,t) - u_2(\cdot,t) \right\|_{C^1(\overline{\Omega})} = 0 \quad \textit{and} \quad \limsup_{t \to \omega_+} \Bigl(\max_{\overline{\Omega}} z(\cdot,t) \Bigr) = +\infty, \tag{2.68}$$

or, for any given $t_0 \in \mathbb{R}$*, there exists a* T*-decreasing solution* $z : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ *of* (2.10) *such that* $z > u_2$ *and*

$$\lim_{t \to +\infty} \left\| z(\cdot, t) - u_2(\cdot, t) \right\|_{C^1(\overline{\Omega})} = 0. \tag{2.69}$$

PROOF. Take $R > 1 + \max_{\overline{Q}_T} u_2$ and define a function $f_R : Q_T \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ by setting, for a.e. $(x, t) \in Q_T$ and every $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$,

$$f_R(x, t, s, \xi) = \begin{cases} f\bigl(x, t, u_2(x, t), \xi\bigr) + \omega\bigl(x, t, u_2(x, t) - s\bigr) & \text{if } s < u_2(x, t), \\ f(x, t, s, \xi) & \text{if } u_2(x, t) \leqslant s \leqslant R, \\ (R + 1 - s) f(x, t, s, \xi) & \text{if } R < s \leqslant R + 1, \\ 0 & \text{if } s > R + 1, \end{cases}$$

where

$$\omega(x, t, \delta) = \max_{|\xi| \leqslant \delta} \Bigl| f\bigl(x, t, u_2(x, t), \nabla_x u_2(x, t) + \xi\bigr) - f\bigl(x, t, u_2(x, t), \nabla_x u_2(x, t)\bigr) \Bigr|.$$

Of course, f_R satisfies the conditions (C) and (N). Consider the modified problem

$$\begin{array}{ll} \partial_t u + A(x, t, \partial_x) u = f_R(x, t, u, \nabla_x u) & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot, 0) = u(\cdot, T) & \text{in } \Omega. \end{array} \tag{2.70}$$

Arguing as in the proof of Theorem 2.15, we see that every solution u of (2.70) is such that $u \geqslant u_2$. Notice that $\beta = R + 2$ is a proper regular upper solution of (2.70). Let us show that it is strict. Otherwise, there exists a solution u of (2.70) satisfying $u < \beta$ but not $u \ll \beta$. Let $v = \beta - u$. Since $v(x, t) > 0$ on $\partial\Omega \times [0, T]$ and $v(x, 0) = v(x, T)$ in Ω, there exists $(x_0, t_0) \in \Omega \times]0, T]$ such that $v(x_0, t_0) = \min_{\overline{Q}_T} v = 0$. Hence we can find an open ball $B \subseteq \Omega$, with $x_0 \in B$, and a point $t_1 \in]0, t_0[$ such that $0 \leqslant v(x, t) \leqslant 1$ in $B \times]t_1, t_0]$ and $v(\bar{x}, \bar{t}) > 0$ for some $(\bar{x}, \bar{t}) \in \bar{B} \times [t_1, t_0]$. Using the definition of f_R and β, we get

$$\partial_t v + A(x, t, \partial_x) v = a_0(R + 2) \geqslant 0,$$

a.e. in $B \times]t_1, t_0[$. Then Proposition 2.2 implies that v is constant in $\bar{B} \times [t_1, t_0]$, which is a contradiction. Assume now that there is no solution u of (2.70) with $u_2 < u \leqslant \beta$. By Lemma 2.19, there exists a proper regular upper solution $\bar{\beta}$ of (2.70) such that $\bar{\beta}|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $u_2 < \bar{\beta} < \beta$. Hence, Proposition 2.13 implies that, for any given $t_0 \in \mathbb{R}$, there exists a T-decreasing solution $z : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ of (2.10), with $z > u_2$, satisfying (2.69).

Next, assume that there exists a solution u of (2.70) with $u > u_2$. Then, as (2.36) has no solution u with $u > u_2$, any possible solution u of (2.70), but u_2, must satisfy

$$\max_{\overline{Q}_T} (u - u_2) \geqslant 1. \tag{2.71}$$

Denote by $\mathcal{S}$ the set of all solutions of (2.70) satisfying (2.71). By Lemma 2.16, there is a minimal solution $v \in \mathcal{S}$ of (2.70). By Theorem 2.22, there exists a T-monotone solution ζ of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f_R(x,t,u,\nabla_x u) & \text{in } \Omega \times \mathbb{R}, \\ u = 0 & \text{on } \partial\Omega \times \mathbb{R}, \end{array}$$

such that either

$$\lim_{t\to-\infty} \left\| \zeta(\cdot,t) - u_2(\cdot,t) \right\|_{C^1(\bar{\Omega})} = 0 \quad \text{and} \quad \lim_{t\to+\infty} \left\| \zeta(\cdot,t) - v(\cdot,t) \right\|_{C^1(\overline{\Omega})} = 0,$$

or

$$\lim_{t\to-\infty} \left\| \zeta(\cdot,t) - v(\cdot,t) \right\|_{C^1(\overline{\Omega})} = 0 \quad \text{and} \quad \lim_{t\to+\infty} \left\| \zeta(\cdot,t) - u_2(\cdot,t) \right\|_{C^1(\bar{\Omega})} = 0.$$

In the former case, there exists $t_0 \in \mathbb{R}$ such that ζ is a solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f(x,t,u,\nabla_x u) & \text{in } \Omega \times \left]-\infty, t_0\right], \\ u = 0 & \text{on } \partial\Omega \times \left]-\infty, t_0\right]. \end{array}$$

Proposition 2.13 implies that ζ can be continued to a T-increasing solution z of (2.10), which is defined in $\overline{\Omega} \times \left]-\infty, \omega_+\right[$ for some $\omega_+ \in \left]-\infty, +\infty\right]$ and satisfies (2.68). In the latter case, there exists $\omega_- \in \mathbb{R}$ such that ζ is a solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f(x,t,u,\nabla_x u) & \text{in } \Omega \times \left]\omega_-, +\infty\right[, \\ u = 0 & \text{on } \partial\Omega \times \left]\omega_-, +\infty\right[. \end{array}$$

Then, for any given $t_0 \in \mathbb{R}$, we take $k_0 \in \mathbb{N}$ such that $\omega_- - k_0 T < t_0$ and we set $z(x,t) = \zeta(x, t + k_0 T)$ in $\overline{\Omega} \times \left[t_0, +\infty\right[$. Clearly, z is a T-decreasing solution of (2.10) satisfying (2.69). □

2.6. *Stability matters*

In this section we study the stability properties of the T-periodic solutions of (2.10) with the aid of lower and upper solutions. It seems natural to use in this context the order norm $\|\cdot\|_e$ introduced in Definition 2.8. We start by recalling the notion of one-sided Lyapunov–Bellmann stability. Since we have to take care of the fact that uniqueness for the initial value problem and validity of comparison principles are not assumed, such a definition does not generally seem the most appropriate to be considered here; indeed, some weaker concept might fit better in order to detect certain residual forms of stability. As an alternative notion to Lyapunov stability we use here order stability; this is common in the frame of order preserving semidynamical systems [93,33,58,60] and appears suited to our approach based on lower and upper solutions. We remark that Lyapunov stability implies order stability, whereas the converse implication is not generally true; however, these concepts are equivalent if a comparison principle holds for solutions of (2.10). Using these notions, we give a precise description of the stability properties of a T-periodic solution in terms of the existence of a lower or an upper solution close to it. Hence, when a pair of lower and upper solutions α, β, with $\alpha \leqslant \beta$, is given, we discuss the stability of the minimum and the maximum T-periodic solutions v and w lying in between, thus getting a completion of Theorem 2.15. Afterwards, we turn to investigate the behaviour of the solutions between v and w, finding in some cases a totally ordered continuum of order stable T-periodic solutions. Finally, in the complementary case where $\alpha \nleqslant \beta$, we see that Lyapunov instability occurs, thus completing Theorem 2.17.

One-sided Lyapunov stability
The following definitions are in the spirit of [17].

DEFINITION 2.14.

- A solution z of (2.36) is said *Lyapunov stable* (briefly, *$\mathcal{L}$-stable*) *from below* if, for every $\varepsilon > 0$, there is $\delta > 0$ such that, for every $t_0 \in [0, T[$ and for every $u_0 \in W_p^{2-2/p}(\Omega)$, with $u_0 = 0$ on $\partial\Omega$, $u_0 < z(\cdot, t_0)$ and $\|u_0 - z(\cdot, t_0)\|_e < \delta$, every nonextendible solution u of (2.13), with $u \leqslant z$, exists in $\overline{\Omega} \times [t_0, +\infty[$ and satisfies

$$\left\|u(\cdot, t) - z(\cdot, t)\right\|_e < \varepsilon \quad \text{in } [t_0, +\infty[. \tag{2.72}$$

- If, further,

$$\lim_{t \to +\infty} \left\|u(\cdot, t) - z(\cdot, t)\right\|_e = 0, \tag{2.73}$$

  z is said *$\mathcal{L}$-asymptotically stable from below*.
- *$\mathcal{L}$-stability from above* and *$\mathcal{L}$-asymptotic stability from above* are defined similarly.
- A solution z of (2.36) is said *$\mathcal{L}$-stable* if, for every $\varepsilon > 0$, there is $\delta > 0$ such that, for every $t_0 \in [0, T[$ and for every $u_0 \in W_p^{2-2/p}(\Omega)$ with $\|u_0 - z(\cdot, t_0)\|_e < \delta$, every nonextendible solution u of (2.13) exists in $\overline{\Omega} \times [t_0, +\infty[$ and satisfies (2.72).

  If, further, (2.73) holds, z is said *$\mathcal{L}$-asymptotically stable*.
- A solution z of (2.36) is said *$\mathcal{L}$-unstable from below* if it is not $\mathcal{L}$-stable from below. *$\mathcal{L}$-instability from above* and *$\mathcal{L}$-instability* are defined similarly.

REMARK 2.19. The notion of $\mathcal{L}$-stability from below given in Definition 2.14 does not require any condition on solutions $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$ of (2.13), satisfying $u_0 < z(\cdot, t_0)$, $\|u(\cdot, t_0) - z(\cdot, t_0)\|_e < \delta$, but not $u \leqslant z$. A similar remark holds for the $\mathcal{L}$-stability from above.

One-sided order stability
The following definitions are in the spirit of [93,33].

DEFINITION 2.15.

- A solution z of (2.36) is said *order stable* (briefly, *$\mathcal{O}$-stable*) (respectively *properly $\mathcal{O}$-stable*, respectively *strictly $\mathcal{O}$-stable*) *from below* if there exists a sequence $(\alpha_n)_n$

of regular lower solutions (respectively proper regular lower solutions, respectively strict regular lower solutions) of (2.36) such that, for every n, $\alpha_n < z$ and $\alpha_n = 0$ on Σ_T, and $\|\alpha_n - z\|_e \to 0$ as $n \to +\infty$.

- A solution z of (2.36) is said *$\mathcal{O}$-stable* (respectively *properly $\mathcal{O}$-stable*, respectively *strictly $\mathcal{O}$-stable*) *from above* if there exists a sequence $(\beta_n)_n$ of regular upper solutions (respectively proper regular upper solutions, respectively strict regular upper solutions) of (2.36) such that, for every n, $\beta_n > z$ and $\beta_n = 0$ on Σ_T, and $\|\beta_n - z\|_e \to 0$ as $n \to +\infty$.
- A solution z of (2.36) is said *$\mathcal{O}$-stable* (respectively *properly $\mathcal{O}$-stable*, respectively *strictly $\mathcal{O}$-stable*) if it is $\mathcal{O}$-stable (respectively properly $\mathcal{O}$-stable, respectively strictly $\mathcal{O}$-stable) from below and from above.
- A solution z of (2.36) is said *$\mathcal{O}$-unstable from below* (respectively *properly $\mathcal{O}$-unstable from below*, respectively *strictly $\mathcal{O}$-unstable from below*) if there exists a sequence $(\beta_n)_n$ of regular upper solutions (respectively proper regular upper solutions, respectively strict regular upper solutions) of (2.36) such that, for every n, $\beta_n < z$ and $\beta_n = 0$ on Σ_T, and $\|\beta_n - z\|_e \to 0$ as $n \to +\infty$.
- A solution z of (2.36) is said *$\mathcal{O}$-unstable from above* (respectively *properly $\mathcal{O}$-unstable from above*, respectively *strictly $\mathcal{O}$-unstable from above*) if there exists a sequence $(\alpha_n)_n$ of regular lower solutions (respectively proper regular lower solutions, respectively strict regular lower solutions) of (2.36) such that, for every n, $\alpha_n > z$ and $\alpha_n = 0$ on Σ_T, and $\|\alpha_n - z\|_e \to 0$ as $n \to +\infty$.

REMARK 2.20. The notion of strict $\mathcal{O}$-stability corresponds to the strong stability introduced by H. Matano in [93] and, as pointed out there, it represents a form of structural stability.

REMARK 2.21. In the frame of Theorem 2.15, if α is a proper lower solution, then the minimum solution v of (2.36) is, by virtue of Proposition 2.20, properly $\mathcal{O}$-stable from below and, if β is a proper upper solution, the maximum solution w of (2.36) is properly $\mathcal{O}$-stable from above. If α is a strict lower solution, v is also not $\mathcal{O}$-unstable from below and, if β is a strict upper solution, w is also not $\mathcal{O}$-unstable from above.

REMARK 2.22. In the definition of $\mathcal{O}$-stability the request that the lower and the upper solutions are regular is not essential at the light of Lemma 2.19.

Lyapunov stability and order stability

We show here that $\mathcal{L}$-stability implies $\mathcal{O}$-stability.

DEFINITION 2.16. A solution u of (2.36) is said *isolated from above* (respectively *isolated from below*) *in* $C^{1,0}(\overline{Q}_T)$ if there is no sequence $(u_n)_n$ of solutions of (2.36) such that $u_n > u$ (respectively $u_n < u$) for every n and $u_n \to u$ in $C^{1,0}(\overline{Q}_T)$.

REMARK 2.23. We point out that, under conditions (D), (A), (C) and (N), it is equivalent to say that a solution z of (2.36) is isolated (from above or from below) in $C^{1,0}(\overline{Q}_T)$, or in $L^\infty(Q_T)$, or in $W_p^{2,1}(Q_T)$. This is a direct consequence of Proposition 2.10.

PROPOSITION 2.24. *Assume* (D), (A), (C) *and* (N). *If a solution* z *of* (2.36) *is* $\mathcal{L}$*-stable from above* (*respectively* $\mathcal{L}$*-stable from below, respectively* $\mathcal{L}$*-stable*), *then it is* $\mathcal{O}$*-stable from above* (*respectively* $\mathcal{O}$*-stable from below, respectively* $\mathcal{O}$*-stable*). *Moreover, if* z *is isolated from above* (*respectively from below*) *in* $C^{1,0}(\overline{Q}_T)$, *then it is properly* $\mathcal{O}$*-stable from above* (*respectively from below*).

PROOF. Let z be a solution of (2.36) which is $\mathcal{L}$-stable from above and suppose that z is isolated from above in $C^{1,0}(\overline{Q}_T)$, because otherwise there is a sequence of solutions $(u_n)_n$, with $u_n > z$, converging to z in $C^{1,0}(\overline{Q}_T)$ and hence z is $\mathcal{O}$-stable from above. Let $\rho > 0$ be such that every solution u of (2.36), with $u > z$, satisfies $\|u - z\|_{C^{1,0}(\overline{Q}_T)} \geqslant \rho$. We then argue as in the proof of Theorem 2.23. We take $R > \rho + \max_{\bar{\Omega}} z$ and we define a function $f_R : Q_T \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ as there, replacing u_2 with z. Consider the modified problem (2.70) and notice that $\beta = R + 2$ is a strict regular upper solution of (2.70).

Assume there is no solution u of (2.70) with $z < u \leqslant \beta$. Then Proposition 2.20 implies that z is properly $\mathcal{O}$-stable from above.

Next, assume that there exists a solution u of (2.70) with $u > z$. Since every solution u of (2.70) is such that $u \geqslant z$ and (2.36) has no solution u, with $\|u - z\|_{C^{1,0}(\overline{Q}_T)} < \rho$, any possible solution u of (2.70), but z, must satisfy

$$\|u - z\|_{C^{1,0}(\overline{Q}_T)} \geqslant \rho. \tag{2.74}$$

Denote by $\mathcal{S}$ the set of all solutions of (2.70) satisfying (2.74). By Lemma 2.16, there is a minimal solution $v \in \mathcal{S}$ of (2.70), with $v > z$. By Proposition 2.18, z is either properly $\mathcal{O}$-stable from above, or properly $\mathcal{O}$-unstable from above. In the latter case, as by Proposition 2.18 $\alpha_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ for each n, Proposition 2.13 imply the existence of a sequence $(\widetilde{\alpha}_n)_n$ of solutions of (2.10) such that $\|\widetilde{\alpha}_n(\cdot,0) - z(\cdot,0)\|_e \to 0$ as $n \to +\infty$, but $\limsup_{t\to+\infty} \|\widetilde{\alpha}_n(\cdot,t) - z(\cdot,t)\|_e \geqslant \frac{1}{2}\|v - z\|_e$, thus contradicting the $\mathcal{L}$-stability from above of z. □

Comparison principles

In order to discuss further relations between $\mathcal{O}$-stability and $\mathcal{L}$-stability, we introduce the following versions of the comparison principle for parabolic equations.

DEFINITION 2.17.

- We say that the *first comparison principle* holds for the equation

$$\partial_t u + A(x,t,\partial_x)u = f(x,t,u,\nabla_x u) \tag{2.75}$$

if for every $t_1, t_2 \in \mathbb{R}$, with $t_1 < t_2$, and for every $u_1, u_2 \in W^{2,1}_p(\Omega \times]t_1,t_2[)$, with $p > N + 2$, satisfying

$$\begin{aligned} &\partial_t u_1 + A(x,t,\partial_x)u_1 - f(x,t,u_1,\nabla_x u_1) \\ &\quad \leqslant \partial_t u_2 + A(x,t,\partial_x)u_2 - f(x,t,u_2,\nabla_x u_2) && \text{in } \Omega \times]t_1,t_2[, \\ &u_1 \leqslant u_2 && \text{on } \partial\Omega \times [t_1,t_2], \\ &u_1(\cdot,t_1) \leqslant u_2(\cdot,t_1), && \text{in } \Omega, \end{aligned} \tag{2.76}$$

we have

$$u_1 \leqslant u_2 \quad \text{in } \overline{\Omega} \times [t_1, t_2].$$

- We say that the *second comparison principle* holds for (2.75) if for every $t_1, t_2 \in \mathbb{R}$, with $t_1 < t_2$, and for every $u_1, u_2 \in W_p^{2,1}(\Omega \times]t_1, t_2[)$, with $p > N+2$, satisfying

$$\begin{aligned} &\partial_t u_1 + A(x,t,\partial_x)u_1 - f(x,t,u_1,\nabla_x u_1) \\ &\quad \leqslant \partial_t u_2 + A(x,t,\partial_x)u_2 - f(x,t,u_2,\nabla_x u_2) && \text{in } \Omega \times]t_1, t_2[, \\ &u_1 \leqslant u_2 && \text{in } \overline{\Omega} \times [t_1, t_2], \\ &u_1(\cdot, t_1) < u_2(\cdot, t_1), && \text{in } \Omega, \end{aligned} \tag{2.77}$$

we have

$$u_1 \ll u_2 \quad \text{in } \overline{\Omega} \times]t_1, t_2].$$

The first or the second comparison principle holds provided that f satisfies certain local one-sided Lipschitz conditions.

PROPOSITION 2.25. *Assume* (D) *and* (A). *Moreover, suppose that*

(L$^+$) *for every* $\rho > 0$ *there is a constant* $L \geqslant 0$ *such that, for a.e.* $(x,t) \in Q_T$, *every* r, $s \in [-\rho, \rho]$, *with* $r < s$, *and every* ξ, $\eta \in [-\rho, \rho]^N$,

$$f(x,t,s,\xi) - f(x,t,r,\eta) \leqslant L\big(s - r + |\xi - \eta|\big).$$

Then the first comparison principle holds.

PROOF. Let t_1, t_2, with $t_1 < t_2$, be given and let $u_1, u_2 \in W_p^{2,1}(\Omega \times]t_1, t_2[)$ satisfy (2.76). Let L be the constant associated with $\rho = \max\{\|u_1\|_{C^{1,0}(\overline{\Omega}\times[t_1,t_2])}, \|u_2\|_{C^{1,0}(\overline{\Omega}\times[t_1,t_2])}\}$ by condition (L$^+$). Set $v(x,t) = \mathrm{e}^{-Lt}(u_2(x,t) - u_1(x,t))$ in $\overline{\Omega} \times [t_1, t_2]$ and suppose that $\min_{\overline{\Omega}\times[t_1,t_2]} v < 0$. Since $v(x,t) \geqslant 0$ on $\partial\Omega \times [t_1, t_2]$ and $v(\cdot, t_1) \geqslant 0$ in Ω, there exists $(x_0, t_0) \in \Omega \times]t_1, t_2]$ such that $v(x_0, t_0) = \min_{\overline{\Omega}\times[t_1,t_2]} v$. Hence there exist an open ball $B \subseteq \Omega$, with $x_0 \in B$, and a point $t^* \in]t_1, t_0]$ such that, for a.e. $(x,t) \in B \times]t^*, t_0]$, $v(x,t) < 0$ and

$$\begin{aligned} &\partial_t v + A(x,t,\partial_x)v \\ &\quad = -L\mathrm{e}^{-Lt}(u_2 - u_1) + \mathrm{e}^{-Lt}\partial_t(u_2 - u_1) + \mathrm{e}^{-Lt}A(x,t,\partial_x)(u_2 - u_1) \\ &\quad \geqslant -L\mathrm{e}^{-Lt}(u_2 - u_1) + \mathrm{e}^{-Lt}\big(f(x,t,u_2,\nabla_x u_2) - f(x,t,u_1,\nabla_x u_1)\big) \\ &\quad \geqslant -L\mathrm{e}^{-Lt}(u_2 - u_1) + \mathrm{e}^{-Lt}L(u_2 - u_1) - \mathrm{e}^{-Lt}L|\nabla_x u_2 - \nabla_x u_1| \\ &\quad = -\sum_{n=1}^{N} \tilde{a}_i(x,t)\partial_{x_i} v, \end{aligned}$$

where, for each $i = 1, \ldots, N$, $\tilde{a}_i \in L^\infty(\Omega \times]t_1, t_2[)$ is defined by

$$\tilde{a}_i(x,t) = \begin{cases} L\dfrac{\partial_{x_i} v(x,t)}{|\nabla_x v(x,t)|} & \text{if } \nabla_x v(x,t) \neq 0, \\ 0 & \text{if } \nabla_x v(x,t) = 0. \end{cases}$$

Proposition 2.2 then implies that $v(x,t) = v(x_0,t_0) < 0$ in $B \times]t^*, t_0]$ and therefore, as $\operatorname{ess\,inf}_{Q_T} a_0 > 0$,

$$\partial_t v + A(x,t,\partial_x)v + \sum_{n=1}^{N} \tilde{a}_i(x,t)\partial_{x_i} v = a_0 v < 0$$

a.e. in $B \times]t_1, t_0]$, thus contradicting the previous inequality. Accordingly, we conclude that $\min_{\overline{\Omega}\times[t_1,t_2]} v \geqslant 0$ and hence $u_1 \leqslant u_2$ in $\overline{\Omega} \times [t_1, t_2]$. □

PROPOSITION 2.26. *Assume* (D) *and* (A). *Moreover, suppose that*

(L$^-$) *for every $\rho > 0$ there is a constant $L \geqslant 0$ such that, for a.e. $(x,t) \in Q_T$, every r, $s \in [-\rho, \rho]$, with $r < s$, and every ξ, $\eta \in [-\rho, \rho]^N$,*

$$f(x,t,s,\xi) - f(x,t,r,\eta) \geqslant -L\big(s - r + |\xi - \eta|\big).$$

Then the second comparison principle holds.

PROOF. Let t_1, t_2, with $t_1 < t_2$, be given and let $u_1, u_2 \in W_p^{2,1}(\Omega \times]t_1, t_2[)$ satisfy (2.77). Let L be the constant associated with $\rho = \max\{\|u_1\|_{C^{1,0}(\overline{\Omega}\times[t_1,t_2])}, \|u_2\|_{C^{1,0}(\overline{\Omega}\times[t_1,t_2])}\}$ by condition (L$^-$). Set $v(x,t) = u_2(x,t) - u_1(x,t)$ in $\overline{\Omega} \times [t_1, t_2]$ and suppose that either there is $(x_0, t_0) \in \Omega \times]t_1, t_2]$ such that $v(x_0,t_0) = 0$, or there is $(x_0,t_0) \in \partial\Omega \times]t_1, t_2]$ such that $v(x_0,t_0) = 0$ and $\partial_\nu v(x_0,t_0) = 0$, where ν is the outer normal to Ω at x_0. Since $v \geqslant 0$ in $\overline{\Omega} \times [t_1, t_2]$, this implies $v(x_0,t_0) = \min_{\overline{\Omega}\times[t_1,t_2]} v$. Moreover we have, for a.e. $(x,t) \in \Omega \times]t_1, t_0]$,

$$\begin{aligned} \partial_t v + A(x,t,\partial_x)v &\geqslant f(x,t,u_2,\nabla_x u_2) - f(x,t,u_1,\nabla_x u_1) \\ &\geqslant -Lv - L|\nabla_x v| = -Lv - \sum_{n=1}^{N} \tilde{a}_i(x,t)\partial_{x_i} v, \end{aligned}$$

where, for each $i = 1, \ldots, N$, $\tilde{a}_i \in L^\infty(\Omega \times]t_1, t_2[)$ is defined by

$$\tilde{a}_i(x,t) = \begin{cases} L\dfrac{\partial_{x_i} v(x,t)}{|\nabla_x v(x,t)|} & \text{if } \nabla_x v(x,t) \neq 0, \\ 0 & \text{if } \nabla_x v(x,t) = 0. \end{cases}$$

If $x_0 \in \Omega$, Proposition 2.2 implies that $v(x,t) = v(x_0,t_0) = 0$ in $\bar{\Omega} \times [t_1, t_0]$, thus contradicting the assumption $v(\cdot, t_1) > 0$ in Ω. If $x_0 \in \partial\Omega$ and $v(x,t) > 0$ in $\Omega \times]t_1, t_0]$, Proposition 2.2 implies that $\partial_\nu v(x_0,t_0) < 0$, which is again a contradiction. Hence we conclude that $u_1 \ll u_2$ in $\overline{\Omega} \times]t_1, t_2]$. □

PROPOSITION 2.27. *Assume* (D) *and* (A). *If the second comparison principle holds, then for every* $t_1, t_2 \in \mathbb{R}$, *with* $t_1 < t_2$, *and for every* $u_1, u_2 \in W_p^{2,1}(\Omega \times]t_1, t_2[)$, *with* $p > N+2$, *satisfying*

$$\begin{aligned} &\partial_t u_1 + A(x,t,\partial_x)u_1 - f(x,t,u_1,\nabla_x u_1) \\ &\quad \leqslant \partial_t u_2 + A(x,t,\partial_x)u_2 - f(x,t,u_2,\nabla_x u_2) && \text{in } \Omega \times]t_1, t_2[, \\ &u_1 \leqslant u_2 && \text{on } \partial\Omega \times [t_1, t_2], \\ &u_1(\cdot, t_1) \ll u_2(\cdot, t_1) && \text{in } \Omega, \end{aligned} \tag{2.78}$$

we have

$$u_1 \ll u_2 \quad \text{in } \overline{\Omega} \times [t_1, t_2].$$

PROOF. Let t_1, t_2 be such that $t_1 < t_2$ and let $u_1, u_2 \in W_p^{2,1}(\Omega \times]t_1, t_2[)$ satisfy (2.78). By continuity, we have $t^* = \sup\{t \in]t_1, t_2] \mid u_1(\cdot, t) \ll u_2(\cdot, t) \text{ in } \Omega\} > t_1$. Suppose that $t^* < t_2$. As $u_1(\cdot, t) \leqslant u_2(\cdot, t)$ in $\Omega \times]t_1, t^*]$ and $u_1(\cdot, t_1) < u_2(\cdot, t_1)$ in Ω, the second comparison principle implies, in particular, that $u_1(\cdot, t^*) \ll u_2(\cdot, t^*)$ in Ω, which is a contradiction. Hence we conclude that $u_1 \ll u_2$ in $\overline{\Omega} \times [t_1, t_2]$. □

REMARK 2.24. The first comparison principle implies uniqueness in the future for the initial value problem (2.13). On the contrary, this is not true for the second comparison principle: examples in this direction can be found, for instance, in [16].

Lyapunov and order stability in case of validity of comparison principles

In this section we discuss some relations between the notion of $\mathcal{O}$-stability and that of $\mathcal{L}$-stability when the first or the second comparison principle holds. We start with the following observation.

PROPOSITION 2.28. *Assume* (D), (A), (C) *and* (N). *Further, suppose that the first or the second comparison principle holds. If a solution* z *of* (2.36) *is* $\mathcal{L}$*-stable* (*respectively* $\mathcal{L}$*-asymptotically stable*) *from above and from below, then it is* $\mathcal{L}$*-stable* (*respectively* $\mathcal{L}$*-asymptotically stable*).

PROOF. Let z be a solution of (2.36) which is $\mathcal{L}$-stable from above and from below. Then, for every $\varepsilon > 0$, there is $\delta > 0$ such that, for every $t_0 \in [0, T[$ and for every $v_0, w_0 \in W_p^{2-2/p}(\Omega)$, with $v_0 = w_0 = 0$ on $\partial\Omega$ and $z(\cdot, t_0) - \delta e < v_0 < z(\cdot, t_0) < w_0 < z(\cdot, t_0) + \delta e$, there exist, by Lemma 2.8 and the $\mathcal{L}$-stability from above and from below, solutions $v, w : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ of (2.13), with $u_0 = v_0$ and $u_0 = w_0$ respectively, such that $z(\cdot, t) - \varepsilon e < v(\cdot, t) < z(\cdot, t) < w(\cdot, t) < z(\cdot, t) + \varepsilon e$ in $[t_0, +\infty[$. Pick now $u_0 \in W_p^{2-2/p}(\Omega)$, with $u_0 = 0$ on $\partial\Omega$ and $z(\cdot, t_0) - \delta e \ll u_0 \ll z(\cdot, t_0) + \delta e$, and let $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$ be a nonextendible solution of (2.13) which exists by Proposition 2.9. Take $v_0, w_0 \in W_p^{2-2/p}(\Omega)$ such that $v_0 = w_0 = 0$ on $\partial\Omega$ and $z(\cdot, t_0) - \delta e < v_0 \ll u_0 \ll w_0 < z(\cdot, t_0) + \delta e$ and let v and w be the corresponding solutions as in the first part of the proof. If the first comparison principle holds, then, by Corollary 2.7 and Remark 2.24,

$\omega = +\infty$, $v < u < w$ and u satisfies (2.72). If the second comparison principle holds, then, by Proposition 2.27, $v \ll u \ll w$, u exists in $\overline{\Omega} \times [t_0, +\infty[$ and satisfies (2.72). $\square$

PROPOSITION 2.29. *Assume* (D), (A), (C) *and* (N). *Further, suppose that the first comparison principle holds. If a solution z of* (2.36) *is strictly $\mathcal{O}$-stable from below* (*respectively strictly $\mathcal{O}$-stable from above*, *respectively strictly $\mathcal{O}$-stable*), *then it is $\mathcal{L}$-stable from below* (*respectively $\mathcal{L}$-stable from above*, *respectively $\mathcal{L}$-stable*).

PROOF. Since z is strictly $\mathcal{O}$-stable from below, there exists a sequence $(\alpha_n)_n$ of strict regular lower solutions of (2.36) such that, for each n, $\alpha_n \ll z$, $\alpha_n = 0$ on Σ_T and $\|\alpha_n - z\|_e \to 0$ as $n \to +\infty$. Fix $\varepsilon > 0$ and take n such that $\|\alpha_n - z\|_e < \varepsilon$. Since $\alpha_n \ll z$, by Proposition 2.1, there exists a constant $\delta > 0$ such that $\delta e \leqslant z(\cdot, t_0) - \alpha_n(\cdot, t_0)$ for every $t_0 \in [0, T]$. Let $u_0 \in W_p^{2-2/p}(\Omega)$, with $u_0 = 0$ on $\partial\Omega$, be such that $z(\cdot, t_0) - \delta e < u_0 \leqslant z(\cdot, t_0)$. Since $\alpha_n(\cdot, t_0) \leqslant z(\cdot, t_0) - \delta e \leqslant u_0 \leqslant z(\cdot, t_0)$, $\alpha_n|_{\bar{\Omega} \times [t_0, +\infty[}$ is a lower solution and $z|_{\overline{\Omega} \times [t_0, +\infty[}$ is an upper solution of (2.13), Corollary 2.7 implies that there exists a solution $u : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ of (2.13), which is unique by Remark 2.24, satisfying $\alpha_n \leqslant u \leqslant z$ in $\overline{\Omega} \times [t_0, +\infty[$ and hence $\|u(\cdot, t) - z(\cdot, t)\|_e < \varepsilon$ in $[t_0, +\infty[$. $\square$

LEMMA 2.30. *Assume* (D), (A), (C) *and* (N). *Further, suppose that the second comparison principle holds. If α is a lower solution of* (2.36) *and z is a solution of* (2.36), *with $\alpha < z$, then $\alpha \ll z$.*

PROOF. According to the definition of lower solution, let $0 = \sigma_0 < \sigma_1 < \cdots < \sigma_k = T$ and for each $h \in \{0, \dots, k-1\}$, $\alpha = \max_{1 \leqslant i \leqslant m_h} \alpha_i^{(h)}$ in $\overline{\Omega} \times]\sigma_h, \sigma_{h+1}[$, with $\alpha_i^{(h)} \in W_p^{2,1}(\Omega \times]\sigma_h, \sigma_{h+1}[)$ for every $i \in \{1, \dots, m_h\}$. Since $\alpha < z$, our definition of a lower solution implies that there is a point $t_1 \in]\sigma_h, \sigma_{h+1}[$, for some $h \in \{0, \dots, k-1\}$, such that $\alpha(\cdot, t_1) < z(\cdot, t_1)$ and hence $\alpha_i^{(h)}(\cdot, t_1) < z(\cdot, t_1)$ for each $i \in \{1, \dots, m_h\}$. The second comparison principle implies that $\alpha_i^{(h)} \ll z$ in $\overline{\Omega} \times]t_1, \sigma_{h+1}]$ for each $i \in \{1, \dots, m_h\}$ and therefore $\alpha(\cdot, \sigma_{h+1}) \leqslant \lim_{t \to \sigma_{h+1}^-} \alpha(\cdot, t) = \lim_{t \to \sigma_{h+1}^-} \max_{1 \leqslant i \leqslant m_h} \alpha_i^{(h)}(\cdot, t) \ll z(\cdot, \sigma_{h+1})$ in Ω. Applying recursively the second comparison principle and using the T-periodicity we conclude that $\alpha \ll z$ in $\overline{\Omega} \times [0, T]$. $\square$

REMARK 2.25. Suppose that the second comparison principle holds. Lemma 2.30 implies that a lower solution of (2.36) is strict if and only if it is proper and hence, in particular, that proper $\mathcal{O}$-stability from below is equivalent to strict $\mathcal{O}$-stability from below. Similar conclusions obviously hold for upper solutions, for proper $\mathcal{O}$-stability from above and for proper $\mathcal{O}$-stability.

PROPOSITION 2.31. *Assume* (D), (A), (C) *and* (N). *Further, suppose that the second comparison principle holds. A solution z of* (2.36) *is $\mathcal{O}$-stable from below* (*respectively $\mathcal{O}$-stable from above*, *respectively $\mathcal{O}$-stable*) *if and only if it is $\mathcal{L}$-stable from below* (*respectively $\mathcal{L}$-stable from above*, *respectively $\mathcal{L}$-stable*).

PROOF. Let z be a solution of (2.36), which is $\mathcal{O}$-stable from below, and let $(\alpha_n)_n$ be a sequence of regular lower solutions of (2.36) such that, for each n, $\alpha_n < z$, $\alpha_n = 0$ on Σ_T and $\|\alpha_n - z\|_e \to 0$ as $n \to +\infty$. Fix $\varepsilon > 0$ and take n such that $\|\alpha_n - z\|_e < \varepsilon$. Since, by Lemma 2.30, $\alpha_n \ll z$, Proposition 2.1 implies the existence of a constant $\delta > 0$ such that $\delta e \leqslant z(\cdot, t_0) - \alpha_n(\cdot, t_0)$ for every $t_0 \in [0, T]$. Let $u_0 \in W_p^{2-2/p}(\Omega)$, with $u_0 = 0$ on $\partial\Omega$, be such that $u_0 \leqslant z(\cdot, t_0)$ and $\|u_0 - z(\cdot, t_0)\|_e < \delta$. Since $\alpha_n(\cdot, t_0) \leqslant z(\cdot, t_0) - \delta e \ll u_0 \leqslant z(\cdot, t_0)$, $\alpha_n|_{\bar{\Omega}\times[t_0,+\infty[}$ is a lower solution and $z|_{\overline{\Omega}\times[t_0,+\infty[}$ is an upper solution of (2.13), Corollary 2.7 implies that there exists a solution $u : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ of (2.13), satisfying $\alpha_n \leqslant u \leqslant z$ in $\bar{\Omega} \times [t_0, +\infty[$ and hence $\|u(\cdot, t) - z(\cdot, t)\|_e < \varepsilon$ in $[t_0, +\infty[$. Moreover, Proposition 2.27 implies that any nonextendible solution $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$ of (2.13), with $\alpha_n(\cdot, t_0) \ll u_0 \leqslant z(\cdot, t_0)$ and $u \leqslant z$, satisfies $\alpha_n(\cdot, t) \ll u(\cdot, t)$ and hence $\|u(\cdot, t) - z(\cdot, t)\|_e < \varepsilon$ in $\overline{\Omega} \times [t_0, \omega[$. Proposition 2.9 finally implies that $\omega = +\infty$. To conclude we use Proposition 2.24 and Proposition 2.28. ☐

Stability via lower and upper solutions

In this section we use lower and upper solutions α, β, with $\alpha \leqslant \beta$, to describe the stability properties of the T-periodic solutions of (2.36) lying in between. We start with an immediate consequence of Proposition 2.13, which shows a certain form of stability of a one-sided isolated solution z of (2.36) in the presence of a lower solution $\alpha < z$, or of an upper solution $\beta > z$.

PROPOSITION 2.32. *Assume* (D), (A), (C) *and* (N). *Moreover, let z be a solution of* (2.36).

(i) *Let α be a proper lower solution of* (2.36) *such that, for some t_0, $\alpha(\cdot, t_0) \in W_p^{2-2/p}(\Omega)$ and $\alpha(\cdot, t_0) = 0$ on $\partial\Omega$. Assume $\alpha < z$ and there is no solution u of* (2.36) *with $\alpha < u < z$. Then there exists the minimum solution $\widetilde{\alpha} : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ in $[\alpha, z]$ of* (2.13), *with $u_0 = \alpha(\cdot, t_0)$. Further, $\widetilde{\alpha}$ is T-increasing and satisfies*

$$\lim_{t\to+\infty} \left\|\widetilde{\alpha}(\cdot, t) - z(\cdot, t)\right\|_{C^1(\overline{\Omega})} = 0.$$

(ii) *Let β be a proper upper solution of* (2.36) *such that, for some t_0, $\beta(\cdot, t_0) \in W_p^{2-2/p}(\Omega)$ and $\beta(\cdot, t_0) = 0$ on $\partial\Omega$. Assume $z < \beta$ and there is no solution u of* (2.36) *with $z < u < \beta$. Then there exists the maximum solution $\widetilde{\beta} : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ in $[z, \beta]$ of* (2.13), *with $u_0 = \beta(\cdot, t_0)$. Further, $\widetilde{\beta}$ is T-decreasing and satisfies*

$$\lim_{t\to+\infty} \left\|\widetilde{\beta}(\cdot, t) - z(\cdot, t)\right\|_{C^1(\overline{\Omega})} = 0.$$

From this proposition we get an attractivity and invariance result, which completes Theorem 2.15. It regards the set of solutions of problem (2.36) lying between α and β. This is a classical topic which has been extensively investigated in the last fifty years (see [124,71,72,33,34,58,37,60]). Unlike all these works, here we require no regularity on the function f besides the Carathéodory conditions. So that our assumptions do not guarantee either uniqueness for the initial value problem or validity of comparison principles, which are the basic tools for applying the theory of order preserving discrete-time semidynamical systems.

THEOREM 2.33. *Assume* (D), (A), (C) *and* (N). *Suppose that* α *is a proper lower solution and* β *is a proper upper solution of* (2.36), *such that, for some* t_0, $\alpha(\cdot, t_0)$, $\beta(\cdot, t_0) \in W_p^{2-2/p}(\Omega)$ *and* $\alpha(\cdot, t_0) = 0$, $\beta(\cdot, t_0) = 0$ *on* $\partial\Omega$. *Finally, assume*

$$\alpha < \beta$$

and denote by v *the minimum solution and by* w *the maximum solution of* (2.36) *in* $[\alpha, \beta]$. *Then the following conclusions hold*:

(i) *there exist a* T*-increasing solution* $\widetilde{\alpha} : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ *of* (2.13), *with* $u_0 = \alpha(\cdot, t_0)$, *and a* T*-decreasing solution* $\widetilde{\beta} : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ *of* (2.13), *with* $u_0 = \beta(\cdot, t_0)$, *such that*

$$\alpha < \widetilde{\alpha} \leqslant v \leqslant w \leqslant \widetilde{\beta} < \beta$$

and

$$\lim_{t \to +\infty} \left\| \widetilde{\alpha}(\cdot, t) - v(\cdot, t) \right\|_{C^1(\bar{\Omega})} = 0 = \lim_{t \to +\infty} \left\| \widetilde{\beta}(\cdot, t) - w(\cdot, t) \right\|_{C^1(\bar{\Omega})};$$

(ii) *if* $u_0 \in W_p^{2-2/p}(\Omega)$ *is such that* $u_0 = 0$ *on* $\partial\Omega$ *and* $\alpha(\cdot, t_0) \leqslant u_0 \leqslant v(\cdot, t_0)$, *then there exists a solution* $u : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ *of* (2.13), *satisfying* $\widetilde{\alpha} \leqslant u \leqslant v$ *and hence*

$$\lim_{t \to +\infty} \left\| u(\cdot, t) - v(\cdot, t) \right\|_e = 0;$$

similarly, if $u_0 \in W_p^{2-2/p}(\Omega)$ *is such that* $u_0 = 0$ *on* $\partial\Omega$ *and* $w(\cdot, t_0) \leqslant u_0 \leqslant \beta(\cdot, t_0)$, *then there exists a solution* $u : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ *of* (2.13), *satisfying* $w \leqslant u \leqslant \widetilde{\beta}$ *and hence*

$$\lim_{t \to +\infty} \left\| u(\cdot, t) - w(\cdot, t) \right\|_e = 0;$$

(iii) *any solution* u *of* (2.10), *such that* $\alpha \leqslant u \leqslant \beta$ *in* $\overline{\Omega} \times [t_0, +\infty[$ *satisfies* $\widetilde{\alpha} \leqslant u \leqslant \widetilde{\beta}$ *in* $\overline{\Omega} \times [t_0, +\infty[$.

PROOF. We first notice that conclusion (i) directly follows from Proposition 2.32. Next, we take $u_0 \in W_p^{2-2/p}(\Omega)$ such that $u_0 = 0$ on $\partial\Omega$ and $\alpha(\cdot, t_0) \leqslant u_0 \leqslant v(\cdot, t_0)$. Since $\widetilde{\alpha}$ and v are, respectively, a lower and an upper solution in $\overline{\Omega} \times [t_0, +\infty[$ of (2.13), Corollary 2.7 implies the existence of a solution $u : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ of (2.13), satisfying $\widetilde{\alpha} \leqslant u \leqslant v$ and hence

$$\lim_{t \to +\infty} \left\| u(\cdot, t) - v(\cdot, t) \right\|_e = 0.$$

Similarly, if $u_0 \in W_p^{2-2/p}(\Omega)$ is such that $u_0 = 0$ on $\partial\Omega$ and $w(\cdot, t_0) \leqslant u_0 \leqslant \beta(\cdot, t_0)$, then there is a solution $u : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ of (2.13), satisfying $w \leqslant u \leqslant \widetilde{\beta}$ and hence

$$\lim_{t\to+\infty} \left\| u(\cdot, t) - w(\cdot, t) \right\|_e = 0.$$

This proves conclusion (ii). Finally, if $u : \overline{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ is any solution of (2.10), satisfying $\alpha \leqslant u \leqslant \beta$, then by the minimality properties of $\widetilde{\alpha}$ described in Proposition 2.13 we conclude that $u \geqslant \widetilde{\alpha}$. In a similar way we prove that $u \leqslant \widetilde{\beta}$. This yields conclusion (iii). □

REMARK 2.26. We assumed in Theorem 2.33 the existence of a common value t_0 such that $\alpha(\cdot, t_0)$, $\beta(\cdot, t_0) \in W_p^{2-2/p}(\Omega)$ and $\alpha(\cdot, t_0) = 0$, $\beta(\cdot, t_0) = 0$ on $\partial\Omega$. Of course, this request is not essential: the above conditions could be satisfied at different values t_0', t_0'', respectively.

The property expressed in Theorem 2.33 is a form of relative attractivity, which is somehow related to the notion of relative stability introduced in [17] and to that of weak positive invariance defined in [15].

Assuming the validity of the second comparison principle, the following stronger stability result holds.

COROLLARY 2.34. *Assume* (D), (A), (C), (N) *and suppose that the second comparison principle holds. Assume that α is a proper lower solution and β is a proper upper solution of* (2.36), *such that, for some t_0, $\alpha(\cdot, t_0)$, $\beta(\cdot, t_0) \in W_p^{2-2/p}(\Omega)$ and $\alpha(\cdot, t_0) = 0$, $\beta(\cdot, t_0) = 0$ on $\partial\Omega$. Further, suppose that $\alpha < \beta$ and denote by v and w the minimum and the maximum solutions of* (2.36) *in* $[\alpha, \beta]$. *Then v is $\mathcal{L}$-asymptotically stable from below and w is $\mathcal{L}$-asymptotically stable from above.*

PROOF. By Remark 2.21, we know that v is properly $\mathcal{O}$-stable from below and hence, by Proposition 2.31, it is $\mathcal{L}$-stable from below. Notice that, by Lemma 2.30, $\alpha \ll v$. Let $\widetilde{\alpha}$ be the solution of (2.13), with $u_0 = \alpha(\cdot, t_0)$, whose existence is guaranteed by Theorem 2.33, satisfying, by Proposition 2.27, $\alpha < \widetilde{\alpha} \ll v$ in $\overline{\Omega} \times [t_0, +\infty[$ and

$$\lim_{t\to+\infty} \left\| \widetilde{\alpha}(\cdot, t) - v(\cdot, t) \right\|_{C^1(\overline{\Omega})} = 0.$$

Then, for every $u_0 \in W_p^{2-2/p}(\Omega)$ such that $u_0 = 0$ on $\partial\Omega$ and $\widetilde{\alpha}(\cdot, t_0) \ll u_0 < v(\cdot, t_0)$, any nonextendible solution u of (2.13) with $u \leqslant v$ satisfies, by Proposition 2.27, $\widetilde{\alpha} \ll u \leqslant v$ in $\overline{\Omega} \times [t_0, +\infty[$ and hence

$$\lim_{t\to+\infty} \left\| u(\cdot, t) - v(\cdot, t) \right\|_e = 0.$$

This implies that v is $\mathcal{L}$-asymptotically stable from below. By a similar argument we get the corresponding conclusion for w. □

REMARK 2.27. Under the assumptions of Corollary 2.34, we have that, for every $u_0 \in W_p^{2-2/p}(\Omega)$ such that $u_0 = 0$ on $\partial\Omega$ and $\widetilde{\alpha}(\cdot, t_0) \ll u_0 \ll \widetilde{\beta}(\cdot, t_0)$, any nonextendible solution $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$ of (2.13) satisfies, by Proposition 2.27 and Proposition 2.9, $\widetilde{\alpha} \ll u \ll \widetilde{\beta}$ in $\overline{\Omega} \times [t_0, \omega[$ with $\omega = +\infty$.

We conclude this section stating some illustrative results where the existence of the pair of lower and upper solutions α, β is replaced by assumptions relating the behaviour of the function f to the principal eigenvalue λ_1 (see Appendix 2.7). Alternative conditions are discussed in [41,42]. We first apply Corollary 2.34 to show the global $\mathcal{L}$-asymptotic stability of the T-periodic solution of a linear problem.

PROPOSITION 2.35. *Assume* (D) *and* (A). *Let* $q \in L^\infty(Q_T)$ *satisfy* $q(x,t) \leqslant \lambda_1$ *a.e. in* Q_T *and* $q(x,t) < \lambda_1$ *on a set of positive measure. Let* $h \in L^p(Q_T)$, *with* $p > N + 2$. *Then the unique solution* z *of*

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = q(x,t)u + h(x,t) & \textit{in } Q_T, \\ u = 0 & \textit{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \textit{in } \Omega, \end{array} \tag{2.79}$$

is globally $\mathcal{L}$*-asymptotically stable.*

PROOF. Let $\varphi_1 \gg 0$ be the principal eigenfunction (see Proposition 2.49). For any $\sigma > 0$ the functions $\alpha = z - \sigma\varphi_1$ and $\beta = z + \sigma\varphi_1$ are respectively a strict lower and a strict upper solution of problem (2.79), with $\alpha \ll z \ll \beta$. Hence, by Corollary 2.34, z is $\mathcal{L}$-asymptotically stable from below and from above. Now, take $t_0 \in [0, T[$ and $u_0 \in W_p^{2-2/p}(\Omega)$ such that $u_0 = 0$ on $\partial\Omega$. Let $\sigma > 0$ be so large that $\alpha(\cdot, t_0) \leqslant u_0 \leqslant \beta(\cdot, t_0)$ in Ω. By Remark 2.27 the solution u of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = q(x,t)u + h(x,t) & \text{in } \Omega \times [t_0, +\infty[, \\ u = 0 & \text{on } \partial\Omega \times [t_0, +\infty[, \\ u(\cdot,t_0) = u_0 & \text{in } \Omega, \end{array}$$

satisfies $\widetilde{\alpha} \ll u \ll \widetilde{\beta}$ in $\overline{\Omega} \times [t_0, +\infty[$, where $\widetilde{\alpha}$, $\widetilde{\beta}$ are the functions whose existence is guaranteed by Theorem 2.33, and hence $\lim_{t\to+\infty} \|u(\cdot,t) - z(\cdot,t)\|_e = 0$. □

For the nonlinear problem (2.36) we state the following stabilization result (see [41,42]). Related, less general, statements can be found in [3,34].

PROPOSITION 2.36. *Assume* (D), (A), (C) *and* (N). *Suppose that there exist functions* $a, b \in L^p(Q_T)$, *with* $p > N + 2$, *and* $\gamma, \Gamma \in L^\infty(Q_T)$, *such that, for a.e.* $(x,t) \in Q_T$ *and all* $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$ *with* $s \neq 0$,

$$\gamma(x,t)|s| + a(x,t) \leqslant f(x,t,s,\xi)\,\mathrm{sgn}(s) \leqslant \Gamma(x,t)|s| + b(x,t),$$

with $\Gamma(x,t) \leqslant \lambda_1$ *a.e. in* Q_T *and* $\Gamma(x,t) < \lambda_1$ *on a subset of positive measure. Then there exist the minimum solution* v *and the maximum solution* w *of* (2.36). *Moreover, for each* $t_0 \in [0, T[$ *and every* $u_0 \in W_p^{2-2/p}(\Omega)$, *with* $u_0 = 0$ *on* $\partial\Omega$, *the following conclusions hold*:

- *if* $u_0 \leqslant v(\cdot, t_0)$, *then there exists a solution* $u : \bar{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ *of* (2.13) *such that* $u \leqslant v$ *and*
$$\lim_{t\to+\infty} \left\| u(\cdot,t) - v(\cdot,t) \right\|_e = 0;$$
- *if* $u_0 \geqslant w(\cdot, t_0)$, *then there exists a solution* $u : \bar{\Omega} \times [t_0, +\infty[\to \mathbb{R}$ *of* (2.13) *such that* $u \geqslant w$ *and*
$$\lim_{t\to+\infty} \left\| u(\cdot,t) - w(\cdot,t) \right\|_e = 0.$$

PROOF. The proof is carried out through some steps.

CLAIM 1. *There exists* $R > 0$ *such that all solutions* u *of* (2.36) *satisfy* $\|u\|_{C^{1,0}(\overline{Q}_T)} < R$.

Without loss of generality, we can suppose $a(x,t) \leqslant 0 \leqslant b(x,t)$ and $\gamma(x,t) \leqslant \Gamma(x,t)$ a.e. in Q_T. Let $q : Q_T \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ be a L^p-Carathéodory function, with $p > N + 2$, such that, for a.e. $(x,t) \in Q_T$ and every $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$,

$$\gamma(x,t) \leqslant q(x,t,s,\xi) \leqslant \Gamma(x,t)$$

and

$$q(x,t,s,\xi) = \max\bigl\{\gamma(x,t), \min\bigl\{s^{-1} f(x,t,s,\xi), \Gamma(x,t)\bigr\}\bigr\}, \quad \text{if } |s| \geqslant 1.$$

Define

$$h(x,t,s,\xi) = f(x,t,s,\xi) - q(x,t,s,\xi)s$$

and observe that there exists a function $c \in L^p(Q_T)$, with $p > N + 2$, such that, for a.e. $(x,t) \in Q_T$ and every $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$,

$$\bigl|h(x,t,s,\xi)\bigr| \leqslant c(x,t).$$

Assume by contradiction that there exists a sequence $(u_n)_n$ of solutions of (2.36) such that $\|u_n\|_{C^{1,0}(\overline{Q}_T)} \to +\infty$. For each n, set $v_n = u_n / \|u_n\|_{C^{1,0}(\overline{Q}_T)}$. Clearly, v_n satisfies

$$\begin{aligned} &\partial_t v_n + A(x,t,\partial_x) v_n = q(x,t,u_n,\nabla u_n) v_n + \frac{h(x,t,u_n,\nabla u_n)}{\|u_n\|_{C^{1,0}(\overline{Q}_T)}} && \text{in } Q_T, \\ &v_n = 0 && \text{on } \Sigma_T, \\ &v_n(\cdot,0) = v_n(\cdot,T) && \text{in } \Omega. \end{aligned}$$

We have that $q(\cdot,\cdot,u_n,\nabla u_n) \to q$ weakly in $L^p(Q_T)$, where $q \in L^\infty(Q_T)$ satisfies $\gamma(x,t) \leqslant q(x,t) \leqslant \Gamma(x,t)$ a.e. in Q_T, and $h_n(\cdot,\cdot,u_n,\nabla u_n)/\|u_n\|_{C^{1,0}(\overline{Q}_T)} \to 0$ strongly in $L^p(Q_T)$. Moreover, $(v_n)_n$ is bounded in $W^{2,1}_p(Q_T)$ and therefore, possibly passing to a subsequence, it converges weakly in $W^{2,1}_p(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$ to some function v with $\|v\|_{C^{1,0}(\overline{Q}_T)} = 1$. Hence v is a nontrivial solution of

$$\begin{array}{ll} \partial_t v + A(x,t,\partial_x)v = q(x,t)v & \text{in } Q_T, \\ v = 0 & \text{on } \Sigma_T, \\ v(\cdot,0) = v(\cdot,T) & \text{in } \Omega, \end{array}$$

thus contradicting Proposition 2.51. This proves the claim.

CLAIM 2. *There exist a strict regular lower solution α and a strict regular upper solution β of* (2.36) *such that every solution u of* (2.36) *satisfies $\alpha \ll u \ll \beta$.*

Denote by $w_\pm$ the solutions of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = \Gamma(x,t)u \pm b(x,t) & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega, \end{array}$$

which exist by Proposition 2.51. Proposition 2.49 and Proposition 2.1 then imply the existence of a constant $k > 0$ such that $w_- - k\varphi_1 \ll u \ll w_+ + k\varphi_1$, for all $u \in C^{1,0}(\overline{Q}_T)$ with $\|u\|_{C^{1,0}(\overline{Q}_T)} < R$. It is now easy to verify that $\alpha = w_- - k\varphi_1$ and $\beta = w_+ + k\varphi_1$ are, respectively, strict regular lower and upper solutions of (2.36). The conclusion then follows using Claim 1.

From the above claims, we deduce by Theorem 2.15 the existence of the minimum solution v and the maximum solution w in $[\alpha,\beta]$; as every solution u of (2.36) satisfies $\alpha \leqslant u \leqslant \beta$, v and w are the minimum and the maximum solution. Possibly modifying the choice of k in the definition of α and β, we can assume that, if u_0 is such that $u_0 \leqslant v(\cdot,t_0)$ (respectively $w(\cdot,t_0) \leqslant u_0$), then $\alpha(\cdot,t_0) \leqslant u_0 \leqslant v(\cdot,t_0)$ (respectively $w(\cdot,t_0) \leqslant u_0 \leqslant \beta(\cdot,t_0)$). The result then follows from Theorem 2.33. □

Continua of stable T-periodic solutions

We now prove a result which generalizes [33, Theorem 1] (see also [34,58]) and provides the existence of $\mathcal{O}$-stable solutions of (2.36) in the presence of lower and upper solutions α, β, with $\alpha \leqslant \beta$. It also provides information about the topological structure of the set of $\mathcal{O}$-stable solutions lying between α and β.

In the proof we shall use a property of compact subsets of $C^{1,0}(\overline{Q}_T)$, which are dense-in-itself with respect to the order, analogous to that one we have described in Lemma 1.3 for compact subsets of $C^0([t_0,t_1])$.

DEFINITION 2.18. A subset $\mathcal{S}$ of $C^{1,0}(\overline{Q}_T)$ is said *dense-in-itself with respect to the order* if for any $u_1, u_2 \in \mathcal{S}$, with $u_1 < u_2$, there exists $u_3 \in \mathcal{S}$ with $u_1 < u_3 < u_2$.

The next result can be proved exactly in the same way as Lemma 1.3.

LEMMA 2.37. *Let $\mathcal{S} \subset C^{1,0}(\overline{Q}_T)$ be a compact set which is dense-in-itself with respect to the order. Let $\mathcal{T} \subseteq \mathcal{S}$ be a maximal nondegenerate totally ordered subset of $\mathcal{S}$. Then $\mathcal{T}$ is homeomorphic to a nondegenerate compact interval of $\mathbb{R}$.*

THEOREM 2.38. *Assume* (D), (A), (C) *and* (N). *Suppose that α is a proper lower solution and β is a proper upper solution of* (2.36) *satisfying $\alpha < \beta$. Denote, respectively, by v and w the minimum and the maximum solution of* (2.36) *in* $[\alpha, \beta]$. *Then there exists a totally ordered continuum $\mathcal{K}$ in $C^{1,0}(\overline{Q}_T)$ such that every $u \in \mathcal{K}$ is an $\mathcal{O}$-stable solution of* (2.36) *satisfying $v \leqslant u \leqslant w$; moreover, $u_1 = \min \mathcal{K}$ is properly $\mathcal{O}$-stable from below and $u_2 = \max \mathcal{K}$ is properly $\mathcal{O}$-stable from above.*

PROOF. The proof closely follows up the proof of Theorem 1.33. We denote by $\mathcal{S}_1$ the set of all solutions u of (2.36), with $\alpha \leqslant u \leqslant \beta$, which are properly $\mathcal{O}$-stable from below. Since, by Remark 2.21, the minimum solution v is properly $\mathcal{O}$-stable from below, $\mathcal{S}_1$ is not empty. Notice also that, since $\mathcal{S}_1$ is bounded in $L^\infty(Q_T)$, the Nagumo condition (N) implies, by Proposition 2.10, that this set is bounded in $W_p^{2,1}(Q_T)$ and therefore it is relatively compact in $C^{1,0}(\overline{Q}_T)$. Arguing as in the proof of Theorem 1.33, with the aid here of Lemma 2.16 instead of Lemma 1.15, we prove the existence of a maximal solution u_1 of (2.36) in $\mathcal{S}_1$. Proceeding further as in that proof, we denote by $\mathcal{S}_2$ the set of all solutions u of (2.36), with $u_1 \leqslant u \leqslant \beta$, which are properly $\mathcal{O}$-stable from above. By Remark 2.21, $w \in \mathcal{S}_2$. Arguing as above, we get a minimal solution u_2 of (2.36) in $\mathcal{S}_2$.

If $u_1 = u_2$, the conclusion is achieved. Otherwise we denote by $\mathcal{S}_3$ the set of all solutions u of (2.36), with $u_1 \leqslant u \leqslant u_2$. Notice that $\mathcal{S}_3$ is compact in $C^{1,0}(\overline{Q}_T)$. To show that $\mathcal{S}_3$ is dense-in-itself with respect to the order we argue as in Theorem 1.33, with the aid here of Remark 2.21, to show that no proper lower or upper solution of (2.36) may exist between u_1 and u_2, and of Proposition 2.18. Finally, we fix a solution $u_0 \in \mathcal{S}_3$ and we denote by $\mathcal{S}(u_0)$ a maximal totally ordered subset of $\mathcal{S}_3$, with $u_0 \in \mathcal{S}(u_0)$. Then we apply Lemma 2.37 to prove that $\mathcal{S}(u_0)$ is compact and connected. Clearly, every $u \in \mathcal{S}(u_0)$ is $\mathcal{O}$-stable and $u_1 = \min \mathcal{S}(u_0)$ and $u_2 = \max \mathcal{S}(u_0)$ are properly $\mathcal{O}$-stable from below and above, respectively. □

Assuming the validity of the second comparison principle stronger conclusions can be achieved.

COROLLARY 2.39. *Assume* (D), (A), (C), (N) *and suppose that the second comparison principle holds. Let α be a proper lower solution and β be a proper upper solution of* (2.36) *satisfying $\alpha < \beta$. Denote, respectively, by v and w the minimum and the maximum solution of* (2.36) *in* $[\alpha, \beta]$. *Then there exists a totally ordered continuum $\mathcal{K}$ in $C^{1,0}(\overline{Q}_T)$ such that every $u \in \mathcal{K}$ is a $\mathcal{L}$-stable solution of* (2.36) *satisfying $v \leqslant u \leqslant w$. Moreover, if $u_1 = \min \mathcal{K}$ is isolated from below in $C^{1,0}(\overline{Q}_T)$, then u_1 is $\mathcal{L}$-asymptotically stable from below and if $u_2 = \max \mathcal{K}$ is isolated from above in $C^{1,0}(\overline{Q}_T)$, then u_2 is $\mathcal{L}$-asymptotically stable from above. In particular, if $u_1 = u_2 = u^*$ is isolated in $C^{1,0}(\overline{Q}_T)$, then u^* is $\mathcal{L}$-asymptotically stable.*

PROOF. The former conclusion follows from Theorem 2.38 and Proposition 2.31. The latter conclusion, concerning the $\mathcal{L}$-asymptotic stability from below of u_1, is a consequence of Corollary 2.34. Indeed, since u_1 is isolated from below in $C^{1,0}(\overline{Q}_T)$ and, by Theorem 2.38, is properly $\mathcal{O}$-stable from below, there exists a proper lower solution α^*, with $\alpha^* < u_1$, such that there is no solution u of (2.36) with $\alpha^* < u < u_1$, i.e. u_1 is the minimum solution of (2.36) in $[\alpha^*, \beta]$. Hence u_1 is $\mathcal{L}$-asymptotically stable from below. The conclusion about u_2 follows from a similar argument. The $\mathcal{L}$-asymptotic stability of u^* is a consequence of Proposition 2.28. □

Instability via lower and upper solutions

We now show that the presence of lower and upper solutions α, β, with $\alpha \nleqslant \beta$, yields the existence of unstable T-periodic solutions of (2.10). We stress that all the results in this section are obtained without requiring the validity of any comparison principle.

PROPOSITION 2.40. *Assume* (D), (A), (C) *and* (N). *Let z be a solution of* (2.36). *If z is $\mathcal{O}$-unstable from above and isolated from above in $C^{1,0}(\overline{Q}_T)$, then it is $\mathcal{L}$-unstable from above. If z is $\mathcal{O}$-unstable from below and isolated from below in $C^{1,0}(\overline{Q}_T)$, then it is $\mathcal{L}$-unstable from below.*

PROOF. According to Remark 2.23, we pick $\varepsilon > 0$ such that any solution v of (2.36), with $v > z$, satisfies $\|z - v\|_{L^\infty(Q_T)} \geqslant \varepsilon$. Further, let $(\alpha_n)_n$ be a sequence of regular lower solutions of (2.36) such that $\alpha_n > z$, $\alpha_n = 0$ on Σ_T and $\|\alpha_n - z\|_e \to 0$, as $n \to +\infty$. Fix $t_0 \in [0, T[$. Then, for every $\delta > 0$, there exists n such that, setting $u_0 = \alpha_n(\cdot, t_0)$, we have $u_0 \in W_p^{2-2/p}(\Omega)$, $u_0 = 0$ on $\partial\Omega$, $u_0 \geqslant z(\cdot, t_0)$ and $\|u_0 - z(\cdot, t_0)\|_e < \delta$. Moreover, problem (2.13) has, by Proposition 2.13, a nonextendible solution $u : \overline{\Omega} \times [t_0, \omega[\to \mathbb{R}$, satisfying $u > z$ in $\overline{\Omega} \times [t_0, \omega[$ and either $\lim_{t\to\omega} \|u(\cdot, t)\|_{L^\infty(\Omega)} = +\infty$, or $\lim_{t\to+\infty} \|u(\cdot, t) - v(\cdot, t)\|_{L^\infty(\Omega)} = 0$, for some solution v of (2.36), with $v > z$ and $\|z - v\|_{L^\infty(Q_T)} \geqslant \varepsilon$. Anyhow we conclude that z is $\mathcal{L}$-unstable from above. □

The following result is a counterpart, for what concerns instability, of Proposition 2.32.

PROPOSITION 2.41. *Assume* (D), (A), (C) *and* (N). *Moreover, let z be a solution of* (2.36).

(i) *If α is a strict lower solution of* (2.36) *and β is an upper solution of* (2.36) *such that $\alpha \leqslant \beta$, $\alpha \nleqslant z$, $z < \beta$ and there is no solution u of* (2.36) *satisfying $z < u \leqslant \beta$ and $u \ngeqslant \alpha$, then z is $\mathcal{L}$-unstable from above.*

(ii) *If β is a strict upper solution of* (2.36) *and α is a lower solution of* (2.36) *such that $\alpha \leqslant \beta$, $\beta \ngeqslant z$, $\alpha < z$ and there is no solution u of* (2.36) *satisfying $\alpha \leqslant u < z$ and $u \nleqslant \beta$, then z is $\mathcal{L}$-unstable from below.*

PROOF. We prove only the former statement; the proof of the latter being similar. Let v be the minimum solution of (2.36) in $[\max\{\alpha, z\}, \beta]$ given by Theorem 2.15. Since α is a strict lower solution, we have $v \gg \alpha$ and hence $v > \max\{\alpha, z\}$. Let us observe that there is no solution u of (2.36) such that $z < u < v$. Indeed, if u were such a solution, by the minimality of v, it should satisfy $u \ngeqslant \max\{\alpha, z\}$ and hence $u \ngeqslant \alpha$. This contradicts the assumptions on z. Then Proposition 2.18 implies that either there exists a sequence $(\alpha_n)_n$

of proper regular lower solutions of (2.36) such that, for each n, $\alpha_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $z < \alpha_n < v$, which converges in $C^{1,0}(\overline{Q}_T)$ to z as $n \to +\infty$, or there exists a sequence $(\beta_n)_n$ of proper regular upper solutions of (2.36) such that, for each n, $\beta_n|_{\overline{Q}_T} \in W^{2,1}_{p,B}(Q_T)$ and $z < \beta_n < v$, which converges in $C^{1,0}(\overline{Q}_T)$ to v as $n \to +\infty$. Let us show that the latter alternative cannot occur. Indeed, otherwise, as $v \gg \alpha$, we could find an upper solution $\hat{\beta}$ of (2.36), with $\max\{\alpha, z\} \leqslant \hat{\beta} < v$. Hence there should exist a solution u of (2.36), with $\max\{\alpha, z\} \leqslant u \leqslant \hat{\beta}$ and therefore $z < u < v$, as $z \ngeq \alpha$. This yields a contradiction with a preceding conclusion. Therefore, the former alternative necessarily occurs, i.e. z is properly $\mathcal{O}$-unstable from above. Arguing as in Proposition 2.40, we get the conclusion. □

An immediate consequence of these statements is the following instability result, in the presence of a lower solution α and an upper solution β satisfying the condition $\alpha \nleq \beta$. It yields a completion of Theorem 2.17; as there, we write

$$\mathcal{V} = \left\{u \in C^{1,0}_B\left(\overline{Q}_T\right) \mid u \ngeq \alpha \text{ and } u \nleq \beta\right\}.$$

THEOREM 2.42. *Assume* (D), (A), (C) *and* (N). *Suppose that α is a strict lower solution and β is a strict upper solution of* (2.36) *satisfying*

$$\alpha \nleq \beta.$$

Further, assume that there exist a lower solution α_1 and an upper solution β_1 of (2.36) *such that $\alpha_1 \leqslant \beta_1$ and $\alpha, \beta \in [\alpha_1, \beta_1]$. Then any minimal solution v and any maximal solution w of* (2.36) *in $\overline{\mathcal{V}} \cap [\alpha_1, \beta_1]$ is, respectively, $\mathcal{L}$-unstable from below and $\mathcal{L}$-unstable from above.*

PROOF. We first notice that, as α is a strict lower solution and β is a strict upper solution, $v, w \in \mathcal{V}$. We then apply the former statement of Proposition 2.41, with β and z replaced respectively by β_1 and w, to show that w is $\mathcal{L}$-unstable from above. Whereas, we apply the latter statement of Proposition 2.41, with α and z replaced respectively by α_1 and v, to show that v is $\mathcal{L}$-unstable from below. □

We conclude this section stating some illustrative results where the existence of the lower and upper solutions α, α_1, β, β_1 is replaced by assumptions relating the behaviour of the function f to the principal eigenvalue λ_1. Alternative conditions can be found in [41,42]. We first discuss the $\mathcal{L}$-instability of the T-periodic solution of a linear problem.

PROPOSITION 2.43. *Assume* (D) *and* (A). *Let $q \in L^\infty(Q_T)$ satisfy $\lambda_1 \leqslant q(x,t) \leqslant \mu$ a.e. in Q_T, with $\mu > \lambda_1$ defined in Proposition* 2.52, *and $\lambda_1 < q(x,t)$ on a subset of positive measure. Let $h \in L^p(Q_T)$, with $p > N + 2$. Then the unique solution z of problem* (2.79) *is $\mathcal{L}$-unstable from below and from above.*

PROOF. Let $\varphi_1 \gg 0$ be the principal eigenfunction. For any $\sigma > 0$ the functions $\alpha = z + \sigma\varphi_1$ and $\beta = z - \sigma\varphi_1$ are respectively a strict lower and a strict upper solution of problem (2.79), with $\alpha \gg z \gg \beta$. Hence z is $\mathcal{O}$-unstable from above and from below. Proposition 2.40 yields the conclusion. □

A nonlinear counterpart of Proposition 2.43 for problem (2.36) is the following instability result (see [41,42]).

PROPOSITION 2.44. *Assume* (D), (A), (C) *and* (N). *Suppose that there exist functions* $a, b \in L^p(Q_T)$, *with* $p > N+2$, *and* $\gamma \in L^\infty(Q_T)$ *such that, for a.e.* $(x,t) \in Q_T$ *and all* $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$ *with* $s \neq 0$,

$$\gamma(x,t)|s| + a(x,t) \leqslant f(x,t,s,\xi)\,\mathrm{sgn}(s) \leqslant \mu|s| + b(x,t),$$

where $\gamma(x,t) \geqslant \lambda_1$ *a.e. in* Q_T, $\gamma(x,t) > \lambda_1$ *on a subset of positive measure and* μ *is defined in Proposition* 2.52. *Then there exist a minimal solution* v *and a maximal solution* w *of* (2.36) *which are, respectively,* $\mathcal{L}$*-unstable from below and* $\mathcal{L}$*-unstable from above.*

PROOF. The proof is carried out through several steps.

CLAIM 1. *There exists a minimal solution* v *and a maximal solution* w *of* (2.36).

Arguing as in Claim 1 of Proposition 2.36 and using Proposition 2.52 instead of Proposition 2.51, we prove the existence of a constant $M > 0$ such that all solutions u of (2.36) satisfy $\|u\|_{C^{1,0}(\overline{Q}_T)} < M$. Hence we conclude by Lemma 2.16.

CLAIM 2. *There exist a strict regular lower solution* α *and a strict regular upper solution* β *of* (2.36) *such that every solution* u *of* (2.36) *satisfies* $\beta \ll u \ll \alpha$.

Using Proposition 2.52, Proposition 2.49 and Proposition 2.1, we define, as in Claim 2 in the proof of Proposition 2.36, $\alpha = w_+ + k\varphi_1$ and $\beta = w_- - k\varphi_1$, where $w_\pm$ are the solutions of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = \gamma(x,t)u \pm a(x,t) & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega, \end{array}$$

and $k > 0$ is so large that $\beta \ll u \ll \alpha$, for all $u \in C^{1,0}(\overline{Q}_T)$ with $\|u\|_{C^{1,0}(\overline{Q}_T)} < M$. It is now easy to verify that α and β are, respectively, strict regular lower and upper solutions of (2.36).

For each $r > 1$, consider the modified problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = f_r(x,t,u,\nabla_x u) & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega, \end{array} \tag{2.80}$$

where $f_r : Q_T \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is defined, for a.e. $(x,t) \in Q_T$ and every $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$, by

$$f_r(x,t,s,\xi) = \begin{cases} f(x,t,s,\xi) & \text{if } |s| < r, \\ \bigl(|s| - r\bigr)\bigl(\lambda_1 - \frac{1}{r}\bigr)s & \\ \quad + \bigl(r + 1 - |s|\bigr) f(x,t,s,\xi) & \text{if } r \leqslant |s| \leqslant r+1, \\ \bigl(\lambda_1 - \frac{1}{r}\bigr)s & \text{if } |s| > r+1. \end{cases}$$

Of course, f_r satisfies condition (C). Let us set

$$\mathcal{V} = \bigl\{ u \in C_B^{1,0}\bigl(\overline{Q}_T\bigr) \mid u \not\geqslant \alpha \text{ and } u \not\leqslant \beta \bigr\}.$$

CLAIM 3. *There exists $K > 0$ such that every solution $u \in \overline{\mathcal{V}}$ of* (2.80) *with $r > K$ satisfies $\|u\|_{C^{1,0}(\overline{Q}_T)} < K$.*

Assume by contradiction that, for every n, there exist $r_n > n$ and $u_n \in \overline{\mathcal{V}}$, solution of (2.80) for $r = r_n$, such that $\|u_n\|_{C^{1,0}(\overline{Q}_T)} \geqslant n$. As in Claim 1 of Proposition 2.36, we can write, for each $r > 1$,

$$f_r(x,t,s,\xi) = q_r(x,t,s,\xi)s + h_r(x,t,s,\xi)$$

for a.e. $(x,t) \in Q_T$ and every $(s,\xi) \in \mathbb{R} \times \mathbb{R}^N$, where q_r and h_r are L^p-Carathéodory functions satisfying

$$\lambda_1 - \frac{1}{r} \leqslant q_r(x,t,s,\xi) \leqslant \mu \quad \text{and} \quad \bigl|h_r(x,t,s,\xi)\bigr| \leqslant c(x,t),$$

for some $c \in L^p(Q_T)$ independent of r. Setting $v_n = u_n / \|u_n\|_{C^{1,0}(\overline{Q}_T)}$, we have that, possibly passing to a subsequence, $(v_n)_n$ converges weakly in $W_p^{2,1}(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$ to a solution v of

$$\begin{array}{ll} \partial_t v + A(x,t,\partial_x)v = q(x,t)v & \text{in } Q_T, \\ v = 0 & \text{on } \Sigma_T, \\ v(\cdot,0) = v(\cdot,T) & \text{in } \Omega, \end{array}$$

where $\|v\|_{C^{1,0}(\overline{Q}_T)} = 1$ and $q \in L^\infty(Q_T)$ satisfies $\lambda_1 \leqslant q(x,t) \leqslant \mu$ a.e. in Q_T. Proposition 2.52 implies that $q = \lambda_1$ a.e. in Q_T and $v = a\varphi_1$ for some $a \neq 0$. Hence, by Proposition 2.1, we conclude that, if $a > 0$, there exists $d > 0$ such that, for all n large enough,

$$u_n \geqslant d\|u_n\|_{C^{1,0}(\overline{Q}_T)} a\varphi_1 \gg \alpha$$

and, if $a < 0$, there exists $d > 0$ such that, for all n large enough,

$$u_n \leqslant d\|u_n\|_{C^{1,0}(\overline{Q}_T)} a\varphi_1 \ll \beta,$$

thus contradicting anyhow the assumption $u_n \in \overline{\mathcal{V}}$. This proves the claim.

Let us set $R = 1 + \max\{K, M, \|\alpha\|_{L^\infty(Q_T)}, \|\beta\|_{L^\infty(Q_T)}\}$. We first notice that α and β are strict lower and upper solutions of (2.80), with $r = R$. Indeed, otherwise there exists a solution $u \in \overline{\mathcal{V}}$ of (2.80), with $r = R$, which, by Claim 3, satisfies $\|u\|_{C^{1,0}(\overline{Q}_T)} < K$. Hence u is a solution of (2.36) too, contradicting the fact that α and β are strict lower and upper solutions of (2.36).

CLAIM 4. *Problem* (2.80), *with* $r = R$, *admits a lower solution* α_1 *and an upper solution* β_1 *such that* $\alpha_1 \leqslant \beta_1$ *and* $\alpha, \beta \in [\alpha_1, \beta_1]$.

We show how to build the upper solution, as the lower solution can be constructed similarly. Let z be the solution of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = \left(\lambda_1 - \frac{1}{R}\right)u + (a_0 - \lambda_1)(R+2) & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega. \end{array}$$

Choose, by Proposition 2.1, a constant $a > 0$ so large that $a\varphi_1 - z \geqslant 0$. We then easily verify that

$$\beta_1 = a\varphi_1 - z + R + 2$$

is an upper solution of (2.80), with $r = R$. This proves the claim.

We finally notice that the minimal and the maximal solution v and w of (2.36), we obtained in Claim 1, are minimal and maximal solutions in $\overline{\mathcal{V}} \cap [\alpha_1, \beta_1]$ of (2.80), with $r = R$. Theorem 2.42 implies that they are, respectively, $\mathcal{L}$-unstable from below and $\mathcal{L}$-unstable from above as solutions of (2.80), with $r = R$, and hence, by our choice of R, of (2.36). □

Stability and multiplicity of periodic solutions

We discuss in this section some relations between the stability properties and the multiplicity of solutions of (2.36). Propositions 2.45 and 2.47, which are direct consequences of Theorem 2.21 and Theorem 2.38, yield some results expressed in terms of $\mathcal{O}$-stability. Afterwards Propositions 2.46 and 2.48 provide an interpretation of these statements in terms of $\mathcal{L}$-stability.

PROPOSITION 2.45. *Assume* (D), (A), (C) *and* (N). *Let* u_1, u_2 *be solutions of* (2.36) *such that* $u_1 < u_2$, u_1 *is not properly* $\mathcal{O}$*-stable from above and* u_2 *is not properly* $\mathcal{O}$*-stable from below. Then there exists a solution* u_3 *of* (2.36) *with* $u_1 < u_3 < u_2$.

If moreover u_1 *is not* $\mathcal{O}$*-stable from above, then it is properly* $\mathcal{O}$*-unstable from above and there exists a solution* u_3 *of* (2.36), *with* $u_1 < u_3 < u_2$, *which is properly* $\mathcal{O}$*-stable from below. Similarly, if moreover* u_2 *is not* $\mathcal{O}$*-stable from below then it is properly* $\mathcal{O}$*-unstable from below and there exists a solution* u_3 *of* (2.36) *with* $u_1 < u_3 < u_2$ *which is properly* $\mathcal{O}$*-stable from above.*

Finally, if moreover both u_1 *is not* $\mathcal{O}$*-stable from above and* u_2 *is not* $\mathcal{O}$*-stable from below, then there exists a solution* u_3 *of* (2.36), *with* $u_1 < u_3 < u_2$, *which is* $\mathcal{O}$*-stable.*

PROOF. If u_1 is not properly $\mathcal{O}$-stable from above and u_2 is not properly $\mathcal{O}$-stable from below, then Theorem 2.21 implies that there exists a solution u_3 of (2.36), with $u_1 < u_3 < u_2$.

If u_1 is not $\mathcal{O}$-stable from above, then there exists $\varepsilon > 0$ such that every solution u of (2.36), with $u_1 < u < u_2$, satisfies $\|u - u_1\|_{C^{1,0}(\overline{Q}_T)} \geqslant \varepsilon$. By Lemma 2.16, there is a solution v of (2.36) such that $u_1 < v < u_2$, $\|v - u_1\|_{C^{1,0}(\overline{Q}_T)} \geqslant \varepsilon$ and there is no solution u of (2.36) with $u_1 < u < v$. Theorem 2.21 then implies that u_1 is properly $\mathcal{O}$-unstable from above and v is properly $\mathcal{O}$-stable from below.

Similarly, if u_2 is not $\mathcal{O}$-stable from below, then there exists a solution w of (2.36), with $u_1 < w < u_2$, such that there is no solution u of (2.36) with $w < u < u_2$. Theorem 2.21 implies that u_2 is properly $\mathcal{O}$-unstable from below and w is properly $\mathcal{O}$-stable from above.

Finally, if u_1 is not $\mathcal{O}$-stable from above and u_2 is not $\mathcal{O}$-stable from below, then there exist solutions v and w of (2.36), with $u_1 < v \leqslant w < u_2$, such that there is no solution u of (2.36) with either $u_1 < u < v$ or $w < u < u_2$. Further, v is properly $\mathcal{O}$-stable from below and w is properly $\mathcal{O}$-stable from above. Theorem 2.38 then implies that there exists a solution u_3 of (2.36), with $v \leqslant u_3 \leqslant w$, which is $\mathcal{O}$-stable. □

PROPOSITION 2.46. *Assume* (D), (A), (C) *and* (N). *Further, suppose that the second comparison principle holds. Let u_1, u_2 be solutions of* (2.36) *such that $u_1 < u_2$, u_1 is $\mathcal{L}$-unstable from above and u_2 is $\mathcal{L}$-unstable from below. Then there exists a solution u_3 of* (2.36) *with $u_1 \ll u_3 \ll u_2$ which is $\mathcal{L}$-stable.*

PROOF. Proposition 2.31 implies that u_1 is not $\mathcal{O}$-stable from above and u_2 is not $\mathcal{O}$-stable from below. Hence Proposition 2.45 yields the existence of a solution u_3 of (2.36) with $u_1 < u_3 < u_2$ which is $\mathcal{O}$-stable and therefore, by Proposition 2.31 again, it is $\mathcal{L}$-stable. □

REMARK 2.28. With reference to Proposition 2.46, we have that $u_1 \ll u_2$, u_1 is isolated from above in $C^{1,0}(\overline{Q}_T)$ and u_2 is isolated from below in $C^{1,0}(\overline{Q}_T)$, by Proposition 2.31. Hence, by Lemma 2.16, there exist solutions v and w of (2.36), with $u_1 \ll v \leqslant w \ll u_2$, such that there is no solution u satisfying either $u_1 < u < v$ or $w < u < u_2$. By Proposition 2.45 and Corollary 2.34, v is $\mathcal{L}$-asymptotically stable from below and w is $\mathcal{L}$-asymptotically stable from above.

The following results are dual versions of Proposition 2.45 and Proposition 2.46.

PROPOSITION 2.47. *Assume* (D), (A), (C) *and* (N). *Let u_1, u_2 be solutions of* (2.36) *such that $u_1 < u_2$, u_1 is not properly $\mathcal{O}$-unstable from above and u_2 is not properly $\mathcal{O}$-unstable from below. Then there exists a solution u_3 of* (2.36) *with $u_1 < u_3 < u_2$.*

If moreover u_1 is not $\mathcal{O}$-unstable from above, then it is properly $\mathcal{O}$-stable from above and there exists a solution u_3 of (2.36), *with $u_1 < u_3 < u_2$, which is properly $\mathcal{O}$-unstable from below. Similarly, if moreover u_2 is not $\mathcal{O}$-unstable from below, then it is properly $\mathcal{O}$-stable from below and there exists a solution u_3 of* (2.36), *with $u_1 < u_3 < u_2$, which is properly $\mathcal{O}$-unstable from above.*

PROOF. If u_1 is not properly $\mathcal{O}$-unstable from above and u_2 is not properly $\mathcal{O}$-unstable from below, then Theorem 2.21 implies that there exists a solution u_3 of (2.36) with $u_1 < u_3 < u_2$.

If u_1 is not $\mathcal{O}$-unstable from above, then there exists $\varepsilon > 0$ such that every solution u of (2.36) with $u_1 < u < u_2$ satisfies $\|u - u_1\|_{C^{1,0}(\overline{Q}_T)} \geqslant \varepsilon$. By Lemma 2.16, there exists a solution v of (2.36) such that $u_1 < v < u_2$, $\|v - u_1\|_{C^{1,0}(\overline{Q}_T)} \geqslant \varepsilon$ and there is no solution u of (2.36) with $u_1 < u < v$. Moreover, Theorem 2.21 implies that u_1 is properly $\mathcal{O}$-stable from above and v is properly $\mathcal{O}$-unstable from below.

Similarly, if u_2 is not $\mathcal{O}$-unstable from below, then there exists a solution w of (2.36) with $u_1 < w < u_2$ and there is no solution u of (2.36) such that $w < u < u_2$. Theorem 2.21 implies then that u_2 is properly $\mathcal{O}$-stable from below and w is properly $\mathcal{O}$-unstable from above. □

PROPOSITION 2.48. *Assume* (D), (A), (C) *and* (N). *Let* u_1, u_2 *be solutions of* (2.36) *such that* $u_1 < u_2$, u_1 *is* $\mathcal{L}$*-stable from above and* u_2 *is* $\mathcal{L}$*-stable from below. Then there exists a solution* u_3 *of* (2.36) *with* $u_1 < u_3 < u_2$.

If moreover u_1 *is isolated from above in* $C^{1,0}(\overline{Q}_T)$, *then there exists a solution* u_3 *of* (2.36), *with* $u_1 < u_3 < u_2$, *which is* $\mathcal{L}$*-unstable from below. Similarly, if moreover* u_2 *is isolated from below in* $C^{1,0}(\overline{Q}_T)$, *then there exists a solution* u_3 *of* (2.36), *with* $u_1 < u_3 < u_2$, *which is* $\mathcal{L}$*-unstable from above.*

PROOF. If we assume that there is no solution u_3 of (2.36), with $u_1 < u_3 < u_2$, then by Theorem 2.21 either u_1 is properly $\mathcal{O}$-unstable from above or u_2 is properly $\mathcal{O}$-unstable from below. In the former case there is a sequence $(\alpha_n)_n$ of proper regular lower solutions of (2.36) such that, for every n, $u_1 < \alpha_n < u_2$, $\alpha_n = 0$ on Σ_T and $\|\alpha_n - u_1\|_e \to 0$, as $n \to +\infty$. Fix $t_0 \in]0, T[$. Then, for every $\delta > 0$, there exists n such that, setting $u_0 = \alpha_n(\cdot, t_0)$, we have $u_0 \in W_p^{2-2/p}(\Omega)$, $u_0 = 0$ on $\partial\Omega$, $u_0 \geqslant u_1(\cdot, t_0)$ and $\|u_0 - u_1(\cdot, t_0)\|_e < \delta$. By Proposition 2.13, problem (2.13) has a nonextendible solution u, satisfying $u > u_1$ in $\overline{\Omega} \times [t_0, +\infty[$ and $\lim_{t\to+\infty} \|u(\cdot, t) - u_2(\cdot, t)\|_e = 0$. Therefore, we get $\limsup_{t\to+\infty} \|u(\cdot, t) - u_1(\cdot, t)\|_e \geqslant \frac{1}{2}\|u_1 - u_2\|_e$, thus contradicting the $\mathcal{L}$-stability from above of u_1. Similarly, we verify that u_2 cannot be properly $\mathcal{O}$-unstable from below. Accordingly, there exists a solution u_3 of (2.36), with $u_1 < u_3 < u_2$.

Suppose further that u_1 is isolated from above in $C^{1,0}(\overline{Q}_T)$, i.e. there is $\delta > 0$ such that every solution u of (2.36), with $u > u_1$, satisfies $\|u - u_1\|_{C^{1,0}(\overline{Q}_T)} \geqslant \delta$, then by Lemma 2.16, there exists a solution u_3 of (2.36), satisfying $u_3 \in [u_1, u_2]$ and $\|u_3 - u_1\|_{C^{1,0}(\overline{Q}_T)} \geqslant \delta$, such that there is no solution u, with $u_1 < u < u_3$. Arguing as in the previous step, we see that u_3 is properly $\mathcal{O}$-unstable from below. Hence, arguing as in Proposition 2.40, we conclude that u_3 is $\mathcal{L}$-unstable from below. Similarly, if we suppose further that u_2 is isolated from below, we derive the existence of a solution u_3 of (2.36) which is $\mathcal{L}$-unstable from above. □

REMARK 2.29. We stress again that the conclusions of Proposition 2.48 are obtained without requiring the validity of any comparison principle. On the other hand, if we assume in Proposition 2.48 that the second comparison principle holds, then u_1 is $\mathcal{L}$-asymptotically stable from above, provided it is isolated from above, and u_2 is $\mathcal{L}$-asymptotically stable from below, provided it is isolated from below. The $\mathcal{L}$-asymptotic stability from above of u_1 easily follows from Corollary 2.34. Indeed, any solution u_3 of (2.36), satisfying $u_1 \ll u_3 \leqslant u_2$ and such that there is no solution u with $u_1 < u < u_3$, is properly $\mathcal{O}$-unstable from below and hence there exist proper regular upper solutions β with $u_1 \ll \beta \ll u_3$. The analogous conclusion about u_2 is obtained in a similar way.

2.7. *Appendix: The principal eigenvalue*

This appendix is devoted to the definition and the discussion of some properties of the principal eigenvalue of the periodic problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = \lambda u & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega. \end{array} \tag{2.81}$$

Our first result is an extension of a similar statement proved in [82] (see also [27,18]) under stronger regularity conditions.

PROPOSITION 2.49. *Assume* (D) *and* (A). *Let $p > N + 2$ be fixed. Then there exist a number $\lambda_1 > 0$ and a function $\varphi_1 \in W^{2,1}_{p,B}(Q_T)$ such that*

$$L\varphi_1 = \lambda_1\varphi_1,$$

where L is defined by (2.47) *and* (2.48). *Moreover, the following conclusions hold*:

(i) $\varphi_1 \gg 0$;
(ii) *if $L\psi = \lambda_1\psi$ for some $\psi \in W^{2,1}_{p,B}(Q_T)$, then $\psi = k\varphi_1$ for some $k \in \mathbb{R}$*;
(iii) *if $L\psi = \lambda\psi$ for some $\psi \in W^{2,1}_{p,B}(Q_T)$, with $\psi > 0$, then $\lambda = \lambda_1$*;
(iv) *for every $\lambda < \lambda_1$ and $h \in L^p(Q_T)$, the problem $Lu - \lambda u = h$ has a unique solution.*

PROOF. Let us introduce the operator $S : C^{1,0}_B(\overline{Q}_T) \to C^{1,0}_B(\overline{Q}_T)$ which sends any $h \in C^{1,0}_B(\overline{Q}_T)$ onto the unique solution $u \in C^{1,0}_B(\overline{Q}_T)$ of (2.46), with $\sigma = 0$. S is compact and strongly positive with respect to the order induced in $C^{1,0}_B(\overline{Q}_T)$ by the cone $K = \{u \in C^{1,0}_B(\overline{Q}_T) \mid u \geqslant 0\}$. Observe that $\overset{\circ}{K} = \{u \in C^{1,0}_B(\overline{Q}_T) \mid u \gg 0\}$. The strong form of Krein–Rutman Theorem (see, e.g., [153, Theorem 7C]) yields the existence of $\varphi_1 \in W^{2,1}_{p,B}(Q_T)$ satisfying $L\varphi_1 = \lambda_1\varphi_1$, with λ_1 the reciprocal of the spectral radius $r(S)$ of S, as well as conditions (i)–(iii). Finally, condition (iv) is a consequence of Proposition 2.14, if $\lambda \leqslant 0$. For $\lambda \in \,]0, \lambda_1[$, let us consider the extension $\tilde{S}$ of S to $L^p(Q_T)$, i.e. the operator $\tilde{S} : L^p(Q_T) \to L^p(Q_T)$ which sends any $h \in L^p(Q_T)$ onto the unique solution $u \in L^p(Q_T)$ of (2.46), with $\sigma = 0$. $\tilde{S}$ is compact and positive with respect to the order induced in $L^p(Q_T)$ by the cone $\tilde{K} = \{u \in L^p(Q_T) \mid u \geqslant 0\}$. By the weak form of Krein–Rutman Theorem (see, e.g., [153, Proposition 7.26]), the spectral radius $r(\tilde{S})$ of $\tilde{S}$ is an eigenvalue of $\tilde{S}$, with eigenfunction $\tilde{\varphi}_1 \in L^p(Q_T)$ satisfying $\tilde{\varphi}_1(x,t) > 0$ a.e. in Q_T. Since, by condition (iii), $r(\tilde{S}) = r(S)$, the result follows from the Fredholm Alternative (see, e.g., [144, Theorem 1.L]) applied to the operator $\tilde{S}$. □

PROPOSITION 2.50. *Assume* (D) *and* (A). *Let* $q_1, q_2 \in L^\infty(Q_T)$ *be such that* $q_1(x,t) \leqslant q_2(x,t)$ *a.e. in* Q_T *and let* u_1, u_2 *be nontrivial solutions of* $Lu_1 = q_1 u_1$ *and* $Lu_2 = q_2 u_2$, *respectively. If* $u_2 \geqslant 0$, *then* $q_1 = q_2$ *and* $u_1 = cu_2$, *for some* $c \in \mathbb{R}$.

PROOF. Since $Lu_2 - q_2 u_2 \geqslant 0$ and $u_2 > 0$, we deduce from Remark 2.7 that $u_2 \gg 0$. If we set $c = \min\{d \in \mathbb{R} \mid du_2 \geqslant u_1\}$ and $v = cu_2 - u_1$, as $v \geqslant 0$, we get $Lv - q_1 v \geqslant 0$ and hence either $v \gg 0$ or $v = 0$. The minimality of c actually yields $v = cu_2 - u_1 = 0$. This finally implies $0 = Lv - q_1 v = (q_2 - q_1)cu_2$ and therefore $q_1 = q_2$. □

PROPOSITION 2.51. *Assume* (D) *and* (A). *Let* $q \in L^\infty(Q_T)$ *satisfy* $q(x,t) \leqslant \lambda_1$ *a.e. in* Q_T *and* $q(x,t) < \lambda_1$ *on a subset of positive measure. Then the problem* (2.79) *has a unique solution, for any given* $h \in L^p(Q_T)$ *with* $p > N+2$.

PROOF. Setting $q_1 = q$ and $q_2 = \lambda_1$, Proposition 2.50 implies that the homogeneous equation $Lu = qu$ has only the trivial solution. The Fredholm Alternative applied to the operator $T = q\tilde{S} : L^p(Q_T) \to L^p(Q_T)$, where $\tilde{S}$ has been defined in the proof of Proposition 2.49, yields the conclusion. □

PROPOSITION 2.52. *Assume* (D) *and* (A). *Then there exists* $\mu > \lambda_1$ *such that, if* $q \in L^\infty(Q_T)$ *satisfies* $\lambda_1 \leqslant q(x,t) \leqslant \mu$ *a.e. in* Q_T *and* $\lambda_1 < q(x,t)$ *on a subset of positive measure, problem* (2.79) *has a unique solution, for any given* $h \in L^p(Q_T)$ *with* $p > N+2$.

PROOF. Let us show that there exists $\mu > \lambda_1$ such that, if $q \in L^\infty(Q_T)$ satisfies $\lambda_1 \leqslant q(x,t) \leqslant \mu$ a.e. in Q_T and $\lambda_1 < q(x,t)$ on a set of positive measure, the problem

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = q(x,t)u & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega, \end{array}$$

has only the trivial solution. Indeed, assume by contradiction that, for every n, there exist $q_n \in L^\infty(Q_T)$, satisfying $\lambda_1 \leqslant q_n(x,t) \leqslant \lambda_1 + 1/n$ a.e. in Q_T and $\lambda_1 < q_n(x,t)$ on a set of positive measure, and a solution u_n of

$$\begin{array}{ll} \partial_t u_n + A(x,t,\partial_x)u_n = q_n(x,t)u_n & \text{in } Q_T, \\ u_n = 0 & \text{on } \Sigma_T, \\ u_n(\cdot,0) = u_n(\cdot,T) & \text{in } \Omega, \end{array}$$

with $\|u_n\|_{C^{1,0}(\overline{Q}_T)} = 1$. Proposition 2.50 implies that each u_n changes sign in Q_T. Possibly passing to a subsequence, $(u_n)_n$ converges weakly in $W_p^{2,1}(Q_T)$ and strongly in $C^{1,0}(\overline{Q}_T)$ to a nontrivial solution u of

$$\begin{array}{ll} \partial_t u + A(x,t,\partial_x)u = \lambda_1 u & \text{in } Q_T, \\ u = 0 & \text{on } \Sigma_T, \\ u(\cdot,0) = u(\cdot,T) & \text{in } \Omega, \end{array}$$

which, therefore, satisfies either $u \gg 0$, or $u \ll 0$. Hence Proposition 2.1 yields, for n large enough, either $u_n \gg 0$, or $u_n \ll 0$, which is a contradiction. The Fredholm Alternative applied to the operator $T = q\tilde{S} : L^p(Q_T) \to L^p(Q_T)$, where $\tilde{S}$ has been defined in the proof of Proposition 2.49, yields the conclusion. □

References

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York (1975).

[2] A. Adje, *Existence et multiplicité des solutions d'équations différentielles ordinaires du premier ordre à non-linéarité discontinue*, Ann. Soc. Sci. Bruxelles Sér I **101** (1987), 69–87.

[3] N.D. Alikakos, P. Hess and H. Matano, *Discrete order preserving semigroups and stability for periodic parabolic differential equations*, J. Differential Equations **82** (1989), 322–341.

[4] M.A.M. Alwash and N.G. Lloyd, *Nonautonomous equations related to polynomial two-dimensional systems*, Proc. Roy. Soc. Edinburgh Sect. A **105** (1987), 129–152.

[5] H. Amann, *On the existence of positive solutions of nonlinear elliptic boundary value problems*, Indiana Univ. Math. J. **21** (1971/72), 125–146.

[6] H. Amann, *Existence and multiplicity theorems for semi-linear elliptic boundary value problems*, Math. Z. **150** (1976), 281–295.

[7] H. Amann, *Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces*, SIAM Rev. **18** (1976), 620–709.

[8] H. Amann, *Periodic solutions of semilinear parabolic equations*, Nonlinear Analysis. Collection of Papers in Honor of E. Rothe, Academic Press, New York (1978), 1–29.

[9] H. Amann, A. Ambrosetti and G. Mancini, *Elliptic equations with noninvertible Fredholm linear part and bounded nonlinearities*, Math. Z. **158** (1978), 179–194.

[10] K.M. Andersen and A. Sandqvist, *On the characteristic exponent function for a first order ordinary differential equation*, Differential Integral Equations **8** (1995), 717–728.

[11] K.M. Andersen and A. Sandqvist, *Existence of closed solutions of an equation $\dot{x} = f(t, x)$, where $f'_x(t, x)$ is weakly convex or concave in x*, J. Math. Anal. Appl. **229** (1999), 480–500.

[12] K.B. Baĭkuzev and D.H. Karimov, *Periodic solutions of a nonlinear equation of parabolic type*, Fergan. Gos. Ped. Inst. Učen. Zap. Ser. Mat. Vyp. **1** (1965), 27–40 (in Russian).

[13] D.W. Bange, *Periodic solutions of a quasilinear parabolic differential equation*, J. Differential Equations **17** (1975), 61–72.

[14] M. Bardi, *An equation of growth of a single species with realistic dependence on crowding and seasonal factors*, J. Math. Biol. **17** (1983), 33–43.

[15] J.W. Bebernes and K. Schmitt, *Invariant sets and the Hukuhara–Kneser property for systems of parabolic partial differential equations*, Rocky Mountain J. Math. **7** (1977), 557–567.

[16] J.W. Bebernes and K. Schmitt, *On the existence of maximal and minimal solutions for parabolic partial differential equations*, Proc. Amer. Math. Soc. **73** (1979), 211–218.

[17] R. Bellman, *On the existence and boundedness of solutions of nonlinear partial differential equation of parabolic type*, Trans. Amer. Math. Soc. **64** (1948), 21–44.

[18] A. Beltramo and P. Hess, *On the principal eigenvalue of a periodic-parabolic operator*, Comm. Partial Differential Equations **9** (1984), 919–941.

[19] F. Brauer and D.A. Sánchez, *Periodic environments and periodic harvesting*, Natur. Resource Modeling **16** (2003), 233–244.

[20] J.L. Bravo, M. Fernández and A. Tineo, *A result of Ambrosetti-Prodi type for first order ODEs with concave and coercive right member*, Nonlinear Stud. **9** (2002), 407–414.

[21] J.L. Bravo, M. Fernández and A. Tineo, *The number of bifurcation points of a periodic one-parameter ODE with at most two periodic solutions*, Nonlinear Anal. **57** (2004), 3–22.

[22] H. Brezis, *Analyse fonctionnelle*, Masson, Paris (1983).

[23] F.E. Browder, *Periodic solutions of nonlinear equations of evolution in infinite dimensional spaces*, Lectures in Differential Equations, Vol. 1, Van Nostrand, New York (1964), 71–96.

[24] M. Calanchi and B. Ruf, *On the number of closed solutions for polynomial ODE's and a special case of Hilbert's 16th problem*, Adv. Differential Equations **7** (2002), 197–216.

[25] R.S. Cantrell and C. Cosner, *Spatial Ecology via Reaction–Diffusion Equations*, Wiley, Chichester (2003).

[26] S. Carl, *Existence of extremal periodic solutions for quasilinear parabolic equations*, Abstr. Appl. Anal. **2** (1997), 257–270.

[27] A. Castro and A.C. Lazer, *Results on periodic solutions of parabolic equations suggested by elliptic theory*, Boll. Un. Mat. Ital. B (6) **1** (1982), 1089–1104.

[28] M. Cherpion and C. De Coster, *Existence of solutions for first order singular problems*, Proc. Amer. Math. Soc. **128** (2000), 1779–1791.

[29] J.Á. Cid, *On uniqueness criteria for systems of ordinary differential equations*, J. Math. Anal. Appl. **281** (2003), 264–275.

[30] B.D. Coleman, *Nonautonomous logistic equations as models of the adjustment of populations to environmental change*, Math. Biosci. **45** (1979), 159–173.

[31] J.M. Cushing, *Stable positive periodic solutions of the time dependent logistic equation under possible hereditary influences*, J. Math. Anal. Appl. **60** (1977), 747–754.

[32] E.N. Dancer, *Upper and lower stability and index theory for positive mappings and applications*, Nonlinear Anal. **17** (1991), 205–217.

[33] E.N. Dancer and P. Hess, *On stable solutions of quasilinear periodic-parabolic problems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **14** (1987), 123–141.

[34] E.N. Dancer and P. Hess, *Stability of fixed points for order-preserving discrete-time dynamical systems*, J. Reine Angew. Math. **419** (1991), 125–139.

[35] E.N. Dancer and P. Hess, *Stable subharmonic solutions in periodic reaction-diffusion equations*, J. Differential Equations **108** (1994), 190–200.

[36] E.N. Dancer and R. Ortega, *The index of Lyapunov stable fixed points in two dimensions*, J. Dynam. Differential Equations **6** (1994), 631–637.

[37] D. Daners and P. Koch Medina, *Abstract Evolution Equations, Periodic Problems and Applications*, Pitman Res. Notes Math. Ser., Vol. 279, Longman Scientific and Technical, Harlow, Essex (1992).

[38] C. De Coster, *La méthode des sur et sous solutions dans l'étude de problèmes aux limites*, Doctoral Dissertation, Université Catholique de Louvain, Louvain-la-Neuve (1994).

[39] C. De Coster and P. Habets, *Two-point Boundary Value Problems: Lower and Upper Solutions*, Mathematics in Science and Engineering, Vol. 205, Elsevier, Amsterdam (2006).

[40] C. De Coster and M. Henrard, *Existence and localization of solutions for second order elliptic BVP in presence of lower and upper solutions without any order*, J. Differential Equations **145** (1998), 420–452.

[41] C. De Coster and P. Omari, *Unstable periodic solutions of a parabolic problem in the presence of non-well-ordered lower and upper solutions*, J. Funct. Anal. **175** (2000), 52–88.

[42] C. De Coster and P. Omari, *Stability and instability in periodic parabolic problems via lower and upper solutions*, Quaderni Matematici (II Serie) **539**, Trieste (2003).

[43] P. De Mottoni and A. Schiaffino, *On logistic equations with time periodic coefficients*, Pubbl. IAC **192**, Roma (1979).

[44] J. Deuel and P. Hess, *Nonlinear parabolic boundary value problems with upper and lower solutions*, Israel J. Math. **29** (1978), 92–104.

[45] G.C. Dong, *Nonlinear Partial Differential Equations of Second Order*, Transl. Math. Monographs, Vol. 95, Amer. Math. Soc., Providence, RI, 1991.

[46] A. El Hachimi and A. Lamrani Alaoui, *Existence of stable periodic solutions for quasilinear parabolic problems in the presence of well-ordered lower and upper-solutions*, Electron. J. Differential Equations **9** (2002), 117–126.

[47] R. Engelking, *General Topology*, Heldermann, Berlin (1989).

[48] P. Fife, *Solutions of parabolic boundary problems existing for all time*, Arch. Rational Mech. Anal. **16** (1964), 155–186.

[49] P. Fife, *Mathematical Aspects of Reacting and Diffusing Systems*, Lecture Notes in Biomath., Vol. 28, Springer, Berlin (1979).

[50] A.F. Filippov, *Differential Equations with Discontinuous Righthand Sides*, Kluwer Academic, Dordrecht (1988).

[51] A. Friedman, *Partial Differential Equations*, Holt, Rinehart and Winston, New York (1969).

[52] M. Frigon and D. O'Regan, *Existence results for some initial and boundary value problems without growth restrictions*, Proc. Amer. Math. Soc. **123** (1995), 207–216.

[53] G.F. Gause, *The Struggle for Existence*, Williams and Wilkins, Baltimore, MD (1934).

[54] J.P. Gossez and P. Omari, *Non-ordered lower and upper solutions in semilinear elliptic problems*, Comm. Partial Differential Equations **19** (1994), 1163–1184.

[55] P. Habets and P. Omari, *Existence and localization of solutions of second order elliptic problems using lower and upper solutions in the reversed order*, Topol. Methods Nonlinear Anal. **8** (1996), 25–56.

[56] J.K. Hale, *Dynamics of a scalar parabolic equation*, Canad. Appl. Math. Quart. **5** (1997), 209–305.

[57] J.K. Hale and H. Koçak, *Dynamics and Bifurcation*, Springer, New York (1991).

[58] P. Hess, *Periodic-Parabolic Boundary Value Problems and Positivity*, Pitman Res. Notes Math. Ser., Vol. 247, Longman Scientific and Technical, Harlow, Essex (1991).

[59] M.W. Hirsch, *Stability and convergence in strongly monotone dynamical systems*, J. Reine Angew. Math. **383** (1988), 1–53.

[60] M.W. Hirsch and H.L. Smith, *Monotone dynamical systems*, Handbook of Differential Equations, Ordinary Differential Equations, Vol. 2, North-Holland, Amsterdam (2005), 239–357.

[61] M.W. Hirsch and H.L. Smith, *Monotone maps: a review*, J. Differential Equations Appl. **11** (2005), 379–398.

[62] Yu.S. Ilyashenko, *Finiteness Theorems for Limit Cycles*, Amer. Math. Soc., Providence, RI (1991).

[63] Yu.S. Ilyashenko, *Hilbert-type numbers for Abel equations, growth and zeros of holomorphic functions*, Nonlinearity **13** (2000), 1337–1342.

[64] D.H. Karimov, *Sur les solutions périodiques des équations différentielles non-linéaires de type parabolique*, Dokl. Akad. Nauk. SSSR **25** (1939), 3–6.

[65] D.H. Karimov, *Über die periodischen Lösungen der nichtlinearen differentialen Gleichungen des parabolischen Typus*, Dokl. Akad. Nauk. SSSR **28** (1940), 403-406.

[66] D.H. Karimov, *On periodic solutions of nonlinear differential equations of parabolic type*, Dokl. Akad. Nauk. SSSR **56** (1947), 119–121 (in Russian).

[67] D.H. Karimov, *On periodic solutions of nonlinear differential equations of parabolic type*, Dokl. Akad. Nauk. SSSR **58** (1947), 969–972 (in Russian).

[68] J.L. Kazdan and F.W. Warner, *Remarks on some quasilinear elliptic equations*, Comm. Pure Appl. Math. **28** (1975), 567–597.

[69] Ju.S. Kolesov, *Certain tests for the existence of stable periodic solutions of quasi-linear parabolic equations*, Soviet Math. Dokl. **5** (1964), 1118–1120.

[70] Ju.S. Kolesov, *A test for the existence of periodic solutions to parabolic equations*, Soviet Math. Dokl. **7** (1966), 1318–1320.

[71] Ju.S. Kolesov, *Study of stability of solutions of second-order parabolic equations in the critical case*, Math. USSR Izv. **3** (1969), 1277–1291.

[72] Ju.S. Kolesov, *Schauder's principle and the stability of periodic solutions*, Soviet Math. Dokl. **10** (1969), 1290–1293.

[73] Ju.S. Kolesov, *Periodic solutions of quasilinear parabolic equations of second order*, Trans. Moscow Math. Soc. **21** (1970-71), 114–146.

[74] V.A. Kostitzin, *Sur la loi logistique et ses généralisations*, Acta Bioth. Ser. A. **5** (1940), 155–159.

[75] M.A. Krasnosel'skiĭ and P.E. Sobolevskiĭ, *On some non-linear problems for the partial differential equations*, Siberian Math. J. (1963), 129–133.

[76] M.A. Krasnosel'skiĭ and P.E. Sobolevskiĭ, *Structure of the set of solutions of an equation of parabolic type*, Amer. Math. Soc. Transl. Ser. 2 **51** (1966), 113–131.

[77] M.A. Krasnosel'skiĭ and P.P. Zabreĭko, *Geometrical Methods of Nonlinear Analysis*, Springer, Berlin (1984).

[78] T. Kusano, *Periodic solutions of parabolic partial differential equations*, Sugaku **18** (1966), 104–106 (in Japanese).

[79] G.S. Ladde, V. Lakshmikantham and A.S. Vatsala, *Monotone Iterative Techniques for Nonlinear Differential Equations*, Pitman, Boston (1985).

[80] O.A. Ladyzhenskaya, V.A. Solonnikov and N.N. Ural'tseva, *Linear and Quasilinear Equations of Parabolic Type*, Transl. Math. Monographs, Vol. 23, Amer. Math. Soc., Providence, RI (1968).

[81] J.M. Lasry and G.M. Troianiello, *On the set of solutions to a semilinear parabolic equation*, Comm. Partial Differential Equations **7** (1982), 1001–1021.

[82] A.C. Lazer, *Some remarks on periodic solutions of parabolic differential equations*, Dynamical Systems II, Academic Press, New York (1982), 227–246.

[83] G.M. Lieberman, *Second Order Parabolic Differential Equations*, World Scientific, Singapore (1996).

[84] G.M. Lieberman, *Time-periodic solutions of linear parabolic differential equations*, Comm. Partial Differential Equations **24** (1999), 631–663.

[85] A. Lins Neto, *On the number of solutions of the equation* $\frac{\mathrm{d}x}{\mathrm{d}t} = \sum_{j=0}^{n} a_j(t)x^j$, $0 \leqslant t \leqslant 1$, *for which* $x(0) = x(1)$, Invent. Math. **59** (1980), 67–76.

[86] E. Liz and R.L. Pouso, *Upper and lower solutions with "jumps"*, J. Math. Anal. Appl. **222** (1998), 484–493.

[87] N.G. Lloyd, *The number of periodic solutions of the equation* $\dot{z} = z^N + p_1(t)z^{N-1} + \cdots + p_N(t)$, Proc. London Math. Soc. (3) **27** (1973), 667–700.

[88] N.G. Lloyd, *On a class of differential equations of Riccati type*, J. London Math. Soc. (2) **10** (1975), 1–10.

[89] C. Marcelli and P. Rubbioni, *A new extension of classical Müller's theorem*, Nonlinear Anal. **28** (1997), 1759–1767.

[90] C. Marcelli and P. Rubbioni, *Some results on the method of lower and upper solutions for non-compact domains*, Nonlinear Anal. **30** (1997), 4903–4908.

[91] J.L. Massera, *The existence of periodic solutions of systems of differential equations*, Duke Math. J. **17** (1950), 457–475.

[92] H. Matano, *Asymptotic behavior and stability of solutions of semilinear diffusion equations*, Publ. Res. Inst. Math. Sci. Kyoto Univ. **15** (1979), 401–454.

[93] H. Matano, *Existence of nontrivial unstable sets for equilibriums of strongly order-preserving systems*, J. Fac. Sci. Tokyo **30** (1984), 645–673 + *Corrections*, J. Fac. Sci. Tokyo **34** (1987), 853–855.

[94] J. Mawhin, *First order ordinary differential equations with several periodic solutions*, Z. Angew. Math. Phys. **38** (1987), 257–265.

[95] J. Mawhin, *Qualitative behavior of the solutions of periodic first order scalar differential equations with strictly convex coercive nonlinearity*, Dynamics of Infinite Dimensional Systems (Lisbon, 1986), NATO Adv. Sci. Inst. Ser. F Comput. Systems Sci., Vol. 37, Springer, Berlin (1987), 151–159.

[96] J. Mawhin, *Riccati type differential equations with periodic coefficients*, Proceedings of the Eleventh International Conference on Nonlinear Oscillations (Budapest, 1987), János Bolyai Math. Soc., Budapest (1987), 157–163.

[97] J. Mawhin, *Les héritiers de Pierre-François Verhulst: une population dynamique*, Acad. Roy. Belg. Bull. Cl. Sci. (6) **13** (2002), 349–378.

[98] J. Mawhin, R. Ortega and A.M. Robles-Pérez, *Maximum principles for bounded solutions of the telegraph equation in space dimensions two and three and applications*, J. Differential Equations **208** (2005), 42–63.

[99] H.P. McKean and J.C. Scovel, *Geometry of some simple nonlinear differential operators*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **13** (1986), 299–346.

[100] W. Mlak, *Parabolic differential inequalities and Chaplighin's method*, Ann. Pol. Math. **8** (1960), 139–152.

[101] S. Moretto, *Sull'esistenza di soluzioni periodiche per l'equazione* $y' = f(x, y)$, Ann. Univ. Ferrara Sez. VII (N.S.) **8** (1958–59), 61–67.

[102] M. Müller, *Über das Fundamentaltheorem in der Theorie der gewöhnlichen Differentialgleichungen*, Math. Z. **26** (1927), 619–645.

[103] J.D. Murray, *Mathematical Biology*, second edition, Springer, Berlin (1993).

[104] P. Murthy, *Periodic solutions of two-dimensional forced systems: The Massera theorem and its extension*, J. Dynam. Differential Equations **10** (1998), 275–302.

[105] M. Nagumo, *Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen*, Proc. Phys. Math. Soc. Japan **24** (1942), 551–559.

[106] M. Nagumo and S. Simoda, *Note sur l'inégalité différentielle concernant les équations du type parabolique*, Proc. Japan Acad. **27** (1951), 536–593.

[107] R. Narasimhan, *On the asymptotic stability of parabolic differential equations*, J. Ration. Mech. Anal. **3** (1954), 303–313.

[108] J.J. Nieto, *Monotone iterates and stability for first-order ordinary differential equations*, Advances in Nonlinear Dynamics, Stability Control Theory Methods Appl., Vol. 5, Gordon and Breach, Amsterdam (1997), 107–115.

[109] F.I. Njoku and P. Omari, *Stability properties of periodic solutions of a Duffing equation in the presence of lower and upper solutions*, Appl. Math. Comput. **135** (2003), 471–490.

[110] M.N. Nkashama, *A generalized upper and lower solutions method and multiplicity results for nonlinear first-order ordinary differential equations*, J. Math. Anal. Appl. **140** (1989), 381–395.

[111] D. Núñez, *The method of lower and upper solutions and the stability of periodic oscillations*, Nonlinear Anal. **51** (2002), 1207–1222.

[112] F. Obersnel and P. Omari, *Old and new results for first order periodic ODEs without uniqueness: a comprehensive study by lower and upper solutions*, Adv. Nonlinear Stud. **4** (2004), 323–376.

[113] F. Obersnel and P. Omari, *On the Ambrosetti-Prodi problem for first order scalar periodic ODEs*, Applied and Industrial Mathematics in Italy, Ser. Adv. Math. Appl. Sci., Vol. 69, World Sci. Publishing, Hackensack, NJ (2005), 404–415.

[114] F. Obersnel and P. Omari, *Period two implies chaos for a class of ODEs*, Proc. Amer. Math. Soc., in press.

[115] A. Okubo, *Diffusion and Ecological Problems: Mathematical Models*, Springer, Berlin (1980).

[116] R. Ortega, *Some applications of the topological degree to stability theory*, Topological Methods in Differential Equations and Inclusions, NATO ASI Ser. C Math. Phys. Sci., Vol. 472, Kluwer Academic, Dordrecht (1995), 377–409.

[117] R. Ortega and A.M. Robles-Perez, *A maximum principle for periodic solutions of the telegraph equation*, J. Math. Anal. Appl. **221** (1998), 625–651.

[118] C.V. Pao, *Nonlinear Parabolic and Elliptic Equations*, Plenum Press, New York (1992).

[119] G. Peano, *Sull'integrabilità delle equazioni differenziali del primo ordine*, Atti Accad. Torino **21** (1885), 677–685.

[120] O. Perron, *Ein neuer Existenzbeweis für die Integrale der Differentialgleichung* $y' = f(x, y)$, Math. Ann. **76** (1915), 471–484.

[121] L.C. Piccinini, G. Stampacchia and G. Vidossich, *Ordinary Differential Equations in* $\mathbb{R}^n$*. Problems and Methods*, Springer, New York (1984).

[122] V.A. Pliss, *Nonlocal Problems of the Theory of Oscillations*, Academic Press, New York (1966).

[123] P. Poláčik, *Parabolic equations: asymptotic behavior and dynamics on invariant manifolds*, Handbook of Dynamical Systems, Vol. 2, North-Holland, Amsterdam (2002), 835–883.

[124] G. Prodi, *Questioni di stabilità per equazioni non lineari alle derivate parziali di tipo parabolico*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat. **10** (1951), 365–370.

[125] G. Prodi, *Soluzioni periodiche di equazioni non lineari di tipo parabolico*, Atti del IV Congresso dell'Unione Matematica Italiana, Taormina 25–31 Ott. (1951), 1–4.

[126] G. Prodi, *Soluzioni periodiche di equazioni alle derivate parziali di tipo parabolico e non lineari*, Rivista Mat. Univ. Parma **3** (1952), 265–290.

[127] G. Prodi, *Problemi al contorno non lineari per equazioni di tipo parabolico non lineari in due variabili—soluzioni periodiche*, Rend. Sem. Mat. Univ. Padova **23** (1954), 25–85.

[128] M.H. Protter and H.F. Weinberger, *Maximum Principles in Differential Equations*, Springer, New York (1984).

[129] S. Rosenblat, *Population models in a periodically fluctuating environment*, J. Math. Biol. **9** (1980), 23–36.

[130] E. Roxin, *Stability in general control systems*, J. Differential Equations **1** (1965), 115–150.

[131] D.A. Sánchez, *A note on periodic solutions of Riccati-type equations*, SIAM J. Appl. Math. **17** (1969), 957–959.

[132] D.A. Sánchez, *Computing periodic solutions of Riccati differential equations*, Appl. Math. Comput. **6** (1980), 283–287.

[133] A. Sandqvist and J. Cheng Fu, *Qualitative behavior of the solutions of periodic first order scalar differential equations with weakly concave nonlinearity*, Differential Integral Equations **5** (1992), 149–164.

[134] D.H. Sattinger, *Monotone methods in nonlinear elliptic and parabolic boundary value problems*, Indiana Univ. Math. J. **21** (1972), 979–1000.

[135] S. Shahshahani, *Periodic solutions of polynomial first order differential equations*, Nonlinear Anal. **5** (1981), 157–165.

[136] H.L. Smith, *Monotone Dynamical Systems. An Introduction to the Theory of Competitive and Cooperative Systems*, Math. Surveys Monographs, Vol. 41, Amer. Math. Soc., Providence, RI (1995).

[137] I.I. Šmulev, *Periodic solutions of boundary-value problems without initial conditions for quasi-linear parabolic equations*, Soviet Math. Dokl. **2** (1961), 1103–1107.

[138] I.I. Šmulev, *Periodic solutions of the first boundary problem for parabolic equations*, Amer. Math. Soc. Transl. Ser. 2 **79** (1969), 215–229.

[139] G. Stampacchia, *Le trasformazioni funzionali che presentano il fenomeno di Peano*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat. **7** (1949), 80–84.

[140] G.P. Szegö and G. Treccani, *Semigruppi di Trasformazioni Multivoche*, Lecture Notes in Math., Vol. 101, Springer, Berlin (1969).

[141] P. Takáč, *Linearly stable subharmonic orbits in strongly monotone time-periodic dynamical systems*, Proc. Amer. Math. Soc. **115** (1992), 691–698.

[142] A. Tineo, *A result of Ambrosetti–Prodi type for first-order ODEs with cubic non-linearities. I, II*, Ann. Mat. Pura Appl. **182** (2003), 113–128, 129–141.

[143] F.G. Tricomi, *Equazioni Differenziali*, second edition, Edizioni Scientifiche Einaudi, Torino (1953).

[144] G.M. Troianiello, *Elliptic Differential Equations and Obstacle Problems*, Plenum Press, New York (1987).

[145] K. Tso, *On an Aleksandrov–Bakel'man type maximum principle for second-order parabolic equations*, Comm. Partial Differential Equations **10** (1985), 543–553.

[146] R.R. Vance and E.A. Coddington, *A nonautonomous model of population growth*, J. Math. Biol. **27** (1989), 491–506.

[147] G. Vidossich, *Towards a theory for periodic solutions to first order differential equations*, ISAS-SISSA Trieste, Report 59/83/M (1983).

[148] G. Vidossich, *Multiple periodic solutions for first-order ordinary differential equations*, J. Math. Anal. Appl. **127** (1987), 459–469.

[149] W. von Wahl, *A-priori-Schranken für semi- und quasilineare parabolische Differentialgleichungen*, Math. Z. **128** (1972), 93–108.

[150] W. Walter, *Differential and Integral Inequalities*, Springer, New York (1970).

[151] H. Westphal, *Zur Abschätzung der Lösung nichtlinearer parabolischer Differentialgleichungen*, Math. Z. **51** (1949), 690–695.

[152] J.A. Yorke, *Invariance for ordinary differential equations*, Math. Systems Theory **1** (1967), 353–372.

[153] E. Zeidler, *Nonlinear Functional Analysis and its Applications I: Fixed Point Theorems*, Springer, New York (1986).

[154] T.I. Zelenyak, M.M. Lavrent'ev Jr., M.P. Vishnevskĭ, *Qualitative Theory of Parabolic Equations. Part 1*, VSP, Utrecht (1997).

[155] X.Q. Zhao, *Dynamical Systems in Population Biology*, CMS Books in Mathematics, Springer, New York (2003).

This page intentionally left blank

CHAPTER 4

Bifurcation Theory of Limit Cycles of Planar Systems

Maoan Han

Department of Mathematics, Shanghai Normal University, Guilin Road 100, Shanghai, 200234, China

Contents

1. Limit cycle and its perturbations 343
   1.1. Basic notations 343
   1.2. Multiplicity, stability and their property 345
   1.3. Perturbations of a limit cycle 355
2. Focus values and Hopf bifurcation 360
   2.1. Poincaré map and focus value 360
   2.2. Normal form and Lyapunov technique 368
   2.3. Hopf bifurcation 381
   2.4. Degenerate Hopf bifurcation 386
3. Perturbations of Hamiltonian systems 399
   3.1. General theory 399
   3.2. Existence of 2 and 3 limit cycles 412
   3.3. Near-Hamiltonian polynomial systems 420
   3.4. Homoclinic bifurcation 423
References 428

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 3
Edited by A. Cañada, P. Drábek and A. Fonda
© 2006 Elsevier B.V. All rights reserved

This page intentionally left blank

In this chapter we present the bifurcation theory of limit cycles of planar systems with relatively simple dynamics. The theory studies the changes of orbital behavior in the phase space, especially the number of limit cycles as we vary the parameters of the system. This theory has been considered by many mathematicians starting with Poincaré who first introduced the notion of limit cycles. A fundamental step towards modern bifurcation theory in differential equations occurred with the definition of structural stability and the classification of structurally stable systems in the plane in 1937 developed by Andronov, Leontovich and Pontryagin. More precisely, Andronov and Pontryagin introduced the notion of a rough system and presented the necessary and sufficient conditions of roughness for systems on the plane. Almost at the same time, Andronov and Leontovich carried out a systematic classification of all principal bifurcations of limit cycles on the plane for the simplest nonrough systems. A further development of the theory had taken yet another direction, namely by selecting bifurcation sets of codimension one for primary bifurcations, and of arbitrary codimension in the general case for degenerate bifurcations. In the two-dimensional case, as was proved in Andronov et al. [2], rough systems compose an open and dense set in the space of all systems on a plane. The nonrough systems fill the boundaries between different regions of structural stability in this space.

In the following sections we concentrate on an in-depth study of limit cycles with general methods of both local and global bifurcations in high codimensional case. Many results are closely related to the second part of Hilbert's 16th problem which concerns with the number and location of limit cycles of a planar polynomial vector field of degree n posed in 1901 by Hilbert [73].

1. Limit cycle and its perturbations

1.1. *Basic notations*

Consider a planar system defined on a region $G \subset \mathbb{R}^2$ of the form

$$\dot{x} = f(x), \tag{1.1}$$

where $f : G \to \mathbb{R}^2$ is a C^r function, $r \geqslant 1$. Then for any point $x_0 \in G$ (1.1) has a unique solution $\varphi(t, x_0)$ satisfying $\varphi(0, x_0) = x_0$. Let $\varphi^t(x_0) = \varphi(t, x_0)$. The family of the transformations $\varphi^t : G \to \mathbb{R}^2$ satisfy the following properties:

(i) $\varphi^0 = \mathrm{Id}$;
(ii) $\varphi^{t+s} = \varphi^t \circ \varphi^s$.

The function φ is called the flow generated by (1.1) or by the vector field f. Let $I(x_0)$ denote the maximal interval of definition of $\varphi(t, x_0)$ in t. If $x_0 \in G$ is such that $\varphi(t, x_0)$ is constant for all $t \in I(x_0)$, then $f(x_0) = 0$. In this case, x_0 is called *a singular point* of (1.1). A point that is not singular is called *a regular point*.

For any regular point $x_0 \in G$, the solution $\varphi(t, x_0)$ defines two planar curves as follows:

$$\gamma^+(x_0) = \{\varphi(t, x_0) \colon t \in I(x_0), t \geqslant 0\}, \quad \gamma^-(x_0) = \{\varphi(t, x_0) \colon t \in I(x_0), t \leqslant 0\},$$

which are called respectively the *positive* and the *negative orbit* of (1.1) through x_0. The union $\gamma(x_0) = \gamma^+(x_0) \cup \gamma^-(x_0)$ is called the *orbit* of (1.1) through x_0. The theorem about the existence and uniqueness of solutions ensures that there is one and only one orbit through any point in G. A *periodic orbit* of (1.1) is an orbit which is a closed curve. The minimal positive number satisfying $\varphi(T, x_0) = x_0$ is said to be the *period* of the periodic orbit $\gamma(x_0)$. Obviously, $\gamma(x_0)$ is a periodic orbit of period T if and only if the corresponding representation $\varphi(t, x_0)$ is a periodic solution of the same period.

DEFINITION 1.1. A periodic orbit of (1.1) is called a *limit cycle* if it is the only periodic orbit in a neighborhood of it. In other words, a limit cycle is an isolated periodic orbit in the set of all periodic orbits.

Now let (1.1) have a limit cycle L: $x = u(t), 0 \leqslant t \leqslant T$. Since (1.1) is autonomous, for any given point $p \in L$ we may suppose $p = u(0)$, and hence, $u(t) = \varphi(t, p)$. Further, for definiteness, let L be oriented clockwise. Introduce a unit vector

$$Z_0 = \frac{1}{|f(p)|}\big(-f_2(p), f_1(p)\big)^{\mathrm{T}}.$$

Then there exists a cross section l of (1.1) which passes through p and is parallel to Z_0. Clearly, a point $x_0 \in l$ near p can be written as $x_0 = p + aZ_0, a = (x_0 - p)^{\mathrm{T}} Z_0 \in \mathbb{R}$.

LEMMA 1.1. *There exist a constant* $\varepsilon > 0$ *and* C^r *functions* P *and* $\tau : (-\varepsilon, \varepsilon) \to \mathbb{R}$ *with* $P(0) = 0$ *and* $\tau(0) = T$ *such that*

$$\varphi\big(\tau(a), p + aZ_0\big) = p + P(a)Z_0 \in l, \quad |a| < \varepsilon. \tag{1.2}$$

PROOF. Define $Q(t, a) = [f(p)]^{\mathrm{T}}(\varphi(t, p + aZ_0) - p)$. We have

$$Q(T, 0) = 0, \qquad Q_t(T, 0) = \big|f(p)\big|^2 > 0.$$

Note that Q is C^r for (t, a) near $(T, 0)$. The implicit function theorem implies that a C^r function $\tau(a) = T + \mathrm{O}(a)$ exists satisfying

$$Q\big(\tau(a), a\big) = 0 \quad \text{or} \quad \big[f(p)\big]^{\mathrm{T}}\big(\varphi\big(\tau(a), p + aZ_0\big) - p\big) = 0.$$

It follows that the vector $\varphi(\tau(a), p + aZ_0) - p$ is parallel to Z_0. Hence, it can be rewritten as $\varphi(\tau(a), p + aZ_0) - p = P(a)Z_0$, where $P(a) = Z_0^{\mathrm{T}}(\varphi(\tau(a), p + aZ_0) - p)$. It is obvious that $P \in C^r$ for $|a|$ small with $P(0) = 0$. This ends the proof. □

The above proof tells us that the function τ is the time of the first return to l. By Definition 1.1, the periodic orbit L is a limit cycle if and only if $P(a) \neq a$ for $|a| > 0$ sufficiently small.

DEFINITION 1.2. The function $P : (-\varepsilon, \varepsilon) \to \mathbb{R}$ defined by (1.2) is called a *Poincaré map* or return map of (1.1) at $p \in l$.

For convenience, we sometimes use the notation $P : l \to l$.

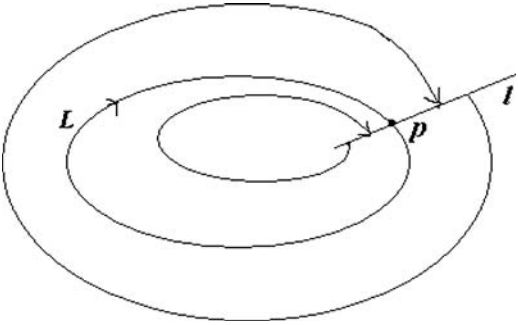


Fig. 1. Behavior of a stable limit cycle.

DEFINITION 1.3. The limit cycle L is said to be *outer stable* (*outer unstable*) if for $a > 0$ sufficiently small,

$$a\bigl(P(a) - a\bigr) < 0\ (> 0).$$

The limit cycle L is said to be *inner stable* (*inner unstable*) if the inequality above holds for $-a > 0$ sufficiently small. A limit cycle is called *stable* if it is both inner and outer stable. A limit cycle is called *unstable* if it is not stable.

For example, if L is stable, then the orbits near it behave like the phase portrait as shown in Fig. 1.

Let $P^k(a)$ denote the kth iterate of a under P. It is evident that $\{P^k(a)\}$ is monotonic in k and $P^k(a) > 0$ (< 0) for $a > 0$ (< 0). Thus, it is easy to see that L is outer stable if and only if $P^k(a) \to 0$ as $k \to \infty$ for all $a > 0$ sufficiently small. Similar conclusions hold for outer unstable, inner stable and inner unstable cases.

REMARK 1.1. If the limit cycle L is oriented anti-clockwise we can define its stability in a similar manner by using the Poincaré map P defined by (1.2). For instance, it is said to be inner stable (inner unstable) if $a(P(a) - a) < 0$ (> 0) for $a > 0$ sufficiently small.

DEFINITION 1.4. The limit cycle L is said to be *hyperbolic or of multiplicity one* if $P'(0) \neq 1$. It is said to have *multiplicity* k, $2 \leqslant k \leqslant r$, if $P'(0) = 1, P^{(j)}(0) = 0$, $j = 2, \dots, k-1$, $P^{(k)}(0) \neq 0$.

By Definition 1.3, one can see that L is stable (unstable) if $|P'(0)| < 1$ (> 1).

1.2. *Multiplicity, stability and their property*

Next, we give formulas for $P'(0)$ and $P''(0)$. For the purpose, let

$$v(\theta) = \frac{u'(\theta)}{|u'(\theta)|} = \bigl(v_1(\theta), v_2(\theta)\bigr)^{\mathrm{T}}, \qquad Z(\theta) = \bigl(-v_2(\theta), v_1(\theta)\bigr)^{\mathrm{T}},$$

and introduce a transformation of coordinates of the form

$$x = u(\theta) + Z(\theta)b, \quad 0 \leqslant \theta \leqslant T,\ |b| < \varepsilon. \tag{1.3}$$

LEMMA 1.2. *The transformation* (1.3) *carries* (1.1) *into the system*

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = 1 + g_1(\theta, b), \qquad \frac{\mathrm{d}b}{\mathrm{d}t} = A(\theta)b + g_2(\theta, b), \tag{1.4}$$

where

$$\begin{aligned}
&A(\theta) = Z^{\mathrm{T}}(\theta) f_x\big(u(\theta)\big)Z(\theta) = \operatorname{tr} f_x\big(u(\theta)\big) - \frac{\mathrm{d}}{\mathrm{d}\theta}\ln\big|f\big(u(\theta)\big)\big|,\\
&g_1(\theta, b) = h(\theta, b)\big[f\big(u(\theta) + Z(\theta)b\big) - f\big(u(\theta)\big)\big] - h(\theta, b)Z'(\theta)b,\\
&g_2(\theta, b) = Z^{\mathrm{T}}(\theta)\big[f\big(u(\theta) + Z(\theta)b\big) - f\big(u(\theta)\big) - f_x\big(u(\theta)\big)Z(\theta)b\big],\\
&h(\theta, b) = \big(\big|f\big(u(\theta)\big)\big| + v^{\mathrm{T}}(\theta)Z'(\theta)b\big)^{-1}v^{\mathrm{T}}(\theta),
\end{aligned}$$

and tr $f_x(u(\theta))$ *denotes the trace of the matrix* $f_x(u(\theta))$, *which is called the divergence of the vector field* f *evaluated at* $u(\theta)$.

PROOF. By (1.3) and (1.1) we have

$$\big(u' + Z'b\big)\frac{\mathrm{d}\theta}{\mathrm{d}t} + Z\frac{\mathrm{d}b}{\mathrm{d}t} = f(u + Zb). \tag{1.5}$$

In order to obtain (1.4) we need to solve $\frac{\mathrm{d}\theta}{\mathrm{d}t}$ and $\frac{\mathrm{d}b}{\mathrm{d}t}$ from (1.5). First, multiplying (1.5) by v^{T} from the left-hand side and using

$$v^{\mathrm{T}}Z = 0, \qquad v^{\mathrm{T}}f(u) = v^{\mathrm{T}}u' = \big|u'\big| = \big|f(u)\big|,$$

we can obtain

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \big[\big|f(u)\big| + v^{\mathrm{T}}Z'b\big]^{-1}v^{\mathrm{T}}f(u + Zb) = h(\theta, b)f(u + Zb).$$

Note that

$$h(\theta, b)f(u) = h(\theta, b)\big[f(u) + Z'b\big] - h(\theta, b)Z'b = 1 - h(\theta, b)Z'b.$$

It follows that

$$h(\theta, b)f(u + Zb) = h(\theta, b)\big[f(u + Zb) - f(u)\big] - h(\theta, b)Z'b + 1.$$

Then the first equation in (1.4) follows.

Now multiplying (1.5) by Z^{T} from the left and using

$$Z^{\mathrm{T}}Z=1, \qquad Z^{\mathrm{T}}f(u)=0, \qquad Z^{\mathrm{T}}Z'=\frac{1}{2}\left(|v|^2\right)'=0,$$

we obtain

$$\frac{\mathrm{d}b}{\mathrm{d}t}=Z^{\mathrm{T}}\left[f(u+Zb)-f(u)-f_x(u)Zb\right]+Z^{\mathrm{T}}f_x(u)Zb.$$

It is direct to prove that

$$Z^{\mathrm{T}}f_x(u)Z=\operatorname{tr}f_x(u)-\frac{\mathrm{d}}{\mathrm{d}\theta}\ln\left|f(u)\right|.$$

Then the second equation of (1.4) follows. This finishes the proof. $\square$

Set

$$B(\theta)=\left[f_x(u+Zb)\right]_b'\Big|_{b=0}, \quad C(\theta)=v^{\mathrm{T}}\left[f_x(u)Z-Z'(\theta)\right], \tag{1.6}$$

and

$$R(\theta,b)=\frac{A(\theta)b+g_2(\theta,b)}{1+g_1(\theta,b)}.$$

Then by Lemma 1.2, we can write

$$\begin{aligned}R(\theta,b)&=A(\theta)b+\frac{1}{2}\left[Z^{\mathrm{T}}BZ-\frac{2AC}{|f(u)|}\right]b^2+\mathrm{O}\left(b^3\right)\\&=A(\theta)b+\frac{1}{2}A_1(\theta)b^2+\mathrm{O}\left(b^3\right).\end{aligned} \tag{1.7}$$

For $|b|$ small we have from (1.4)

$$\frac{\mathrm{d}b}{\mathrm{d}\theta}=R(\theta,b) \tag{1.8}$$

which is a T-periodic equation. From Lemma 1.2 we know that the function R is C^{r-1} in (θ,b) and C^r in b. Let $b(\theta,a)$ denote the solution of (1.8) with $b(0,a)=a$. We have:

LEMMA 1.3. *$P(a)=b(T,a)$.*

PROOF. Consider the equation

$$\frac{\mathrm{d}\theta}{\mathrm{d}t}=1+g_1\left(\theta,b(\theta,a)\right).$$

It has a unique solution $\theta = \theta(t, a)$ satisfying $\theta(0, a) = 0$. From (1.7) it implies $b(\theta, 0) = 0$. This yields $\theta(T, 0) = T$, $\frac{\partial\theta}{\partial t}(T, 0) = 1$. Hence, by the implicit function theorem a unique function $\tilde{\tau}(a) = T + \mathrm{O}(a)$ exists such that $\theta(\tilde{\tau}, a) = T$.

For $x_0 = u(0) + Z(0)a$, we have by (1.3)

$$\varphi(t, x_0) = u\big(\theta(t, a)\big) + Z\big(\theta(t, a)\big)b\big(\theta(t, a), a\big).$$

In particular,

$$\varphi(\tilde{\tau}, x_0) = u(T) + Z(T)b(T, a) = u(0) + Z(0)b(T, a).$$

Thus, it follows from Lemma 1.1 that $\tau = \tilde{\tau}$ and $P(a) = b(T, a)$.

The proof is completed. □

For $|a|$ small we can write

$$b(\theta, a) = b_1(\theta)a + b_2(\theta)a^2 + \mathrm{O}\big(a^3\big),$$

where $b_1(0) = 1, b_2(0) = 0$. By (1.7) and (1.8) one can obtain $b_1' = Ab_1, b_2' = Ab_2 + \frac{1}{2}A_1b_1^2$, which give

$$b_1(\theta) = \exp\int_0^\theta A(s)\,\mathrm{d}s, \qquad b_2(\theta) = b_1(\theta)\int_0^\theta \frac{1}{2}A_1(s)b_1(s)\,\mathrm{d}s.$$

Then by Lemma 1.3 we have

$$P'(0) = b_1(T) = \exp\int_0^T A(s)\,\mathrm{d}s = \exp\int_0^T \operatorname{tr} f_x\big(u(t)\big)\,\mathrm{d}t,$$

$$P''(0) = 2b_2(T) = b_1(T)\int_0^T A_1(s)b_1(s)\,\mathrm{d}s.$$

Thus, noting (1.7) we obtain the following theorem.

THEOREM 1.1. *Suppose P is a Poincaré map of* (1.1) *at $p \in L$. Then*

$$\text{(i)}\quad P'(0) = \exp\oint_L \operatorname{div} f\,\mathrm{d}t, \quad \operatorname{div} f = \operatorname{tr} f_x,$$

$$\text{(ii)}\quad P''(0) = P'(0)\int_0^T \mathrm{e}^{\int_0^t A(s)\,\mathrm{d}s}\left[Z^T(t)B(t)Z(t) - \frac{2A(t)C(t)}{|f(u(t))|}\right]\mathrm{d}t.$$

Hence, L is stable (*unstable*) *if $I(L) = \oint_L \operatorname{div} f\,\mathrm{d}t < 0$ (> 0).*

We remark that Theorem 1.1 remains true in the case of counter clockwise orientation of L.

EXAMPLE 1.1. Consider the quadratic system

$$
\begin{aligned}
\dot{x} &= -y(1+cx) - \bigl(x^2+y^2-1\bigr),\\
\dot{y} &= x(1+cx), \quad 0<c<1.
\end{aligned}
$$

This system has the circle L: $x^2+y^2=1$ as its limit cycle. We claim that the cycle is unstable.

In fact, we have

$$
\begin{aligned}
I(L) &= \oint_L (-2x-cy)\,\mathrm{d}t = \oint_L \left(\frac{c\,\mathrm{d}x}{1+cx} - \frac{2\,\mathrm{d}y}{1+cx}\right)\\
&= \iint_{x^2+y^2\leqslant 1} \frac{\mathrm{d}x\,\mathrm{d}y}{(1+cx)^2} > 0.
\end{aligned}
$$

EXAMPLE 1.2. The system

$$
\begin{aligned}
\dot{x} &= -y - x\bigl(x^2+y^2-1\bigr)^2,\\
\dot{y} &= x - y\bigl(x^2+y^2-1\bigr)^2
\end{aligned}
$$

has a unique limit cycle given by L: $(x,y)=(\cos t,\sin t), 0\leqslant t\leqslant 2\pi$. For the system, it is easy to see that $v(\theta)=(-\sin\theta,\cos\theta)^{\mathrm{T}}$, $Z(\theta)=(-\cos\theta,-\sin\theta)^{\mathrm{T}}$. By Lemma 1.2 and (1.6) we then have

$$
A(\theta)=0, \quad B(\theta)=\begin{pmatrix} 8\cos^2\theta & 8\sin\theta\cos\theta\\ 8\sin\theta\cos\theta & 8\sin^2\theta \end{pmatrix}.
$$

Thus from Theorem 1.1 it follows $P'(0)=1$, $P''(0)=16\pi$. This shows that L is a limit cycle of multiplicity 2.

From (1.6) and formulas for $P'(0)$ and $P''(0)$ in Theorem 1.1 the derivatives $P'(0)$ and $P''(0)$ are independent of the choice of the cross section l. This fact suggests that the stability and the multiplicity of a limit cycle should have the same property. Below we will prove this in detail even if the cross section l is taken as a C^r smooth curve.

To do this, let L be a limit cycle of (1.1) as before and let l_1 be a C^r curve which has an intersection point $p_1\in L$ with L and is not tangent to L at p_1. Then it can be represented as

$$
l_1\colon\ x=p_1+q(a), \quad q(0)=0,\ \det\bigl(f(p_1),q'(0)\bigr)>0,
$$

where $q:(-\varepsilon,\varepsilon)\to\mathbb{R}$ is C^r for a constant $\varepsilon>0$. The condition $\det(f(p_1),q'(0))>0$ means that the point $p_1+q(a)$ is outside L if and only if $a>0$.

In the same way as Lemma 1.1 we can prove that there exist two C^r functions

$$
P_1,\tau_1:(-\varepsilon,\varepsilon)\to\mathbb{R}, \quad P_1(0)=0, \quad \tau_1(0)=T
$$

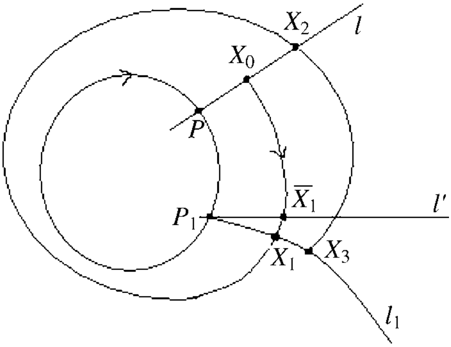


Fig. 2. Two Poincaré maps.

such that

$$\varphi\big(\tau_1(a), p_1 + q(a)\big) = p_1 + q\big(P_1(a)\big) \in l_1. \tag{1.9}$$

This yields another Poincaré map $P_1 : (-\varepsilon, \varepsilon) \to \mathbb{R}$.

LEMMA 1.4. *Let P and P_1 be two Poincaré maps defined by* (1.2) *and* (1.9), *respectively. Then there exists a C^r function $h_1 : (-\varepsilon, \varepsilon) \to \mathbb{R}$ with $h_1(0) = 0$, $h_1'(0) > 0$ such that $h_1 \circ P = P_1 \circ h_1$.*

PROOF. Since $p = u(0)$ we can suppose $p_1 = u(t_1)$ for some $t_1 \in [0, T)$. Similar to Lemma 1.1 again, there exist two C^r functions h_1 and τ^*, both from $(-\varepsilon, \varepsilon)$ to $\mathbb{R}$, with $h_1(0) = 0$ and $\tau^*(0) = t_1$ such that

$$\varphi\big(\tau^*(a), p + aZ_0\big) = p_1 + q\big(h_1(a)\big) \in l_1. \tag{1.10}$$

See Fig. 2. Let $x_0 = p + aZ_0$, $x_1 = \varphi(\tau^*(a), x_0)$, $x_2 = \varphi(\tau(a), x_0)$. By (1.10) and (1.2) we have $x_1 = p_1 + q(a_1)$, $a_1 = h_1(a)$ and $x_2 = p + P(a)Z_0$. Hence, by (1.9) and (1.10) we have

$$\begin{aligned}&\varphi\big(\tau_1(a_1), x_1\big) = p_1 + q\big(P_1(a_1)\big),\\ &\varphi\big(\tau^*\big(P(a)\big), x_2\big) = p_1 + q(a_2), \quad a_2 = h_1\big(P(a)\big).\end{aligned}$$

On the other hand, by the flow property of φ we have

$$\begin{aligned}x_3 &= \varphi\big(\tau_1(a_1), x_1\big) = \varphi\big(\tau_1(a_1) + \tau^*(a), x_0\big) = \varphi\big(\tau^*\big(P(a)\big) + \tau(a), x_0\big)\\ &= \varphi\big(\tau^*\big(P(a)\big), x_2\big),\end{aligned}$$

which, together with the above, follows that $q(P_1(a_1)) = q(a_2)$ or $a_2 = P_1(a_1)$. Hence $h_1 \circ P = P_1 \circ h_1$.

It needs only to prove $h_1'(0) > 0$. Let $a \geqslant 0$. Introduce one more cross section below

$$l'\colon\ x = u(t_1) + Z(t_1)a, \quad 0 \leqslant a \leqslant \varepsilon.$$

Let $\tilde{\tau}_1(a) = t_1 + \mathrm{O}(a)$ be such that $\theta(\tilde{\tau}_1, a) = t_1$. By (1.3) we have

$$\overline{x}_1 = \varphi(\tilde{\tau}_1, x_0) = u(t_1) + Z(t_1)b(t_1, a) \in l'.$$

Then $b(t_1, a) = |p_1\overline{x}_1|$. By the proof of Lemma 1.3,

$$\frac{\partial b}{\partial a}(t_1, 0) = \exp\int_0^{t_1} A(s)\,\mathrm{d}s > 0.$$

Consider the triangle formed by points p_1, x_1 and $\overline{x}_1$. There exists a point x^* on the orbital arc $\widehat{x_1\overline{x}_1}$ such that $f(x^*)$ is parallel to the side $x_1\overline{x}_1$. Since the arc $\widehat{x_1\overline{x}_1}$ approaches p_1 as $a \to 0$ we have $x^* \to p_1$, $f(x^*) \to f(p_1)$ as $a \to 0$. Hence, if we let α_1 denote the angle between sides $p_1\overline{x}_1$ and $\overline{x}_1x_1$, and α_2 the angle between sides p_1x_1 and $\overline{x}_1x_1$, then we have $\alpha_1 \to \frac{\pi}{2}, \alpha_2 \to \alpha_0$ as $a \to 0$, where $\alpha_0 \in (0, \frac{\pi}{2}]$ is the angle between the vectors $f(p_1)$ and $q'(0)$. That is, α_0 is the angle between L and l_1 at p_1. By the Sine theorem, it follows

$$\frac{|p_1\overline{x}_1|}{\sin\alpha_2} = \frac{|p_1x_1|}{\sin\alpha_1}, \quad \text{or}$$

$$\bigl|q\bigl(h_1(a)\bigr)\bigr| = \frac{\sin\alpha_1}{\sin\alpha_2}b(t_1, a) = \frac{a}{\sin\alpha_0}\exp\int_0^{t_1} A(s)\,\mathrm{d}s\bigl(1 + \mathrm{O}(a)\bigr).$$

On the other hand, $q(h_1(a)) = q'(0)h_1'(0)a + \mathrm{O}(a^2)$ which gives

$$\bigl|q\bigl(h_1(a)\bigr)\bigr| = \bigl|q'(0)\bigr| \cdot \bigl|h_1'(0)\bigr|a + \mathrm{O}\bigl(a^2\bigr), \quad a > 0.$$

Hence, we obtain

$$\bigl|h_1'(0)\bigr| = \frac{1}{|q'(0)|\sin\alpha_0}\exp\int_0^{t_1} A(s)\,\mathrm{d}s \neq 0.$$

Noting that $h_1(a) > 0$ for $a > 0$ we have $h_1'(0) > 0$. The proof is completed. □

COROLLARY 1.1. *The stability and the multiplicity of the limit cycle L are independent of the choice of cross sections.*

PROOF. By Lemma 1.4 we have

$$h_1'(\overline{a})\bigl[P(a) - a\bigr] = P_1\bigl(h_1(a)\bigr) - h_1(a), \tag{1.11}$$

where $\overline{a}$ lies between a and $P(a)$. By (1.11) and Definition 1.3, the stability of L under P is the same as that of it under P_1. Also, if the limit cycle L has multiplicity k under P, then $P(a) - a = c_k a^k + \mathrm{O}(a^{k+1})$ for some $c_k \neq 0$. It follows from (1.11) that

$$P_1(a) - a = \overline{c}_k a^k + \mathrm{O}\big(a^{k+1}\big), \quad \overline{c}_k = \frac{c_k}{[h_1'(0)]^{k-1}}.$$

Therefore, L has the same multiplicity under P_1. This ends the proof. □

In the following we discuss the relation of equivalence between two planar systems. Consider (1.1) and another system of the form

$$\dot{y} = g(y), \tag{1.12}$$

where $g : D \to \mathbb{R}^2$ is C^r $(r \geqslant 1)$ on a region $D \subset \mathbb{R}^2$.

DEFINITION 1.5. Let $U \subset G$, $V \subset D$ be two regions. Two planar systems defined by vector fields $f|_U$ and $g|_V$ are said to be C^k $(1 \leqslant k \leqslant r)$ *equivalent* if there exists a C^k diffeomorphism $h : U \to V$ which takes orbits of (1.1) on U to orbits of (1.12) on V preserving their orientation.

Let $\psi(t, y)$ denote the flow generated by (1.12). Then under the condition of Definition 1.5 a C^k function $s(t, x)$ exists with $s(0, x) = 0$, $\frac{\partial s}{\partial t} > 0$ such that

$$h\big(\varphi(t, x)\big) = \psi\big(s(t, x), h(x)\big), \tag{1.13}$$

as long as $\varphi(t, x) \in U$.

If the functions f and g have the following relationship:

$$g\big(h(x)\big) = \mathrm{D}h(x) f(x),$$

then the system (1.12) is obtained from (1.1) by making the coordinate transformation $y = h(x)$. In this case, (1.13) becomes $h \circ \varphi^t = \psi^t \circ h$. If f and g satisfy $g(x) = K(x) f(x)$ where $K : \mathbb{R}^2 \to \mathbb{R}$ is a C^r positive function, then orbits of (1.1) and (1.12) are identical, and the flows φ and ψ satisfy

$$K\big(\varphi(t, x)\big)\frac{\mathrm{d}\varphi}{\mathrm{d}t} = g\big(\varphi(t, x)\big), \qquad \frac{\mathrm{d}\psi}{\mathrm{d}s} = g\big(\psi(s, x)\big).$$

Hence, we are to sink a function $s(t, x)$ with $s(0, x) = 0$ and satisfying

$$\varphi(t, x) = \psi\big(s(t, x), x\big).$$

Differentiating the equality in t gives

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \frac{\mathrm{d}\psi}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t}.$$

Substituting the equality above into the previous equations yields

$$\frac{\mathrm{d}\psi}{\mathrm{d}s} = K\big(\varphi(t,x)\big)\frac{\mathrm{d}\varphi}{\mathrm{d}t} = K\big(\varphi(t,x)\big)\frac{\mathrm{d}\psi}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t}.$$

Thus, the function s should satisfy $\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{1}{K(\varphi(t,x))}$, which has the solution

$$s(t,x) = \int_0^t \frac{\mathrm{d}t}{K(\varphi(t,x))}.$$

With this choice of $s(t,x)$ above one can prove $\varphi(t,x) = \psi(s(t,x),x)$ by the uniqueness of initial solutions.

For C^k equivalent systems we have the following lemma.

LEMMA 1.5. *Suppose* (1.1) *and* (1.12) *are C^k equivalent under a C^k diffeomorphism $h: U \to V$, $1 \leqslant k \leqslant r$. Let $L \subset U$ and $L_1 = h(L) \subset V$ be limit cycles of* (1.1) *and* (1.12) *respectively. Then*:

(i) *The cycles L and L_1 have the same multiplicity.*

(ii) *The inner* (*outer*) *stability of L is the same as the inner* (*outer*) *stability of L_1 if $h((Int.L) \cap U) \subset Int.L_1$, and the inner* (*outer*) *stability of L is the same as the outer* (*inner*) *stability of L_1 if $h((Int.L) \cap U) \subset Ext.L_1$.*

PROOF. For the limit cycle L, take a cross section l at $p \in L$ as before. Then the curve $l_1 = h(l)$ is a cross section of L_1 at $p_1 = h(p)$. Let

$$Z_1 = \frac{1}{|g(p_1)|}\big(-g_2(p_1), g_1(p_1)\big)^{\mathrm{T}}, \quad Z_1' = h_x(p)Z_0.$$

Note that $h(p + aZ_0) = p_1 + Z_1'a + \mathrm{O}(a^2)$. It is easy to see that the vector Z_1' is tangent to l_1 at p_1. According to the orientation of L_1 and the direction of Z_1' there are four cases to consider as follows:

- Case 1. $Z_1 \cdot Z_1' > 0$ with L_1 clockwise oriented;
- Case 2. $Z_1 \cdot Z_1' < 0$ with L_1 clockwise oriented;
- Case 3. $Z_1 \cdot Z_1' > 0$ with L_1 counter clockwise oriented;
- Case 4. $Z_1 \cdot Z_1' < 0$ with L_1 counter clockwise oriented.

See Fig. 3.

By (1.13) the periods T of L and T_1 of L_1 have the relation $T_1 = s(T, p)$. Let P be the Poincaré map near L defined by (1.2). The cross section l_1 has a representation

$$y = h(p + aZ_0) = p_1 + q(a),$$

where $q(a) = h(p + aZ_0) - h(p) = Z_1'a + \mathrm{O}(a^2)$. Introduce a function q_1 as follows:

$$q_1(a) = \begin{cases} q(a) & \text{for case 1 or 3,} \\ q(-a) & \text{for case 2 or 4.} \end{cases}$$

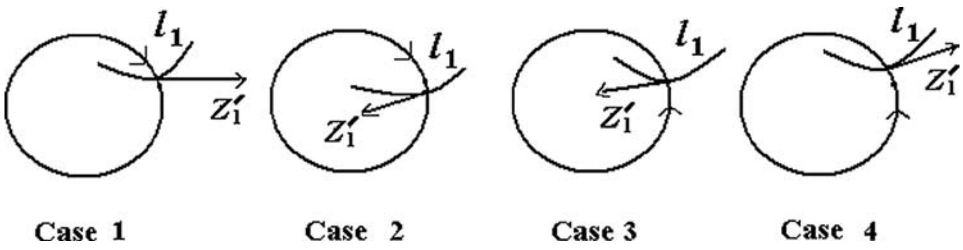


Fig. 3. Four possible cases for L_1 and Z_1'.

Then l_1 can be rewritten as l_1: $y = p_1 + q_1(a)$, $|a| \ll 1$ with q_1 satisfying $Z_1' \cdot q_1'(0) > 0$. On l_1 we can define a Poincaré map P_1 near L_1 by

$$\psi\big(\tau_1(a), p_1 + q_1(a)\big) = p_1 + q_1\big(P_1(a)\big) \in l_1, \tag{1.14}$$

where $\tau_1(a) = T_1 + \mathrm{O}(a)$ is the time of the first return to l_1.

Since $h(p + aZ_0) = p_1 + q(a) = p_1 + q_1((-1)^{i-1}a)$ for case i, we have

$$h\big(p + P(a)Z_0\big) = p_1 + q\big(P(a)\big) = p_1 + q_1\big((-1)^{i-1}P(a)\big) \in l_1$$

and hence by (1.2)

$$h\big(\varphi\big(\tau(a), p + aZ_0\big)\big) = p_1 + q_1\big((-1)^{i-1}P(a)\big) \in l_1. \tag{1.15}$$

On the other hand, by (1.13) we have

$$\begin{aligned} h\big(\varphi\big(\tau(a), p + aZ_0\big)\big) &= \psi\big(s\big(\tau(a), p + aZ_0\big), h(p + aZ_0)\big) \\ &= \psi\big(\tau^*(a), p_1 + q_1\big((-1)^{i-1}a\big)\big) \in l_1, \end{aligned} \tag{1.16}$$

for case i, where $\tau^*(a) = s(\tau(a), p + aZ_0) = T_1 + \mathrm{O}(a)$.

Hence, comparing (1.14) with (1.16) we obtain

$$\tau^*(a) = \tau_1\big((-1)^{i-1}a\big), \qquad h\big(\varphi\big(\tau(a), p + aZ_0\big)\big) = p_1 + q_1\big(P_1\big((-1)^{i-1}a\big)\big).$$

Therefore, it follows from (1.15) that $(-1)^{i-1}P(a) = P_1((-1)^{i-1}a)$ for case i. Thus for case i we have

$$a\big(P(a) - a\big) = (-1)^{i-1}a\big[P_1\big((-1)^{i-1}a\big) - (-1)^{i-1}a\big].$$

Hence, similar to Corollary 1.1, one can prove easily that L and L_1 have the same multiplicity. Furthermore, they have the same stability for cases 1 and 4. However, for cases 2 and 3, the inner (outer) stability of L is the same as the outer (inner) stability of L_1.

Then noting that $h((Int.L) \cap U) \subset Int.L_1$ for cases 1 and 4, and $h((Int.L) \cap U) \subset Ext.L_1$ for cases 2 and 3, the proof follows from Corollary 1.1. $\square$

EXAMPLE 1.3. Consider the system

$$\dot{x} = y, \qquad \dot{y} = -x + y\bigl(x^2 + y^2 - 1\bigr).$$

By Theorem 1.1, the system has a unique limit cycle L: $x^2 + y^2 = 1$, and it is hyperbolic and unstable.

Define four functions

$$h_1(x, y) = (x, y)^{\mathrm{T}}, \quad h_2(x, y) = \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2}\right)^{\mathrm{T}},$$
$$h_3(x, y) = \left(\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right)^{\mathrm{T}}, \quad h_4(x, y) = (y, x)^{\mathrm{T}},$$

and a region $U = \{(x, y)\colon x^2 + y^2 > \frac{1}{2}\}$.

Let

$$\dot{u} = g_{i1}(u, v), \qquad \dot{v} = g_{i2}(u, v)$$

denote the system obtained from the previous cubic system by making the coordinate transformation $(u, v)^{\mathrm{T}} = h_i(x, y)$, $(x, y) \in U$. Note that the unit circle is invariant under each h_i. It is evident that the case i in Fig. 3 occurs with $L_1 = L$ for each $i \in \{1, 2, 3, 4\}$.

This example shows that each case in Fig. 3 can happen. In particular, the orientation of a limit cycle may be changed under a coordinate transformation.

1.3. *Perturbations of a limit cycle*

In the rest, we return to the perturbations of a limit cycle.

Consider the following system

$$\dot{x} = f(x) + F(x, \mu) \tag{1.17}$$

where $\mu \in \mathbb{R}^m$ is a vector parameter with $m \geqslant 1$, and $F : G \times \mathbb{R}^m \to \mathbb{R}^2$ is a C^r function with $F(x, 0) = 0$. Thus, system (1.1) is the unperturbed system of (1.17). As before, let (1.1) have a limit cycle L: $x = u(t)$, $0 \leqslant t \leqslant T$. Then completely similar to Lemma 1.2 we have the following lemma.

LEMMA 1.6. *The transformation* (1.3) *carries* (1.17) *into the system*

$$\begin{aligned} \dot{\theta} &= 1 + g_1(\theta, b) + h(\theta, b)F\bigl(u(\theta) + Z(\theta)b, \mu\bigr), \\ \dot{b} &= A(\theta)b + g_2(\theta, b) + Z^{\mathrm{T}}(\theta)F\bigl(u(\theta) + Z(\theta)b, \mu\bigr), \end{aligned} \tag{1.18}$$

where the functions A, g_1, g_2 and h are the same as those given in Lemma 1.2.

By (1.18) we have

$$\frac{\mathrm{d}b}{\mathrm{d}\theta} = R(\theta, b, \mu), \tag{1.19}$$

where $R(\theta, b, \mu) = A(\theta)b + Z^{\mathrm{T}}(\theta)F_\mu(u(\theta), 0)\mu + \mathrm{O}(|b, \mu|^2)$.

Let $b(0, a, \mu)$ denote the solution of (1.19) with $b(0, a, \mu) = a$. We can write

$$b(\theta, a, \mu) = b_1(\theta)a + b_0(\theta)\mu + \mathrm{O}\big(|a, \mu|^2\big),$$

where $b_1(0) = 1$, $b_0(0) = 0$. Inserting the above into (1.19) yields

$$b_1' = Ab_1, \quad b_0' = Ab_0 + Z^{\mathrm{T}}F_\mu\big(u(s), 0\big).$$

It follows that $b_1(\theta) = \exp\int_0^\theta A(s)\,\mathrm{d}s$ as before, and

$$b_0(\theta) = b_1(\theta)\int_0^{\mathrm{T}} b_1^{-1}(s)Z^{\mathrm{T}}(s)F_\mu\big(u(s), 0\big)\,\mathrm{d}s.$$

Therefore,

$$\begin{aligned} b(T, a, \mu) = \exp\int_0^{\mathrm{T}} A(s)\,\mathrm{d}s\left[a + \int_0^{\mathrm{T}} b_1^{-1}(s)Z^{\mathrm{T}}(s)F_\mu\big(u(s), 0\big)\,\mathrm{d}s\mu\right] \\ + \mathrm{O}\big(|a, \mu|^2\big). \end{aligned} \tag{1.20}$$

On the other hand, using (1.2) we can define a Poincaré map $P(a, \mu)$ similarly. By Lemma 1.3 it follows

$$P(a, \mu) = b(T, a, \mu). \tag{1.21}$$

Obviously, for $|\mu|$ small (1.17) has a limit cycle near L_1 if and only if P has a fixed point in a near $a = 0$.

The simplest case is that L is hyperbolic. In this case, L will persist under perturbations. In other words, we have:

THEOREM 1.2. *Let L be hyperbolic. Then there exist $\varepsilon > 0$ and a neighborhood U of L such that* (1.17) *has a unique limit cycle in U for $|\mu| < \varepsilon$. Moreover, the limit cycle is also hyperbolic and has the same stability as L.*

PROOF. Since L is hyperbolic we have $P_a'(0, 0) \neq 1$, or equivalently $I(L) = \oint_L \operatorname{div} f\,\mathrm{d}t \neq 0$. By (1.20) and (1.21),

$$P(a, \mu) - a = \big(\mathrm{e}^{I(L)} - a\big)a + N_0\mu + \mathrm{O}\big(|a, \mu|^2\big), \tag{1.22}$$

where $N_0 = \mathrm{e}^{I(L)}\int_0^{\mathrm{T}} \exp(-\int_0^t A(s)\,\mathrm{d}s)Z^{\mathrm{T}}(s)F_\mu(u(s), 0)\,\mathrm{d}s$.

For (a,μ) near zero applying the implicit function theorem to the equation $P(a,\mu)-a=0$ we find a unique fixed point $a=a^*(\mu)=\mathrm{O}(\mu)$ of P. This means that for $|\mu|$ small (1.17) has a unique limit cycle, denoted by L_μ, near L. The cycle has a representation $\varphi(t,p+a^*(\mu)Z_0,\mu)$ where $\varphi(t,x,\mu)$ denotes the flow of (1.17) with $\varphi(0,x,\mu)=x$. Moreover, by (1.22) we have $P'_a(a^*(\mu),\mu)=\mathrm{e}^{I(L)}+\mathrm{O}(|a^*(\mu),\mu|)$, which gives

$$\big(\mathrm{e}^{I(L)}-1\big)\big[P'_a\big(a^*(\mu),\mu\big)-1\big]>0$$

for all $|\mu|$ small. Then the conclusion follows from Theorem 1.1. This ends the proof. □

Further, for the nonhyperbolic case we have

THEOREM 1.3. *Let L be a nonhyperbolic limit cycle of the C^r system* (1.1) *with $r\geqslant 2$.*

(i) *If L has multiplicity* 2, *then there exist $\varepsilon>0$, a neighborhood U of L and a C^r function $\Delta(\mu)=\mathrm{O}(\mu)$ such that* (1.17) *has no limit cycles* (*respectively a limit cycle of multiplicity* 2, *two hyperbolic limit cycles*) *in U as $\Delta(\mu)<0$* (*respectively* $=0$, >0) *for $|\mu|<\varepsilon$.*

(ii) *If L has multiplicity k with $3\leqslant k\leqslant r$, then there exist $\varepsilon>0$ and a neighborhood U of L such that* (1.17) *has at most k limit cycles in U for $|\mu|<\varepsilon$. Moreover,* (1.17) *has at least a limit cycle in U for $|\mu|<\varepsilon$ if k is odd.*

PROOF. Since $r\geqslant 2$, $I(L)=0$ and L has multiplicity 2 we can rewrite (1.22) as

$$P(a,\mu)-a=Q_0(\mu)+Q_1(\mu)a+Q_2(\mu)a^2\big(1+\mathrm{o}(1)\big),$$

where $Q_0(0)=Q_1(0)=0$, $Q_2(0)\neq 0$.

Let $G(a,\mu)=P(a,\mu)-a$, which is called a bifurcation function of (1.17). By the implicit function theorem a unique C^r function $q(\mu)=\mathrm{O}(\mu)$ exists such that $G_a(q(\mu),\mu)=0$ for $|\mu|$ small. Then Taylor's formula yields

$$G(a,\mu)=G\big(q(\mu),\mu\big)+\frac{1}{2}G_{aa}\big(q(\mu),\mu\big)\big(a-q(\mu)\big)^2\big(1+\mathrm{o}(1)\big).$$

Now it is clear that the conclusion (i) follows by taking $\Delta(\mu)=-G(q(\mu),\mu)Q_2(0)$.

To prove the second conclusion let us assume (1.17) has $k+1$ limit cycles in an arbitrary neighborhood of L for some sufficiently small $\mu\neq 0$. Then the function G has $k+1$ zero in a for this small μ. From Rolle's theorem, $\frac{\partial G}{\partial a}$ has k zeros in a. Using the same theorem repeatedly we see that $\frac{\partial^k G}{\partial a^k}$ has a zero $a_0(\mu)$ which can be arbitrarily small as μ goes to zero.

On the other hand, we have

$$\frac{\partial^k G}{\partial a^k}=\frac{\partial^k P}{\partial a^k},\qquad P(a,0)-a=q_k a^k+\mathrm{o}\big(a^k\big),\quad q_k\neq 0,$$

which implies $\frac{\partial^k G}{\partial a^k}=k!q_k+\mathrm{o}(1)\neq 0$ for (a,μ) near zero. This is a contradiction. Hence, (1.17) has at most k limit cycles near L for $|\mu|$ small.

Finally, if k is odd, we have then $q_k a[P(a,0) - a] > 0$ for $|a| = \varepsilon$, where $\varepsilon > 0$ is a small constant. Thus, there exists a constant $\delta > 0$ such that $q_k a[P(a,\mu) - a] > 0$ for $|a| = \varepsilon$, $|\mu| \leqslant \delta$. Therefore, a function $a^*(\mu)$ exists with $|a^*(\mu)| < \varepsilon$ such that $P(a^*(\mu),\mu) - a^*(\mu) = 0$, $|\mu| < \delta$. The proof is completed. $\square$

EXAMPLE 1.4. Consider

$$\begin{aligned}\dot{x} &= -y + x\big(x^2+y^2-1\big)^2 - x\big[\mu_1\big(x^2+y^2\big)+\mu_2\big],\\ \dot{y} &= x + y\big(x^2+y^2-1\big)^2 - y\big[\mu_1\big(x^2+y^2\big)+\mu_2\big].\end{aligned} \tag{1.23}$$

For $\mu_1 = \mu_2 = 0$, (1.23) has a unique limit cycle

$$L: \quad x = \cos t, \quad y = \sin t, \quad 0 \leqslant t \leqslant 2\pi.$$

Similar to Example 1.2, the limit cycle L has multiplicity 2 with $P_a''(0,0) = -16\pi$. We make a transformation of the form $(x,y) = (1-b)(\cos\theta, \sin\theta)$, $0 \leqslant \theta \leqslant 2\pi$, so that (1.23) becomes

$$\dot{\theta} = 1, \qquad \dot{b} = -(1-b)\big[h^2 - \mu_1 h - (\mu_1+\mu_2)\big],$$

where $h = (1-b)^2 - 1$. Hence we have

$$\frac{\mathrm{d}b}{\mathrm{d}\theta} = -(1-b)\big[h^2 - \mu_1 h - (\mu_1+\mu_2)\big].$$

Obviously, the solution $b(\theta, a, \mu_1, \mu_2)$ of the above equation with $b(0, a, \mu_1, \mu_2) = a$ is 2π-periodic near $b = 0$ if and only if the initial data a satisfies

$$G^*(a, \mu_1, \mu_2) \equiv a^2 - \mu_1 a - (\mu_1+\mu_2) = 0.$$

Let $\Delta(\mu_1,\mu_2) = \mu_1^2 + 4(\mu_1+\mu_2)$. Then for (μ_1,μ_2) near $(0,0)$, (1.23) has no limit cycles (respectively a unique multiple 2 limit cycle, two hyperbolic limit cycles) if $\Delta(\mu_1,\mu_2) < 0$ (respectively $= 0$, > 0). The equation $\Delta(\mu_1,\mu_2) = 0$ defines a curve $\mu_2 = -\mu_1 - \frac{1}{4}\mu_1^2$ on the (μ_1,μ_2)-plane, which is called a bifurcation curve of saddle-node type. See Fig. 4.

Turn back to system (1.17). Note that

$$A(\theta) = \operatorname{tr} f_x(u) - \frac{\mathrm{d}}{\mathrm{d}\theta}\ln\big|f(u)\big|, \qquad Z(\theta) = \frac{1}{|f(u)|}\big(-f_2(u), f_1(u)\big).$$

The constant N_0 in (1.22) can be written as $N_0 = \frac{1}{|f(u(0))|}\mathrm{e}^{I(L)}M$, where

$$M = \int_0^{\mathrm{T}} \mathrm{e}^{-\int_0^t \operatorname{tr} f_x(u(s))\,\mathrm{d}s} f\big(u(t)\big) \wedge F_\mu\big(u(t),0\big)\,\mathrm{d}t, \tag{1.24}$$

with $(a_1,a_2) \wedge (b_1,b_2) = a_1b_2 - a_2b_1$, $\mu \in \mathbb{R}$.

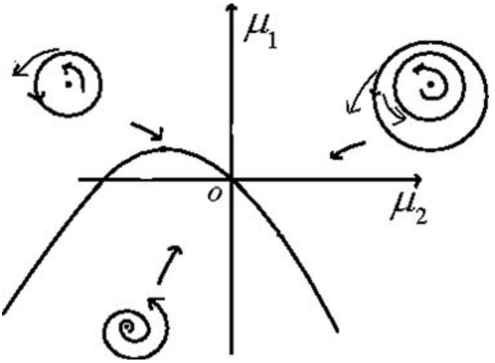


Fig. 4. Bifurcation diagram (saddle-node type) for (1.23).

Thus, if $P(a,0)-a=q_k a^k+\mathrm{o}(a^k)$, we can rewrite (1.22) as

$$G(a,\mu)=P(a,\mu)-a=q_k a^k+\frac{\mathrm{e}^{I(L)}}{|f(u(0))|}M\mu+\mathrm{o}\big(\big|\mu,a^k\big|\big). \tag{1.25}$$

Using (1.25) we can prove

THEOREM 1.4. *Let* (1.17) *be analytic with* $\mu\in\mathbb{R}$. *If* $M\neq 0$, *then exist* $\varepsilon>0$ *and a neighborhood* U *of* L *such that for* $|\mu|<\varepsilon$ *and in the region* U, (1.17) *has*

(i) *a unique limit cycle with multiplicity one if* L *is of odd multiplicity*;
(ii) *two limit cycles with each having multiplicity one for* μ *lying one side of* $\mu=0$ *and no limit cycles for* μ *lying the opposite side if* L *is of even multiplicity*;
(iii) *no limit cycles if* L *is nonisolated.*

PROOF. By Theorem 1.2 we may suppose $I(L)=0$. By (1.25), a unique function of the form

$$\mu=-\frac{q_k}{M}\big|f\big(u(0)\big)\big|a^k+\mathrm{O}\big(a^{k+1}\big)=\mu^*(a) \tag{1.26}$$

exists such that $G(a,\mu^*(a))=0$. If k is odd with $q_k\neq 0$ the function μ^* has a unique inverse

$$a=\left(\frac{-M\mu}{q_k|f(u(0))|}\right)^{1/k}\big(1+\mathrm{o}\big(\mu^{1/k}\big)\big)=a^*(\mu),$$

which satisfies $\frac{\partial G}{\partial a}(a^*(\mu),\mu)\neq 0$.

Then the conclusion (i) follows. The conclusion (ii) follows just similarly.

In the case that L is a nonisolated periodic orbit, we have $G(a,0)\equiv 0$ for $|a|$ small. This implies that $\mu^*(a)\equiv 0$ since $G(a,\mu)=0$ if and only if $\mu=\mu^*(a)$. Hence, for all $|\mu|>0$ we have $G(a,\mu)\neq 0$. This means that (1.17) has no limit cycles near L for $|\mu|>0$ small. This ends the proof. □

EXAMPLE 1.5. Consider (1.23) again with (μ_1, μ_2) varying on a straight line. That is, let $\mu_1 = \mu$, $\mu_2 = c\mu$, where c is a constant. Then (1.23) becomes

$$\begin{aligned} \dot{x} &= -y + x\bigl(x^2 + y^2 - 1\bigr)^2 - \mu x\bigl(x^2 + y^2 + c\bigr), \\ \dot{y} &= x + y\bigl(x^2 + y^2 - 1\bigr)^2 - \mu y\bigl(x^2 + y^2 + c\bigr). \end{aligned} \tag{1.27}$$

By (1.24) we have $M = (1 + c)2\pi$. Note that $q_2 = \frac{1}{2}P_a''(0, 0) = -8\pi$. The formula (1.26) becomes $\mu^*(a) = \frac{4}{1+c}a^2 + O(a^3)$, which has inverse functions

$$a = a_j(\mu) = \left[\frac{1+c}{4}\mu\right]^{1/2}\bigl[(-1)^j + O\bigl(|\mu|^{1/2}\bigr)\bigr], \quad j = 1, 2.$$

Thus, when $1 + c \neq 0$, (1.27) has no limit cycles (2 limit cycles) for $(1 + c)\mu < 0$ (> 0). When $(1 + c) = 0$, then (1.27) has always two limit cycles given by $x^2 + y^2 = 1$ and $x^2 + y^2 = 1 - \mu$ for all $|\mu| > 0$ small. This shows that the condition $M \neq 0$ in Theorem 1.4 is somehow necessary.

One can give the bifurcation diagram of (1.27) and compare it with Fig. 4.

REMARK 1.2. A typical result on qualitative theory of differential equations is the so-called Poincaré–Bendixson theorem which says that the positive limit set of a bounded positive semi-orbit is a connected curve consisting of singular points and orbits connecting them. One can find a proof of the theorem in Hale [35], Chicone [14] and Zhang et al. [127]. An important corollary of the theorem is that for an analytic system a positively invariant set with no singular points contains at least one limit cycle. This corollary has many applications to various planar systems to study the existence of a limit cycle. On the other hand, there are also many results on the nonexistence of a limit cycle or the existence of two or more limit cycles. The reader can consult Zhang et al. [127], Ye [116], Ye et al. [117] and Luo et al. [101] for a general theory of limit cycles. For a given planar system with or without parameters it is usually very significant and also difficult to find the number of limit cycles of it. Thousands of papers have been published on this aspect. For recent works, one can see [1,6,8,11,15,18,19,22–26,28–34,97–99].

Results in Theorems 1.1–1.4 are fundamental which can be found in Andronov [2], Chow and Hale [16] and Han [51]. The conclusions in Lemma 1.5 seem very natural, but not obvious. The proof presented here is new.

2. Focus values and Hopf bifurcation

2.1. *Poincaré map and focus value*

In this section we consider local behavior of a planar C^∞ systems near an elementary focus.

After removing the focus to the origin, the system can be written as

$$\begin{aligned}\dot{x} &= ax + by + F(x, y) = f(x, y),\\ \dot{y} &= cx + dy + G(x, y) = g(x, y),\end{aligned} \tag{2.1}$$

where $c \neq 0$, F and $G : \mathbb{R}^2 \to \mathbb{R}$ are C^∞ functions with

$$F(0,0) = G(0,0) = 0, \qquad \frac{\partial(F,G)}{\partial(x,y)}(0,0) = 0,$$

and the eigenpolynomial $h(\lambda) = \lambda^2 - (a+d)\lambda + ad - bc$ has a pair of conjugate complex zero $\alpha \pm \mathrm{i}\beta \neq 0$.

Let

$$C = \begin{pmatrix} 1 & \dfrac{\alpha - a}{\beta} \\ 0 & -\dfrac{c}{\beta} \end{pmatrix}, \qquad C^{-1} = \begin{pmatrix} 1 & \dfrac{\alpha - a}{c} \\ 0 & -\dfrac{\beta}{c} \end{pmatrix}.$$

By using $h(\alpha \pm \mathrm{i}\beta) = 0$, or

$$\alpha^2 - \beta^2 - (a+d)\alpha + ad - bc = 0, \qquad \big(2\alpha - (a+d)\big)\beta = 0,$$

it is easy to see that

$$\begin{aligned} &C^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix} C = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix},\\ &\alpha = \frac{1}{2}(a+d), \quad |\beta| = \frac{1}{2}\sqrt{-4bc - (a-d)^2}. \end{aligned} \tag{2.2}$$

Then the linear transformation $(x, y)^{\mathrm{T}} = C(u, v)^{\mathrm{T}}$ carries (2.1) into the form

$$\dot{u} = \alpha u + \beta v + \widetilde{F}(u, v), \qquad \dot{v} = -\beta u + \alpha v + \widetilde{G}(u, v),$$

where $\widetilde{F}, \widetilde{G} = \mathrm{O}(|u, v|^2)$. We call the above system a first order normal form of (2.1).

Let us define a Poincaré map of (2.1) near the origin. For any given $\theta_0 \in [0, 2\pi)$, let

$$Z_0 = (\cos\theta_0, \sin\theta_0)^{\mathrm{T}}, \qquad Z_0^{\perp} = (\sin\theta_0, -\cos\theta_0)^{\mathrm{T}}.$$

The unit vector Z_0 determines a cross section l below

$$l: \quad (x, y)^{\mathrm{T}} = rZ_0, \quad 0 < r < r_0$$

for $r_0 > 0$ small. For $(x_0, y_0)^{\mathrm{T}} = rZ_0 \in l$, let $(x(t, x_0, y_0), y(t, x_0, y_0))$ be the solution of (2.1) with $x(0, x_0, y_0) = x_0$, $y(0, x_0, y_0) = y_0$. Then the formula of constant variation follows:

$$\begin{pmatrix} x(t, x_0, y_0) \\ y(t, x_0, y_0) \end{pmatrix} = \left[\exp\begin{pmatrix} at & bt \\ ct & dt \end{pmatrix}\right]\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \mathrm{O}\big(|x_0, y_0|^2\big). \tag{2.3}$$

Thus,

$$\begin{pmatrix} x(t, x_0, y_0) \\ y(t, x_0, y_0) \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \left[\exp\begin{pmatrix} at & bt \\ ct & dt \end{pmatrix} - I\right] Z_0 r + \mathrm{O}\big(r^2\big) \equiv rV(t, r).$$

By (2.2) we have

$$V(t, r) = C\left[\exp\begin{pmatrix} \alpha t & \beta t \\ -\beta t & \alpha t \end{pmatrix} - I\right] C^{-1} Z_0 + \mathrm{O}(r).$$

Let $H(t, r) = [V(t, r)]^{\mathrm{T}} Z_0^{\perp}$. Define

$$V(t, 0) = \lim_{r \to 0} V(t, r) = C\left[\exp\begin{pmatrix} \alpha t & \beta t \\ -\beta t & \alpha t \end{pmatrix} - I\right] C^{-1} Z_0.$$

Also, if we allow the variable r to be negative in the definition of l then H is well defined for $|r| < r_0$. It is easy to see that

$$H(T, 0) = \big(\mathrm{e}^{\alpha T} - 1\big) Z_0 Z_0^{\perp} = 0, \qquad H_t(T, 0) = \mathrm{e}^{\alpha T} \beta \neq 0,$$

where $T = \frac{2\pi}{|\beta|}$. Hence, the implicit function theorem yields that a unique C^∞ function $\tau(r) = T + \mathrm{O}(r)$ exists such that $H(\tau(r), r) = 0$ for $|r| < r_0$. That is to say, for $0 < r < r_0$, $\tau(r)$ is the time of the first return to l by a circle. Therefore, we can write

$$\begin{pmatrix} x(\tau, x_0, y_0) \\ y(\tau, x_0, y_0) \end{pmatrix} = P(r) Z_0 \in l, \quad 0 < r < r_0, \tag{2.4}$$

where $P \in C^\infty$ for $|r| < r_0$ with $rP(r) > 0$ for $r \neq 0$.

DEFINITION 2.1. Let (2.1) satisfy (2.2). The function P defined by (2.4) is called a *Poincaré map* of (2.1) near the origin. The map is often written as $P : l \to l$.

DEFINITION 2.2. If there exists $r_0' > 0$ such that $P(r) - r < 0 (> 0)$ for $0 < r < r_0'$, we say the origin is a *stable* (*unstable*) *focus*. If $P(r) - r = 0$ for $0 < r < r_0'$, we say the origin is a *center*.

REMARK 2.1. Similar to Corollary 1.1 we can prove that the stability of the origin is independent of the choice of θ_0 which is the angle defining the section l.

In fact, for any given $\theta_0, \theta_1 \in [0, 2\pi)$ with $\theta_0 \neq \theta_1$, let the corresponding cross sections and Poincaré maps be l_0, l_1 and $P_0 : l_0 \to l_0$ and $P_1 : l_1 \to l_1$, respectively. Similarly, there exist unique C^∞ functions $\tau^*(r) = t_1 + \mathrm{O}(r)$ and $h_1(r) = \mathrm{O}(r)$ such that

$$\begin{pmatrix} x(\tau^*, rZ_0) \\ y(\tau^*, rZ_0) \end{pmatrix} = h_1(r) Z_1, \quad t_1 \in (0, T).$$

Differentiating the equation above in r and using formula of constant variation (2.3) we obtain

$$\left[\exp\begin{pmatrix} at_1 & bt_1 \\ ct_1 & dt_1 \end{pmatrix}\right] Z_0 = h_1'(0) Z_1,$$

which yields $h_1'(0) \neq 0$. It is clear that $h_1(r) > 0$ for $r > 0$. Thus, $h_1'(0) > 0$. Then in the same way as in Lemma 1.4 we have

$$h_1 \circ P_0 = P_1 \circ h_1,$$

which gives the conclusion in Remark 2.1 easily.

LEMMA 2.1. *For any given* $\theta_0 \in [0, 2\pi)$ *the origin is a stable* (*unstable*) *focus of* (2.1) *if and only if*

$$r\big[P(r) - r\big] < 0 \ (> 0) \quad \text{for } 0 < |r| < r_0.$$

Hence, if for some $k \geqslant 1$ *it holds that*

$$P(r) - r = 2\pi v_k r^k + \mathrm{O}\big(r^{k+1}\big), \quad v_k \neq 0, \tag{2.5}$$

then k is odd, and the origin is stable (*unstable*) *if* $v_k < 0$ (> 0).

PROOF. Let P_0 denote the function P defined by (2.4). For any $\theta \in [0, 2\pi)$ there exists a unique $\theta_1 \in [0, 2\pi)$ such that

$$\theta_1 = \theta_0 + \pi \quad \text{if } 0 \leqslant \theta_0 < \pi; \qquad \theta_0 = \theta_1 + \pi \quad \text{if } \pi \leqslant \theta_0 < 2\pi.$$

Let P_1 denote the Poincaré map associated to θ_1. We claim that

$$P_0(r) = -P_1(-r), \quad |r| < r_0. \tag{2.6}$$

For definiteness, assume $0 \leqslant \theta_0 < \pi$ and $0 < r < r_0$. Let $Z_1 = (\cos\theta_1, \sin\theta_1)$. Then $Z_1 = -Z_0$ and $(x_0, y_0)^{\mathrm{T}} = rZ_0 = (-r)Z_1$. Hence, by (2.4) we have

$$\begin{pmatrix} x(\tau, x_0, y_0) \\ y(\tau, x_0, y_0) \end{pmatrix} = P_0(r) Z_0 = P_1(-r) Z_1,$$

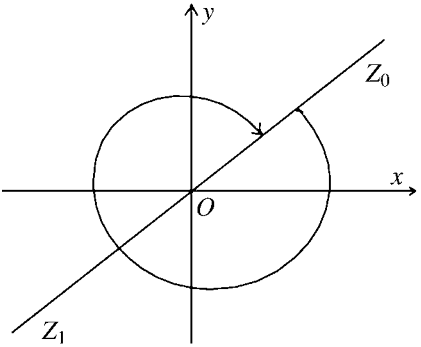


Fig. 5. The Poincaré maps P_0 and $P_1 (b > 0)$.

which gives (2.6). See Fig. 5.

By Definition 2.2, the origin is stable if and only if

$$P_0(r) - r < 0 \quad \text{for } 0 < r < r_0. \tag{2.7}$$

By Remark 2.1, the above inequality is equivalent to

$$P_1(r) - r < 0 \quad \text{for } 0 < r < r_0.$$

Note that by (2.6)

$$P_1(r) - r = -\big[P_0(-r) - (-r)\big].$$

Hence, (2.7) is equivalent to

$$P_0(r) - r > 0 \quad \text{for } 0 < -r < r_0.$$

Then combining (2.7) and the above together we see that (2.7) holds if and only if

$$r\big[P_0(r) - r\big] < 0 \quad \text{for } 0 < |r| < r_0.$$

Thus, if (2.5) holds, the number k must be odd and the sign of v_k determines the stability of the origin. This ends the proof. □

DEFINITION 2.3. Let (2.5) hold with $k = 2m + 1$, $m \geqslant 0$. We call the origin to be *a focus of order m*, and v_k the *mth Lyapunov constant* or *focus value*. The focus at the origin is called *rough or hyperbolic* (*weak or fine*) as $m = 0$ ($m \geqslant 1$).

REMARK 2.2. Similar to Remark 2.1 and Corollary 1.1, one can prove that the order and the first nonzero Lyapunov constant of the origin are independent of the choice of θ_0 and l.

For the sake of convenience, we will take $\theta_0 = 0$ below. In this case we have $(x_0, y_0) = (r, 0)$, and hence from (2.4) and (2.2)–(2.3) we obtain

$$\begin{aligned} P(r) &= \big(x(\tau, x_0, y_0), y(\tau, x_0, y_0)\big) Z_0 \\ &= \left[C \mathrm{e}^{\alpha\tau} \begin{pmatrix} \cos\beta\tau & \sin\beta\tau \\ -\sin\beta\tau & \cos\beta\tau \end{pmatrix} C^{-1} \begin{pmatrix} r \\ 0 \end{pmatrix} + \mathrm{O}\big(r^2\big) \right]^{\mathrm{T}} Z_0, \end{aligned}$$

which yields $P(r) = r\mathrm{e}^{\alpha T} + \mathrm{O}(r^2)$ since $\tau(0) = T$. Comparing with (2.5) gives that

$$v_1 = \frac{1}{2\pi}\big[\mathrm{e}^{\alpha T} - 1\big] = \frac{1}{2\pi}\big[\mathrm{e}^{\frac{a+d}{2}T} - 1\big].$$

Thus, the origin is stable (unstable) if

$$\operatorname{div}(f, g)|_0 = f_x(0, 0) + g_y(0, 0) < 0 \ (> 0).$$

Next, we suppose $\alpha = 0$ and give a computation formula for v_3. For the purpose it is convenient to use the first order normal form of (2.1). Without loss of generality we may suppose (2.1) has been of the form. In other words, we may assume (2.2) holds with C being the identity. Introducing the polar coordinate to (2.1) we obtain

$$\dot{r} = \cos\theta F + \sin\theta G, \qquad \dot{\theta} = -b + (\cos\theta G - \sin\theta F)/r \tag{2.8}$$

and

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{\cos\theta F + \sin\theta F}{-b + (\cos\theta G - \sin\theta F)/r} \equiv R(\theta, r). \tag{2.9}$$

Let $r(\theta, r_0)$ be the solution of (2.9) with $r(0, r_0) = r_0$. We have

LEMMA 2.2. *Let $a = d = 0, b = -c \neq 0$. Then*

(i) $P(r_0) = r(2\pi, r_0)$ *if* $b < 0$, *and* $P(r(2\pi, r_0)) = r_0$ *if* $b > 0$,

(ii) $v_3 = \frac{1}{2\pi}\operatorname{sgn}(-b)\int_0^{2\pi} R_3(\theta)\,\mathrm{d}\theta$, *where* $R_3(\theta) = \frac{1}{3!}\frac{\partial^3 R}{\partial r^3}(\theta, 0)$.

PROOF. The conclusion (i) can be proved in a similar manner to Lemma 1.3. Let

$$R(\theta, r) = R_2(\theta)r^2 + R_3(\theta)r^3 + \cdots \tag{2.10}$$

and

$$r(\theta, r_0) = r_1(\theta)r_0 + r_2(\theta)r_0^2 + r_3(\theta)r_0^3 + \cdots,$$

where $r_1(0) = 1, r_2(0) = r_3(0) = 0$. Inserting the above solution and (2.10) into (2.9), and then comparing the like powers of r_0 we obtain

$$r_1'(\theta) = 0, \qquad r_2'(\theta) = R_2 r_1^2, \qquad r_3'(\theta) = R_3 r_1^2 + 2R_2 r_1 r_2,$$

or

$$r_1(\theta) = 1, \qquad r_2(\theta) = \int_0^\theta R_2(\theta)\,\mathrm{d}\theta,$$
$$r_3(\theta) = \int_0^\theta \big[R_3(\theta) + 2R_2(\theta)r_2(\theta)\big]\,\mathrm{d}\theta.$$

Hence, if we let

$$\widetilde{P}(r_0) = r(2\pi, r_0) = r_0 + 2\pi\,\tilde{v}_2 r_0^2 + 2\pi\,\tilde{v}_3 r_0^3 + \mathrm{O}\big(r_0^4\big),$$

then

$$\begin{aligned}\tilde{v}_2 &= \frac{1}{2\pi}\int_0^{2\pi} R_2(\theta)\,\mathrm{d}\theta,\\ \tilde{v}_3 &= \frac{1}{2\pi}\int_0^{2\pi}\big[R_3(\theta) + 2R_2(\theta)r_2(\theta)\big]\,\mathrm{d}\theta\\ &= \frac{1}{2\pi}\bigg[\int_0^{2\pi} R_3(\theta)\,\mathrm{d}\theta + \int_0^{2\pi} 2r_2(\theta)r_2'(\theta)\,\mathrm{d}\theta\bigg]\\ &= \frac{1}{2\pi}\int_0^{2\pi} R_3(\theta)\,\mathrm{d}\theta + 2\pi(\tilde{v}_2)^2.\end{aligned}$$

If $b < 0$, we have $P(r_0) = \widetilde{P}(r_0)$, and by Lemma 2.1, $\tilde{v}_2 = 0$. Thus

$$v_3 = \tilde{v}_3 = \frac{1}{2\pi}\int_0^{2\pi} R_3(\theta)\,\mathrm{d}\theta.$$

If $b > 0$, we have $P(r_0) = \widetilde{P}^{-1}(r_0)$. Note that if $\widetilde{P}(r_0) = r_0 + 2\pi\,\tilde{v}_k r_0^k + \mathrm{O}(r_0^{k+1})$ then

$$\widetilde{P}^{-1}(r_0) = r_0 - 2\pi\,\tilde{v}_k r_0^k + \mathrm{O}\big(r_0^{k+1}\big).$$

Hence, using Lemma 2.1 we have $\tilde{v}_2 = 0$ and

$$v_3 = -\tilde{v}_3 = -\frac{1}{2\pi}\int_0^{2\pi} R_3(\theta)\,\mathrm{d}\theta.$$

This ends the proof. □

Based on the above lemma we can prove:

THEOREM 2.1. *Let* $a = d = 0, b = -c \neq 0$, *and*

$$F(x,y) = \sum_{i+j=2}^{3} a_{ij}x^i y^j + \mathrm{o}\big(|x,y|^3\big),$$

$$G(x,y) = \sum_{i+j=2}^{3} b_{ij}x^i y^j + \mathrm{o}\big(|x,y|^3\big).$$

Then

$$\begin{aligned} v_3 &= \frac{1}{8|b|}\bigg\{3(a_{30}+b_{03}) + a_{12} + b_{21} - \frac{1}{b}\big[a_{11}(a_{20}+a_{02}) - b_{11}(b_{20}+b_{02}) \\ &\quad + 2(a_{02}b_{02} - a_{20}b_{20})\big]\bigg\} \\ &= \frac{1}{16|b|}\bigg\{F_{xxx} + F_{xyy} + G_{yyy} - \frac{1}{b}\big[F_{xy}(F_{xx}+F_{yy}) - G_{xy}(G_{xx}+G_{yy}) \\ &\quad - F_{xx}G_{xx} + F_{yy}G_{yy}\big]\bigg\}\bigg|_{(0,0)}. \end{aligned}$$

PROOF. By (2.9) and (2.10),

$$R_2(\theta) = -b^{-1}P_2(\theta), \qquad R_3(\theta) = -b^{-1}\big[P_3(\theta) + b^{-1}P_2(\theta)S_2(\theta)\big],$$

where

$$\begin{aligned} &P_k(\theta) = \cos\theta F_k(\cos\theta,\sin\theta) + \sin\theta G_k(\cos\theta,\sin\theta), \\ &S_k(\theta) = \cos\theta G_k(\cos\theta,\sin\theta) - \sin\theta F_k(\cos\theta,\sin\theta), \\ &F_k(x,y) = \sum_{i+j=k} a_{ij}x^i y^j, \qquad G_k(x,y) = \sum_{i+j=k} b_{ij}x^i y^j. \end{aligned}$$

It follows directly that

$$\begin{aligned} &P_3(\theta) = (a_{12}+b_{21})\sin^2\theta\cos^2\theta + a_{30}\cos^4\theta + b_{03}\sin^4\theta + K_0(\theta), \\ &P_2(\theta)S_2(\theta) = -a_{02}b_{02}\sin^6\theta + a_{20}b_{20}\cos^6\theta - N_1\sin^4\theta\cos^2\theta \\ &\quad - N_2\sin^2\theta\cos^4\theta + K_1(\theta), \end{aligned}$$

where K_0 and K_1 are 2π-periodic functions with zero mean value over the interval $[0, 2\pi]$, and

$$\begin{aligned} &N_1 = 2a_{02}a_{11} - 2b_{02}b_{11} + a_{02}b_{20} + a_{11}b_{11} + a_{20}b_{02} - a_{02}b_{02}, \\ &N_2 = 2a_{20}a_{11} - 2b_{20}b_{11} - a_{02}b_{20} - a_{11}b_{11} + a_{20}b_{20} - a_{20}b_{02}. \end{aligned}$$

Then we have

$$\frac{1}{2\pi}\int_0^{2\pi} R_3(\theta)\,\mathrm{d}\theta = -\frac{\pi}{2\pi b}\biggl[(a_{30}+b_{03})\frac{3}{4}+(a_{12}+b_{21})\frac{1}{4} \\ -\frac{1}{b}\biggl[(a_{02}b_{02}-a_{20}b_{20})\frac{5}{8}+(N_1+N_2)\frac{1}{8}\biggr]\biggr].$$

Thus, the conclusion follows from Lemma 2.2(ii). □

Along the above line one can find formulas for computing v_5, v_7, etc. See [31] for more detail.

2.2. *Normal form and Lyapunov technique*

Next, we introduce another method (called normal form method) to find focus values.

LEMMA 2.3. *Let* $a = d = 0$, $b = -c \neq 0$. *Then for any integer* $m \geqslant 1$ *there exists a polynomial change of variables of the form*

$$(x, y)^{\mathrm{T}} = (u, v)^{\mathrm{T}} + \mathrm{O}\bigl(|u, v|^2\bigr) = Q(u, v)$$

which transforms the system (2.1) *into*

$$\begin{aligned} \dot{u} &= bv + \sum_{j=1}^{m}(a_j u + b_j v)\bigl(u^2+v^2\bigr)^j + \mathrm{O}\bigl(|u, v|^{2m+2}\bigr), \\ \dot{v} &= -bu + \sum_{j=1}^{m}(-b_j u + a_j v)\bigl(u^2+v^2\bigr)^j + \mathrm{O}\bigl(|u, v|^{2m+2}\bigr). \end{aligned} \tag{2.11}$$

PROOF. For convenience we introduce complex variables $z = x + \mathrm{i}y$, $\bar{z} = x - \mathrm{i}y$, or

$$\begin{pmatrix} z \\ \bar{z} \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \tag{2.12}$$

so that (2.1) becomes

$$\begin{aligned} \dot{z} &= f\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2\mathrm{i}}\right) + \mathrm{i}g\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2\mathrm{i}}\right) \equiv h(z, \bar{z}), \\ \dot{\bar{z}} &= f\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2\mathrm{i}}\right) - \mathrm{i}g\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2\mathrm{i}}\right) \equiv \bar{h}(z, \bar{z}). \end{aligned} \tag{2.13}$$

It is easy to verify that $\bar{h}$ is the conjugate of h. Hence, we may neglect the second equation in (2.13), since it can be obtained from the first one by taking complex conjugation.

By (2.3) we can write

$$h(z,\bar{z}) = -\mathrm{i}bz + \sum_{2\leqslant j+k\leqslant 2m+1} A_{jk} z^j \bar{z}^k + \mathrm{O}\big(|z|^{2m+2}\big). \tag{2.14}$$

We would like to make a change of variables of the form

$$z = \omega + \sum_{2\leqslant j+k\leqslant 2m+1} C_{jk}\omega^j \overline{\omega}^k = \omega + p(\omega,\overline{\omega}), \tag{2.15}$$

and expect that the resulting equation which has the form in general

$$\dot{\omega} = -\mathrm{i}b\omega + \sum_{2\leqslant j+k\leqslant 2m+1} B_{jk}\omega^j \overline{\omega}^k + \mathrm{O}\big(|\omega|^{2m+2}\big) \tag{2.16}$$

is as simple as possible.

For the purpose, differentiating both sides of (2.15) in t and using (2.13), (2.14) and (2.16) we obtain

$$\begin{aligned}
&\left[1 + \sum_{j+k=2}^{2m+1} C_{jk} j\omega^{j-1}\overline{\omega}^k\right]\left[-\mathrm{i}b\omega + \sum_{j+k=2}^{2m+1} B_{jk}\omega^j\overline{\omega}^k\right] \\
&\quad + \sum_{j+k=2}^{2m+1} C_{jk} k\omega^j\overline{\omega}^{k-1}\left[\mathrm{i}b\overline{\omega} + \sum_{j+k=2}^{2m+1} \overline{B}_{jk}\omega^k\overline{\omega}^j\right] \\
&= -\mathrm{i}b\left[\omega + \sum_{2\leqslant j+k\leqslant 2m+1} C_{jk}\omega^j\overline{\omega}^k\right] \\
&\quad + \sum_{j+k=2}^{2m+1} A_{jk}\big(\omega + p(\omega,\overline{\omega})\big)^j \big(\overline{\omega} + \bar{p}(\omega,\overline{\omega})\big)^k + \mathrm{O}\big(|\omega|^{2m+1}\big).
\end{aligned}$$

By considering terms of $\omega^j\overline{\omega}^k$ for $j+k=2$ we obtain

$$-\mathrm{i}bC_{jk} j\omega^j\overline{\omega}^k + B_{jk}\omega^j\overline{\omega}^k + \mathrm{i}bC_{jk} k\omega^j\overline{\omega}^k = -\mathrm{i}bC_{jk}\omega^j\overline{\omega}^k + A_{jk}\omega^j\overline{\omega}^k,$$

which yields

$$B_{jk} = A_{jk} + \mathrm{i}bC_{jk}(j-k-1). \tag{2.17}$$

Hence, to nullify the coefficients B_{jk} we can choose

$$C_{jk} = \frac{A_{jk}}{\mathrm{i}b(1+k-j)}$$

if

$$j \neq k+1, \tag{2.18}$$

which always holds for $j+k=2$. Thus it appears that $B_{jk}=0$ for the chosen C_{jk} for $j+k=2$. In other words, all quadratic terms will be nullified in the new equation (2.16) so far if we choose the coefficient of quadratic terms in (2.15) proper.

By equating the coefficients of $\omega^j\overline{\omega}^k$ for $j+k=3$, we obtain

$$B_{jk} = A_{jk} + \mathrm{i}bC_{jk}(j-k-1) + Z_{jk}, \tag{2.19}$$

where Z_{jk} depends on $C_{j'k'}$ with $j'+k'=2$. Thus, in this case, for j and k satisfying (2.18) we can also choose C_{jk} by

$$C_{jk} = \frac{1}{\mathrm{i}b(1+k-j)}[A_{jk}+Z_{jk}]$$

so that $B_{jk}=0$ for the new equation. The only monomial left up to now is $B_{21}\omega^2\overline{\omega}$ which is called a resonant term, and by (2.19) the coefficient C_{21} can be chosen freely, say $C_{21}=0$.

For higher values of $j+k$, the expression (2.19) remains valid, where Z_{jk} depends only on $C_{j'k'}$ with $2 \leqslant j'+k' < j+k$. Hence, by continuing the way above, an appropriate change of variables can be found that eliminates all those monomials $B_{jk}\omega^j\overline{\omega}^k$ for which (2.18) are satisfied. The only monomials which survive, called resonant terms, have the form $B_{k+1,k}\omega^{k+1}\overline{\omega}^k$. Therefore, eventually, Eq. (2.16) takes the form

$$\dot{\omega} = -\mathrm{i}b\omega + \sum_{k=1}^{m} B_{k+1,k}\omega^{k+1}\overline{\omega}^k + \mathrm{O}\big(|\omega|^{2m+2}\big) = R(\omega,\overline{\omega}).$$

More precisely, there exists a change of variables

$$z = \omega + p(\omega,\overline{\omega}), \qquad \overline{z} = \overline{\omega} + \bar{p}(\omega,\overline{\omega})$$

which carries (2.13) into the form of

$$\dot{\omega} = R(\omega,\overline{\omega}), \qquad \dot{\overline{\omega}} = \overline{R}(\omega,\overline{\omega}).$$

Substituting $\omega = u+\mathrm{i}v$, $\overline{\omega} = u-\mathrm{i}v$ into the above gives the desired system (2.11) with

$$a_j = \operatorname{Re} B_{j+1,j}, \quad b_j = -\operatorname{Im} B_{j+1,j}.$$

The coordinate change from (2.1) to (2.11) is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix}^{-1} \begin{pmatrix} \omega + p(\omega,\overline{\omega}) \\ \overline{\omega} + \bar{p}(\omega,\overline{\omega}) \end{pmatrix}, \qquad \begin{pmatrix} \omega \\ \overline{\omega} \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

where (x, y) and (u, v) are both real variables, and

$$\begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{\mathrm{i}}{2} & \frac{\mathrm{i}}{2} \end{pmatrix}.$$

The proof is completed. □

We call the system (2.11) the normal form of (2.1) of order $2m+1$. It is easy to see that in polar coordinates the normal form equation takes the form

$$\begin{aligned} \dot{r} &= a_1 r^3 + \cdots + a_m r^{2m+1} + \mathrm{O}\big(r^{2m+2}\big), \\ \dot{\theta} &= -\big(b + b_1 r^2 + \cdots + b_m r^{2m}\big) + \mathrm{O}\big(r^{2m+2}\big). \end{aligned} \tag{2.20}$$

For a relationship between focus values appeared in (2.5) and the constants a_j in (2.20) we have

LEMMA 2.4. *Let* $a = d = 0$, $b = -c \neq 0$. *Then* (2.5) *holds with* $k = 2m+1 \geqslant 3$ *if and only if* $a_j = 0$, $j = 1, \ldots, m-1$, $a_m \neq 0$. *Moreover,* $v_{2m+1} = a_m/|b|$ *as* $a_j = 0$, $j = 1, \ldots, m-1$.

PROOF. Let P and P^* denote respectively the Poincaré maps of (2.1) and (2.11) both associated to $\theta_0 = 0$. For $r_0 > 0$ small, let

$$\begin{aligned} &l\colon\ (x, y) = (r, 0), \quad 0 < r < r_0, \\ &l_1 = Q^{-1}(l) = \big\{\big(q_1(r), q_2(r)\big), 0 < r < r_0\big\}, \end{aligned}$$

where Q^{-1} is the inverse of the transformation Q appeared in Lemma 2.3, and $q_1(r) = r + \mathrm{O}(r^2)$, $q_2(r) = \mathrm{O}(r^2)$. Let $P_1 : l_1 \to l_1$ denote the Poincaré map of (2.11) defined by a similar way to (1.14). Then by the proof of Lemma 1.5 with the case 1 we have

$$P_1 = P. \tag{2.21}$$

Let $l' = \{(u, v)\colon\ u = r, v = 0, 0 < r < r_0\}$ which is tangent to l_1 at the origin. Denote by $(u(t, r),\ v(t, r))$ the solution of (2.11) with $(u(0, r),\ v(0, r)) = (r, 0)$. It is easy to see from (2.11) that

$$u(t, r) = r \cos bt + \mathrm{O}\big(r^2\big), \qquad v(t, r) = -r \sin bt + \mathrm{O}\big(r^2\big).$$

Note that l_1 can be represented as

$$l_1 = \big\{\big(r', c\big(r'\big)\big),\ 0 < r' < q_1(r_0)\big\}$$

where c is a C^∞ function satisfying $c(0) = c'(0) = 0$. Consider the function

$$K(t, r) = c\big(u(t, r)\big) - v(t, r) = r K_1(t, r),$$

where $K_1(t,r)=\sin bt+\mathrm{O}(r)$. By the implicit function theorem, there exists a C^∞ function $\tau^*(r)=\mathrm{O}(r)$ such that $K_1(\tau^*(r),r)\equiv 0$. Then similar to (1.10), there exists a unique function $h_1(r)$ such that

$$\big(u\big(\tau^*(r),r\big),v\big(\tau^*(r),r\big)\big)=\big(q_1\big(h_1(r)\big),q_2\big(h_1(r)\big)\big)\in l_1. \tag{2.22}$$

Moreover, the function satisfies

$$h_1\circ P^*=P_1\circ h_1. \tag{2.23}$$

By (2.22) it follows that

$$q_1(h_1)=u(\tau^*,r)=r\cos b\tau^*+\mathrm{O}\big(r^2\big)=r+\mathrm{O}\big(r^2\big),$$

which gives that

$$h_1=q_1^{-1}\big(r+\mathrm{O}\big(r^2\big)\big)=r+\mathrm{O}\big(r^2\big).$$

Therefore, by (2.21) and (2.23) we obtain

$$h_1\circ P^*=P\circ h_1,\quad h_1(r)=r+\mathrm{O}\big(r^2\big). \tag{2.24}$$

Then, it is easy to verify that (2.5) holds if and only if

$$P^*-r=2\pi v_k r^k+\mathrm{O}\big(r^{k+1}\big),\quad v_k\neq 0. \tag{2.25}$$

On the other hand, let $a_j=0,\ j=1,\dots,n-1, a_n\neq 0$. We then have from (2.20)

$$\frac{\mathrm{d}r}{\mathrm{d}\theta}=-\frac{a_n}{b}r^{2n+1}+\mathrm{O}\big(r^{2n+2}\big)$$

which has the solution

$$r(\theta,r_0)=r_0-\frac{a_n}{b}\theta r_0^{2n+1}+\mathrm{O}\big(r_0^{2n+2}\big).$$

Thus,

$$r(2\pi,r_0)=r_0-\frac{2\pi a_n}{b}r_0^{2n+1}+\mathrm{O}\big(r_0^{2n+2}\big).$$

Hence, by Lemma 2.2 we have

$$P^*(r)=r+\frac{2\pi a_n}{|b|}r^{2n+1}+\mathrm{O}\big(r^{2n+2}\big).$$

Obviously, for $n\geqslant 1$ the above holds if and only if (2.25) holds with $k=2m+1=2n+1$. This ends the proof. □

We remark that some algorithms have been established for calculating the coefficients $a_1, \dots, a_m$ in (2.20). An efficient one can be found in Yu [119].

We finally give a third method to determine the stability and the order of a focus originated by Lyapunov.

LEMMA 2.5. *For any given integer* $N > 1$ *and expansions of* f *and* g *of the form*

$$f = by + f_1 + \mathrm{O}\big(|x, y|^{2N+2}\big), \qquad g = -bx + g_1 + \mathrm{O}\big(|x, y|^{2N+2}\big),$$

where

$$f_1(x, y) = \sum_{2 \leqslant i+j \leqslant 2N+1} a_{ij} x^i y^j, \quad g_1(x, y) = \sum_{2 \leqslant i+j \leqslant 2N+1} b_{ij} x^i y^j,$$

there exist constants $L_2, \dots, L_{N+1}$ *and a polynomial*

$$V(x, y) = \sum_{k=2}^{2N+2} V_k(x, y),$$

where

$$V_2(x, y) = x^2 + y^2, \qquad V_k(x, y) = \sum_{i+j=k} c_{ij} x^i y^j, \quad 3 \leqslant k \leqslant 2N+2$$

such that

$$V_x f + V_y g = \sum_{k=2}^{N+1} L_k \big(x^2 + y^2\big)^k + \mathrm{O}\big(|x, y|^{2N+3}\big). \tag{2.26}$$

Moreover, for $2 \leqslant k \leqslant N+1$, L_{k+1} *depends only on* a_{ij} *and* b_{ij} *with* $i + j \leqslant 2k + 1$.

PROOF. We want to find a polynomial V with the supposed form and constants $L_2, \dots, L_{N+1}$ satisfying (2.26). Note that for the given form of f, g and V we have

$$\begin{aligned} V_x f + V_y g &= b(yV_x - xV_y) + (V_x f_1 + V_y g_1) + \mathrm{O}\big(|x, y|^{2N+3}\big) \\ &= b \sum_{k=3}^{2N+2} (yV_{kx} - xV_{ky}) - \sum_{k=3}^{2N+2} G_k + \mathrm{O}\big(|x, y|^{2N+3}\big), \end{aligned}$$

where

$$\sum_{k=3}^{2N+2} G_k = -(V_x f_1 + V_y g_1) + \mathrm{O}\big(|x, y|^{2N+3}\big) \tag{2.27}$$

with G_k being a homogeneous polynomial of degree k depending only on the coefficients a_{ij}, b_{ij} and c_{ij} with $2 \leqslant i + j < k$.

Then neglecting terms of degree more than $2N+2$ Eq. (2.26) is equivalent to

$$b \sum_{k=3}^{2N+2} (yV_{kx} - xV_{ky}) = \sum_{k=3}^{2N+2} G_k + \sum_{k=2}^{N+1} L_k \bigl(x^2 + y^2\bigr)^k.$$

In order to find V_k and L_k satisfying this equation it suffices to solve the following equation for $3 \leqslant k \leqslant 2N+2$:

$$b(yV_{kx} - xV_{ky}) = G_k \qquad \text{for } k \text{ odd,} \tag{2.28}$$

$$b(yV_{kx} - xV_{ky}) = G_k + L_{k/2}\bigl(x^2 + y^2\bigr)^{k/2} \quad \text{for } k \text{ even.} \tag{2.29}$$

The above equations can be rewritten as

$$b\frac{\mathrm{d}}{\mathrm{d}\theta} V_k(\cos\theta, \sin\theta) = G_k(\cos\theta, \sin\theta) \qquad \text{for } k \text{ odd,} \tag{2.30}$$

$$b\frac{\mathrm{d}}{\mathrm{d}\theta} V_k(\cos\theta, \sin\theta) = L_{k/2} + G_k(\cos\theta, \sin\theta) \quad \text{for } k \text{ even} \tag{2.31}$$

since

$$b(yV_{kx} - xV_{ky})(r\cos\theta, r\sin\theta) = b\,r^k \frac{\mathrm{d}}{\mathrm{d}\theta} V_k(\cos\theta, \sin\theta).$$

Let us solve (2.30) and (2.31) by induction in k. First, for $k = 3$, we have

$$\int_0^{2\pi} G_3(\cos\theta, \sin\theta)\,\mathrm{d}\theta = 0.$$

Hence there is a function $\widetilde{G}_3$ of the form

$$\widetilde{G}_3(\theta) = \sum_{i+j=3} \tilde{g}_{ij} \cos^i\theta \sin^j\theta$$

such that $\widetilde{G}_3'(\theta) = G_3(\cos\theta, \sin\theta)$. Denote $\widetilde{G}_3$ by

$$\widetilde{G}_3(\theta) = \int G_3(\cos\theta, \sin\theta)\,\mathrm{d}\theta.$$

We can write

$$\int G_3(\cos\theta, \sin\theta)\,\mathrm{d}\theta = \sum_{i+j=3} \tilde{g}_{ij} \cos^i\theta \sin^j\theta.$$

Then by (2.30) we have a solution for V_3 below

$$V_3(x,y)=\frac{1}{b}\sum_{i+j=3}\tilde{g}_{ij}x^i y^j.$$

That is, we have $c_{ij}=\frac{1}{b}\tilde{g}_{ij}$ for $i+j=3$.

For $k=4$, G_4 will be known when V_3 is definite. Thus, we can choose L_2 to be such that

$$L_2+\frac{1}{2\pi}\int_0^{2\pi}G_4(\cos\theta,\sin\theta)\,\mathrm{d}\theta=0.$$

It then follows that

$$\int\bigl[L_2+G_4(\cos\theta,\sin\theta)\bigr]\,\mathrm{d}\theta=\sum_{i+j=4}\tilde{g}_{ij}\cos^i\theta\sin^j\theta$$

as before. By (2.31), V_4 will be determined by taking $c_{ij}=\frac{1}{b}\tilde{g}_{ij}$ for $i+j=4$.

For higher values of k, V_k can be determined in the same procedure. This ends the proof. □

The following lemma says that the first nonzero constants of $L_2,\dots,L_{N+1}$ will determine the stability and the order of the focus of (2.11) at the origin if $a=0$.

LEMMA 2.6. *Let $a=d=0$, $b=-c\neq 0$. Then* (2.5) *holds with $k=2m+1\geqslant 3$ if and only if*

$$L_j=0,\quad j=2,\dots,m,\qquad L_{m+1}\neq 0.$$

Moreover, $v_{2m+1}=\frac{L_{m+1}}{2|b|}$ as $L_j=0$, $j=2,\dots,m$.

PROOF. Let $r>0$ be small and L denote the orbit arc of (2.1) from $(r,0)$ to $(P(r),0)$. Note that for any $\mu\in(0,1)$,

$$r+\mu\bigl[P(r)-r\bigr]=r+\mathrm{O}\bigl(r^3\bigr).$$

The mean value theorem implies that

$$\begin{aligned}V\bigl(P(r),0\bigr)-V(r,0)&=V_x\bigl(r+\mu\bigl[P(r)-r\bigr],0\bigr)\bigl(P(r)-r\bigr)\\&=2r\bigl[P(r)-r\bigr]\bigl(1+\mathrm{O}\bigl(r^2\bigr)\bigr),\end{aligned}$$

where V is the function in Lemma 2.5.

On the other hand, by the formula of constant variation we have $x^2+y^2=r^2(1+\mathrm{O}(r))$ along L. Thus, by (2.26)

$$V\big(P(r),0\big) - V(r,0) = \int_L \mathrm{d}V = \int_L (V_x f + V_y g)\,\mathrm{d}t$$
$$= \int_0^{\tau(r)} \left[\sum_{k=2}^{N+1} L_k r^{2k}\big(1+\mathrm{O}(r)\big) + \mathrm{O}\big(r^{2N+3}\big)\right]\mathrm{d}t$$
$$= \frac{2\pi}{|b|}\sum_{k=2}^{N+1} L_k r^{2k}\big(1+\mathrm{O}(r)\big) + \mathrm{O}\big(r^{2N+3}\big),$$

where $\tau(r) = \frac{2\pi}{|b|} + \mathrm{O}(r)$ is the time along L. It follows that

$$P(r) - r = \frac{\pi}{|b|}\sum_{k=2}^{N+1} L_k r^{2k-1}\big(1+\mathrm{O}(r)\big) + \mathrm{O}\big(r^{2N+2}\big).$$

Then the conclusion follows easily and the proof is completed. □

The proof of Lemma 2.6 presents an algorithm to compute constants $L_2, L_3, \ldots$. It includes the following three steps in a loop:

(i) Find G_{2m+1} by (2.27),
(ii) Find V_{2m+1} by (2.30),
(iii) Find G_{2m+2}, L_{m+1} and V_{2m+2} by (2.31).

To begin with $m = 1$ we can get L_2 by executing the 3 steps. We can get L_3 further by doing the same procedure for $m = 2$. See Chicone [14] for more detail.

By (2.3) and (2.5) we can prove the following corollary in a similar way to Lemma 2.6.

COROLLARY 2.1. *Let* $a = d = 0$, $b = -c \neq 0$. *Suppose there exists a function* $\widetilde{V}(x, y) = x^2 + y^2 + \mathrm{O}(|x, y|^3)$ *such that*

$$\dot{\widetilde{V}} = \widetilde{V}_x f + \widetilde{V}_y g = H_{2k}(x, y) + \mathrm{O}\big(|x, y|^{2k+1}\big), \quad k \geqslant 2,$$

where H_{2k} *is a homogeneous polynomial of order* $2k$ *satisfying*

$$\widetilde{L}_k = \frac{|b|}{2\pi}\int_0^{2\pi/|b|} H_{2k}(\cos bt, -\sin bt)\,\mathrm{d}t < 0 \ (> 0).$$

Then the origin is a stable (*unstable*) *focus of order* $k-1$ *of Eq.* (2.1).

Now let us recall some facts obtained so far in this section. We will outline them for analytic systems. Suppose f and g in (2.1) are analytic functions near the origin and (2.3) holds with $a + d = 0$. Then by Lemma 2.1, the proof of Lemma 2.6 and the definition of the Poincaré map in (2.4) we have

$$P(r) - r = \frac{\pi}{|b|}\sum_{k=2}^{N+1} L_k r^{2k-1}\big(1+\mathrm{O}(r)\big) + \mathrm{O}\big(r^{2N+2}\big) = 2\pi\sum_{k=3}^{\infty} v_k r^k, \qquad (2.32)$$

where

$$v_{2m} = \mathrm{O}\big(|v_3, v_5, \ldots, v_{2m-1}|\big), \quad m \geqslant 2.$$

The implicit function theorem ensures the convergence of the series in the right-hand side of (2.32).

If $a = d = 0,\ b = -c \neq 0$, by Lemma 2.3, there exists a formal change of variables given by the formal series

$$(x, y)^{\mathrm{T}} = (u, v)^{\mathrm{T}} + \sum_{i+j\geqslant 2} q_{ij} u^i v^j = Q(u, v)$$

which carries (2.1) formally into

$$\begin{aligned} \dot{u} &= bv + \sum_{j\geqslant 1} (a_j u + b_j v)\big(u^2 + v^2\big)^j, \\ \dot{v} &= -bu + \sum_{j\geqslant 1} (-b_j u + a_j v)\big(u^2 + v^2\big)^j. \end{aligned} \tag{2.33}$$

By Lemma 2.5, there exist a formal series

$$V(x, y) = x^2 + y^2 + \sum_{i+j\geqslant 3} c_{ij} x^i y^j$$

and constants $L_2, L_3, \ldots$, such that

$$V_x f + V_y g = \sum_{k\geqslant 2} L_k \big(x^2 + y^2\big)^k. \tag{2.34}$$

Lemmas 2.4 and 2.6 show that the following three statements are equivalent to each other:

(i) $v_{2j+1} = 0,\ j = 1, \ldots, m-1,\ v_{2m+1} \neq 0$;

(ii) $a_j = 0,\ j = 1, \ldots, m-1,\ a_m \neq 0$;

(iii) $L_j = 0,\ j = 1, \ldots, m,\ L_{m+1} \neq 0$.

Moreover, when one of the conditions (i)–(iii) holds, we have

$$v_{2m+1} = \frac{a_m}{|b|} = \frac{L_{m+1}}{2|b|}.$$

Based on this relationship, we call either v_{2m+1}, a_m or L_{m+1} the mth Lyapunov quantity or focus value.

Note that the series in (2.32) is always convergent for $|r|$ small if (2.1) is analytic. We therefore obtain

THEOREM 2.2. *Suppose* (2.1) *is an analytic system satisfying* $a = d = 0,\ b = -c \neq 0$. *Then* (2.1) *has a center at the origin if and only if one of the following holds*:

(a) $v_{2k+1} = 0$ *for all* $k \geqslant 1$;
(b) $a_k = 0$ *for all* $k \geqslant 1$;
(c) $L_k = 0$ *for all* $k \geqslant 2$.

Lyapunov proved that the formal series for $V(x, y)$ is convergent near the origin if $L_k = 0$ for all $k \geqslant 2$. Hence, by (2.34) in this case the function V represents a first integral whose level sets are orbits of the system (2.1).

Bryuno [9] proved that the formal series for $Q(u, v)$ is convergent near the origin if and only if $a_k = 0$ for all $k \geqslant 1$. In this case, system (2.33) becomes

$$\dot{u} = v\left[b + \sum_{j\geqslant 1} b_j\bigl(u^2 + v^2\bigr)^j\right],$$
$$\dot{v} = -u\left[b + \sum_{j\geqslant 1} b_j\bigl(u^2 + v^2\bigr)^j\right]$$

which has a first integral of the form $u^2 + v^2$. This implies that the original system (2.1) has a first integral of the form $W_1^2(x, y) + W_2^2(x, y)$, where $W(x, y) = (W_1(x, y), W_2(x, y))$ is the inverse of the coordinate change $(x, y)^{\mathrm{T}} = Q(u, v)$.

Thus, by Theorem 2.2 and the next theorem we know that under the condition of Theorem 2.2 system (2.1) has a center at the origin if and only if it has an analytic first integral of the form $x^2 + y^2 + \mathrm{O}(|x, y|^3)$.

The following theorem gives a sufficient condition for (2.1) to have a center at the origin.

THEOREM 2.3. *Consider the* C^∞ *system* (2.1). *Let* (2.2) *hold. If one of the conditions below is satisfied*:
(i) $f(-x, y) = f(x, y)$, $g(-x, y) = -g(x, y)$;
(ii) *there exists a* C^∞ *function* $H(x, y)$ *for* (x, y) *near the origin satisfying* $H(0, 0) = 0$, $H(x, y) \neq 0$ *for* $0 < x^2 + y^2 \ll 1$ *such that* $H_x f + H_y g = 0$,
then (2.1) *has a center at the origin.*

PROOF. Let (i) hold first. Without loss of generality we assume that the orbits of (2.1) near the origin are oriented clockwise. For $y_0 > 0$ small let $(x(t), y(t))$ be the solution of (2.1) with $(x(0), y(0)) = (0, y_0)$. Set

$$L_1 = \bigl\{\bigl(x(t), y(t)\bigr) \mid 0 \leqslant t \leqslant t_1\bigr\}, \qquad L_2 = \bigl\{\bigl(x(t), y(t)\bigr) \mid t_1 \leqslant t \leqslant t_2\bigr\},$$
$$L_1' = \bigl\{\bigl(-x(-t), y(-t)\bigr) \mid -t_1 \leqslant t \leqslant 0\bigr\},$$

where

$$t_1 = \min\bigl\{t > 0 \mid x(t) = 0, y(t) < 0\bigr\}, \qquad t_2 = \min\bigl\{t > 0 \mid x(t) = 0, y(t) > 0\bigr\}.$$

By our assumption, L_1' is an orbit arc of (2.1) starting at $(0, y(t_1))$. This implies that $L_2 = L_1'$ and hence $y(t_2) = y_0$. In other words, $(x(t), y(t))$ is a periodic solution.

Now let (ii) hold. For definiteness, we suppose $H(x, y) > 0$ for $0 < x^2 + y^2 \ll 1$. Then the function $z = H(x, y)$ takes a minimal value at $(x, y) = (0, 0)$. This implies that for sufficiently small $h > 0$, the equation $H(x, y) = h$ defines a closed curve near the origin. Our assumption ensures that the curve is a periodic orbit which approaches the origin as $h \to 0$. The proof is completed. □

EXAMPLE 2.1. Consider a C^∞ system of the form

$$\dot{x} = y - xh(x, y), \qquad \dot{y} = -x - yh(x, y),$$

where

$$h(x, y) = \begin{cases} 0, & \text{for } (x, y) = (0, 0), \\ \mathrm{e}^{-\frac{1}{x^2+y^2}}, & \text{otherwise.} \end{cases}$$

Let $V(x, y) = x^2 + y^2$. Then for the system, (2.34) becomes

$$V_x f + V_y g = -2\left(x^2 + y^2\right)h(x, y) < 0$$

for $x^2 + y^2 > 0$. By the proof of Lemma 2.6, the origin is a stable focus. However, in this case we have $L_k = 0$ for all $k \geqslant 2$.

This example shows that Theorem 2.2 is no longer true for C^∞ systems.

EXAMPLE 2.2. Consider the following cubic Liénard system:

$$\begin{aligned} &\dot{x} = y - \left(a_3x^3 + a_2x^2 + a_1x\right), \\ &\dot{y} = -x, \end{aligned} \tag{2.35}$$

where a_1, a_2 and a_3 are real constants. The divergence of (2.35) takes value $-a_1$ at the origin. Hence, the origin is a stable (unstable) focus if $a_1 > 0$ (< 0). Further, by Theorem 2.1 we have $V_3 = -\frac{3}{8}a_3$ as $a_1 = 0$. Thus, in this case, the origin is stable (unstable) if $a_3 > 0$ (< 0).

Let $a_1 = a_3 = 0$. Theorem 2.3(i) implies that (2.35) has a center at the origin now.

EXAMPLE 2.3. For a given C^∞ function $H : \mathbb{R}^2 \to \mathbb{R}$, it induces a planar system of the form

$$\dot{x} = H_y, \qquad \dot{y} = -H_x$$

which is called a Hamiltonian system with the Hamiltonian function H. The level sets of the function give orbits of the system. By Theorem 2.3(ii), if there is a point $(x_0, y_0) \in \mathbb{R}^2$

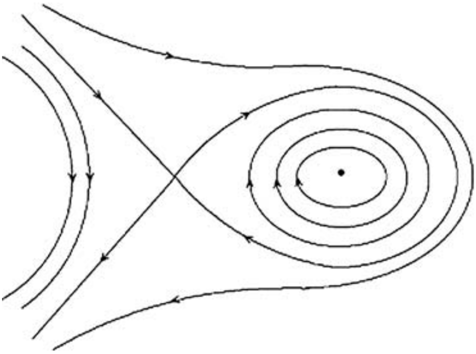


Fig. 6. The phase portrait of (2.36) with $g(x) = x^2 - x$.

satisfying

$$H_x(x_0, y_0) = H_y(x_0, y_0) = 0,$$
$$-\big[H_{xy}(x_0, y_0)\big]^2 + H_{xx}(x_0, y_0)H_{yy}(x_0, y_0) > 0$$

then the above Hamiltonian system has a center at the point.

For example, let

$$H(x, y) = \frac{1}{2}y^2 + G(x).$$

Then the induced system is

$$\dot{x} = y, \qquad \dot{y} = -g(x), \tag{2.36}$$

where $g(x) = G'(x)$. In mechanics, the function H is called the total energy of the system (2.36), while the term $\frac{1}{2}y^2$ is called the kinetic energy and the function G is the potential energy. It is evident that (2.36) has a center at point $(x_0, 0)$ if $g(x_0) = 0,\ g'(x_0) > 0$. A singular point $(x_0, 0)$ satisfying $g'(x_0) < 0$ is a saddle point. Taking $g(x) = x^2 - x$, (2.34) becomes

$$\dot{x} = y, \qquad \dot{y} = x - x^2$$

which has two singular points: center (1,0) and saddle (0,0). Note that $H(0, 0) = 0$. There is a nontrivial orbit γ defined by the equation $H(x, y) = 0$. The orbit approaches the saddle point both positively and negatively. An orbit with this property is called a homoclinic orbit. The phase portrait of the above quadratic Hamiltonian system has been drawn in Fig. 6.

2.3. *Hopf bifurcation*

In the next part, we consider a planar C^∞ system with a vector parameter of the form

$$\dot{x} = f(x, y, \mu), \qquad \dot{y} = g(x, y, \mu), \tag{2.37}$$

where $\mu \in \mathbb{R}^m$, $m \geqslant 1$. Suppose (2.37) has an elementary focus for $\mu = 0$. Without loss of generality we can assume that for all small $|\mu|$ (2.37) has a focus at the origin. Then we can write near the origin

$$\begin{aligned} f(x, y, \mu) &= a(\mu)x + b(\mu)y + \mathrm{O}\big(|x, y|^2\big), \\ g(x, y, \mu) &= c(\mu)x + d(\mu)y + \mathrm{O}\big(|x, y|^2\big), \end{aligned} \tag{2.38}$$

where $a(0) = d(0)$, $b(0) = -c(0) \neq 0$.

Let $P(r, \mu)$ denote the Poincaré map of (2.37) near the origin. By Lemma 2.1, similar to (2.32) for any integer $N > 0$, P has the following expansion:

$$P(r, \mu) = r + 2\pi \sum_{k=1}^{N} v_k(\mu) r^k + \mathrm{O}\big(r^{N+1}\big), \tag{2.39}$$

where

$$v_{2m} = \mathrm{O}\big(|v_1, v_3, \ldots, v_{2m-1}|\big), \quad 1 \leqslant m \leqslant \frac{N}{2}. \tag{2.40}$$

Sometimes, for convenience, we call $v_{2m+1}(\mu)$ in (2.39) the mth Lyapunov constant of (2.37) at the origin, $m \geqslant 1$. Introduce

$$d(r, \mu) = P(r, \mu) - r,$$

which is called a succession function or bifurcation function of (2.37). This function is also called the displacement function.

By using an analogous formula to (2.6) it is easy to obtain

LEMMA 2.7. *For* $|\mu|$ *small, Eq.* (2.37) *has a limit cycle near the origin if and only if there exist* $r_1(\mu) > 0$, $r_2(\mu) < 0$ *near* $r = 0$ *such that*

$$d\big(r_j(\mu), \mu\big) = 0, \quad j = 1, 2.$$

By (2.38) and the discussion after Remark 2.2 we know that $v_1(\mu)$ has the same sign as $a(\mu) + d(\mu)$. In fact, $v_1(\mu) = \frac{1}{2\pi}[\mathrm{e}^{2\pi\alpha/|\beta|} - 1]$ with $\alpha = \frac{1}{2}(a + d)$, $|\beta| = \frac{1}{2}\sqrt{-4bc - (a - d)^2}$. Hence, if $a(0) + d(0) \neq 0$ then $v_1(0) \neq 0$ and $d(r, \mu)$ has no positive zero near $r = 0$ for all $|\mu|$ small. Thus, if $a(0) + d(0) \neq 0$, there is no limit cycle near the origin for all $|\mu|$ small.

Let $a(0)+d(0)=0$, and let v_{30} denote the first order focus value of (2.37) for $\mu=0$ at the origin. Then by (2.39) and (2.40) we have

$$d(r,\mu)=2\pi r\bigl[v_1(\mu)+v_2(\mu)r+v_3(\mu)r^2+\mathrm{O}\bigl(r^3\bigr)\bigr],$$

where $v_1(0)=v_2(0)=0,\ v_3(0)=v_{30}$. Similar to Theorem 1.3(i) we can prove that if $v_{30}\neq 0,\ d(r,\mu)$ has two zeros in r with one positive and the other negative for $v_1(\mu)v_{30}<0$ and has no nontrivial zero for $v_1(\mu)v_{30}>0$. Thus, by Lemma 2.7 and noting

$$v_1=\frac{a(\mu)+d(\mu)}{2|b(0)|}\bigl(1+\mathrm{O}(\mu)\bigr),$$

we have proved the following

THEOREM 2.4. *If for $\mu=0$ (2.37) has a first order focus at the origin, then it has at most one limit cycle near the origin for all $|\mu|$ small. Moreover, the limit cycle exists if and only if $(a(\mu)+d(\mu))v_{30}<0$.*

In general, similar to Theorem 1.3(ii), we have:

THEOREM 2.5. *If for $\mu=0$ (2.37) has a kth order focus at the origin ($k\geqslant 2$), then it has at most k limit cycles near the origin for $|\mu|$ small. Moreover, k limit cycles can appear by suitable perturbations.*

EXAMPLE 2.4. Consider a cubic Liénard system

$$\begin{aligned}\dot{x}&=y-\bigl(x^5+\mu_1x^3+\mu_2x\bigr),\\ \dot{y}&=-x,\end{aligned}\tag{2.41}$$

where μ_1 and μ_2 are small parameters.

First, by (2.39) and (2.40), we have

$$\begin{aligned}v_1&=-\frac{1}{2}\mu_2+\mathrm{O}\bigl(\mu_2^2\bigr),\quad &v_2&=\mathrm{O}(\mu_2),\\ v_3&=-\frac{3}{8}\mu_1+\mathrm{O}(\mu_2),\quad &v_4&=\mathrm{O}\bigl(|v_1,v_3|\bigr).\end{aligned}$$

It is easy to check that for

$$V(x,y)=x^2+y^2-\frac{5}{8}xy^5-\frac{5}{3}x^3y^3-\frac{11}{8}x^5y$$

it holds along the orbits of (2.41)

$$\left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{\mu_1=\mu_2=0}=-\frac{5}{8}\bigl(x^2+y^2\bigr)^3+\mathrm{O}\bigl(|x,y|^{10}\bigr).$$

Hence, by Lemma 2.6, it follows that

$$v_5 = -\frac{5}{16} + \mathrm{O}\big(|\mu_1, \mu_2|\big).$$

This shows that for $\mu_1 = \mu_2 = 0$ the origin is a stable focus of order 2 for (2.41).

We claim that there is a function

$$\delta(\mu_1) = \frac{9}{40}\mu_1^2 + \mathrm{O}\big(|\mu_1|^{5/2}\big)$$

such that for $|\mu_1| + |\mu_2|$ small Eq. (2.41) has

(i) no limit cycle if either $\mu_1 \geqslant 0$ and $\mu_2 \geqslant 0$ or $\mu_1 < 0$ and $\mu_2 > \delta(\mu_1)$;
(ii) a unique simple limit cycle if either $\mu_1 \geqslant 0$ and $\mu_2 < 0$ or $\mu_1 < 0$ and $\mu_2 \leqslant 0$;
(iii) a unique double limit cycle if $\mu_1 < 0$ and $\mu_2 = \delta(\mu_1)$;
(iv) two simple limit cycles if $\mu_1 < 0$ and $\mu_2 < \delta(\mu_1)$.

In fact, by $v_2 = \mathrm{O}(v_1)$, $v_3 = -\frac{3}{8}\mu_1 + \mathrm{O}(v_1)$ and $v_4 = \mathrm{O}(|v_1| + |\mu_1|)$ we can write

$$\begin{aligned} d(r,\mu) &= 2\pi r P_1(r,u)\left[v_1 - \frac{3}{8}\mu_1 r^2 P_2(r,\mu) + v_5^* r^4 P_3(r,\mu)\right] \\ &= 2\pi r P_1(r,u) d_1(r,\mu), \end{aligned}$$

where $P_i = 1 + \mathrm{O}(r) \in C^\infty$, $i = 1, 2, 3$, $v_5^* = -\frac{5}{16} + \mathrm{O}(|\mu_1|)$. The function d_1 can be written further as

$$d_1(r,\mu) = v_1 - \frac{3}{8}\mu_1\rho^2 + v_5^*\rho^4 P_3^*(\rho,\mu) \equiv d_2(\rho,\mu),$$

where $\rho = r\sqrt{P_2(r,\mu)} = r + \mathrm{O}(r^2) \in C^\omega$, $P_3^* = 1 + \mathrm{O}(\rho) \in C^\omega$.

Clearly, for $\mu_1 \geqslant 0$ we have $\frac{\partial d_2}{\partial\rho} < 0$ for $0 < \rho \ll 1$. This implies that for $\mu_1 \geqslant 0$, d_1 has a unique zero if and only if $v_1 > 0$. Note that Eq. (2.41) has a double limit cycle near the origin if and only if the function $d_2(\rho,\mu)$ has a double zero in ρ near $\rho = 0$. Let us consider the equations below

$$d_2(\rho,\mu) = 0, \qquad \frac{\partial d_2}{\partial\rho}(\rho,\mu) = 0.$$

We can solve form the above equations

$$\begin{aligned} \rho &= \sqrt{-\frac{3}{5}\mu_1}\big(1 + \delta_1\big(\sqrt{-\mu_1}\big)\big), \\ v_1 &= \frac{9}{256 v_5^*}\mu_1^2\big(1 + \delta_2\big(\sqrt{-\mu_1}\big)\big) \end{aligned}$$

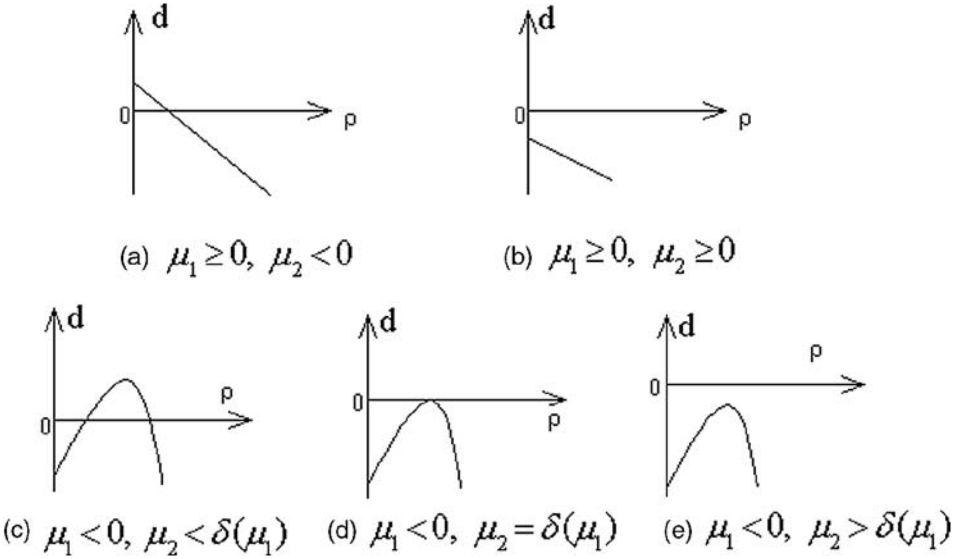


Fig. 7. Curves determined by $d_2(\rho, \mu)$.

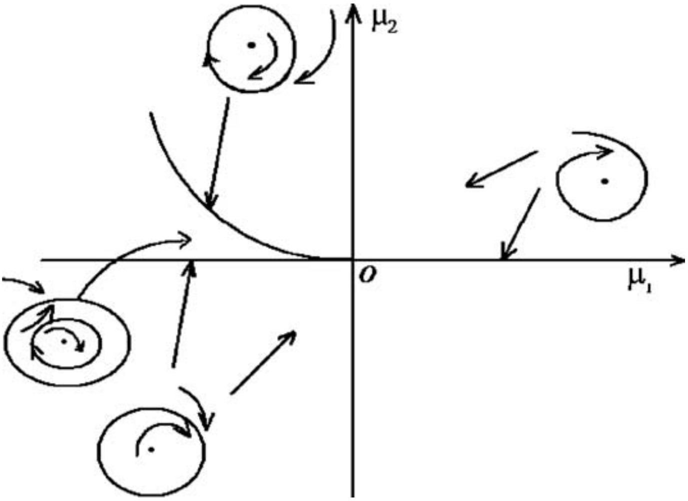


Fig. 8. The bifurcation diagram of (2.41).

for $\mu_1 < 0$, where $\delta_i(u) = \mathrm{O}(u) \in C^\omega,\ i = 1, 2$. The second equation above determines a unique function

$$\mu_2 = \delta(\mu_1) = \frac{9}{40}\mu_1^2 + \mathrm{O}\big(\mu_1^{5/2}\big).$$

Now it is easy to draw curves determined by the function $d = d_2(\rho, \mu)$ on (d, ρ) plane, see Fig. 7.

Then the claim follows easily. For the bifurcation diagram of (2.41) see Fig. 8.

On the (μ_1, μ_2) plane in Fig. 8, the line $\mu_2 = 0$ (near the origin) is Hopf bifurcation curve, and the curve $\mu_2 = \delta(\mu_1)$ is the double limit cycle bifurcation curve, or in other words, the saddle-node bifurcation curve for limit cycles.

If (2.37) has a center at the origin for $\mu = 0$, we can also study the bifurcation of limit cycles for $|\mu|$ small. We give an example to illustrate this phenomenon.

EXAMPLE 2.5. Consider the cubic system

$$\begin{aligned} \dot{x} &= y + x^2 + \varepsilon\big(x^3 - \varepsilon x\big), \\ \dot{y} &= -\big(x - x^3\big). \end{aligned}$$

We prove that the system has a unique limit cycle near the origin for $0 < \varepsilon \ll 1$.

For the purpose, let us consider

$$\begin{aligned} \dot{x} &= y + x^2 + \varepsilon\big(x^3 - \delta x\big), \\ \dot{y} &= -\big(x - x^3\big), \end{aligned} \tag{2.42}$$

and prove that the system has a unique limit cycle near the origin for $0 < |\varepsilon| < \varepsilon_0,\ 0 < \delta < \varepsilon_0$ for a small constant $\varepsilon_0 > 0$.

First, by Theorem 2.3(i), (2.42) has a center at the origin for $\varepsilon = 0$. Hence, the succession function $d(r, \varepsilon, \delta)$ of (2.42) can be written as

$$\begin{aligned} d(r, \varepsilon, \delta) &= 2\pi\varepsilon r\big[v_1^*(\varepsilon, \delta) + v_2^*(\varepsilon, \delta)r + v_3^*(\varepsilon, \delta)r^2 + \mathrm{O}\big(r^3\big)\big] \\ &= 2\pi\varepsilon r d^*(r, \varepsilon, \delta), \end{aligned}$$

where

$$\varepsilon v_1^* = v_1 = \frac{1}{2\pi}\big[\mathrm{e}^{-\varepsilon\delta\pi} - 1\big] = -\frac{\varepsilon\delta}{2}\big(1 + \mathrm{O}(\varepsilon\delta)\big),$$

$$v_2^* = \mathrm{O}\big(v_1^*\big), \quad \varepsilon v_3^* = v_3 = \frac{3}{8}\varepsilon + \mathrm{O}(v_1).$$

Therefore,

$$d^* = -\frac{\delta}{2}\big(1 + \mathrm{O}(\varepsilon\delta)\big) + \mathrm{O}(\delta)r + \left(\frac{3}{8} + O(\delta)\right)r^2 + \mathrm{O}\big(r^3\big)$$

which has a positive zero $r = \frac{2}{\sqrt{3}}\sqrt{\delta}(1 + \mathrm{O}(\varepsilon\delta))$ near $r = 0$ if $0 < \delta \ll 1$. Then the conclusion for (2.42) follows directly.

2.4. *Degenerate Hopf bifurcation*

We now turn to the degenerate Hopf bifurcation near a center and establish a general theory by using the first order Melnikov function.

Consider a C^∞ planar system of the form

$$\begin{aligned} \dot{x} &= f(x,y) + \varepsilon p(x,y,\varepsilon,\delta), \\ \dot{y} &= g(x,y) + \varepsilon q(x,y,\varepsilon,\delta), \end{aligned} \tag{2.43}$$

where $\varepsilon \in \mathbb{R}$, $\delta \in D \subset \mathbb{R}^m$ $(m \geqslant 1)$ with D compact, and

$$\left.\begin{aligned} &\big(f(x,y), g(x,y)\big) = \mu(H_y, -H_x), \quad \mu = \pm 1, \\ &H(x,y) = K\big(x^2+y^2\big) + \mathrm{O}\big(|x,y|^3\big), \quad K > 0, \\ &p(0,0,\varepsilon,\delta) = q(0,0,\varepsilon,\delta) = 0. \end{aligned}\right\} \tag{2.44}$$

Then for $0 < h \ll 1$, the equation $H(x,y) = h$ defines a periodic orbit L_h of (2.43) $(\varepsilon = 0)$ which intersects the positive x-axis at $A(h) = (a(h), 0)$. Let $B(h,\varepsilon,\delta)$ denote the first intersection point of the positive orbit of (2.43) starting at $A(h)$ with the positive x-axis. Then by (2.44) we have

$$H(B) - H(A) = \int_{AB} \mathrm{d}H = \varepsilon\big[M(h,\delta) + \mathrm{O}(\varepsilon)\big], \tag{2.45}$$

where

$$\begin{aligned} M(h,\delta) &= \frac{1}{\mu}\oint_{L_h} (fq - gp)|_{\varepsilon=0}\,\mathrm{d}t = \frac{1}{\mu}\oint_{L_h} (q\,\mathrm{d}x - p\,\mathrm{d}y)|_{\varepsilon=0} \\ &= \frac{1}{|\mu|}\iint_{H\leqslant h} (p_x + q_y)|_{\varepsilon=0}\,\mathrm{d}x\,\mathrm{d}y. \end{aligned} \tag{2.46}$$

We call M the first order Melnikov function. We will see that the function plays an important role in the study of the number of limit cycles.

Let $u(t,c)$ be the solution of (2.43) $(\varepsilon = 0)$ with initial value $(c, 0)$. Then we have by (2.44)

$$\begin{aligned} &H\big(u(t,c)\big) = H(c,0) = r^2(c), \quad t \in \mathbb{R}, \\ &r(c) = c\sqrt{K + S(c)}, \quad S(c) = \mathrm{O}(c) \in C^\infty. \end{aligned}$$

Denote by $c = c(r) = \frac{r}{\sqrt{K}} + \mathrm{O}(r^2)$ the inverse of $r = r(c)$. Then

$$H\big(v(t,r)\big) = r^2, \qquad v\big(T(r), r\big) = \big(c(r), 0\big),$$

where $v(t,r) = u(t, c(r))$, and $T(r)$ is the period of the function v in t. Clearly, the functions $v(t,r)$ and $T(r)$ are both C^∞ with $T(0) = \frac{\pi}{K|\mu|} > 0$. Introduce a C^∞ function as follows:

$$G(\theta, r) = v\left(\frac{T(r)}{2\pi}\theta, r\right).$$

Obviously, $G = \mathrm{O}(r)$ is 2π-periodic in θ, and

$$H\big(G(\theta, r)\big) = r^2, \quad \theta \in \mathbb{R}. \tag{2.47}$$

LEMMA 2.8. *The change of variables*

$$(x, y)^{\mathrm{T}} = G(\theta, r) \tag{2.48}$$

transforms the system (2.43) *into the* C^∞ *system*

$$\begin{aligned} \dot\theta &= \frac{2\pi}{T(r)}\left[1 - \frac{\varepsilon}{2\mu r} G_r \wedge \big(p(G,\varepsilon,\delta), q(G,\varepsilon,\delta)\big)\right], \\ \dot r &= \frac{\varepsilon}{2r} DH(G) \cdot \big(p(G,\varepsilon,\delta), q(G,\varepsilon,\delta)\big)^{\mathrm{T}}. \end{aligned} \tag{2.49}$$

PROOF. Differentiating (2.48) in t yields

$$G_\theta \dot\theta + G_r \dot r = \big(f(G) + \varepsilon p(G,\varepsilon,\delta), g(G) + \varepsilon q(G,\varepsilon,\delta)\big)^{\mathrm{T}}. \tag{2.50}$$

By (2.47) and the definition of G, we have

$$\mathrm{D}H(G)G_\theta = 0, \qquad \mathrm{D}H(G)G_r = 2r, \qquad G_\theta = \frac{T(r)}{2\pi}\big(f(G), g(G)\big).$$

Multiplying (2.50) by $\mathrm{D}H(G)$ from the left-hand side gives

$$2r\dot r = \varepsilon \mathrm{D}H(G) \cdot \big(p(G,\varepsilon,\delta), q(G,\varepsilon,\delta)\big),$$

which gives the second equation in (2.49).

Further, noting that

$$G_r \wedge G_\theta = \frac{T(r)}{2\pi} G_r \wedge \big(f(G), g(G)\big) = -\frac{\mu T(r)}{2\pi} \mathrm{D}H(G)G_r = -\frac{\mu T(r)}{\pi} r,$$

it follows from (2.50) that

$$(G_r \wedge G_\theta)\dot\theta = G_r \wedge \big(f(G), g(G)\big) + \varepsilon G_r \wedge (p, q),$$

or

$$-\frac{\mu T(r)}{\pi} r\dot\theta = -2\mu r + \varepsilon G_r \wedge (p, q)$$

which gives the first equation in (2.49). The proof is completed. □

By (2.49) we obtain

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \varepsilon R(\theta, r, \varepsilon, \delta), \tag{2.51}$$

where R is C^∞ and 2π-periodic in θ with $R = \mathrm{O}(r)$ and

$$R(\theta, r, \theta, \delta) = \frac{T(r)}{4\pi r} \mathrm{D}H(G) \cdot \big(p(G, 0, \delta), q(G, 0, \delta)\big)^{\mathrm{T}}. \tag{2.52}$$

Let $r(\theta, \rho, \varepsilon, \delta)$ denote the solution of (2.51) with $r(0, \rho, \varepsilon, \delta) = \rho$, and $(x(t, c, \varepsilon, \delta), y(t, c, \varepsilon, \delta))$ the solution of (2.43) with $(x(0, c, \varepsilon, \delta), y(0, c, \varepsilon, \delta)) = (c, 0)$. Then the Poincaré map $P(c, \varepsilon, \delta)$ of (2.43) is given by

$$P(c, \varepsilon, \delta) = x(\tau, c, \varepsilon, \delta) = P(c)$$

where

$$\tau = \min\big\{t > 0 \mid cx(t, c, \varepsilon, \delta) > 0,\ y(t, c, \varepsilon, \delta) = 0\big\}.$$

By the definition of G we have $G(0, r) = (c(r), 0)$ which yields

$$G\big(0, P_1(\rho)\big) = \big(c\big(P_1(\rho)\big), 0\big),$$

where

$$P_1(\rho) = r(2\pi, \rho, \varepsilon, \delta) = P_1(\rho, \varepsilon, \delta)$$

which is called the Poincaré map of (2.51).

On the other hand, by (2.48) for $c = c(\rho)$ we have

$$\big(P(c, \varepsilon, \delta), 0\big) = G\big(2\pi, r(2\pi, \rho, \varepsilon, \delta)\big) = G\big(0, P_1(\rho)\big).$$

Hence

$$\big(P\big(c(\rho), \varepsilon, \delta\big), 0\big) = \big(c\big(P_1(\rho)\big), 0\big),$$

or

$$P \circ c = c \circ P_1. \tag{2.53}$$

THEOREM 2.6. *Let* (2.43) *satisfy* (2.44). *Then*:

(i) *The function M is C^∞ at $h = 0$. It is analytic at $h = 0$ if* (2.43) *is analytic.*

(ii) *If there exist a compact subset* D_0 *of* D *and an integer* $k \geqslant 0$ *such that for* $0 < h \ll 1$

$$M(h,\delta) = B_k(\delta)h^{k+1} + \mathrm{O}\big(h^{k+2}\big), \quad B_k(\delta) \neq 0,\ \delta \in D_0, \tag{2.54}$$

then there exist $\varepsilon_0 > 0$, *an open set* $U(D_0) \supset D_0$ *and a neighborhood* V *of the origin such that* (2.43) *has at most* k *limit cycles in* V *for* $0 < |\varepsilon| < \varepsilon_0$, $\delta \in U(D_0)$.

PROOF. It is easy to see that the Poincaré map P_1 of (2.51) can be written as

$$P_1(r,\varepsilon,\delta) = r + \varepsilon r F(r,\varepsilon,\delta), \tag{2.55}$$

where

$$rF(r,0,\delta) = \int_0^{2\pi} R(\theta,r,0,\delta)\,\mathrm{d}\theta \equiv R_0(r,\delta).$$

By (2.44) and (2.52) we have

$$\begin{aligned} R_0(r,\delta) &= \frac{1}{2r\mu}\oint_{H=r^2} (f,g)\wedge(p,q)|_{\varepsilon=0}\,\mathrm{d}t \\ &= \frac{1}{2r\mu}\oint_{H=r^2} (q\,\mathrm{d}x - p\,\mathrm{d}y)|_{\varepsilon=0} \end{aligned} \tag{2.56}$$

which immediately follows $R_0(-r,\delta) = -R_0(r,\delta)$. Hence, the function $M^*(r)$ defined by

$$M^*(r) = 2rR_0(r,\delta)$$

is even. Since $M^* \in C^\infty$, for any integer $j > 1$, we have

$$M^*(r) = \sum_{i=1}^{j} A_i r^{2i} + N(r),$$

where $N \in C^\infty$ is even and $N^{(i)}(0) = 0$, $i = 0,1,\dots,2j$.

Let $\overline{N}(h) = N(\sqrt{h})$. We claim that

$$\overline{N}^{(i)}(0) = 0, \quad i = 0,1,\dots,j.$$

In fact, we can prove

$$\overline{N}^{(i)}(h) = h^{j-i}\widetilde{N}_i\big(\sqrt{h}\big), \quad i = 0,1,\dots,j, \tag{2.57}$$

by induction in i, where $\widetilde{N}_i(r)$ is C^∞ in r and $\widetilde{N}_i(0) = 0$.

First, it is not hard to see that

$$N^{(i)}(r) = r^{2j-i}N_i(r), \quad N_i \in C^\infty, \qquad N_i(0) = 0, \quad i = 0,1,\dots,2j.$$

Hence, (2.57) holds for $i = 0$. That is,

$$\overline{N}(h) = h^j N_0\big(\sqrt{h}\big) = h^j \widetilde{N}_0\big(\sqrt{h}\big).$$

It then follows that

$$\big(\overline{N}(h)\big)' = h^{j-1}\left[jN_0\big(\sqrt{h}\big) + \frac{1}{2}\sqrt{h}N_0'\big(\sqrt{h}\big)\right].$$

Let

$$\widetilde{N}_1(r) = jN_0(r) + \frac{1}{2}rN_0'(r).$$

Then $\widetilde{N}_1 \in C^\infty$ and $\widetilde{N}_1(0) = 0$. Thus, (2.57) holds for $i = 1$. Let (2.57) hold for $i = k$. Then we have

$$\overline{N}^{(k+1)}(h) = \big[h^{j-k}\widetilde{N}_k\big(\sqrt{h}\big)\big]' = h^{j-k-1}\left[(j-k)\widetilde{N}_k\big(\sqrt{h}\big) + \frac{1}{2}\sqrt{h}\widetilde{N}_k'\big(\sqrt{h}\big)\right].$$

Set

$$\widetilde{N}_{k+1}(r) = (j-k)\widetilde{N}_k(r) + \frac{1}{2}r\widetilde{N}_k'(r).$$

It follows that (2.57) holds for $i = k+1$. Hence, (2.57) has been proved.

By (2.57) it is immediate that $\overline{N} \in C^j$ for $0 \leqslant h \ll 1$. Let

$$\overline{M}(h) = \sum_{i=1}^{j} A_i h^i + \overline{N}(h).$$

Then $\overline{M} \in C^j$. Therefore $\overline{M} \in C^\infty$ since j is arbitrarily large.

Note that by (2.46) and (2.56) we have

$$\overline{M}(h) = M^*\big(\sqrt{h}\big) = 2\sqrt{h}R_0\big(\sqrt{h}, \delta\big) = M(h, \delta).$$

Thus, $M \in C^\infty$ in h at $h = 0$, and if (2.54) holds, then

$$R_0(r, \delta) = \frac{1}{2}B_k(\delta)r^{2k+1} + \mathrm{O}\big(r^{2k+3}\big).$$

By (2.53) and (2.55) and Lemma 2.7, Eq. (2.43) has a limit cycle near the origin if and only if F has two zeros correspondingly with one positive and the other negative. Hence, by Rolle theorem, (2.54) ensures that at most k limit cycles can appear near the origin for $|\varepsilon|$ small and δ in a neighborhood of D_0.

Finally, if (2.43) is analytic, then (2.51) is analytic. Hence the functions R_0 and M^* are also analytic. This implies that the series

$$M^*(r) = \sum_{i \geqslant 1} A_i r^{2j}$$

is convergent for $|r|$ small. Therefore M is analytic since $M(h, \delta) = M^*(\sqrt{h})$. The proof is completed. □

By Theorem 2.6, for any $k \geqslant 1$ we have the following expansion for M:

$$M(h, \delta) = h\bigl[b_0(\delta) + b_1(\delta)h + \cdots + b_k(\delta)h^k + \mathrm{O}\bigl(h^{k+1}\bigr)\bigr], \quad 0 < h \ll 1. \tag{2.58}$$

We can use the coefficients $b_0, b_1, \ldots, b_k$ to study the Hopf bifurcation of limit cycles.

COROLLARY 2.2. *Suppose that there exist $k \geqslant 1$, $\delta_0 \in D$ such that $b_k(\delta_0) \neq 0$ and*

$$b_j(\delta_0) = 0, \quad j = 0, 1, \ldots, k-1, \qquad \det \frac{\partial(b_0, \ldots, b_{k-1})}{\partial(\delta_1, \ldots, \delta_k)}(\delta_0) \neq 0, \tag{2.59}$$

where $\delta = (\delta_1, \ldots, \delta_m)$, $m \geqslant k$. Then for any $\varepsilon_0 > 0$ and any neighborhood V of the origin there exist $0 < \varepsilon < \varepsilon_0$ and $|\delta - \delta_0| < \varepsilon_0$ such that (2.43) *has precisely k limit cycles in V.*

PROOF. Fix $\delta_j = \delta_{j0}$ for $j = k+1, \ldots, m$. By (2.59) the change of parameters

$$b_j = b_j(\delta), \quad j = 0, \ldots, k-1$$

has the inverse $\delta_j = \delta_j(b_0, \ldots, b_{k-1})$, $j = 1, \ldots, k$. Then (2.58) becomes

$$M(h, \delta) = h\bigl[b_0 + b_1 h + \cdots + b_{k-1}h^{k-1} + b_k h^k + \mathrm{O}\bigl(h^{k+1}\bigr)\bigr]$$

where $b_k = b_k(\delta_0) \neq 0$ as $b_0 = \cdots = b_{k-1} = 0$. By changing the sign of $b_{k-1}, b_{k-2}, \ldots, b_0$ in turn such that

$$b_{j-1}b_j < 0, \quad j = k, k-1, \ldots, 1, \qquad 0 < |b_0| \ll |b_1| \ll \cdots \ll |b_{k-1}| \ll 1,$$

we can find k simple positive zeros $h_1, h_2, \ldots, h_k$ with $0 < h_k < h_{k-1} < \cdots < h_1 \ll 1$. Let $r_j = \sqrt{h_j}$, $j = 1, \ldots, k$. Then $r_1, \ldots, r_k$ are simple positive zeros of $R_0(r, \delta)$. By (2.55) and the implicit function theorem, the function F has k zeros $r_j + \mathrm{O}(\varepsilon)$, $j = 1, \ldots, k$. This ends the proof. □

The following is evident from Theorem 2.6(ii).

COROLLARY 2.3. *If for any* $\varepsilon_0 > 0$ *and any neighborhood* V *of the origin* (2.43) *has* k *limit cycles in* V *for some* (ε, δ) *satisfying* $0 < |\varepsilon| < \varepsilon_0$, $|\delta - \delta_0| < \varepsilon_0$, *then* $M(h, \delta_0) = \mathrm{O}(h^{k+1})$.

In the multiple parameter case it often occurs that $M(h, \delta_0) \equiv 0$, $M(h, \delta) \not\equiv 0$ for $\delta \neq \delta_0$. In this case (2.54) fails to hold, and we have further the following result.

THEOREM 2.7. *Let* (2.44) *hold. Suppose*

(i) *the coefficients* $b_0, \ldots, b_{k-1}$ *in* (2.58) *satisfy* (2.59);

(ii) *there exists a* k*-dimensional vector function* $\varphi(\varepsilon, \delta_{k+1}, \ldots, \delta_m)$ *such that* (2.43) *has a center at the origin when* $(\delta_1, \ldots, \delta_k) = \varphi(\varepsilon, \delta_{k+1}, \ldots, \delta_m)$ *for* $|\varepsilon| + |\delta - \delta_0|$ *small.*

Then there exist $\varepsilon_0 > 0$ *and a neighborhood* V *of the origin such that* (2.43) *has at most* $k - 1$ *limit cycles in* V *for* $0 < |\varepsilon| < \varepsilon_0$, $|\delta - \delta_0| < \varepsilon_0$. *Moreover,* $k - 1$ *limit cycles can appear in an arbitrary neighborhood of the origin for some* (ε, δ) *sufficiently near* $(0, \delta_0)$.

PROOF. By (2.58) and noting that $M(h, \delta) = 2rR_0(r, \delta)$, $r = \sqrt{h}$, we have

$$R_0(r, \delta) = r\left[\sum_{j=0}^{k} b_j(\delta) r^{2j} + \mathrm{O}\big(r^{2k+2}\big)\right].$$

Then the function F in (2.55) has the following expansion:

$$F(r, \varepsilon, \delta) = \sum_{j=0}^{2k-1} c_j(\varepsilon, \delta) r^j + r^{2k} Q(r, \varepsilon, \delta), \tag{2.60}$$

where $Q \in C^\infty$, and

$$c_{2j}(0, \delta) = b_j(\delta), \quad c_{2j+1}(0, \delta) = 0, \quad j = 0, \ldots, k - 1.$$

By (2.59), the equations

$$b_j = c_{2j}(\varepsilon, \delta), \quad j = 0, \ldots, k - 1$$

have the solution

$$(\delta_1, \ldots, \delta_k) = \bar{\varphi}(\varepsilon, b_0, \ldots, b_{k-1}, \delta_{k+1}, \ldots, \delta_m). \tag{2.61}$$

By condition (ii) we have $F(r, \varepsilon, \delta) = 0$ and hence $c_{2j}(\varepsilon, \delta) = 0$, $j = 0, \ldots, k - 1$ as long as $(\delta_1, \ldots, \delta_k) = \varphi(\varepsilon, \delta_{k+1}, \ldots, \delta_m)$. The uniqueness of the solution $\bar{\varphi}$ implies that

$$\bar{\varphi} = \varphi \quad \text{if and only if} \quad b_0 = \cdots = b_{k-1} = 0. \tag{2.62}$$

Inserting (2.61) into (2.60) we obtain

$$\begin{aligned} F &= \sum_{j=0}^{k-1}\big[b_j r^{2j} + c_{2j+1}(\varepsilon,\bar{\varphi},\delta_{k+1},\dots,\delta_m)r^{2j+1}\big] + r^{2k}Q(r,\varepsilon,\bar{\varphi},\delta_{k+1},\dots,\delta_m) \\ &\equiv \widetilde{F}\big(r,\varepsilon,\tilde{\delta}\big), \quad \tilde{\delta} = (b_0,\dots,b_{k-1},\delta_{k+1},\dots,\delta_m). \end{aligned} \tag{2.63}$$

It follows from (2.61) and (2.62) that $\widetilde{F} = 0$ as $b_j = 0$, $j = 0,\dots,k-1$. Hence, by the mean value theorem we can write

$$c_{2j+1}(\varepsilon,\bar{\varphi},\delta_{k+1},\dots,\delta_m) = \varepsilon\sum_{i=0}^{k-1} b_i A_{ij}\big(\varepsilon,\tilde{\delta}\big),$$

$$Q(r,\varepsilon,\bar{\varphi},\delta_{k+1},\dots,\delta_m) = \sum_{i=0}^{k-1} b_i Q_i\big(r,\varepsilon,\tilde{\delta}\big),$$

where $Q_i \in C^\infty$, $i = 0,\dots,k-1$. Note that $\varepsilon r^2 F = r(P_1 - r)$ keeps sign for $0 < |r| \ll 1$ by Lemma 2.1 and (2.53). By (2.63), it follows that

$$c_{2j+1}(\varepsilon,\bar{\varphi},\delta_{k+1},\dots,\delta_m) = 0, \quad \text{if } b_0 = \cdots = b_j = 0,\ j = 0,\dots,k-1.$$

This yields that

$$A_{ij} = 0, \quad j+1 \leqslant i \leqslant k-1,\ j = 0,\dots,k-2.$$

Therefore, the function $\widetilde{F}$ in (2.63) can be written as

$$\widetilde{F}\big(r,\varepsilon,\tilde{\delta}\big) = \sum_{j=0}^{k-1} b_j r^{2j} P_j\big(r,\varepsilon,\tilde{\delta}\big), \tag{2.64}$$

where

$$P_j = 1 + \varepsilon\sum_{i=j}^{k-1} A_{ji}\big(\varepsilon,\tilde{\delta}\big)\, r^{2(i-j)+1} + r^{2(k-j)} Q_j, \quad 0 \leqslant j \leqslant k-1.$$

We particularly have

$$\widetilde{F}\big(r,\varepsilon,\tilde{\delta}\big) = \sum_{j=0}^{k-1} b_j r^{2j} + \mathrm{O}\big(|r|^{2k} + |\varepsilon|\big).$$

By using this form we can prove, in a similar manner to Corollary 2.2, that $k-1$ limit cycles can appear in an arbitrary neighborhood of the origin for some (ε,δ) sufficiently close to $(0,\delta_0)$.

It needs only to prove that $\widetilde{F}$ has at most $k-1$ positive zeros near $r=0$ for $|b_0|+\cdots+|b_{k-1}|>0$ small. We do this by induction in k. For convenience, denote by F_{k-1} the right-hand side function of (2.64).

First, for $k=1$, we have

$$F_0 = b_0 P_0 = b_0\big(1+\mathrm{O}\big(|\varepsilon|+|r|\big)\big) \neq 0 \quad \text{for } |b_0|>0.$$

Suppose for $k=n$ the function

$$F_{n-1} = \sum_{j=0}^{n-1} b_j r^{2j} P_j$$

has at most $n-1$ positive zeros in r near $r=0$ for $|\varepsilon|$ and $|b_0|+\cdots+|b_{n-1}|>0$ small, where $P_j = 1+\mathrm{O}(|\varepsilon|+|r|) \in C^\infty$, $j=0,\ldots,n-1$. Consider the function

$$F_n = \sum_{j=0}^{n} b_j r^{2j} P_j,$$

where $P_j = 1+\mathrm{O}(|\varepsilon|+|r|) \in C^\infty$. We have

$$F_n = P_0 \widetilde{F}_n, \quad \widetilde{F}_n = \sum_{j=0}^{n} b_j r^{2j} \widetilde{P}_j,$$

where $\widetilde{P}_0 = 1$, $\widetilde{P}_j = P_j/P_0 = 1+\mathrm{O}(|\varepsilon|+|r|) \in C^\infty$, $j=1,\ldots,n$. Then

$$\frac{\mathrm{d}\widetilde{F}_n}{\mathrm{d}r} = \sum_{j=1}^{n} 2jb_j r^{2j-1}\left(\widetilde{P}_j + \frac{r}{2}\frac{\mathrm{d}\widetilde{P}_j}{\mathrm{d}r}\right) = r\sum_{j=0}^{n-1} \bar{b}_j r^{2j} \overline{P}_j = r\overline{F}_{n-1},$$

where

$$\bar{b}_j = 2(j+1)b_{j+1}, \quad \overline{P}_j = \widetilde{P}_{j+1} + \frac{r}{2}\frac{\mathrm{d}\widetilde{P}_{j+1}}{\mathrm{d}r} = 1+\mathrm{O}\big(|\varepsilon|+|r|\big) \in C^\infty.$$

By the induction assumption, the function $\overline{F}_{n-1}$ has at most $n-1$ positive zeros in r near $r=0$ for $|\varepsilon|$ and $|\bar{b}_0|+\cdots+|\bar{b}_{n-1}|>0$ small. Hence, by Rolle's theorem it follows that $\widetilde{F}_n$ has at most n positive zeros near $r=0$ in r for $|\varepsilon|$ and $|b_0|+\cdots+|b_n|>0$ small. This finishes the proof. □

Observe that in both Theorems 2.6 and 2.7 the parameter δ is required to vary near δ_0. The results are local in this sense. In many cases the functions p and q in (2.43) depend on δ linearly. This enables us to obtain a global result based on Theorems 2.6 and 2.7 as follows.

THEOREM 2.8. *Let* (2.44) *hold, and let the functions* p *and* q *in* (2.43) *be linear in* δ. *Suppose further that for an integer* $k \geqslant 1$

(i) $\operatorname{rank} \dfrac{\partial(b_0, \ldots, b_{k-1})}{\partial(\delta_1, \ldots, \delta_m)} = k,\ m \geqslant k$;

(ii) *Eq.* (2.43) *has a center at the origin as* $b_j(\delta) = 0,\ j = 0, 1, \ldots, k-1$.

Then for any given $N > 0$, *there exist* $\varepsilon_0 > 0$ *and a neighborhood* V *of the origin such that Eq.* (2.43) *has at most* $k-1$ *limit cycles in* V *for* $0 < |\varepsilon| < \varepsilon_0$, $|\delta| \leqslant N$. *Moreover,* $k-1$ *limit cycles can appear in an arbitrary neighborhood of the origin for some* (ε, δ).

PROOF. Since p and q are linear in δ, each coefficient b_j in (2.58) is also linear in δ. Hence, by condition (i) we can suppose

$$\det \frac{\partial(b_0, \ldots, b_{k-1})}{\partial(\delta_1, \ldots, \delta_k)} \neq 0.$$

It follows that the equations $b_j(\delta) = 0,\ j = 0, \ldots, k-1$ have a solution $(\delta_1, \ldots, \delta_k) = \varphi(\delta_{k+1}, \ldots, \delta_m)$.

For $\delta_0^* = (\varphi(0), 0) \in \mathbb{R}^m$, by Theorem 2.7, Eq. (2.43) can have $k-1$ limit cycles near the origin for some (ε, δ) near $(0, \delta_0^*)$. Then we need to prove that $k-1$ is the maximal number of limit cycles. If it is not the case, then there exist $N > 0$ and sequences $\varepsilon_n \to 0$, $\delta^{(n)} \in \mathbb{R}^m$ with $|\delta^{(n)}| \leqslant N$ such that for $(\varepsilon, \delta) = (\varepsilon_n, \delta^{(n)})$ Eq. (2.43) has k limit cycles which approach the origin as $n \to \infty$. We may suppose $\delta^{(n)} \to \delta_0$ for some $\delta_0 \in \mathbb{R}^m$ as $n \to \infty$. First, by Theorem 2.6 we must have $b_j(\delta_0) = 0,\ j = 0, \ldots, k-1$. By our assumption (ii), Eq. (2.43) has a center at the origin for $(\delta_1, \ldots, \delta_k) = \varphi(\delta_{k+1}, \ldots, \delta_m)$. Hence, it follows from Theorem 2.7 that (2.43) has at most $k-1$ limit cycles near the origin for all (ε, δ) sufficiently close to $(0, \delta_0)$. This is a contradiction since Eq. (2.43) has k limit cycles approaching the origin as $(\varepsilon, \delta) = (\varepsilon_n, \delta^{(n)}) \to (0, \delta_0)$. This ends the proof. □

Theorems 2.5 and 2.6 tell us that the coefficients in the expansion of M act like focus values. The following lemma gives a relation between the two groups of the values.

LEMMA 2.9. *Let* (2.44) *and* (2.58) *hold. Then*

$$b_0 = 4\pi v_1^*, \qquad b_j = \frac{4\pi}{K^j}\left[v_{2j+1}^* + \mathrm{O}\left(\left|v_1^*, v_3^*, \ldots, v_{2j-1}^*\right|\right)\right], \quad j = 1, \ldots, k-1,$$

where

$$v_{2j+1}^* = \frac{\partial v_{2j+1}}{\partial \varepsilon}\bigg|_{\varepsilon=0}, \quad j = 0, \ldots, k-1,$$

and v_{2j+1} *is the* j*th Lyapunov constant of* (2.43) *at the origin.*

PROOF. Let $P(r, \varepsilon, \delta)$ denote the Poincaré map of (2.43). Then the point B in (2.45) is given by $(P(a, \varepsilon, \delta), 0)$, where $a = \sqrt{\frac{h}{K}}(1 + \mathrm{O}(\sqrt{h}))$. Note that $P - a = \mathrm{O}(|\varepsilon a|)$. By the mean value theorem we have

$$\begin{aligned} H(B)-H(A) &= H_x\big(a\big(1+\mathrm{O}(\varepsilon)\big),0\big)(P-a) \\ &= 2Ka\big[1+\mathrm{O}\big(|\varepsilon|+|a|\big)\big](P-a). \end{aligned}$$

By (2.39) and (2.40) we can write formally

$$P(a,\varepsilon,\delta)-a = 2\pi\sum_{i\geqslant 0} v_{2i+1}a^{2i+1}P_i,$$

where $P_i = 1+\mathrm{O}(a)$, $i\geqslant 0$. Hence

$$H(B)-H(A) = 4\pi h\sum_{i\geqslant 0}\frac{v_{2i+1}}{K^i}h^i\overline{P}_i,$$

where $\overline{P}_i = 1+\mathrm{O}(|\varepsilon|+\sqrt{h})$, $i\geqslant 0$. Inserting the above into (2.45) and noting $v_{2i+1} = \varepsilon v^*_{2i+1}+\mathrm{O}(\varepsilon^2)$ we obtain

$$4\pi h\sum_{i\geqslant 0}\frac{v^*_{2i+1}}{K^i}h^i\big(1+\mathrm{O}\big(\sqrt{h}\big)\big) = M(h,\delta).$$

Since M is C^∞ at $h=0$ it follows from the above that

$$M(h,\delta) = 4\pi h\bigg[v_1^* + \frac{1}{K}\big(v_3^*+O\big(v_1^*\big)\big)h + \frac{1}{K^2}\big(v_5^*+\mathrm{O}\big(v_1^*,v_3^*\big)\big)h^2+\cdots\bigg].$$

Thus the conclusion follows by comparing with (2.58). This completes the proof. □

EXAMPLE 2.6. Consider

$$\begin{aligned} \dot{x} &= y-\big(\mu_0x^5+\mu_1x^3+\mu_2x\big), \\ \dot{y} &= -x. \end{aligned} \tag{2.65}$$

We claim that there exist $\varepsilon_0>0$ and a neighborhood V of the origin such that (2.65) has at most 2 limit cycles in V for $|\mu_0|+|\mu_1|+|\mu_2|<\varepsilon_0$, and 2 limit cycles can appear.

In fact, by the discussion to (2.41) we know that

$$\begin{aligned} &v_1 = -\frac{1}{2}\mu_2+\mathrm{O}\big(\mu_2^2\big), \qquad v_3 = -\frac{3}{8}\mu_1+\mathrm{O}(\mu_2), \\ &v_5 = -\frac{5}{16}\mu_0+\mathrm{O}\big(|\mu_1|+|\mu_2|\big). \end{aligned}$$

Let $\mu_i=\varepsilon\delta_i$, $\varepsilon=\sqrt{\mu_0^2+\mu_1^2+\mu_2^2}$, $\delta_0^2+\delta_1^2+\delta_2^2=1$. Then by Lemma 2.9 (taking $\mu=1$ and $K=1/2$) we have

$$b_0=-2\pi\delta_2, \qquad b_1=-3\pi\delta_1+\mathrm{O}(\delta_2), \qquad b_2=-5\pi\delta_0+\mathrm{O}\big(|\delta_1|+|\delta_2|\big).$$

For the sake of convenience, we take $\delta_0, \delta_1, \delta_2$ as free parameters with $|\delta_i| \leqslant 1$. Then the conditions of Theorem 2.8 are satisfied with $m = k = 3$. Therefore, the claim follows.

Note that for $\varepsilon = 0$ the origin is a linear center of (2.65). Hence, in this case the function M can be obtained directly by using (2.46).

DEFINITION 2.4. We say that Eq. (2.43) has Hopf cyclicity $k - 1$ at the origin if the conclusion of Theorem 2.8 holds.

Thus, both systems (2.41) and (2.65) have Hopf cyclicity 2 at the origin.

EXAMPLE 2.7. Consider a Liénard system of the form

$$\dot{x} = y - \varepsilon \sum_{i=1}^{2n+1} a_i x^i, \qquad \dot{y} = -x \tag{2.66}$$

where $n \geqslant 1$, $|a_i| \leqslant 1$, and $\varepsilon > 0$ is small. We claim that the system has Hopf cyclicity n at the origin.

In fact, we have $H(x, y) = \frac{1}{2}(x^2 + y^2)$ with $\mu = 1$, $K = \frac{1}{2}$. The curve L_h given by $H(x, y) = h$ has the representation $(x, y) = \sqrt{2h}(\cos t, -\sin t)$. Hence, by (2.46) we have

$$M(h) = -\sum_{j=0}^{n} 2^{j+1} N_j a_{2j+1} h^{j+1},$$

where

$$N_j = \int_0^{2\pi} \cos^{2(j+1)} t \, \mathrm{d}t > 0.$$

Set $\delta = (a_1, a_2, \ldots, a_{2n+1})$, $b_j = -2^{j+1} N_j a_{2j+1}$, $j = 0, \ldots, n$ and $k = n + 1$. Note that by Theorem 2.3(i) Eq. (2.66) has a center at the origin as $b_j = 0$, $j = 0, \ldots, n$. Thus, the claim follows from Theorem 2.8.

In the proof of Corollary 2.2 we have given a way to find limit cycles near the origin. In the case of (2.66), $M(h)$ can have n zeros $h_1 > h_2 > \cdots > h_n > 0$ for some $(a_1, a_3, \ldots, a_{2n+1})$. Then the corresponding function F in (2.55) has n zeros $\sqrt{h_j} + \mathrm{O}(\varepsilon)$, $j = 1, \ldots, n$. It follows that the n limit cycles of (2.66) approach the curves $\frac{1}{2}(x^2 + y^2) = h_j$, $j = 1, \ldots, n$, respectively, as $\varepsilon \to 0$.

In the following we give another way to obtain limit cycles near the origin. For convenience, take $a_{2j} = 0$, $j = 1, \ldots, n$.

First, from Lemma 2.9, we have

$$v_{2j+1} = \varepsilon v_{2j+1}^* + \mathrm{O}(\varepsilon^2), \quad j = 0, \ldots, n,$$

where

$$v_1^* = \frac{1}{4\pi} b_0 = -\frac{N_0}{2\pi} a_1,$$
$$v_{2j+1}^* = \frac{1}{2^{j+2}\pi} b_j + \mathrm{O}\big(|b_0, \ldots, b_{j-1}|\big) = -\frac{N_j}{2\pi} a_{2j+1} + \mathrm{O}\big(|a_1, a_3, \ldots, a_{2j-1}|\big).$$

On the other hand, for $\widetilde{V} = x^2 + y^2$ we have along the orbit of (2.66)

$$\dot{\widetilde{V}} = -2\varepsilon \sum_{j=0}^{n} a_{2j+1} x^{2j+2}.$$

By Corollary 2.1, it holds that

$$\widetilde{L}_{k+1} = -\frac{\varepsilon}{\pi} a_{2k+1} N_k$$

if $a_{2j+1} = 0$ for $j = 0, \ldots, k-1$. It then follows that

$$v_{2j+1} = -\varepsilon \left[\frac{N_j}{2\pi} a_{2j+1} + \mathrm{O}\big(|a_1, a_3, \ldots, a_{2j-1}|\big) \right] \big(1 + \mathrm{O}(\varepsilon)\big), \quad j = 0, \ldots, n.$$

For $\varepsilon > 0$, let us vary $a_1, a_3, \ldots, a_{2n+1}$ such that

$$a_{2j-1} a_{2j+1} < 0, \qquad 0 < |a_{2j-1}| \ll |a_{2j+1}| \ll \varepsilon, \quad j = 1, \ldots, n$$

which yields

$$v_{2j-1} v_{2j+1} < 0, \qquad 0 < |v_{2j-1}| \ll |v_{2j+1}| \ll \varepsilon^2, \quad j = 1, \ldots, n.$$

Thus, by (2.32), the corresponding function $P(r) - r$ will have n zeros which give n limit cycles of (2.66). Geometrically, the n limit cycles are obtained by changing the stability of the origin n times. Moreover, unlike above they approach the origin as $\varepsilon \to 0$.

REMARK 2.3. There are different ways to prove Lemma 2.3 for the normal form system (2.11). Here the proof is given by following [111]. The conclusions in Lemmas 2.4 and 2.6 are well known. However, the definite relationship between constants a_m, v_{2m+1} and L_m given in the lemmas appears for the first time here. For more discussions on the level set of a Hamiltonian system, the reader can consult [36]. On the bifurcation of limit cycles, Theorems 2.4 and 2.5 are basic tools, while Theorems 2.6–2.8 were recently obtained by [47]. More discussions similar to Lemma 2.9 can be found in [80]. Under condition (2.44), it was proved in [64] that the right-hand side function in (2.45) is C^∞ in $\sqrt{h}$ at $h = 0$, but not C^2 in h at $h = 0$ generally.

As we know, in 1901, Hilbert [73] posed 23 mathematical problems of which the second part of the 16th one is to find the maximal number and relative position of limit cycles of planar polynomial systems. Many works have been done on the study of the problem, especially for quadratic and cubic systems, see [3,4,8,10–13,15,18–34,38–51,54–72,74–110, 113–126,128,129]. A detailed introduction and related literatures can be found in Li [85], Schlomiuk [109] and Ilyashenko [78]. As was showed in Examples 2.4 and 2.7, a typical way to find limit cycles in Hopf bifurcation is to compute focal values and change the stability of the focus by using the values. Certain nice results have been obtained for quadratic and cubic systems. Bautin [5] proved that a focus of a quadratic system has order at most three and that for this system a focus or center can generate at most three limit cycles under perturbations of its coefficients. Then Chen and Wang [12] (by bifurcation method) and Shi [110] (by using Poincaré–Bendixson theorem) separately found a quadratic system having four limit cycles. More and more mathematicians believe that quadratic system have at most four limit cycles. However, up to now the problem is still open. Li and Li [89], Li and Huang [88], Li and Liu [91–93], and Liu, Yang and Jiang [96] found different cubic systems having eleven limit cycles by using Melnikov function method. James and Lloyd [79] gave a cubic system having eight limit cycles in a neighborhood of a focus. Recently, Han, Lin and Yu [60] and Yu and Han [120,121] obtained sufficient conditions for a cubic system to have 10 or 12 limit cycles, respectively (all limit cycles having small amplitude). It seems that the maximal number of limit cycle for cubic system is 12.

For Hopf bifurcation in higher dimension, the reader can see Hale [35], Chow and Hale [16], Chow, Li and Wang [17] and Han [51] etc. For bifurcation of periodic solutions of delay-differential equations, see [7,37,52,53] et al.

3. Perturbations of Hamiltonian systems

3.1. *General theory*

In this section we will study a C^∞ system of the form

$$\dot{x} = H_y + \varepsilon p(x, y, \varepsilon, \delta), \qquad \dot{y} = -H_x + \varepsilon q(x, y, \varepsilon, \delta), \tag{3.1}$$

where $H(x, y), p(x, y, \varepsilon, \delta), q(x, y, \varepsilon, \delta)$ are C^∞ functions, $\varepsilon \geqslant 0$ is small and $\delta \in D \subset \mathbb{R}^m$ is a vector parameter with D compact. For $\varepsilon = 0$ (3.1) becomes

$$\dot{x} = H_y, \qquad \dot{y} = -H_x \tag{3.2}$$

which is Hamiltonian. Hence, Eq. (3.1) is called a near-Hamiltonian system.

For Eq. (3.2) we suppose there exist a family of periodic orbits given by

$$L_h\colon\ H(x, y) = h, \quad h \in (\alpha, \beta)$$

such that L_h approaches an elementary center point, denoted by L_α, as $h \to \alpha$, and an invariant curve, denoted by L_β, as $h \to \beta$. Without loss of generality, we can assume that each L_h is oriented clockwise.

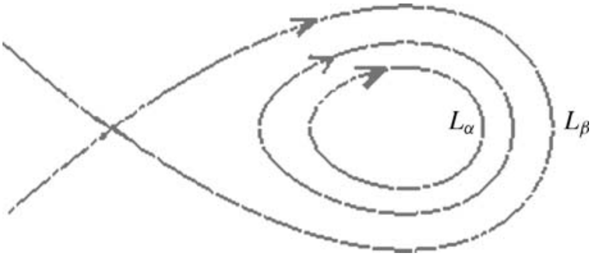


Fig. 9. The phase portrait of (3.2) with a L_β homoclinic loop.

If L_β is bounded, it usually is a homoclinic loop consisting of a saddle and a connection or a heteroclinic loop consisting of at least two saddles and connections between them. In the homoclinic case, the phase portrait of the family $\{L_h\colon \alpha \leqslant h \leqslant \beta\}$ is given in Fig. 9.

Introduce an open set G as follows:

$$G = \bigcup_{\alpha<h<\beta} L_h.$$

Our main purpose is to study the number of limit cycles of Eq. (3.1) in a neighborhood of the closure $\overline{G}$ of G for $\varepsilon > 0$ small and $\delta \in D$.

Note that if Eq. (3.1) has a limit cycle $L(\varepsilon, \delta)$ for $\varepsilon > 0$ small and $\delta \in D_0 \subset D$, then the limit of the cycle as $\varepsilon \to 0$ is either the center L_α, a periodic orbit L_h with $h \in (\alpha, \beta)$ or the boundary L_β. That is,

$$\lim_{\varepsilon\to 0} L(\varepsilon, \delta) = L_h, \qquad h \in [\alpha, \beta].$$

In this case, we say that the limit cycle $L(\varepsilon, \delta)$ is generated from L_h. Thus, in order to study the number of limit cycles, we first need to study the number of limit cycles generated from each L_h.

For the purpose, similar to Definition 2.4 we first introduce a notation below.

DEFINITION 3.1. We say that Eq. (3.1) has *cyclicity* k at a given L_h, $h \in [\alpha, \beta]$, if there exist $\varepsilon_0 > 0$ and a neighborhood V of L_h such that Eq. (3.1) has at most k limit cycles in V for $0 < \varepsilon < \varepsilon_0$, $\delta \in D$ and if k limit cycles can appear in an arbitrary neighborhood of L_h for some (ε, δ) with $\varepsilon > 0$ sufficiently small. More specifically, k is said to be *Hopf* (*Poincaré or homoclinic*) *cyclicity* when $h = \alpha$ ($h \in (\alpha, \beta)$ or $h = \beta$ with L_β being a homoclinic loop).

In the last section we gave some method to find Hopf cyclicity. This section concerns with global bifurcations of limit cycles and presents further methods to find Hopf, Poincaré and homoclinic cyclicity.

Take $h = h_0 \in (\alpha, \beta)$ and $A(h_0) \in L_{h_0}$. Let l be a cross section of Eq. (3.2) passing through $A(h_0)$. Then for h near h_0 the periodic orbit L_h has a unique intersection point with l, denoted by $A(h)$. That is, $A(h) = L_h \cap l$. Consider the positive orbit $\gamma(h, \varepsilon, \delta)$

of Eq. (3.1) starting at $A(h)$. Let $B(h,\varepsilon,\delta)$ denote the first intersection point of the orbit with l. Then similar to (2.45) we have

$$H(B)-H(A)=\varepsilon\bigl[M(h,\delta)+\mathrm{O}(\varepsilon)\bigr]=\varepsilon F(h,\varepsilon,\delta), \tag{3.3}$$

where

$$\begin{aligned} M(h,\delta)&=\oint_{L_h}(H_y q+H_x p)|_{\varepsilon=0}\,\mathrm{d}t\\ &=\oint_{L_h}(q\,\mathrm{d}x-p\,\mathrm{d}y)|_{\varepsilon=0}=\iint_{H\leqslant h}(p_x+q_y)|_{\varepsilon=0}\,\mathrm{d}x\,\mathrm{d}y. \end{aligned} \tag{3.4}$$

The function $F(h,\varepsilon,\delta)$ in (3.3) is called a bifurcation function of Eq. (3.1). It has the following property.

LEMMA 3.1. *For $\varepsilon>0$ small and $\delta\in D$, Eq.* (3.1) *has a limit cycle near L_{h_0}, $h_0\in(\alpha,\beta)$, if and only if the equation $F(h,\varepsilon,\delta)=0$ has a zero in h near h_0.*

PROOF. Since the orbit $\gamma(h,\varepsilon,\delta)$ starting at $A(h)$ is closed if and only if $A=B$, we need only to prove that $A=B$ if and only if $H(A)=H(B)$ by (3.3). It is easy to see that $B=A+\mathrm{O}(\varepsilon)$. Hence, by Taylor formula for $\varepsilon>0$ small we have

$$H(B)-H(A)=\bigl(H_x(A),H_y(A)\bigr)\cdot(B-A)+\mathrm{O}\bigl(|B-A|^2\bigr).$$

Note that the cross section l can be taken to be parallel to the gradient $(H_x(A),H_y(A))$. It follows that

$$H(B)-H(A)=\Bigl[\pm\sqrt{H_x^2(A)+H_y^2(A)}+\mathrm{O}\bigl(|B-A|\bigr)\Bigr]\cdot|B-A|,$$

which gives the desired conclusion. The proof is completed. ☐

As in the situation of Hopf bifurcation, the Melnikov function $M(h,\delta)$ can also be used to determine the cyclicity at a periodic orbit. First, by (3.3) we have

THEOREM 3.1. *Let $h_0\in(\alpha,\beta)$, $\delta_0\in D$.*

(i) *There is no limit cycle near L_{h_0} for $\varepsilon+|\delta-\delta_0|$ small if $M(h_0,\delta_0)\neq 0$.*

(ii) *There is exactly one* (*at least one*, *respectively*) *limit cycle $L(\varepsilon,\delta)$ for $\varepsilon+|\delta-\delta_0|$ small which approaches L_{h_0} as $(\varepsilon,\delta)\to(0,\delta_0)$ if $M(h_0,\delta_0)=0$, $M_h(h_0,\delta_0)\neq 0$* (*h_0 is a zero of $M(h,\delta_0)$ with odd multiplicity*, *respectively*).

(iii) *The cyclicity of Eq.* (3.1) *at L_{h_0} is at most k if for any $\delta\in D$ there exists $0\leqslant j\leqslant k$ such that*

$$M_h^{(j)}(h_0,\delta)\neq 0.$$

PROOF. By (3.3) and Lemma 3.1, the first conclusion is clear. For the second one, the conclusion follows from the implicit function theorem if h_0 is a simple zero of $M(h, \delta_0)$. Let h_0 be a multiple zero of $M(h, \delta_0)$ with odd multiplicity. Then for $\varepsilon_0 > 0$ small we have

$$M(h_0 - \varepsilon_0, \delta_0) \cdot M(h_0 + \varepsilon_0, \delta_0) < 0.$$

Hence, by (3.3) we have

$$F(h_0 - \varepsilon_0, \varepsilon, \delta) \cdot F(h_0 + \varepsilon_0, \varepsilon, \delta) < 0$$

for $0 < \varepsilon < \varepsilon_0$, $|\delta - \delta_0| < \varepsilon_0$ as long as ε_0 is sufficiently small. Thus the function $F(h, \varepsilon, \delta)$ has a zero $h^* \in (h_0 - \varepsilon_0, h_0 + \varepsilon_0)$.

For the third conclusion, if Eq. (3.1) has cyclicity at least $k + 1$ at L_{h_0}, then there exist some $\varepsilon_n \to 0$, $\delta^{(n)} \in D$ such that for $(\varepsilon, \delta) = (\varepsilon_n, \delta^{(n)})$ Eq. (3.1) has at least $k + 1$ limit cycles which approach L_{h_0} as $n \to \infty$. We can suppose $\delta^{(n)} \to \delta_0$ as $n \to \infty$. On the other hand, by our assumption $M(h, \delta_0)$ has h_0 as a zero of multiplicity at most k. By (3.3) and Rolle's theorem for $\varepsilon + |\delta - \delta_0|$ small the function $F(h, \varepsilon, \delta)$ has at most k zeros in h near h_0, which follows that at most k limit cycles exist near L_{h_0} for $\varepsilon + |\delta - \delta_0|$ small. Thus a contradiction appears if $(\varepsilon, \delta) = (\varepsilon_n, \delta^{(n)})$ with n sufficiently large.

The proof is completed. ☐

By the above proof we have immediately

COROLLARY 3.1. *Suppose for some* $\delta_0 \in D$, $M(h, \delta_0)$ *has* k *zeros in* $h \in (\alpha, \beta)$ *with each having odd multiplicity. Then for* $\varepsilon + |\delta - \delta_0|$ *small Eq.* (3.1) *has* k *limit cycles in a compact subset of the open set* G.

COROLLARY 3.2. *If there exist* $h_0 \in (\alpha, \beta), \delta_0 \in D$ *such that for an arbitrary neighborhood of* L_{h_0}, *Eq.* (3.1) *has* k *limit cycles in the neighborhood for some* (ε, δ) *with* $\varepsilon + |\delta - \delta_0|$ *sufficiently small, then* $M(h, \delta_0) = \mathrm{O}(|h - h_0|^k)$.

Let $L(\varepsilon, \delta)$ be the limit cycle appeared in Theorem 3.1. To determine its stability we need to consider the sigh of the integral

$$\varepsilon \oint_{L(\varepsilon,\delta)} (p_x + q_y)\,\mathrm{d}t = \varepsilon \left[\oint_{L_{h_0}} (p_x + q_y)|_{\varepsilon=0}\,\mathrm{d}t + \mathrm{O}(\varepsilon) \right].$$

Obviously, if

$$\sigma(h_0, \delta_0) = \oint_{L_{h_0}} (p_x + q_y)|_{\varepsilon=0,\delta=\delta_0}\,\mathrm{d}t \neq 0$$

then the stability will be determined easily. The following lemma gives a relation between $\sigma(h, \delta)$ and $M_h(h, \delta)$.

LEMMA 3.2. *Suppose L_h is oriented clockwise. Then*

$$M_h(h,\delta) = \pm \oint_{L_h} (p_x + q_y)|_{\varepsilon=0}\, \mathrm{d}t = \pm\sigma(h,\delta),$$

where "+" (respectively "−") is taken when L_h expands (respectively shrinks) with h increasing.

PROOF. Fix $h_0 \in (\alpha, \beta)$. For definiteness, suppose that L_h expands with h increasing. Then for $h > h_0$, applying Green's formula we have

$$M(h,\delta) - M(h_0,\delta) = \iint_{\Delta(h)} (p_x + q_y)|_{\varepsilon=0}\, \mathrm{d}x\, \mathrm{d}y, \tag{3.5}$$

where $\Delta(h)$ denotes the annulus bounded by L_h and L_{h_0}. Let $u(t,h)$ denote a representation of L_h satisfying

$$H\bigl(u(t,h)\bigr) = h, \quad 0 \leqslant t \leqslant T(h),\ h \in (\alpha,\beta).$$

Here $T(h)$ denotes the period of L_h. Consider the integral transformation of variables given by

$$(x,y) = u(t,r), \quad 0 \leqslant t \leqslant T(r),\ h_0 < r < h.$$

Note that

$$\mathrm{D}H(u) \cdot \mathrm{D}_r u = \det \frac{\partial u(t,r)}{\partial(t,r)} = 1.$$

We obtain from (3.5)

$$M(h,\delta) - M(h_0,\delta) = \int_{h_0}^{h} \mathrm{d}r \int_0^{T(r)} (p_x + q_y)\bigl(u(t,r), 0, \delta\bigr)\, \mathrm{d}t.$$

Then differentiating the above in h yields

$$M_h(h,\delta) = \int_0^{T(h)} (p_x + q_y)\bigl(u(t,h), 0, \delta\bigr)\, \mathrm{d}t.$$

This ends the proof. □

EXAMPLE 3.1. Consider van der Pol equation

$$\ddot{x} + \varepsilon\bigl(x^2 - 1\bigr)\dot{x} + x = 0.$$

The equation is equivalent to

$$\dot{x} = y, \qquad \dot{y} = -x - \varepsilon\big(x^2 - 1\big)y. \tag{3.6}$$

For $\varepsilon = 0$, Eq. (3.6) has periodic orbits L_h: $\frac{1}{2}(x^2 + y^2) = h$, $h > 0$. By (3.4) we have

$$M(h) = \iint_{x^2+y^2 \leqslant 2h} \big(1 - x^2\big)\,\mathrm{d}x\,\mathrm{d}y = \pi h(2 - h).$$

The function M has a unique positive zero $h = 2$. By Lemma 3.2 we have

$$\sigma_0 = \oint_{L_2} \big(1 - x^2\big)\,\mathrm{d}t = M'(2) = -2\pi.$$

Hence, by Theorem 3.1, Eq. (3.6) has a unique limit cycle $L(\varepsilon)$ for $\varepsilon > 0$ which is stable, simple and approaches the circle $x^2 + y^2 = 4$ as $\varepsilon \to 0$.

In many cases (polynomial systems for example), the first order Melnikov function has the form

$$M(h, \delta) = \sum_{i=1}^{k} b_i(\delta)\, I_i(h), \quad k \geqslant 2.$$

Let

$$I_1'(\alpha) \neq 0, \qquad I_1(h) \neq 0, \quad \alpha < h < \beta. \tag{3.7}$$

Then we can write

$$M(h, \delta) = I_1(h) \sum_{i=1}^{k} b_i\, J_i(h) = I_1(h) N(h, b), \tag{3.8}$$

where $b = (b_1, \dots, b_n)$, $b_i = b_i(\delta)$, $J_i(h) = I_i(h)/I_1(h)$, $i = 1, 2, \dots, k$.

Introduce the Wronskian of the functions $J_1, J_2, \dots, J_k$

$$W(h) = \begin{vmatrix} J_1(h) & J_2(h) & \dots & J_k(h) \\ J_1'(h) & J_2'(h) & \dots & J_k'(h) \\ \dots & \dots & \dots & \dots \\ J_1^{(k-1)}(h) & J_2^{(k-1)}(h) & \dots & J_k^{(k-1)}(h) \end{vmatrix}. \tag{3.9}$$

Clearly, by a property of determinant we have

$$W'(h) = \begin{vmatrix} J_1(h) & J_2(h) & \dots & J_k(h) \\ J_1'(h) & J_2'(h) & \dots & J_k'(h) \\ \cdots & \cdots & \cdots & \cdots \\ J_1^{(k-2)}(h) & J_2^{(k-2)}(h) & \dots & J_k^{(k-2)}(h) \\ J_1^{(k)}(h) & J_2^{(k)}(h) & \dots & J_k^{(k)}(h) \end{vmatrix}. \tag{3.10}$$

We have

THEOREM 3.2. *Suppose that* (3.7) *and* (3.8) *hold. Let* $|b| > 0$ *for all* $\delta \in D$ *and* $h_0 \in [\alpha, \beta)$.

(i) *If* $W(h_0) \neq 0$, *then Eq.* (3.1) *has cyclicity at most* $k-1$ *at* L_{h_0}.

(ii) *If* $W(h_0) = 0$, $W'(h_0) \neq 0$, *then Eq.* (3.1) *has cyclicity at most* k *at* L_{h_0}.

PROOF. For the first conclusion, let us suppose Eq. (3.1) has cyclicity at least k at L_{h_0}. Then there exist $\varepsilon_n \to \infty$, $\delta_n \in D$ such that for $(\varepsilon, \delta) = (\varepsilon_n, \delta_n)$ Eq. (3.1) has k limit cycles $L_j^{(n)}$ with $L_j^{(n)} \to L_{h_0}$ as $n \to \infty$, $j = 1, 2, \dots, k$. Without loss of generality, we can suppose

$$\delta_n \to \delta_0, \quad b(\delta_n) \to b_0 = (b_{10}, \dots, b_{k0}) \quad \text{as } n \to \infty.$$

Near h_0 we have the following expansion for $N(h, b)$:

$$N(h,b) = \bar{b}_1 + \bar{b}_2(h-h_0) + \dots + \bar{b}_k(h-h_0)^{k-1} + \bar{b}_{k+1}(h-h_0)^k + \cdots, \tag{3.11}$$

where, by (3.8),

$$\bar{b}_{j+1} = \left[\frac{M(h,\delta)}{I_1(h)}\right]^{(j)}_{h=h_0} = \frac{1}{j!}\big[b_1 J_1^{(j)}(h_0) + \dots + b_k J_k^{(j)}(h_0)\big],$$
$$j = 0, 1, \dots, k-1, k. \tag{3.12}$$

It follows that

$$\det \frac{\partial(\bar{b}_1, \dots, \bar{b}_k)}{\partial(b_1, \dots, b_k)} = \frac{W(h_0)}{1!2!\cdots(k-1)!} \neq 0. \tag{3.13}$$

Note that $|b_0| > 0$ by our assumption. By (3.12) and (3.13) we obtain

$$(\bar{b}_1, \dots, \bar{b}_k)\big|_{\delta=\delta_0} = (\bar{b}_{10}, \dots, \bar{b}_{k0}) \neq 0.$$

Thus, an integer l satisfying $1 \leqslant l \leqslant k$ exists such that

$$\bar{b}_{j0} = 0, \quad j = 1, \dots, l-1, \qquad \bar{b}_{l0} \neq 0.$$

Hence, we have by (3.11)

$$N(h, b_0) = \bar{b}_{l0}(h - h_0)^{l-1} + \mathrm{O}\big(|h - h_0|^l\big), \quad \bar{b}_{l0} \neq 0. \tag{3.14}$$

Therefore, by (3.7) and Corollary 2.3 (if $h_0 = \alpha$) or Corollary 3.2 (if $h_0 > \alpha$), there exist $\varepsilon_0 > 0$ and a neighborhood U of L_{h_0} such that for $0 < \varepsilon < \varepsilon_0$, $|\delta - \delta_0| < \varepsilon_0$ Eq. (3.1) has at most $l - 1$ limit cycles in U. This contradicts that Eq. (3.1) has k limit cycles approaching L_{h_0} for $(\varepsilon, \delta) = (\varepsilon_n, \delta_n)$ and $n \to \infty$. This finishes the proof of conclusion (i).

Let $W(h_0) = 0$, $W'(h_0) \neq 0$. If the conclusion (ii) is not true, then, as before, there exists a sequence $(\varepsilon_n, \delta_n)$ approaching $(0, \delta_0)$ such that for $(\varepsilon, \delta) = (\varepsilon_n, \delta_n)$, Eq. (3.1) has $k + 1$ limit cycles approaching L_{h_0} as $n \to \infty$. In this case, formulas (3.11) and (3.12) remain true with

$$\det \frac{\partial(\bar{b}_1, \ldots, \bar{b}_{k-1}, \bar{b}_{k+1})}{\partial(b_1, \ldots, b_k)} = \frac{W'(h_0)}{1!2! \cdots (k-2)!k!} \neq 0, \tag{3.15}$$

and hence

$$\big(\bar{b}_1, \ldots, \bar{b}_{k-1}, \bar{b}_{k+1}\big)\big|_{\delta=\delta_0} = \big(\bar{b}_{10}, \ldots, \bar{b}_{k-1,0}, \bar{b}_{k+1,0}\big) \neq 0.$$

It follows that there exists $1 \leqslant l \leqslant k + 1$ and $l \neq k$ such that (3.14) holds. Thus, a contradiction appears in the same way as above. This ends the proof. ☐

It often happens that $b = 0$ for some $\delta \in D$. In this case, the condition of Theorem 3.2 fails and the following one can apply further.

THEOREM 3.3. *Suppose the following conditions are satisfied.*

(a) *The vector* b *is linear in* δ *with rank* $\frac{\partial b}{\partial \delta} = k$ *and* $b(\delta_0) = 0$ *for some* $\delta_0 \in D$.

(b) *Eq.* (3.1) *is analytic on the closure* $\overline{G}$ *and has a center near* L_α *when* $b = 0$.

(c) *There exists* $h_0 \in [\alpha, \beta)$ *such that* $|W(h_0)| + |W'(h_0)| \neq 0$.

Then

(i) *When* $W(h_0) \neq 0$, *Eq.* (3.1) *has cyclicity* $k - 1$ *at* L_{h_0}.

(ii) *When* $W(h_0) = 0$, $W'(h_0) \neq 0$, *Eq.* (3.1) *has cyclicity* k (*respectively* $k - 1$ *or* k) *at* L_{h_0} *if* $h_0 > \alpha$ (*respectively* $h_0 = \alpha$).

(iii) *For each* $h \in [\alpha, \beta)$, *Eq.* (3.1) *has cyclicity at least* $k - 1$ *at* L_h.

PROOF. Let $W(h_0) \neq 0$ first. By Definition 3.1, it suffices to prove the following two points:

(1) There are at most $k - 1$ limit cycles near L_{h_0} for $\varepsilon > 0$ small and $\delta \in D$.

(2) There can appear $k - 1$ limit cycles in any neighborhood of L_{h_0} for some arbitrarily small $\varepsilon + |\delta - \delta_0|$.

We proceed the proof by contradiction. If the conclusion (1) is not true, then there exists a sequence $(\varepsilon_n, \delta_n) \to (0, \delta^*)$ with $\delta^* \in D$ such that for $(\varepsilon, \delta) = (\varepsilon_n, \delta_n)$ Eq. (3.1) has k

limit cycles which approach L_{h_0} as $n \to \infty$. The proof of Theorem 3.2 implies $b(\delta^*) = 0$. By condition (a) we may assume

$$\det \frac{\partial(b_1, \ldots, b_k)}{\partial(\delta_1, \ldots, \delta_k)} \neq 0.$$

Then for δ near δ^* the linear equation $b = b(\delta)$ has a unique set of solutions $\delta_j = \tilde{\delta}_j(b, \delta_{k+1}, \ldots, \delta_m)$, $j = 1, \ldots, k$. By (3.12) we have

$$\det \frac{\partial(\bar{b}_1, \ldots, \bar{b}_k)}{\partial(\delta_1, \ldots, \delta_k)} = \det \frac{\partial(\bar{b}_1, \ldots, \bar{b}_k)}{\partial(b_1, \ldots, b_k)} \det \frac{\partial(b_1, \ldots, b_k)}{\partial(\delta_1, \ldots, \delta_k)} \neq 0.$$

Further, by condition (b) Eq. (3.1) has a center near L_α when $(\bar{b}_1, \ldots, \bar{b}_k) = 0$. Thus, in the case of $h_0 = \alpha$, by (3.7), (3.8), (3.11) and Theorem 2.7 we know that Eq. (3.1) has Hopf cyclicity $k - 1$ at L_α for $\varepsilon + |\delta - \delta^*|$ small. This contradicts to the existence of k limit cycles near L_α for $(\varepsilon, \delta) = (\varepsilon_n, \delta_n)$. The conclusion (1) above is proved for $h_0 = \alpha$. Since $b(\delta_0) = 0$, using δ_0 instead of δ^*, the above discussion implies that $k - 1$ limit cycles can appear in any neighborhood of L_α for arbitrarily small $\varepsilon + |\delta - \delta_0|$. Then the conclusion (2) follows for $h_0 = \alpha$.

For the case of $h_0 > \alpha$, by (3.3), (3.7), (3.8) and (3.11) we have

$$\begin{aligned} F(h, \varepsilon, \delta) &= I_1(h)\left[\sum_{j=1}^{k} b_j J_j(h) + \mathrm{O}(\varepsilon)\right] \\ &= I_1(h)\left[\sum_{j=1}^{k} \tilde{b}_j (h - h_0)^{j-1} + \mathrm{O}\big(|h - h_0|^k\big)\right], \end{aligned} \tag{3.16}$$

where $\tilde{b}_j = \bar{b}_j + \mathrm{O}(\varepsilon)$, $j = 1, \ldots, k$.

Also, by (3.13) and condition (b) Eq. (3.1) has a center near L_α for $(\bar{b}_1, \ldots, \bar{b}_k) = 0$ and hence for h near h_0,

$$F(h, \varepsilon, \delta) = 0 \quad \text{if } \big(\bar{b}_1, \ldots, \bar{b}_k\big) = 0.$$

Thus (3.16) can be rewritten as

$$F(h, \varepsilon, \delta) = I_1(h) \sum_{j=1}^{k} \tilde{b}_j (h - h_0)^{j-1}\big[1 + P_j(h, \varepsilon, \delta)\big], \tag{3.17}$$

where $\tilde{b}_j = \bar{b}_j + \mathrm{O}(\varepsilon|\bar{b}_1, \ldots, \bar{b}_k|)$, $P_j(h, \varepsilon, \delta) = \mathrm{O}(|h - h_0|^{k-j+1})$, $j = 1, \ldots, k$. Then using the form of (3.17) and similar to the proof of Theorem 2.7 we can prove that F has at most $k - 1$ zeros near $h = h_0$ for $\varepsilon + |\delta - \delta^*|$ small. A contradiction occurs too as before. Also, as before, by using (3.17) F must have $k - 1$ zeros in h near $h = h_0$ for $(\tilde{b}_1, \ldots, \tilde{b}_k)$

satisfying $0 < |\tilde{b}_j| \ll |\tilde{b}_{j+1}| \ll |\tilde{b}_k|$, $\tilde{b}_j \tilde{b}_{j+1} < 0$, $j = 1, \dots, k-1$. Hence, the proof of the first conclusion of the theorem is completed.

To prove the second one, suppose $W(h_0) = 0$, $W'(h_0) \neq 0$ with $h_0 > \alpha$. Then by (3.9) and (3.10) there exist constants $\alpha_0, \dots, \alpha_{k-2}$ such that

$$\left(J_1^{(k-1)}(h_0), \dots, J_k^{(k-1)}(h_0)\right) = \sum_{j=0}^{k-2} \alpha_j \left(J_1^{(j)}(h_0), \dots, J_k^{(j)}(h_0)\right).$$

Hence, by (3.12) we have

$$\bar{b}_k = \sum_{j=1}^{k-1} \bar{\alpha}_j \bar{b}_j, \tag{3.18}$$

for some constants $\bar{\alpha}_1, \dots, \bar{\alpha}_{k-1}$. Therefore, by (3.15), (3.18), similar to (3.17) we obtain

$$\begin{aligned} F(h, \varepsilon, \delta) = I_1(h) \Bigg[&\sum_{j=1, j\neq k}^{k+1} \tilde{b}_j (h-h_0)^{j-1} \left(1 + \mathrm{O}\left(|h-h_0|^{k-j+2}\right)\right) \\ &+ \left(\sum_{j=1}^{k-1} \bar{\alpha}_j \bar{b}_j + \mathrm{O}\left(\varepsilon \left|\bar{b}_1, \dots, \bar{b}_{k-1}, \bar{b}_{k+1}\right|\right) \right) \cdot (h-h_0)^{k-1} \Bigg], \end{aligned}$$

where

$$\tilde{b}_j = \bar{b}_j + \mathrm{O}\left(\varepsilon \left|\bar{b}_1, \dots, \bar{b}_{k-1}, \bar{b}_{k+1}\right|\right), \quad j = 1, \dots, k-1, k+1.$$

Note that

$$\bar{b}_j = \tilde{b}_j + \mathrm{O}\left(\varepsilon \left|\tilde{b}_1, \dots, \tilde{b}_{k-1}, \tilde{b}_{k+1}\right|\right), \quad j = 1, \dots, k-1, k+1.$$

We have further

$$F = I_1(h) \sum_{j=1}^{k+1} \tilde{b}_j (h-h_0)^{j-1} \left(1 + \widetilde{P}_j\right), \tag{3.19}$$

where

$$\begin{aligned} &\widetilde{P}_j = \left(\bar{\alpha}_j + \mathrm{O}(\varepsilon)\right)(h-h_0)^{k-j}, \quad j = 1, \dots, k-1, \\ &\widetilde{P}_k = 0, \qquad \widetilde{P}_{k+1} = \mathrm{O}\left(|h-h_0|\right), \qquad \tilde{b}_k = \mathrm{O}\left(\varepsilon \tilde{b}_{k+1}\right). \end{aligned}$$

Using (3.19), similar to Theorem 2.7, one can prove that Eq. (3.1) has at most k limit cycles near L_{h_0} for ε small and $\delta \in D$.

Furthermore, we can prove that k limit cycles can appear in any neighborhood of L_{h_0}. In fact, let us take

$$h - h_0 = \lambda\varepsilon, \qquad \tilde{b}_j = B_j, \varepsilon^{k+1-j}, \quad j = 1, \ldots, k-1, k+1,$$

where B_j are such constants that the polynomial in λ

$$g(\lambda) = B_1 + B_2\lambda + \cdots + B_{k-1}\lambda^{k-2} + B_{k+1}\lambda^k$$

has k simple zeros $\lambda_j \neq 0$, $j = 1, \ldots, k$. Then (3.19) becomes

$$F = \varepsilon^k I_1(h)\bigl[g(\lambda) + \mathrm{O}(\varepsilon)\bigr],$$

which has k simple zeros $\tilde{\lambda}_j = \lambda_j + \mathrm{O}(\varepsilon)$ in λ, $j = 1, \ldots, k$. Thus F has k simple zeros $h_j = h_0 + \tilde{\lambda}_j\varepsilon$ in h, $j = 1, \ldots, k$.

For $h_0 = \alpha$, we may assume $\alpha = 0$. Then noting (3.18), similar to (2.64) we have

$$\begin{aligned} F = {} & \bar{b}_1\bigl(1 + \mathrm{O}\bigl(|\varepsilon, r|\bigr)\bigr) + \bar{b}_2 r^2\bigl(1 + \mathrm{O}\bigl(|\varepsilon, r|\bigr)\bigr) + \cdots + \bar{b}_{k-1} r^{2(k-2)} \\ & \times \bigl(1 + \mathrm{O}\bigl(|\varepsilon, r|\bigr)\bigr) + \bar{b}_{k+1} r^{2(k-1)}\mathrm{O}(\varepsilon) + \bar{b}_{k+1} r^{2k}\bigl(1 + \mathrm{O}\bigl(|\varepsilon, r|\bigr)\bigr). \end{aligned}$$

It is easy to see that F has at most k zeros in $r > 0$ and $k - 1$ zeros can appear.

For conclusion (iii), let $h \in [\alpha, \beta)$ and U be any neighborhood of L_h. Since W is analytic on $[\alpha, \beta)$ there exists $\bar{h} > h$ such that $W(\bar{h}) \neq 0$ and $L_{\bar{h}} \subset U$. By the conclusion (i) Eq. (3.1) has $k - 1$ limit cycles in U for some (ε, δ). Since U is arbitrary it follows that L_h has cyclicity at least $k - 1$. The proof is completed. ☐

By (3.17) and (3.19) one can prove easily

COROLLARY 3.3. *Suppose the conditions* (a)–(c) *hold with* $h_0 > \alpha$. *If*

(i) $W(h_0) \neq 0$, $1 \leqslant l \leqslant k-1$, *or*

(ii) $W(h_0) = 0$, $W'(h_0) \neq 0$, $1 \leqslant l \leqslant k$, $l \neq k-1$,

then there exists a function $\delta = \delta(\varepsilon)$ *with* $\delta(0) \in D$ *such that for* $\delta = \delta(\varepsilon)$ *and* $\varepsilon > 0$ *small Eq.* (3.1) *has a limit cycle of multiplicity* l *which approaches* L_{h_0} *as* $\varepsilon \to 0$.

REMARK 3.1. Under conditions of Theorem 3.3, the function W is analytic and has only isolated zeros on $[\alpha, \beta)$. Thus for any $\bar{h} \in (\alpha, \beta)$, there exist $\alpha \leqslant h_1 < h_2 < \cdots < h_l < \bar{h}$, $l \geqslant 0$ such that Eq. (3.1) has cyclicity $k - 1$ at L_h for $h \in [\alpha, \bar{h}) - \{h_1, \ldots, h_l\}$.

EXAMPLE 3.2. Consider the Liénard system (2.66) discussed in Example 2.7

$$\dot{x} = y - \varepsilon \sum_{i=1}^{2n+1} a_i x^i, \qquad \dot{y} = -x.$$

As before, suppose that $\varepsilon > 0$ is small and $|a_i| \leqslant 1$ for $i = 1, \ldots, 2n+1$ with $n \geqslant 1$.

We claim that for each $h \geqslant 0$ the above system has cyclicity n at the circle $x^2 + y^2 = 2h$.

In fact, by the discussion in Example 2.7 we have

$$M(h)=\sum_{j=0}^{n} b_{j+1}h^{j+1}, \quad b_{j+1}=-2^{j+1}N_j a_{2j+1},\ N_j>0,\ j=0,\dots,n.$$

Comparing with (3.8) we may take

$$I_1(h)=h, \qquad J_j(h)=h^{j-1}, \quad j=1,\dots,n+1.$$

Hence, by (3.9) we have $W(h)=1$ for all $h\geqslant 0$. Thus the claim follows from Theorem 3.3 and Theorem 2.3(i).

EXAMPLE 3.3. For the system

$$\dot{x}=y-\varepsilon\sum_{i=0}^{n} a_{2i+1}x^{2i+1}, \qquad \dot{y}=-x(1-x),$$

we have

$$M(h)=\sum_{j=0}^{n} b_{j+1}h^{j+1}\bigl(1+\mathrm{O}(h)\bigr)$$

for $0<h\ll 1$. It follows $W(0)=1$. Thus, by Remark 3.1 there may exist finitely many values $0<h_1<\cdots<h_l$ in the interval $(0,\frac{1}{6})$ such that for any $h\in[0,\frac{1}{6})-\{h_1,\dots,h_l\}$ (respectively $h\in\{h_1,\dots,h_l\}$) the above system has cyclicity n (respectively at least n) at L_h.

DEFINITION 3.2. Let $U\in\mathbb{R}^2$ be a bounded set and N_0 a positive integer. We call N_0 *the cyclicity of* U for Eq. (3.1) with $\delta\in D$ and $\varepsilon>0$ small if the following are satisfied:

(i) For any given compact set $V\in U$ Eq. (3.1) has at most N_0 limit cycles in V for all $\delta\in D$ and $0<\varepsilon<\varepsilon_0$, where $\varepsilon_0=\varepsilon_0(V)>0$.

(ii) Eq. (3.1) has N_0 limit cycles for some (ε,δ) with $\delta\in D$ whose limits as $\varepsilon\to 0$ are in U.

The cyclicity of the set $G=\bigcup_{h\in(\alpha,\beta)}L_h$ for Eq. (3.1) is also called the cyclicity of the period annulus $\{L_h\colon \alpha<h<\beta\}$.

The following theorem gives a sufficient condition for finding the cyclicity of the open sets G and $G\cup L_\alpha$.

THEOREM 3.4. *Suppose* (3.7) *and* (3.8) *hold. Let the conditions* (a) *and* (b) *of Theorem* 3.3 *are satisfied. Let further, for each* $b\neq 0$ *the function* $N(h,b)$ *has at most* N_0 *zeros* (*taking multiplicity into account*) *in* $h\in[\alpha,\beta)$ (*respectively* $h\in(\alpha,\beta)$) *and for some* $b\neq 0$ *it has* N_0 *simple zeros in* $h\in(\alpha,\beta)$. *Then the cyclicity of the set* $G\cup L_\alpha$ (*respectively* G) *for Eq.* (3.1) *is* N_0 *as* G *is bounded.*

PROOF. For the sake of simplicity, we suppose that the singular point of Eq. (3.1) near L_α is at the origin with $\alpha = 0$. Then we can take a cross section l which is on the positive x-axis with an endpoint at the origin. In this case we have

$$A(h) = L_h \cap l = \bigl(a(h), 0\bigr), \quad B(h, \varepsilon, \delta) = \bigl(b^*(h, \varepsilon, \delta), 0\bigr)$$

for $h > 0$ small. By (2.4), the function b^* is analytic in a. Thus, $b^* = \mathrm{O}(a)$ and $H(B) - H(A) = \mathrm{O}(a^2) = \mathrm{O}(h)$. Then it follows from (3.3) that $F(h, \varepsilon, \delta) = \mathrm{O}(h)$. As before, suppose

$$\det \frac{\partial(b_1, \ldots, b_k)}{\partial(\delta_1, \ldots, \delta_k)} \neq 0.$$

Then we can solve $\delta_i = \tilde{\delta}_i(b, \delta_{k+1}, \ldots, \delta_m)$, $i = 1, \ldots, k$ from $b = b(\delta)$. Hence, for $h \in [0, \beta)$ we have

$$F(h, \varepsilon, \delta) = I_1(h)\bigl[N(h, b) + \mathrm{O}(\varepsilon)\bigr] \equiv F^*(h, \varepsilon, b, \delta_{k+1}, \ldots, \delta_m). \tag{3.20}$$

By our assumption and Corollary 3.1 Eq. (3.1) can have N_0 limit cycles for some (ε, δ). What we need to do is to prove that for any given constants $0 < \lambda < \beta$ (respectively $0 < \mu < \lambda < \beta$) there exists $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$, $\delta \in D$ the function F has at most N_0 zeros in $h \in (0, \lambda]$ (respectively $[\mu, \lambda]$). If the conclusion is not true, then there exists a sequence $\{(\varepsilon_j, \delta_j)\}$ with $\varepsilon_j \to 0, \delta_j \to \delta_0 \in D$ as $j \to \infty$ such that $F(h, \varepsilon_j, \delta_j)$ has $N_0 + 1$ zeros h_{ij}, $i = 1, \ldots, N_0 + 1$ in $h \in (0, \lambda]$ (respectively $[\mu, \lambda]$). We can suppose $h_{ij} \to h_{i0} \in [0, \lambda]$ (respectively $[\mu, \lambda]$) as $j \to \infty$. Then by (3.20)

$$N(h_{i0}, b_0) = 0, \quad i = 1, \ldots, N_0 + 1,$$

where $b_0 = b(\delta_0)$. Hence, by Corollaries 2.3 and 3.2 the function $N(h, b_0)$ has at least $N_0 + 1$ zeros (multiplicity taken into account) in $[0, \lambda]$ (respectively $[\mu, \lambda]$). This implies $b_0 = 0$.

On the other hand, we have $F^*|_{b=0} = 0$ on $[0, \lambda]$. This yields

$$\begin{aligned} F^* &= \sum_{i=1}^{k} b_i F_i(h, \varepsilon, b, \delta_{k+1}, \ldots, \delta_m) \\ &= I_1(h)\bigl[N(h, b) + \mathrm{O}\bigl(\varepsilon|b|\bigr)\bigr] \\ &= |b| I_1(h)\bigl[N(h, c) + \mathrm{O}(\varepsilon)\bigr] \\ &\equiv |b| I_1(h) G(h, \varepsilon, c, \delta_{k+1}, \ldots, \delta_m), \end{aligned} \tag{3.21}$$

where $c = \frac{b}{|b|} = (c_1, \ldots, c_k)$. Let

$$c^{(j)} = \frac{b(\delta_j)}{|b(\delta_j)|}, \quad j \geqslant 1.$$

Then $b(\delta_j) \to b(\delta_0) = 0$, $j \to \infty$. We can assume $c^{(j)} \to c_0$ with $|c_0| = 1$. Note that

$$F(h_{ij}, \varepsilon_j, \delta_j) = 0, \quad j \geqslant 1,\ i = 1, \dots, N_0 + 1.$$

It follows from (3.20) and (3.21) that $N(h_{i0}, c_0) = 0$, $i = 1, \dots, N_0 + 1$, which contradicts to our assumption. The proof is completed. ☐

REMARK 3.2. Let the conditions of Theorem 3.4 be satisfied. If G is unbounded, then there exists a compact set $V_0 \subset G \cup L_\alpha$ (respectively $\subset G$) such that for any compact set V satisfying $V_0 \subset V \subset G \cup L_\alpha$ (respectively $V_0 \subset V \subset G$) Eq. (3.1) has cyclicity N_0 on V.

EXAMPLE 3.4. Consider Eq. (2.66) again. Since

$$N(h, b) = \sum_{j=0}^{n} b_{j+1} h^j,$$

in this case, by the discussion in Example 3.2 for any compact set V containing the origin its cyclicity is n.

An open problem for Eq. (3.1) is: What is the maximal number of limit cycles on the plane? The above example suggests that the answer be n. The most difficult part is to study the number of limit cycles which disappear into infinity as $\varepsilon \to 0$.

3.2. *Existence of 2 and 3 limit cycles*

By (3.8), the function N can be written in the form

$$N(h, b) = b_1 - P(h, b_2, \dots, b_k). \tag{3.22}$$

The function P here is called a detection function and its graph on the (h, b_1) plane a detection curve. For a given system, an interesting problem is to find a point $(b_{10}, b_{20}, \dots, b_{k0})$ such that the line $b_1 = b_{10}$ and the curve $b_1 = P(h, b_{20}, \dots, b_{k0})$ have as many intersection points as possible.

The simplest case is that P is monotonic in $h \in (\alpha, \beta)$. Some sufficient conditions for some special systems on the monotonicity of P were obtained by Li and Zhang [84], and Han [49]. Obviously, if $P(\alpha, b_2, \dots, b_k) \neq P(\beta, b_2, \dots, b_k)$ for some $(b_2, \dots, b_k)$ then Eq. (3.1) can have a limit cycle.

Next, we give some conditions for Eq. (3.1) to have 2 or 3 limit cycles. We will suppose $\beta < \infty$ and L_β is a homoclinic loop with a hyperbolic saddle S on it. First, we prove

LEMMA 3.3. *Let $\phi(x, y)$ be a C^∞ function with $\phi(S) = 0$. Then along the orbit L_h of Eq.* (3.2) *the limit* $\lim_{h\to\beta} \oint_{L_h} \phi(x, y)\, \mathrm{d}t = \oint_{L_\beta} \phi(x, y)\, \mathrm{d}t$ *exists finitely.*

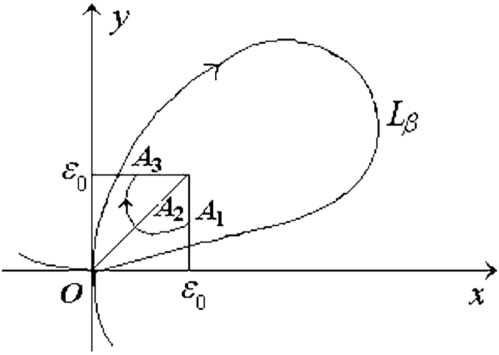


Fig. 10.

PROOF. Without loss of generality, we may suppose the saddle S is at the origin and the Eq. (3.2) has the form

$$\dot{x} = \lambda x + f(x, y), \quad \dot{y} = -\lambda y + g(x, y) \tag{3.23}$$

where $f,\ g = \mathrm{O}(|x, y|^2)$, which implies that $H(x, y) = \lambda xy + \mathrm{O}(|x, y|^3)$, $\lambda \neq 0$ with $\beta = 0$. For definiteness, assume $\lambda < 0$ and L_h locates in the first quadrant. See Fig. 10.

For $\varepsilon_0 > 0$ small take points $A_1 \in L_h \cap \{x = \varepsilon_0\}$, $A_2 \in L_h \cap \{x = y\}$, $A_3 \in L_h \cap \{y = \varepsilon_0\}$. Then the coordinate (a, a) of A_2 satisfies $a = a(h) = \sqrt{\frac{h}{\lambda}} + \mathrm{O}(h)$. Let $y = xu$. Then

$$H(x, y) = x^2\big[\lambda u + x\phi_0(x, u)\big],$$

where, by the integral mean value theorem, $\phi_0 \in C^\infty$. The orbit arc $\widehat{A_1 A_2}$ satisfies the equation

$$H(x, y) = h, \quad a(h) \leqslant x \leqslant \varepsilon_0,$$

which is equivalent to

$$V(x, u, v) = 0, \quad a(h) \leqslant x \leqslant \varepsilon_0, \tag{3.24}$$

where $V(x, u, v) = \lambda u + x\phi_0(x, u) - v$, $v = \frac{h}{x^2} \in [\frac{h}{\varepsilon_0^2}, \frac{h}{a^2}]$. Note that

$$V\left(0, \frac{v}{\lambda}, v\right) = 0, \qquad V_u\left(0, \frac{v}{\lambda}, v\right) = \lambda.$$

Taking v as a parameter we can solve uniquely from (3.24)

$$u = u(x, v) = \frac{v}{\lambda} + \mathrm{O}(x) \quad \text{with } u_v = \frac{1}{\lambda} + \mathrm{O}(x).$$

It is clear that along $\widehat{A_1A_2}$

$$\frac{\phi(x,y)}{\lambda x+f(x,y)}=\frac{\phi(x,xu(x,v))}{\lambda x+f(x,xu(x,v))}=\frac{\phi(S)}{\lambda x}+R_0(x)+R_1(x,v)v, \tag{3.25}$$

where $R_0, R_1 \in C^\infty$. Hence, by (3.23) and $\phi(S)=0$

$$\int_{\widehat{A_1A_2}}\phi(x,y)\,\mathrm{d}t=\int_{\varepsilon_0}^{a}(R_0+R_1v)\,\mathrm{d}x=\int_{\varepsilon_0}^{a}R_0\,\mathrm{d}x+\int_{\varepsilon_0}^{a}R_1v\,\mathrm{d}x.$$

Since

$$\left|\int_{\varepsilon_0}^{a}v\,\mathrm{d}x\right|=\left|h\int_{\varepsilon_0}^{a}\frac{\mathrm{d}x}{x^2}\right|=\left|\frac{h}{a}-\frac{h}{\varepsilon_0}\right|=\mathrm{O}\big(|h|^{1/2}\big),$$

it follows that

$$\lim_{h\to 0}\int_{\widehat{A_1A_2}}\phi(x,y)\,\mathrm{d}t=\int_{\varepsilon_0}^{0}R_0\,\mathrm{d}x=\int_{\widehat{A_{10}A_{20}}}\phi(x,y)\,\mathrm{d}t\in\mathbb{R},$$

where $A_{i0}=\lim_{h\to 0}A_i$, $i=1,2,3$.

Similarly,

$$\lim_{h\to 0}\int_{\widehat{A_2A_3}}\phi(x,y)\,\mathrm{d}t=\int_{\widehat{A_{20}A_{30}}}\phi(x,y)\,\mathrm{d}t\in\mathbb{R}.$$

Also, it is obvious that

$$\lim_{h\to 0}\int_{\widehat{A_3A_1}}\phi(x,y)\,\mathrm{d}t=\int_{\widehat{A_{30}A_{10}}}\phi(x,y)\,\mathrm{d}t\in\mathbb{R}.$$

Thus

$$\begin{aligned}\lim_{h\to 0}\oint_{L_h}\phi(x,y)\,\mathrm{d}t&=\lim_{h\to 0}\int_{\widehat{A_1A_2}\cup\widehat{A_2A_3}\cup\widehat{A_3A_1}}\phi(x,y)\,\mathrm{d}t\\&=\int_{\widehat{A_{10}A_{20}}}\phi\,\mathrm{d}t+\int_{\widehat{A_{20}A_{30}}}\phi\,\mathrm{d}t+\int_{\widehat{A_{30}A_{10}}}\phi\,\mathrm{d}t\\&=\oint_{L_\beta}\phi(x,y)\,\mathrm{d}t\in\mathbb{R}.\end{aligned}$$

This finishes the proof. □

By Lemmas 3.2 and 3.3 we have immediately

COROLLARY 3.4. *Let L_h be oriented clockwise and expand with h increasing. Then for any C^∞ functions $\bar{p}(x,y)$ and $\bar{q}(x,y)$ the Abelian*

$$I(h)=\oint_{L_h}\bar{q}\,\mathrm{d}x-\bar{p}\,\mathrm{d}y$$

has the derivative

$$I'(h)=\oint_{L_h}(\bar{p}_x+\bar{q}_y)\,\mathrm{d}t,$$

and

$$I'(\alpha)=(\bar{p}_x+\bar{q}_y)(L_\alpha)T_\alpha,\qquad I'(h)=c_0T_h+c_1+c_2(h),$$

where T_h denotes the period of L_h, $T_\alpha=\lim_{h\to\alpha}T_h$, $c_0=(\bar{p}_x+\bar{q}_y)(S)$, $c_1=\oint_{L_\beta}[\bar{p}_x+\bar{q}_y-c_0]\,\mathrm{d}t$, $\lim_{h\to\beta}c_2(h)=0$. Further, by (3.25) *it is easy to see that*

$$\lim_{h\to\beta}\frac{T_h}{\ln|h-\beta|}=p_0<0.\tag{3.26}$$

The following lemma gives formulas for computing the value of the function P and its derivative at $h=\alpha$.

LEMMA 3.4. *Let* (3.7), (3.8) *and* (3.22) *hold. Suppose*

$$(p_x+q_y)(L_\alpha,0,\delta)=\sum_{j=1}^{k}d_{j0}b_j,\quad d_{10}\neq 0,\tag{3.27}$$

and

$$v_3^*=\sum_{j=2}^{k}v_{3j}b_j\quad \text{as } b_1=-\frac{1}{d_{10}}\sum_{j=2}^{k}d_{j0}b_j,\tag{3.28}$$

where $v_3^*=\frac{\partial v_3}{\partial\varepsilon}|_{\varepsilon=0}$ and v_3 is the first focus value of Eq. (3.1) *at the focus near L_α obtained by using Theorem* 2.1, *and d_{j0} $(1\leqslant j\leqslant k)$ and $v_{3j}(2\leqslant j\leqslant k)$ are constants. Then*

$$P(\alpha,b_2,\dots,b_k)=-\frac{1}{d_{10}}\sum_{j=2}^{k}d_{j0}b_j,$$

$$P_h'(\alpha,b_2,\dots,b_k)=-\frac{4\pi}{Kb_{10}}\sum_{j=2}^{k}v_{3j}b_j,$$

where $K>0$, and $b_{10}=I_1'(\alpha)$ which can be obtained by Corollary 3.4.

PROOF. Let

$$I_j(h) = b_{j0}(h-\alpha) + b_{j1}(h-\alpha)^2 + \mathrm{O}\big(|h-\alpha|^3\big), \quad j = 1, \ldots, k.$$

Then

$$M(h,\delta) = b_0^*(h-\alpha) + b_1^*(h-\alpha)^2 + \mathrm{O}\big(|h-\alpha|^3\big),$$

where $b_i^* = \sum_{j=1}^k b_{ji} b_j$, $i = 0, 1$. By Lemma 2.9 and Corollary 3.4 we have

$$b_0^* = 0 \quad \text{if and only if} \quad (p_x + q_y)(L_\alpha, 0, \delta) = 0,$$

and

$$v_3^* = \frac{K}{4\pi} b_1^* \quad \text{when } b_0^* = 0.$$

Therefore, by (3.27) $b_0^* = 0$ implies

$$b_1 = -\frac{1}{b_{10}} \sum_{j=2}^k b_{j0} b_j = -\frac{1}{d_{10}} \sum_{j=2}^k d_{j0} b_j,$$

and hence

$$v_3^* = \frac{K}{4\pi}\left[b_{11} b_1 + \sum_{j=2}^k b_{j1} b_j \right] = \frac{K}{4\pi b_{10}} \sum_{j=2}^k (b_{10} b_{j1} - b_{11} b_{j0}) b_j,$$

when $b_0^* = 0$. Hence, by (3.22) and (3.8)

$$\begin{aligned}
P(h, b_2, \ldots, b_k) &= -\sum_{j=2}^k \frac{b_j I_j(h)}{I_1(h)} \\
&= -\sum_{j=2}^k \frac{b_j[b_{j0} + b_{j1}(h-\alpha) + \mathrm{O}(|h-\alpha|^2)]}{b_{10} + b_{11}(h-\alpha) + \mathrm{O}(|h-\alpha|^2)} \\
&= -\frac{1}{b_{10}^2} \sum_{j=2}^k b_j \big[b_{10} b_{j0} + (b_{10} b_{j1} - b_{11} b_{j0})(h-\alpha) \\
&\quad + \mathrm{O}\big(|h-\alpha|^2\big) \big] \\
&= -\frac{1}{b_{10}} \sum_{j=2}^k b_{j0} b_j - \frac{4\pi}{K b_{10}} v_3^*(h-\alpha) + \mathrm{O}\big(|h-\alpha|^2\big)
\end{aligned}$$

$$= -\frac{1}{d_{10}}\sum_{j=2}^{k} d_{j0}b_j - \frac{4\pi}{Kb_{10}}\sum_{j=2}^{k} v_{3j}b_j(h-\alpha) + \mathrm{O}\big(|h-\alpha|^2\big).$$

Then the conclusion follows from (3.28) and the above easily. The proof is ended. □

Now we can give a condition for the existence of 2 limit cycles.

THEOREM 3.5. *Let L_h be oriented clockwise and expand with h increasing. Suppose* (3.7), (3.8), (3.22), (3,27) *and* (3.28) *are satisfied. If there exists $\delta_0 \in D$ such that*

$$\text{(i)}\quad (p_x + q_y)(S, 0, \delta_0) = \sum_{j=1}^{k} d_j b_j(\delta_0),$$

$$\sigma(\delta_0) = \sum_{j=2}^{k}\left(d_j - \frac{\mathrm{d}_1 I_j(\beta)}{I_1(\beta)}\right)b_j(\delta_0) \neq 0;$$

$$\text{(ii)}\quad \sigma(\delta_0)\sum_{j=2}^{k} v_{3j}b_j(\delta_0) < 0;$$

$$\text{(iii)}\quad b_1'(\delta_0) \neq 0 \quad \textit{and}$$

$$b_1(\delta_0) = \begin{cases} \min\big\{P\big(\alpha, b_2(\delta_0), \dots, b_k(\delta_0)\big), P\big(\beta, b_2(\delta_0), \dots, b_k(\delta_0)\big)\big\} \\ \quad \textit{as } \sigma(\delta_0)I_1(\beta) < 0, \\ \max\big\{P\big(\alpha, b_2(\delta_0), \dots, b_k(\delta_0)\big), P\big(\beta, b_2(\delta_0), \dots, b_k(\delta_0)\big)\big\} \\ \quad \textit{as } \sigma(\delta_0)I_1(\beta) > 0 \end{cases}$$

then Eq. (3.1) *has at least two limit cycles for some* (ε, δ) *near* $(0, \delta_0)$.

PROOF. By (3.8) and (3.22) we have $M = I_1(b_1 - P)$. Hence $M' = I_1'(b_1 - P) - I_1 P'$ and

$$\frac{P'}{T_h} = \frac{I_1'}{T_h}\cdot\frac{b_1 - P}{I_1} - \frac{M'}{T_h}\cdot\frac{1}{I_1}.$$

Take $b_1 = P(\beta, b_2, \dots, b_k)$ and apply Corollary 3.4 to function I_1 and M so that

$$\begin{aligned}\lim_{h\to\beta}\frac{P'}{T_h} &= -\frac{1}{I_1(\beta)}\lim_{h\to\beta}\frac{M'}{T_h}\bigg|_{b_1 = P(\beta, b_2, \dots, b_k)} \\ &= -\frac{1}{I_1(\beta)}(p_x + q_y)(S, 0, \delta)|_{b_1 = P(\beta, b_2, \dots, b_k)} \\ &= -\frac{1}{I_1(\beta)}\sigma(\delta).\end{aligned} \tag{3.29}$$

Then noting $I_1'(\alpha)I_1(\beta) > 0$, by Lemma 3.4 we know that for h near β the product $P_h'(\alpha, b_2, \ldots, b_k)P_h'(h, b_2, \ldots, b_k)$ has the same sign as $\sigma(\delta)\sum_{j=2}^{k} v_{3j}b_j$. Then it follows from (ii) that for $\delta = \delta_0$ the function $P(h, b_2, \ldots, b_k)$ has a minimum or maximum in the interval (α, β). Therefore, by (3.29) and (iii) for some δ near δ_0 the function $b_1 - P(h, b_2, \ldots, b_k)$ has two zeros with odd multiplicity in the interval (α, β). Then the conclusion follows from Theorem 3.1. This ends the proof. □

EXAMPLE 3.5. Consider the system

$$\dot{x} = y(1 - y) - \varepsilon\bigl(x^3 - \delta x\bigr), \qquad \dot{y} = -x, \tag{3.30}$$

where $\varepsilon > 0$ is small and $\delta \in \mathbb{R}$ is bounded. For $\varepsilon = 0$, Eq. (3.30) has a family of periodic orbits giving by

$$L_h\colon\ \frac{1}{2}\bigl(x^2 + y^2\bigr) - \frac{1}{3}y^3 = h, \quad 0 < h < \frac{1}{6},\ y < 1.$$

By (3.4) we have

$$M(h, \delta) = \delta I_1(h) - 3I_2(h) = I_1(h)\bigl(\delta - P(h)\bigr)$$

where

$$P(h) = \frac{3I_2(h)}{I_1(h)},$$
$$I_j(h) = -\frac{1}{2j - 1}\oint_{L_h} x^{2j-1}\,\mathrm{d}y = \iint_{Int.L_h} x^{2j-2}\,\mathrm{d}x\,\mathrm{d}y, \quad j = 1, 2.$$

By Corollary 3.4, $I_1(h) = 2\pi h + \mathrm{O}(h^2)$, $I_2(h) = \mathrm{O}(h^2)$. By Theorem 2.1, we have $v_3^* = -\frac{3}{8}$ as $\delta = 0$. Let

$$\sigma = \frac{3I_2(\frac{1}{6})}{I_1(\frac{1}{6})} = \delta_0.$$

Since $L_{1/6}$ can be represented as $x^2 = \frac{2}{3}(y - 1)^2(y + \frac{1}{2})$, $-\frac{1}{2} \leqslant y \leqslant 1$, it is easy to see that

$$I_1\left(\frac{1}{6}\right) = 2\int_{-1/2}^{1}(1 - y)\sqrt{1 - \frac{2}{3}(1 - y)}\,\mathrm{d}y = \frac{6}{5},$$
$$I_2\left(\frac{1}{6}\right) = \frac{2}{3}\int_{-1/2}^{1}(1 - y)^3\left(1 - \frac{2}{3}(1 - y)\right)^{3/2}\mathrm{d}y = \frac{108}{385}.$$

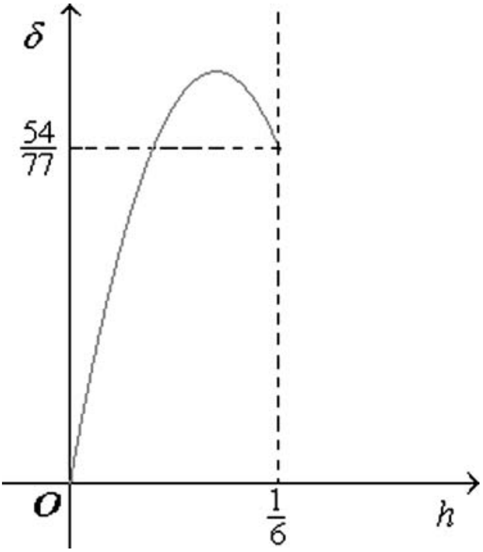


Fig. 11. The graph of $P(h)$ for Eq. (3.27).

Thus, $\delta_0 = \frac{54}{77}$. By Theorem 3.5 for Eq. (3.30) the system has two limit cycles for (ε, δ) near $(0, \frac{54}{77})$ with $\delta > \frac{54}{77}$. The graph of the function $\delta = P(h)$ on the (h, δ) plan is as shown in Fig. 11.

For the existence of three limit cycles, we have:

THEOREM 3.6. *Let L_h be oriented clockwise and expand with h increasing. Suppose* (3.7), (3.8), (3.22), (3.27) *and* (3.28) *are satisfied. Assume further*

(i) *there exists $\delta_0 \in D$ such that*

$$(p_x + q_y)(S, 0, \delta_0) \sum_{j=2}^{k} v_{3j} b_j(\delta_0) > 0, \quad b_1'(\delta_0) \neq 0,$$

$$b_1(\delta_0) = P\big(\alpha, b_2(\delta_0), \dots, b_k(\delta_0)\big) = P\big(\beta, b_2(\delta_0), \dots, b_k(\delta_0)\big);$$

(ii) *there exists δ^* near δ_0 such that*

$$\begin{aligned} &I_1(\beta)(p_x + q_y)(S, 0, \delta_0) \\ &\quad \times \big[P\big(\alpha, b_1(\delta^*), \dots, b_k(\delta^*)\big) - P\big(\beta, b_1(\delta^*), \dots, b_k(\delta^*)\big)\big] > 0. \end{aligned}$$

Then for $\varepsilon > 0$ small and $\delta = \delta^*$ Eq. (3.1) *has* 3 *limits cycles.*

We will give an example to show the way to prove the above theorem.

EXAMPLE 3.6. Consider a system of the form

$$\begin{cases} \dot{x} = y(1-y) - \varepsilon\big(x^3 - \delta_1 x + \delta_2 xy\big), \\ \dot{y} = -x. \end{cases} \tag{3.31}$$

We claim that for $\varepsilon > 0$, $\delta_1 < 0$, $\delta_2 < -54/11$ and $\varepsilon + |\delta_1| + |\delta_2 + 54/11|$ small Eq. (3.31) has 3 limit cycles.

For this system we have

$$M(h) = b_1 I_1(h) + b_2 I_2(h) + b_3 I_3(h),$$

where $b_1 = \delta_1$, $b_2 = -3$, $b_3 = \delta_2$ and $I_1(h)$, $I_2(h)$ are the same as in Example 3.5, and $I_3(h) = \oint_{L_h} xy\,\mathrm{d}y$ with $I_3(\frac{1}{6}) = -\frac{6}{35}$. Let

$$P(h, b_2, b_3) = -\frac{1}{I_1(h)}\big[b_2 I_2(h) + b_3 I_3(h)\big].$$

Also, let $b_{10} = \delta_{10} = 0$, $b_{20} = -3$, $b_{30} = \delta_{20} = 3I_2(\frac{1}{6})/I_3(\frac{1}{6}) = -\frac{54}{11}$. Then

$$b_{10} = P(0, b_{20}, b_{30}) = P\left(\frac{1}{6}, b_{20}, b_{30}\right) = 0.$$

Further, by Lemma 3.4 and (3.29) we have

$$P'_h(0, b_{20}, b_{30}) = \frac{1}{2}(3 + \delta_{20}) < 0,$$

$$\lim_{h \to \frac{1}{6}} \frac{P'_h(h, b_{20}, b_{30})}{T_h} = \frac{\delta_{20}}{I_1(\frac{1}{6})} < 0,$$

$$P\left(\frac{1}{6}, b_2, b_3\right) = -\frac{I_3(\frac{1}{6})}{I_1(\frac{1}{6})}(\delta_2 - \delta_{20}).$$

Denote by Γ^0 and Γ^* respectively the graph of the function $P(h, b_2, b_3)$ on the (b_1, h) plane for $\delta_2 = \delta_{20}$ and $\delta_2 = \delta_2^*$, where $\delta_2^* < \delta_{20}$ with $|\delta_2^* - \delta_{20}|$ small. See Fig. 12.

It is clear that for $\delta_1^* < 0$ and $|\delta_1^*|$ sufficiently small, the line $b_1 = \delta_1^*$ and Γ^*: $b_1 = P(h, -3, \delta_2^*)$ have at least 3 intersection points. Then the conclusion follows easily.

3.3. *Near-Hamiltonian polynomial systems*

We now give an important application of Theorem 3.3 to polynomial systems.

Consider a near-Hamiltonian polynomial system of the form

$$\begin{aligned} \dot{x} &= H_y(x, y, \mu) + \varepsilon p(x, y, a), \\ \dot{y} &= -H_x(x, y, \mu) + \varepsilon q(x, y, a), \end{aligned} \tag{3.32}$$

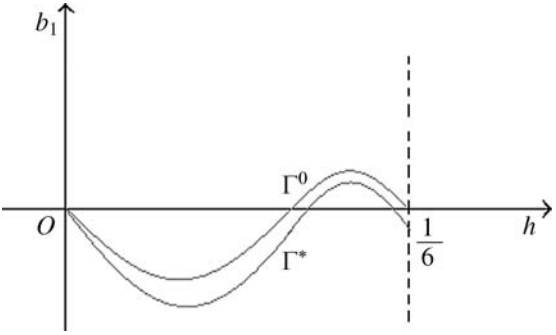


Fig. 12. Curves Γ^0 and Γ^*.

where

$$H(x,y,\mu)=\frac{1}{2}\left(x^2+y^2\right)+\sum_{i+j=3}^{m} h_{ij}x^i y^j,\quad m\geqslant 3,$$

$$p(x,y,a)=\sum_{i+j=1}^{n} a_{ij}x^i y^j,\qquad q(x,y,a)=\sum_{i+j=1}^{n} b_{ij}x^i y^j,$$

and $\mu=\{h_{ij}\}\in\mathbb{R}^r$, $r=\frac{1}{2}(m+1)(m+2)-6$, $a=\{a_{ij},b_{ij}\}\in\mathbb{R}^{n^2+3n}$, $n\geqslant 2$. For each μ, there exists $\beta=\beta(\mu)\in(0,+\infty]$ such that for $h\in(0,\beta)$ the equation $H(x,y,\mu)=h$ defines a smooth closed curve L_h which surrounds a unique singular point (the origin) of the system (3.32) ($\varepsilon=0$). By (3.4) we can write

$$M(h,a)=\sum_{i+j=0}^{n-1} c_{ij}I_{ij}(h)=\sum_{j=1}^{k} b_j I_j(h),$$

where

$$I_{ij}=\iint_{H\leqslant h} x^i y^j\,\mathrm{d}x\,\mathrm{d}y,$$

$$\{I_1,\dots,I_k\}=\{I_{ij},\ 0\leqslant i+j\leqslant n-1\},\quad I_1=I_{00},$$

$$\{b_1,\dots,b_k\}=\{c_{ij},\ 0\leqslant i+j\leqslant n-1\},\quad k=\frac{1}{2}n(n+1).$$

Denote by $W(h,\mu)$ the Wronskian defined by (3.9). Then the equation $W(0,\mu)=0$ defines an $(r-1)$-dimensional surface $\Sigma_{r-1}^{(1)}$ in $\mathbb{R}^r$ and the equations $W(h,\mu)=W_h'(h,\mu)=0$ with $0\leqslant h<\beta(\mu)$ define another $(r-1)$-dimensional surface $\Sigma_{r-1}^{(2)}$. Let

$$B_r=\mathbb{R}^r-\Sigma_{r-1}^{(1)}-\Sigma_{r-1}^{(2)}.$$

Then B_r is an open set in $\mathbb{R}^r$ with boundary $\Sigma_{r-1}^{(1)} \cup \Sigma_{r-1}^{(2)}$. For each $\mu \in B_r$, we have

$$W(0,\mu) \neq 0, \qquad \left|W(h,\mu)\right| + \left|W_h'(h,\mu)\right| \neq 0, \quad h \in [0,\beta).$$

Note that for each $\mu \in B_r$ and $\bar{h} \in (0,\beta)$, the analytic function $W(h,\mu)$ has only finitely many zeros $h_1,\ldots,h_l$ on $(0,\bar{h}]$. Then by Theorem 3.3 we have immediately:

THEOREM 3.7. *There exists an open set B_r in $\mathbb{R}^r$ whose boundary consists of $r-1$ dimensional surfaces such that for each $\mu \in B_r$, and $\bar{h} \in (0,\beta)$ there may exist constants $0 < h_1 < \cdots < h_l$ with $l \geqslant 0$ such that*

(i) *for $h \in [0,\bar{h}] - \{h_1,\ldots,h_l\}$, Eq.* (3.32) *has cyclicity $\frac{1}{2}n(n+1)-1$ at L_h;*

(ii) *for $h \in \{h_1,\ldots,h_l\}$, Eq.* (3.32) *has cyclicity $\frac{1}{2}n(n+1)$ at L_h.*

In particular, for each $\mu \in B_r$, Eq. (3.32) *has Hopf cyclicity $\frac{1}{2}n(n+1)-1$ at the origin.*

Roughly speaking, for almost all $\mu \in \mathbb{R}^r$, Eq. (3.32) has cyclicity $\frac{1}{2}n(n+1)-1$ or $\frac{1}{2}n(n+1)$ at each L_h with $h \in [0,\beta)$.

REMARK 3.3. It is easy to see that in some cases, Theorem 3.7 is still valid if μ is a vector parameter of dimension r with $r < \frac{1}{2}(m+1)(m+2)-6$.

As an application, let us consider quadratic systems. By Ye et al. [117], a quadratic system having a focus or center can be changed into the form

$$\begin{aligned}\dot{x} &= -y + \delta x + lx^2 + mxy + ny^2,\\ \dot{y} &= x(1 + ax + by).\end{aligned}$$

The system is Hamiltonian if and only if $\delta = m = b + 2l = 0$. Then taking $\delta = \varepsilon\delta_1$, $m = \varepsilon m_1$, $b = -2l_0 + \varepsilon b_1$, $l = l_0 + \varepsilon l_1$, $a = a_0 + \varepsilon a_1$, it becomes

$$\begin{aligned}\dot{x} &= H_y + \varepsilon\left(\delta_1 x + l_1 x^2 + m_1 xy + n_1 y^2\right),\\ \dot{y} &= -H_x + \varepsilon\left(a_1 x^2 + b_1 xy\right),\end{aligned} \tag{3.33}$$

where

$$H(x,y) = -\frac{1}{2}\left(x^2+y^2\right) + l_0 x^2 y + \frac{1}{3}n_0 y^3 - \frac{1}{3}a_0 x^3.$$

For (3.33), we have

$$M(h) = \delta_1 I_1(h) + \bar{b}_1 I_2(h) + m_1 I_3(h), \quad \bar{b}_1 = b_1 + 2l_1,$$

where

$$\begin{aligned}I_1(h) &= \iint_{-H\leqslant h} \mathrm{d}x\,\mathrm{d}y, \quad I_2(h) = \iint_{-H\leqslant h} x\,\mathrm{d}x\,\mathrm{d}y,\\ I_3(h) &= \iint_{-H\leqslant h} y\,\mathrm{d}x\,\mathrm{d}y, \quad h \in [0,\beta).\end{aligned}$$

By [117], up to a positive constant, the first four focus values of the origin can be taken as

$$W_0 = \varepsilon V_0, \qquad W_i = \varepsilon V_i + \mathrm{O}\big(\varepsilon^2\big), \quad i = 1, 2, 3,$$

where

$$\begin{aligned} &V_0 = \delta_1, \quad V_1 = m_1(n_0 + l_0) - a_0\bar{b}_1, \\ &V_2 = 5m_1a_0^2\big[(l_0 + n_0)^2(n_0 - 2l_0) - a_0^2n_0\big]. \end{aligned}$$

Thus,

$$\det \frac{\partial(V_0, V_1, V_2)}{\partial(\delta_1, \bar{b}_1, m_1)} = -5a_0^3\big[(l_0 + n_0)^2(n_0 - 2l_0) - a_0^2n_0\big].$$

Hence, by Lemma 2.9 and (3.12), it is easy to see that

$$W(0, a_0, l_0, n_0) = N_0a_0^3\big[(l_0 + n_0)^2(n_0 - 2l_0) - a_0^2n_0\big], \qquad N_0 \neq 0.$$

Then from the discussion before Theorem 3.7 it follows that for any $(a_0, l_0, n_0) \in \mathbb{R}^3$ satisfying

$$a_0\big[(l_0 + n_0)^2(n_0 - 2l_0) - a_0^2n_0\big] \neq 0,$$

the Hopf cyclicity of (3.33) at the origin is 2.

3.4. *Homoclinic bifurcation*

In the rest of the chapter, we introduce a way to find limit cycles in a neighborhood of the homoclinic loop L_β.

As a preliminary, we first discuss the stability of an isolated homoclinic loop. Consider a C^∞ planar system of the form

$$\dot{x} = f(x, y), \qquad \dot{y} = g(x, y). \tag{3.34}$$

Suppose that (3.34) has a homoclinic loop L with a hyperbolic saddle. Assume that the saddle is at the origin. Then on the stability of L we have

LEMMA 3.5. *Let*

$$c_1 = (f_x + g_y)(0), \qquad c_2 = \oint_L (f_x + g_y - c_1)\,\mathrm{d}t. \tag{3.35}$$

Then L is stable (*unstable*) *if* $c_1 < 0$, *or* $c_1 = 0$, $c_2 < 0$ ($c_1 > 0$, *or* $c_1 = 0$, $c_2 > 0$).

For a proof of the lemma, see Han [47,51], Han and Chen [57], and Feng and Qian [27].

Further, let $c_1 = c_2 = 0$. We can assume the functions f and g have the following form:

$$f(x, y) = \lambda x + \mathrm{O}\bigl(|x, y|^2\bigr), \quad g(x, y) = -\lambda y + \mathrm{O}\bigl(|x, y|^2\bigr), \quad \lambda > 0.$$

In this case, let

$$c_3 = \frac{1}{2\lambda}\left[f_{xxy} + g_{xyy} - \frac{1}{\lambda}(f_{xx} f_{xy} - g_{xy} g_{yy}) \right]\Bigg|_{x=y=0}, \tag{3.36}$$

which is called the first saddle value of (3.34) at the origin.

If f and g satisfy

$$f(x, y) = \lambda y + \mathrm{O}\bigl(|x, y|^2\bigr), \qquad g(x, y) = \lambda x + \mathrm{O}\bigl(|x, y|^2\bigr),$$

then, instead of (3.36), let

$$\begin{aligned} c_3 = -\frac{1}{2\lambda}\biggl[& f_{xxx} - f_{xyy} + g_{xxy} - g_{yyy} + \frac{1}{\lambda}\bigl(f_{xy}(f_{yy} - f_{xx}) \\ & + g_{xy}(g_{yy} - g_{xx}) - f_{xx} g_{xx} + f_{yy} g_{yy}\bigr)\biggr]\Bigg|_{x=y=0}. \end{aligned} \tag{3.37}$$

Then we have the following lemma obtained by Han, Hu and Liu [58].

LEMMA 3.6. *Let* $c_1 = c_2 = 0$, *and* $c_3 \neq 0$. *Then* L *is stable if and only if one of the following occurs*:

(i) $c_3 < 0$, *a Poincaré map is well-defined near inside* L *and* L *is oriented anti-clockwise*;
(ii) $c_3 < 0$, *a Poincaré map is well-defined near outside* L *and* L *is oriented clockwise*;
(iii) $c_3 > 0$, *a Poincaré map is well-defined near outside* L *and* L *is oriented anti-clockwise*;
(iv) $c_3 > 0$, *a Poincaré map is well-defined near inside* L *and* L *is oriented clockwise.*

We remark that the formula (3.37) can be obtained from (3.36) by introducing a liner transformation $x = \frac{1}{\sqrt{2}}(u - v)$ and $y = \frac{1}{\sqrt{2}}(u + v)$ where u and v are new variables.

Going back to (3.1) we introduce the following four functions:

$$\begin{aligned} d_0(\delta) &= \oint_{L_\beta} (q\,\mathrm{d}x - p\,\mathrm{d}y)\Bigg|_{\varepsilon=0} = M(\beta, \delta), \\ d_1(\delta) &= (p_x + q_y)(S, 0, \delta), \\ d_2(\delta) &= \oint_{L_\beta} \bigl[p_x + q_y - d_1(\delta)\bigr]\,\mathrm{d}t, \\ d_3(\delta) &= \frac{\partial c_3}{\partial \varepsilon}(\varepsilon, \delta)\Bigg|_{\varepsilon=0}, \end{aligned} \tag{3.38}$$

where c_3 is obtained by (3.36) or (3.37).

The following theorem was obtained by Han [43].

THEOREM 3.8. (i) *Let there exist* $\delta_0 = (\delta_{10}, \dots, \delta_{m0}) \in D$, $m \geqslant 2$ *such that*

$$d_0(\delta_0) = d_1(\delta_0) = 0, \qquad d_2(\delta_0) \neq 0, \qquad \det \frac{\partial(d_0, d_1)}{\partial(\delta_1, \delta_2)}(\delta_0) \neq 0.$$

Then for any $\varepsilon_0 > 0$ *and neighborhood* U *of* δ_0 *there exists an open subset* $V_\varepsilon \subset U$ *for* $0 < \varepsilon < \varepsilon_0$ *such that Eq.* (3.1) *has* 2 *limit cycles near* L_β *for* $a \in V_\varepsilon$.

(ii) *Let there exist* $\delta_0 = (\delta_{10}, \dots, \delta_{m0}) \in D$, $m \geqslant 3$, *such that*

$$d_i(\delta_0) = 0, \quad i = 0, 1, 2, \qquad d_3(\delta_0) \neq 0, \qquad \det \frac{\partial(d_0, d_1, d_2)}{\partial(\delta_1, \delta_2, \delta_3)}(\delta_0) \neq 0.$$

Then for any $\varepsilon_0 > 0$ *and neighborhood* U *of* δ_0 *there exists an open subset* $V_\varepsilon \subset U$ *for* $0 < \varepsilon < \varepsilon_0$ *such that Eq.* (3.1) *has* 3 *limit cycles near* L_β *for* $a \in V_\varepsilon$.

We briefly outline the proof of the first conclusion. For definiteness, suppose L_β is oriented clockwise and the Poincaré map is well defined inside it as before. For $\varepsilon > 0$ small a unique saddle S_ε near S and two separatrices L_ε^u and L_ε^s near L_β exist. The directed distance between L_ε^s and L_ε^u on a cross section is given by (see [14,51,112])

$$d(\varepsilon, \delta) = \varepsilon N\big[d_0(\delta) + \mathrm{O}(\varepsilon)\big], \quad N > 0.$$

By the assumption, we can suppose $d'_{00} = \frac{\partial d_0}{\partial \delta_1}(\delta_0) \neq 0$. The implicit function theorem implies that a unique function $\delta_1 = \varphi_1(\varepsilon, \delta_2, \dots, \delta_m) = \varphi_{10}(\delta_2, \dots, \delta_m) + \mathrm{O}(\varepsilon)$ exists such that for $\varepsilon > 0$ and $|\delta - \delta_0|$ small $d(\varepsilon, \delta) \geqslant 0$ if and only if $d'_{00}[\delta_1 - \varphi_1] \geqslant 0$. Hence, a homoclinic loop L_ε^* appears near L_β if $\delta_1 = \varphi_1$.

Let $\delta_1 = \varphi_1$ and define

$$c_1(\varepsilon, \delta_2, \dots, \delta_m) = \varepsilon(p_x + q_y)(S_\varepsilon, 0, \delta) = \varepsilon\big[c_{10}(\delta_2, \dots, \delta_m) + \mathrm{O}(\varepsilon)\big],$$

where

$$c_{10} = (p_x + q_y)(S, 0, \delta)|_{\delta_1 = \varphi_{10}}.$$

Then our assumption implies that

$$c_{10}(\delta_{20}, \dots, \delta_{m0}) = 0, \qquad d'_{10} = \frac{\partial c_{10}}{\partial \delta_2}(\delta_{20}, \dots, \delta_{m0}) \neq 0.$$

Thus a unique function $\delta_2 = \varphi_2(\varepsilon, \delta_3, \dots, \delta_m) = \varphi_{20}(\delta_3, \dots, \delta_m) + \mathrm{O}(\varepsilon)$ exists such that

$$c_1(\varepsilon, \delta_2, \dots, \delta_m) \geqslant 0 \quad \text{if and only if} \quad d'_{10}[\delta_2 - \varphi_2] \geqslant 0.$$

Let $\delta_1 = \varphi_1, \delta_2 = \varphi_2$ and define

$$c_2(\varepsilon, \delta_3, \ldots, \delta_m) = \varepsilon \oint_{L_\varepsilon^*} (p_x + q_y)\,\mathrm{d}t = \varepsilon\big[c_{20}(\delta_3, \ldots, \delta_m) + \mathrm{O}(\varepsilon)\big],$$

where

$$c_{20}(\delta_3, \ldots, \delta_m) = \oint_{L_\beta} (p_x + q_y)|_{\delta_1=\varphi_{10},\, \delta_2=\varphi_{20}}\,\mathrm{d}t.$$

It is easy to see that

$$c_{20}(\delta_{30}, \ldots, \delta_{m0}) = d_2(\delta_0) \neq 0.$$

Without loss of generality, suppose $d_2(\delta_0) > 0$. Then for $\varepsilon > 0$, $\delta_1 = \varphi_1, \delta_2 = \varphi_2$ and $\varepsilon + |\delta - \delta_0|$ small L_ε^* is unstable by Lemma 3.6. Fix $\varepsilon > 0$ and δ_j near δ_{j0} for $j = 3, \ldots, m$, and vary δ_1 and δ_2 such that

$$\delta_1 = \varphi_1, \qquad 0 < |\delta_2 - \varphi_2| \ll 1, \qquad c_1(\varepsilon, \delta_2, \ldots, \delta_m) < 0.$$

Then, L_ε^* has changed its stability from unstable into stable and therefore an unstable limit cycle has appeared near it at the same time. Next, noting that we have assumed that L_β is oriented clockwise and the Poincaré map is well-defined inside it, we then change δ_1 such that $0 < |\delta_1 - \varphi_1| \ll |\delta_2 - \varphi_2|$, $d(\varepsilon, \delta) < 0$. Clearly, L_ε^* has broken and a stable limit cycle has appeared. Therefore, 2 limit cycles can appear near L_β.

We can summarize the method used above into 3 steps:

1. For $\varepsilon > 0$ fixed, a homoclinic loop L_ε^* appears for δ on a codimension 1 surface in $\mathbb{R}^m$ when $d(\varepsilon, \delta) = 0$.
2. Keep L_ε^* to appear and study its stability and then change the stability of L_ε^* in turn to produce limit cycles.
3. Find a final limit cycle by making the homoclinic loop broken.

This method was first used by Han [43] and then developed by Han and Chen [57] and Han, Hu and Liu [58] to study the number of limit cycles near a double homoclinic loop.

General theorems like Theorem 3.8 on double homoclinic bifurcations can be found in Han and Zhang [71]. Interesting applications of the method to quadratic and cubic systems et al. for the existence limit cycles are given in [65–68,70,72,122–126].

EXAMPLE 3.7. Consider

$$\begin{aligned} \dot{x} &= y - \varepsilon\big(a_1 x + a_2 x^2 + x^4\big), \\ \dot{y} &= x - x^2. \end{aligned} \tag{3.39}$$

For $\varepsilon = 0$, (3.39) has a first integral of the form

$$H(x, y) = \frac{1}{2}\big(y^2 - x^2\big) + \frac{1}{3}x^3,$$

which gives a family of periodic orbits

$$L_h\colon\ H(x,y)=h,\quad -\frac{1}{6}<h<0.$$

The limit of L_h as $h\to 0$ is a homoclinic loop L_0. We can prove that (3.39) has 2 limit cycles near L_0 for some (a_1,a_2).

In fact, for (3.39) we have

$$d_0(a_1,a_2)=-a_1I_{00}-2a_2I_{01}-4I_{03},$$

where

$$I_{0j}=\oint_{L_0}x^jy\,\mathrm{d}x=2\int_0^{3/2}x^{j+1}\sqrt{1-\frac{2}{3}x}\,\mathrm{d}x,\quad j=0,1,3.$$

It easy to get that

$$I_{00}=\frac{6}{5},\qquad I_{01}=\frac{36}{35},\qquad I_{03}=\frac{72}{77}I_{00}.$$

Thus,

$$d_0(a_1,a_2)=-I_{00}\left[a_1+\frac{12}{7}a_2+\frac{288}{77}\right].$$

By (3.38), we have further

$$\begin{aligned}d_1(a_1,a_2)&=-a_1,\\ d_2(a_1,a_2)&=-\oint_{L_0}\left(2a_2x+4x^3\right)\mathrm{d}t\\ &=-2\int_0^{3/2}\frac{2a_2x+4x^3}{\sqrt{1-\frac{2}{3}x}}\,\mathrm{d}x=-6\left[2a_2+\frac{24}{5}\right].\end{aligned}$$

Hence, if $a_1=0, a_2=-\frac{24}{11}$, then

$$d_0=d_1=0,\qquad d_2\neq 0.$$

It follows from Theorem 2.8 that Eq. (3.39) has 2 limit cycles for $\varepsilon>0$ small and some (a_1,a_2) near $(0,-\frac{24}{11})$.

From Han [51] we know Eq. (3.39) has at most 2 limit cycles on the plane for all $\varepsilon>0$ small and (a_1,a_2) bounded.

REMARK 3.4. Study of the number of limit cycles for near-Hamiltonian systems has been a significant and important part of bifurcation theory for decades. Most results can be divided into two aspects. One is to give a lower bound of the number for a given system with parameters. In other words, this aspect mainly concentrates on finding limit cycles as many as possible by choosing suitable parameters. The other is to study the maximal number of limit cycles and give an upper bound of it. In this section we mainly concern with the theory and methods in the first aspect. Theorem 3.1 is well known as Poincaré–Pontryagin–Andronov theorem. The formula in Lemma 3.2 was first obtained by Han [39]. Theorems 3.2, 3.3 and 3.7 are just recently obtained by Han, Chen and Sun [56]. Theorem 3.4 is from Han [51]. Results in Lemma 3.3 were given by Luo, Han and Zhu [100]. Here we present a new and simple proof. Conclusions of Lemma 3.4 were first established in Han and Ye [69]. The homoclinic bifurcation under the condition of Lemma 3.6 was studied in Zhu [130] in detail. Theorems 3.5 and 3.6 are new and obtained by the author. There have been many interesting results in the second aspect on the estimate of an upper bound for the number of limit cycles planar systems. For the theory and methods on this aspect, the reader can consult [13,15–18,23–26,42–44,49,74,76,83,107,109,115–118,127, 128]. For the bifurcation of periodic solutions of higher dimensional systems, see [14,16, 35,37,39,41,51–53,111,112,131].

References

[1] M.U. Akhmet, *Perturbations and Hopf bifurcation of the planar discontinuous dynamical system*, Nonlinear Anal. **60** (2005), 163–178.

[2] A.A. Andronov, E.A. Leontovich, I.I. Gordon and A.G. Maier, *Theory of Bifurcations of Dynamic Systems on a Plane*, Israel Program for Scientific Translations, Jerusalem–London (1973).

[3] V.I. Arnold, *Loss of stability of self-oscillations close to resonance and versal deformations of equivariant vector fields*, Funct. Anal. Appl. **11** (1977), 85–92.

[4] V.I. Arnold, *Geometric Methods in Theory of Ordinary Differential Equations*, Springer, New York (1983).

[5] N.N. Bautin, *On the number of limit cycles which appear with the variation of coefficients from an equilibrium position of focus or center type*, Mat. Sb. (N.S.) **30** (1952), 181–196 (in Russian); Transl. Amer. Math. Soc. **100** (1954), 397–413.

[6] J.D. Bejarano and J. Garcia-Margallo, *The greatest number of limit cycles of the generalized Rayleigh–Linard oscillator*, J. Sound Vibration **221** (1999), 133–142.

[7] P. Bi, M. Han and Y. Wu, *Bifurcation of periodic solutions of delay differential equation with two delays*, J. Math. Anal. Appl. **284** (2003), 548–563.

[8] T.R. Blows and C. Rousseau, *Bifurcation at infinity in polynomial vector fields*, J. Differential Equations **104** (1993), 215–242.

[9] A.D. Bryuno, *Analytical form of differential equations*, Trudy Moskov. Mat. Obsc. **25** (1971), 120–262.

[10] S. Cai and G. Guo, *Saddle values and limit cycles generated by separatrix of quadratic systems*, Proceedings of Asian Mathematical Conference, Hong Kong, 1990, World Sci. Publishing, River Edge, NJ (1992), 25–31.

[11] J. Chavarriga, J. Llibre and J. Sorolla, *Algebraic limit cycles of degree 4 for quadratic systems*, J. Differential Equations **200** (2004), 206–244.

[12] L. Chen and M. Wang, *On relative locations and the number of limit cycles for quadratic systems*, Acta Math. Sinica **22** (1979), 751–758.

[13] L.A. Cherkas, *Bifurcation of limit cycles of quadratic system on the plane*, International Conference on Differential Equations, Vols. 1, 2 (Barcelona, 1991), World Sci. Publishing, River Edge, NJ (1993), 390–392.

[14] C. Chicone, *Ordinary Differential Equations with Applications*, Springer, New York (1999).

[15] C. Chicone and M. Jacobs, *Bifurcation of limit cycles from quadratic isochrones*, J. Differential Equations **91** (1991), 268–326.

[16] S.N. Chow and J. Hale, *Methods of Bifurcation Theory*, Springer, New York (1982).

[17] S.N. Chow, C. Li and D. Wang, *Normal Form and Bifurcations of Planar Vector Fields*, Cambridge University Press, Cambridge (1994).

[18] S.N. Chow, C. Li and Y. Yi, *The cyclicity of period annulus of degenerate quadratic Hamiltonian system with elliptic segment*, Ergodic Theory Dynamical Systems **22** (2002), 1233–1261.

[19] C.J. Christopher and N.G. Lloyd, *Polynomial systems: A lower bound for the Hilbert numbers*, Proc. Roy. Soc. London Ser. A **450** (1995), 219–224.

[20] C.J. Christopher and N.G. Lloyd, *Small-amplitude limit cycles in polynomial Linard systems*, Nonlinear Differential Equations Appl. **3** (1996), 183–190.

[21] C.J. Christopher and S. Lynch, *Small-amplitude limit cycle bifurcations for Linard systems with quadratic or cubic damping or restoring forces*, Nonlinearity **12** (1999), 1099–1112.

[22] B. Coll, A. Gasull and R. Prohens, *Degenerate Hopf bifurcations in discontinuous planar systems*, J. Math. Anal. Appl. **253** (2001), 671–690.

[23] F. Dumortier and C. Li, *Perturbations from an elliptic Hamiltonian of degree four. I. Saddle loop and two saddle cycle*, J. Differential Equations **176** (2001), 114–157.

[24] F. Dumortier and C. Li, *Perturbations from an elliptic Hamiltonian of degree four. II. Cuspidal loop*, J. Differential Equations **175** (2001), 209–243.

[25] F. Dumortier and C. Li, *Perturbation from an elliptic Hamiltonian of degree four. III. Global centre*, J. Differential Equations **188** (2003), 473–511.

[26] F. Dumortier and C. Li, *Perturbation from an elliptic Hamiltonian of degree four. IV. Figure eight-loop*, J. Differential Equations **188** (2003), 512–554.

[27] B. Feng and M. Qian, *The stability of a saddle point separatrix loop and a criterion for its bifurcation limit cycles*, Acta Math. Sinica **28** (1985), 53–70 (in Chinese).

[28] B. Fiedler, S. Liebscher and J.C. Alexander, *Generic Hopf bifurcation from lines of equilibria without parameters. I. Theory*, J. Differential Equations **167** (2000), 16–35.

[29] B. Fiedler, S. Liebscher and J.C. Alexander, *Generic Hopf bifurcation from lines of equilibria without parameters. III. Binary oscillations*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **10** (2000), 1613–1621.

[30] B. Fiedler and S. Liebscher, *Generic Hopf bifurcation from lines of equilibria without parameters. II. Systems of viscous hyperbolic balance laws*, SIAM J. Math. Anal. **31** (2000), 1396–1404 (electronic).

[31] A. Gasull, A. Guillamon and V. Manosa, *An explicit expression of the first Liapunov and period constants with application*, J. Math. Appl. **211** (1997), 190–212.

[32] A. Gasull and J. Torregrosa, *Small-amplitude limit cycles in Linard systems via multiplicity*, J. Differential Equations **159** (1999), 186–211.

[33] A. Gasull and J. Torregrosa, *A new approach to the computation of the Lyapunov constants. The geometry of differential equations and dynamical systems*, Comput. Appl. Math. **20** (2001), 149–177.

[34] H. Giacomini, J. Llibre and M. Viano, *The shape of limit cycles that bifurcate from non-Hamiltonian centers*, Nonlinear Anal. **43** (2001), 837–859.

[35] J.K. Hale, *Ordinary Differential Equations*, second edition. Krieger Publishing Co., Inc., Huntington, New York (1980).

[36] J.K. Hale and H. Kocak, *Dynamics and Bifurcations*, Springer, New York (1991).

[37] J.K. Hale and M. Weedermann, *On perturbations of delay-differential equations with periodic orbits*, J. Differential Equations **197** (2004), 219–246.

[38] M. Han, *Stability of the equator and bifurcation of limit cycles, Dynamical Systems, Tianjin, 1990/1991*, Nankai Ser. Pure Appl. Math. Theoret. Phys., Vol. 4, World Sci. Publishing, River Edge, NJ (1993), 79–85.

[39] M. Han, *Bifurcations of invariant tori and subharmonic solutions for periodic perturbed systems*, Sci. China Ser. A **37** (1994), 1325–1336.

[40] M. Han, *Uniqueness of limit cycles in codimension two bifurcations*, Chinese Sci. Bull. **40** (1995), 1766–1771.

[41] M. Han, *Bifurcation theory of invariant tori of planar periodic perturbed systems*, Sci. China Ser. A **39** (1996), 509–519.

[42] M. Han, *Bifurcation of limit cycles and the cusp of order n*, Acta Math. Sinica (N.S.) **13** (1997), 64–75.

[43] M. Han, *Cyclicity of planar homoclinic loops and quadratic integrable systems*, Sci. China Ser., A **40** (1997), 1247–1258.

[44] M. Han, *Bifurcations of limit cycles from a heteroclinic cycle of Hamiltonian systems*, Chin. Ann. Math. B **19** (1998), 189–196.

[45] M. Han, *Global behavior of limit cycles in rotated vector fields*, J. Differential Equations **151** (1999), 20–35.

[46] M. Han, *Liapunov constants and Hopf cyclicity of Linard systems*, Ann. Differential Equations **15** (1999), 113–126.

[47] M. Han, *On Hopf cyclicity of planar systems*, J. Math. Anal. Appl. **245** (2000), 404–422.

[48] M. Han, *The Hopf cyclicity of Liénard systems*, Appl. Math. Lett. **14** (2001), 183–188.

[49] M. Han, *Existence of at most 1, 2, or 3 zeros of a Melnikov function and limit cycles*, J. Differential Equations **170** (2001), 325–343.

[50] M. Han, *The number and distributions of limit cycles in a cubic system*, Chinese J. Contemp. Math. **23** (2002), 111–122.

[51] M. Han, *Periodic Solutions and Bifurcation Theory of Dynamical Systems*, Science Press, Beijing (2002).

[52] M. Han, *Bifurcations of periodic solutions of delay differential equations*, J. Differential Equations **189** (2003), 396–411.

[53] M. Han and P. Bi, *Existence and bifurcation of periodic solutions of high-dimensional delay differential equations*, Chaos Solitons Fractals **20** (2004), 1027–1036.

[54] M. Han and P. Bi, *Bifurcation of limit cycles from a double homoclinic loop with a rough saddle*, Chinese Ann. Math. Ser. B **25** (2004), 233–242.

[55] M. Han, P. Bi and D. Xiao, *Bifurcation of limit cycles and separatrix loops in singular Liénard systems*, Chaos Solitons Fractals **20** (2004), 529–546.

[56] M. Han, G. Chen and C. Sun, *On the number of limit cycles of near-Hamiltonian polynomial systems*, preprint.

[57] M. Han and J. Chen, *The number of limit cycles bifurcating from a pair of homoclinic loops*, Sci. China Ser. A **30** (2000), 401–414.

[58] M. Han, S. Hu and X. Liu, *On the stability of double homoclinic and heteroclinic cycles*, Nonlinear Anal. **53** (2003), 701–713.

[59] M. Han and K. Jiang, *Normal forms of integrable systems at a resonant saddle*, Ann. Differential Equations **14** (1998), 150–155.

[60] M. Han, Y. Lin and P. Yu, *A study on the existence of limit cycles of a planar system with third-degree polynomials*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **14** (2004), 41–60.

[61] M. Han and D. Luo, *Bifurcation of singular closed orbit with a saddle-node*, Ann. Differential Equations **8** (1992), 455–462.

[62] M. Han, D. Luo and D. Zhu, *Uniqueness of limit cycles bifurcating from a singular closed orbit*, II, Acta Math. Sinica **35** (1992), 541–548 (in Chinese).

[63] M. Han, D. Luo and D. Zhu, *Uniqueness of limit cycles bifurcating from a singular closed orbit*, III, Acta Math. Sinica **35** (1992), 673–684 (in Chinese).

[64] M. Han and Z. Wang, *The differentiability problem for Melnikov functions at a center*, J. Systems Sci. Math. Sci. **17** (1997), 269–274 (in Chinese).

[65] M. Han and Y. Wu, *The stability of double homoclinic loops*, Appl. Math. Lett. **17** (2004), 1291–1298.

[66] M. Han, Y. Wu and P. Bi, *Bifurcation of limit cycles near polycycles with n vertices*, Chaos Solitons Fractals **22** (2004), 383–394.

[67] M. Han, Y. Wu and P. Bi, *A new cubic system having eleven limit cycles*, Disc. & Cont. Dynam. Syst. **12** (2005), 675–686.

[68] M. Han and C. Yang, *On the cyclicity of a 2-polycycle for quadratic systems*, Chaos, Solitons and Fractals, **23** (2005) 1787–1794.

[69] M. Han and Y. Ye, *On the coefficients appearing in the expansion of Melnikov function in homoclinic bifurcations*, Ann. Diff. Eqns. **14** (1998), 156–162.

[70] M. Han, Y. Ye and D. Zhu, *Cyclicity of homoclinic loops and degenerate cubic Hamiltonian*, Sci. China Ser. A **42** (1999), 607–617.

[71] M. Han and T. Zhang, *Some bifurcation methods of finding limit cycles*, Math. Biosci. Engrg. **3** (2006), 67–77.

[72] M. Han and Z. Zhang, *Cyclicity 1 and 2 conditions for a 2-polycycle of integrable systems on the plane*, J. Differential Equations **155** (1999), 245–261.

[73] D. Hilbert, *Mathematical problems* (M. Newton, Transl.), Bull. Amer. Math. Soc. **8** (1902), 437–479.

[74] E. Horozov and I.D. Iliev, *On saddle-loop bifurcations of limit cycles in perturbations of quadratic Hamiltonian systems*, J. Differential Equations **113** (1994), 84–105.

[75] W. Huang and Y. Liu, *A polynomial differential system with nine limit cycles at infinity*, Comput. Math. Appl. **48** (2004), 577–588.

[76] I.D. Iliev, *Inhomogeneous Fuchs equations and the limit cycles in a class of near-integrable quadratic systems*, Proc. Roy. Soc. Edinburgh Sect. A **127** (1997), 1207–1217.

[77] I.D. Iliev and L.M. Perko, *Higher order bifurcations of limit cycles*, J. Differential Equations **154** (1999), 339–363.

[78] Yu. Ilyashenko, *Centennial history of Hilbert's 16th problem*, Bull. Amer. Math. Soc. (N.S.) **39** (2002), 301–354.

[79] E.M. James and N.G. Lloyd, *A cubic system with eight small-amplitude limit cycles*, IMA J. Appl. Math. **47** (1991), 163–171.

[80] Q. Jiang and M. Han, *Melnikov functions and perturbation of a planar Hamiltonian system*, Chinese Ann. Math. Ser. B **20** (1999), 233–246.

[81] P. Joyal and C. Rousseau, *Saddle quantities and applications*, J. Differential Equations **78** (1989), 374–399.

[82] C. Li, W. Li, J. Llibre and Z. Zhang, *Linear estimation of the number of zeros of abelian integrals for some cubic isochronous centers*, J. Differential Equations **180** (2002), 307–333.

[83] C. Li and R. Roussarie, *The cyclicity of the elliptic segment loops of the reversible quadratic Hamiltonian systems under quadratic perturbations*, J. Differential Equations **205** (2004), 488–520.

[84] C. Li and Z. Zhang, *A criterion for determining the monotonicity of the ratio of two Abelian integrals*, J. Differential Equations **124** (1996), 407–424.

[85] J. Li, *Hilbert's 16th problem and bifurcations of planar Polynomial vector fields*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **13** (2003), 47–106.

[86] J. Li, H. Chan and K. Chung, *Investigations of bifurcations of limit cycles in Z_2-equivariant planar vector fields of degree 5*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **12** (2002), 2137–2157.

[87] J. Li, H. Chan and K. Chung, *Bifurcations of limit cycles in a Z_6-equivariant planar vector field of degree 5*, Sci. China Ser. A **45** (2002), 817–826.

[88] J. Li and Q. Huang, *Bifurcations of limit cycles forming compound eyes in the cubic system*, Chinese Ann. Math. Ser. B **8** (1987), 391–403.

[89] J. Li and C. Li, *Planar cubic Hamiltonian systems and distribution of limit cycles of (E3)*, Acta Math. Sinica **28** (1985), 509–521.

[90] J. Li and Y. Lin, *Global bifurcations in a perturbed cubic system with Z_2-symmetry*, Acta Math. Appl. Sinica (English Ser.) **8** (1992), 131–143.

[91] J. Li and Z. Liu, *Bifurcation set and limit cycles forming compound eyes in a perturbed Hamiltonian system*, Publ. Math. **35** (1991), 487–506.

[92] J. Li and Z. Liu, *Bifurcation set and compound eyes in a perturbed cubic Hamiltonian system*, Ordinary and Delay Differential Equations, Pitman Res. Notes Math. Ser., Vol. 272, Longman, England (1991) 116–128.

[93] J. Li and Z. Liu, *Bifurcation set and limit cycles forming compound eyes in a perturbed Hamiltonian system*, Publ. Mat. **35** (1991), 487–506.

[94] C. Liu, G. Chen and C. Li, *Integrability and linearizability of the Lotka–Volterra systems*, J. Differential Equations **198** (2004), 301–320.

[95] Z. Liu, B. Hu and J. Li, *Bifurcation sets and distributions of limit cycles in a Hamiltonian system approaching the principal deformation of a Z_4-field*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **5** (1995), 809–818.

[96] Z. Liu, Z. Yang and T. Jiang, *The same distribution of limit cycles in five perturbed cubic Hamiltonian systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **13** (2003), 243–249.

[97] J. Llibre and G. Rodriguez, *Configurations of limit cycles and planar polynomial vector fields*, J. Differential Equations **198** (2004), 374–380.

[98] N.G. Lloyd and J.M. Pearson, *Symmetry in planar dynamical systems*, J. Symbolic Comput. **33** (2002), 357–366.

[99] N.G. Lloyd, J.M. Pearson, E. Saz and I. Sznt, *A cubic Kolmogorov system with six limit cycles*, Comput. Math. Appl. **44** (2002), 445–455.

[100] D. Luo, M. Han and D. Zhu, *Uniqueness of limit cycles bifurcating from a singular closed orbit. I*, Acta Math. Sinica **35** (1992), 407–417 (in Chinese).

[101] D. Luo, X. Wang, D. Zhu and M. Han, *Bifurcation Theory and Methods of Dynamical Systems*, Adv. Ser. Dyn. Syst., Vol. 15, World Sci. Publishing, Inc., River Edge, NJ (1997).

[102] S. Lynch, *Generalized cubic Linard equations*, Appl. Math. Lett. **12** (1999), 1–6.

[103] S. Lynch and C.J. Christopher, *Limit cycles in highly non-linear differential equations*, J. Sound Vibration **224** (1999), 505–517.

[104] Z. Ma and E. Wang, *The stability of a loop formed the separatrix of a saddle point and the condition to produce a limit cycle*, Chin. Ann. Math. A **4** (1983), 105–110 (in Chinese).

[105] V.K. Melnikov, *On the stability of the center for time periodic perturbations*, Trans. Moscow Math. Soc. **12** (1963), 1–57.

[106] L.M. Perko, *Global families of limit cycles of planar analytic systems*, Trans. Amer. Math. Soc. **322** (1990), 627–656.

[107] R. Roussarie, *On the number of limit cycles which appear by perturbation of separatrix loop of planar vector fields*, Bol. Soc. Brasil. Mat. **17** (1986), 67–101.

[108] C. Rousseau, D. Schlomiuk and P. Thibaudeau, *The centers in the reduced Kukles system*, Nonlinearity **8** (1995), 541–569.

[109] D. Schlomiuk, *Algebraic and geometric aspects of the theory of polynomial vector fields*, Bifurcations and Periodic Orbits of Vector Fields, D. Schlomiuk, ed., NATO ASI Series C, Vol. 408, Kluwer Academic, London (1993), 429–467.

[110] S. Shi, *A concrete example of a quadratic system of the existence of four limit cycles for plane quadratic systems*, Sci. Sinica A **23** (1980), 153–158.

[111] L.P. Shilnikov, A.L. Shilnikov, D.V. Turaev and L.O. Chua, *Methods of Qualitative Theory in Nonlinear Dynamics*, Parts I and II, World Scientific, Singapore (2001).

[112] S. Wiggins, *Introduction to Applied Nonlinear Dynamics and Chaos*, Springer (1990).

[113] Y. Wu and M. Han, *On the study of limit cycles of the generalized Rayleigh-Liénard oscillator*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **14** (2004), 2905–2914.

[114] G. Xiang and M. Han, *Global bifurcation of limit cycles in a family of multiparameter system*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **14** (2004), 3325–3335.

[115] D. Xiao and Z. Zhang, *On the uniqueness and nonexistence of limit cycles for predator-prey systems*, Nonlinearity **16** (2003), 1185–1201.

[116] Y. Ye, *Qualitative Theory of Polynomial Differential Systems*, Shanghai Scientific and Technical Publishers (1995).

[117] Y. Ye et al., *Theory of Limit Cycles*, Trans. Math. Monogr., Vol. 66, Amer. Math. Soc., Providence, RI (1986).

[118] J. Yu and C. Li, *Bifurcation of a class of planar non-Hamiltonian integrable systems with one center and one-homoclinic loop*, J. Math. Anal. Appl. **269** (2002), 227–243.

[119] P. Yu, *Computation of normal forms via a perturbation technique*, J. Sound Vibration **211** (1998), 19–38.

[120] P. Yu and M. Han, *Twelve limit cycles in a cubic order planar system with Z_2-symmetry*, Comm. Pure Appl. Anal. **3** (2004), 515–526.

[121] P. Yu and M. Han, *Small limit cycles bifurcating from fine focus points in cubic order Z_2-equivariant vector fields*, Chaos, Solitons & Fractals **24** (2005), 329–348.

[122] H. Zang, T. Zhang and W. Chen, *Bifurcations of limit cycles from quintic Hamiltonian systems with a double figure eight loop*, J. Complexity **20** (2004), 544–560.

[123] T. Zhang, W. Chen, M. Han and H. Zang, *The abelian integrals of a one-parameter Hamiltonian system under polynomial perturbations*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **14** (2004), 2449–2456.

[124] T. Zhang, M. Han and H. Zang, *Perturbation from an asymmetric cubic Hamiltonian*, J. Math. Anal. Appl. **305** (2005), 617–630.

[125] T. Zhang, M. Han, H. Zang and X. Meng, *Bifurcations of limit cycles for a cubic Hamiltonian system under quartic perturbations*, Chaos Solitons Fractals **22** (2004), 1127–1138.

[126] T. Zhang, H. Zang and M. Han, *Bifurcations of limit cycles in a cubic system*, Chaos Solitons Fractals **20** (2004), 629–638.

[127] Z. Zhang, T. Ding, W. Huang and Z. Dong, *Qualitative Theory of Differential Equations*, Transl. Math. Monogr., Vol. 101, Amer. Math. Soc., Providence, RI (1992).

[128] S. Zhao and P. Zhang, *Uniqueness of limit cycle for the quadratic systems with weak saddle and focus*, Acta Math. Sin. (Engl. Ser.) **20** (2004), 647–652.

[129] D. Zhu, *Saddle values and integrability of quadratic differential systems*, Chinese Ann. Math. Ser. B **8** (1987), 466–478.

[130] D. Zhu, *Homoclinic bifurcation with codimension 3*, Chinese Ann. Math. Ser. B **15** (1994), 205–216.

[131] D. Zhu and Z. Xia, *Bifurcations of heteroclinic loops*, Sci. Chinese Ser. A **8** (1998), 837–848.

This page intentionally left blank

CHAPTER 5

Functional Differential Equations with State-Dependent Delays: Theory and Applications

Ferenc Hartung

Department of Mathematics and Computing, University of Veszprém, P.O. Box 158, H 8201 Veszprém, Hungary
E-mail: hartung@szt.vein.hu

Tibor Krisztin

Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H 6720 Szeged, Hungary
E-mail: krisztin@math.u-szeged.hu

Hans-Otto Walther

Mathematisches Institut, Universität Gießen, Arndtstr. 2, D35392 Gießen, Germany
E-mail: Hans-Otto.Walther@math.uni-giessen.de

Jianhong Wu

Department of Mathematics and Statistics, York University, Toronto, Ontario, Canada M3J 1P3
E-mail: wujh@mathstat.yorku.ca

Contents

1. Introduction . . . 437
2. Models and applications . . . 441
    2.1. A two-body problem of classical electrodynamics . . . 441
    2.2. Position control . . . 444
    2.3. Mechanical models . . . 447
    2.4. Delay adaptation in neural networks and distributed systems . . . 450
    2.5. Disease transmission and threshold phenomena . . . 454
    2.6. Population models and size-dependent interaction . . . 457
    2.7. Cell biology and hematological disorders . . . 462
3. A framework for the initial value problem . . . 464

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 3
Edited by A. Cañada, P. Drábek and A. Fonda

© 2006 Elsevier B.V. All rights reserved

- 3.1. Preliminaries . . . 464
- 3.2. The semiflow on the solution manifold . . . 467
- 3.3. Compactness . . . 470
- 3.4. Linearization at equilibria . . . 472
- 3.5. Local invariant manifolds at stationary points . . . 476
- 3.6. The principle of linearized stability . . . 482

4. Center manifolds . . . 483
- 4.1. Preliminaries . . . 483
- 4.2. The linear inhomogeneous equation . . . 485
- 4.3. Construction of a center manifold . . . 488
- 4.4. Discussion . . . 494

5. Hopf bifurcation . . . 495

6. Differentiability of solutions with respect to parameters . . . 497
- 6.1. Preliminaries . . . 497
- 6.2. Pointwise differentiability with respect to parameters . . . 499
- 6.3. Differentiability with respect to parameters in norm . . . 501

7. Periodic solutions via fixed point theory . . . 504
- 7.1. A general result by continuation . . . 505
- 7.2. Periodic solutions when delay is described by a differential equation . . . 507
- 7.3. Attracting periodic orbits . . . 513

8. Attractors, singular perturbation, small delay, generic convergence, stability and oscillation . . . 515
- 8.1. An attracting disk . . . 515
- 8.2. Limiting profiles for a singular perturbation problem . . . 518
- 8.3. Small delay . . . 521
- 8.4. Generic convergence . . . 522
- 8.5. Stability and oscillation . . . 524

9. Numerical methods . . . 526
- 9.1. Preliminaries . . . 526
- 9.2. Continuous Runge–Kutta methods for ODEs . . . 526
- 9.3. The standard approach to approximation of SD-DDEs . . . 529
- 9.4. Tracking of derivative discontinuities . . . 533
- 9.5. Concluding remarks . . . 535

Acknowledgements . . . 536

References . . . 536

1. Introduction

Studies of single differential equations with state-dependent delayed or advanced arguments go back at least to Poisson [181], but as an object of a broader mathematical activity the area is rather young.

Most work during the past 50 years is devoted to equations with state-dependent delays, which arise as models in applications. A prominent example is the two-body problem of electrodynamics, which remained a mathematical terra incognita until R.D. Driver's work began to appear in the sixties of the last century.

The present survey reports about the more recent work on equations with state-dependent delays, with emphasis on particular models and on the emerging theory from the dynamical systems point of view. Several new results are presented. It will also become obvious that challenging problems remain to be solved.

State-dependent delays were addressed earlier in survey papers on the larger area of functional differential equations, notably by Halanay and Yorke [94] and Myshkis [165]. It is tempting to borrow as a motto from Halanay and Yorke [94] their nice statement

> *This … proved once more, if necessary, that the delay existing at the present time between the moment results are obtained and the moment of publication (as well as the great number of publications which are very difficult to follow) makes it necessary to present from time to time such reports on yet unpublished results and unsolved problems.*

The simplest example of a differential equation with constant delay is the linear equation

$$y'(t) = ay(t-h)$$

with the fixed delay $h > 0$ and a parameter $a \in \mathbb{R}$. Analogues with state-dependent delay like

$$x'(t) = ax\big(t - r\big(x(t)\big)\big),$$

with a bounded delay map $r : \mathbb{R} \to [0, h]$, are already nonlinear in general. Both equations can be written in the same general form

$$x'(t) = f(x_t) \tag{1.0.1}$$

of a delay differential equation. Here $f : U \to \mathbb{R}^n$ is defined on a subset U of the set $(\mathbb{R}^n)^{[-h,0]}$ of all functions $\phi : [-h, 0] \to \mathbb{R}^n$. The *solution segment* $x_t : [-h, 0] \to \mathbb{R}^n$ is given by

$$x_t(s) = x(t+s), \quad -h \leqslant s \leqslant 0.$$

In case of the examples above, we have

$$n = 1, \quad U = \mathbb{R}^{[-h,0]}, \quad f(\phi) = a\phi(-h) \quad \text{and} \quad f(\phi) = a\phi\big(-r\big(\phi(0)\big)\big),$$

respectively.

The maps $f : U \to \mathbb{R}^n$ describing equations with state-dependent delay have in general less smoothness properties than those representing equations with constant delay, and the theory of Retarded Functional Differential Equations (RFDEs) which has been developed since the fifties of the last century (see, e.g., [54,98]) is not applicable to equations with state-dependent delay. This concerns already basic questions of existence, uniqueness and smooth dependence on initial data for the initial value problem (IVP)

$$x'(t) = f(x_t), \quad x_{t_0} = \phi \tag{1.0.2}$$

which for data $\phi \in U$ and for $t_0 \in \mathbb{R}$ is associated with Eq. (1.0.1). When in the sequel we speak of a solution $x : [t_0 - h, T) \to \mathbb{R}^n$, $t_0 < T \leqslant \infty$, to the IVP (1.0.2) it is always understood that x at least satisfies

$$x_t \in U \quad \text{for all } t \in [0, T), \quad x_{t_0} = \phi,$$

and that x is differentiable on (t_0, T) with

$$x'(t) = f(x_t) \quad \text{for all } t \in (0, T).$$

In Section 3 solutions of well-posed IVPs will even be continuously differentiable.

The following Section 2 describes examples of differential equations with state-dependent delays which arise in physics, automatic control, neural networks, infectious diseases, population growth, and cell production. Some of these models differ considerably from others, and most of them do not look simple. Typically the delay is not given explicitly as a function of what seems to be the natural state variable; the delay may be defined implicitly by a functional, integral or differential equation and should often be considered as part of the state variables.

Modelling systems with state-dependent delays seems to require extra care, perhaps because there is not much experience with this phenomenon. We tried to avoid models for which as yet no consistent motivation can be given.

The models in Section 2 indicate the types of equations to which the subsequent sections are confined. Not covered are, for example, nonautonomous systems, delays of the form $r = r(t) = a\,t + b\,x(t)$ or $r(t) = at + bx''(t)$ like in [186–188], and constructions of explicit solutions as in [104]. Also we do not say much about state-dependent delays in control theory. Early work in this area is found in [80–82,160].

Section 3 presents a framework for the study of the IVP (1.0.2). We analyze how state-dependent delays prevent the IVP (1.0.2) from being well-posed on open subsets U of familiar Banach spaces, and see that under mild smoothness hypotheses the IVP (1.0.2) *is* well-posed for data only in a submanifold of finite codimension. This *solution manifold* is given by the equation considered and generalizes the familiar domain of the generator of the semigroup given by a linear autonomous RFDE (as in [54,98]). On the solution manifold the IVP (1.0.2) with $t_0 = 0$ defines a semiflow of continuously differentiable solution operators. This resolves the problem of linearization for equations with state-dependent delay, which had been pointed out earlier by Cooke and Huang [48]. The widely known heuristic technique of *freezing the delay at equilibrium and then linearizing the resulting*

RFDE, often skilfully applied, can now be understood in terms of the true linearization. It is clarified why familiar characteristic equations for linear autonomous RFDEs can be used to analyze local dynamics generated by differential equations with state-dependent delay.

At stationary points the continuously differentiable solution operators have local center, stable, and unstable manifolds. It is shown that these stable and unstable manifolds of maps yield local stable and unstable manifolds also for the semiflow. In particular there is a convenient Principle of Linearized Stability. Center manifolds for the semiflow can not be immediately obtained as just described; they are constructed in Section 4.

Examples to which the results of Section 3 apply are given in [206,208]; for the proof in [208] that hyperbolic stable periodic orbits exist the smoothness results of Section 3 are indispensable. Most other work about which we report was accomplished before the basic theory of Section 3 was developed; some further work was done in parallel. We do not attempt to present these other results in the framework of Section 3. In fact, for many models it remains to be studied whether they fit into this framework or not.

In Section 4 a new result is proved, namely existence of Lipschitz continuous local center manifolds for the semiflow found in Section 3, at stationary points. The more technical proof that these center manifolds actually are continuously differentiable will appear elsewhere. An important open problem is to obtain more smoothness, as it was established for local unstable manifolds [129].

Section 5 is about local Hopf bifurcation, i.e., about the appearance of small periodic orbits close to a stationary point when a parameter in the underlying differential equation is varied and passes a critical value. We state a Hopf bifurcation theorem recently obtained by M. Eichmann [65], which seems to be the first such result for differential equations with state-dependent delays.

Section 6 presents results about differentiability of solutions with respect to parameters and initial data, for a certain class of nonautonomous differential equations with state-dependent delay. The framework is different from the one developed in Section 3. The IVP is considered for Lipschitz continuous initial data, and a quasi-normed space derived from Sobolev spaces turns out to be useful for studying differentiability under relaxed smoothness assumptions, which may be convenient for applications.

Section 7 deals with periodic orbits. The search for periodic solutions has been an important topic in the study of nonlinear autonomous delay differential equations since the sixties of the last century. By now, several methods have been developed in this area. The most general results on existence and global bifurcation employ fixed point theorems and the fixed point index. Others are based on the study of 2-dimensional invariant sets, or on Poincaré–Bendixson type analysis of plane curves which are obtained from evaluations like $x_t \mapsto (x(t), x(t-1))$ along certain solutions. There are local and global Hopf bifurcation theorems for RFDEs, the Fuller index counting periodic orbits is used, and certain symmetric periodic solutions can be obtained from associated ordinary differential systems. Not all approaches mentioned here have been tried for equations with state-dependent delays, which cause complications. We describe results which use the topological concept of ejectivity, and an approach which yields stable periodic orbits. Let us add here that topological tools (fixed point theorems, degree, coincidence degree) have also been employed to prove existence of periodic solutions to nonautonomous, *periodic* differential equations with state-dependent delay [43,138,139,221,225]; related work on nonautonomous equations is found in [50,55].

The topic of Section 8 is limiting behaviour with respect to the independent variable and with respect to parameters. Section 8.1 presents a study of a two-dimensional attractor with periodic orbits, which extends a part of a result for equations with constant delay. Section 8.2 reports about results of Mallet-Paret and Nussbaum on the precise asymptotic shape of periodic solutions in a singular perturbation problem. These results are genuine for equations with state-dependent delay, and involve work on unusual eigenvalue problems for so-called *max-plus operators*. Section 8.3 describes an approach to periodic solutions in case of small delay. Section 8.4 contains an application of the monotone dynamical systems theory which yields a generic convergence result, and Section 8.5 comments on further results about stability and oscillation properties.

Section 9 deals with numerical methods. The study of numerical approximation of solutions to differential equations with state-dependent delay goes back at least to the mid-sixties of the last century, and since then it has been an intensively investigated area. The section begins with a brief summary of continuous Runge–Kutta methods for ordinary differential equations and reports about modifications and extensions which are necessary in case of state-dependent delays.

Let us mention here a few out of many open questions, in addition to those addressed in the subsequent sections. Do equations with state-dependent delay generate semiflows with better smoothness properties than obtained in Section 3, on suitable invariant sets? The results from [129] on unstable manifolds point in this direction. A suspicion is that periodic solutions of equations with state-dependent delay may have stronger stability properties than their counterparts in related equations with constant delay. Can this be made precise and established, for suitable classes of equations? Complicated motion, like chaos, has not yet been rigorously shown to exist for equations with state-dependent delay. Very little is known about the general two-body problem of electrodynamics with two charged particles in a 3-dimension configuration space. Vanishing state-dependent delays, like in a collision in the two-body problem, are also limiting cases of advanced arguments; this indicates that a better understanding of more general differential equations with both delayed and advanced state-dependent arguments may be needed.

It is convenient to end this introduction with notation, for function spaces which occur frequently in the sequel. The Banach spaces of continuous, Lipschitz continuous, and continuously differentiable maps $\phi : [-h, 0] \to \mathbb{R}^n$ are denoted by

$$C = C\bigl([-h,0];\mathbb{R}^n\bigr), \quad C^{0,1} = C^{0,1}\bigl([-h,0];\mathbb{R}^n\bigr), \quad \text{and}$$
$$C^1 = C^1\bigl([-h,0];\mathbb{R}^n\bigr),$$

respectively. The norms on these spaces are given by

$$\|\phi\|_C = \max_{-h\leqslant t\leqslant 0}\bigl|\phi(t)\bigr|, \quad \|\phi\|_{C^{0,1}} = \|\phi\|_C + \sup_{t\neq s}\frac{|\phi(t)-\phi(s)|}{|t-s|},$$
$$\|\phi\|_{C^1} = \|\phi\|_C + \bigl\|\phi'\bigr\|_C,$$

respectively.

If X and Y are real or complex Banach spaces then $L(X, Y)$ denotes the Banach space of continuous linear mappings $T : X \to Y$, with the norm given by $\|T\| = \sup_{\|x\| \leqslant 1} \|Tx\|$.

2. Models and applications

A remark in [206] says that *state-dependent delays arise in various circumstances, but it seems not obvious how to single out a tractable class of equations which contains a large set of examples which are well motivated.* The difficulty of singling out a tractable class of equations to include many interesting models may prove to be an extremely valuable source to stimulate new mathematical techniques and theories. In this section we describe differential equations with state-dependent delay that arise from electrodynamics, automatic and remote control, machine cutting, neural networks, population biology, mathematical epidemiology and economics.

2.1. *A two-body problem of classical electrodynamics*

In Driver [58] (see also [57,59]), a mathematical model for a two-body problem of classical electrodynamics incorporating retarded interaction is proposed and analyzed. He considers the motion for two charged particles moving along the x-axis and substituted the expressions for the field of a moving charge, calculated from the Liénard–Wiechert potential, into the Lorentz–Abraham force law. Radiation reaction is omitted, but time delays are incorporated due to the finite speed of propagation, c, of electrical effects. As a result, the model is a system of delay differential equations involving time delays, which depend on the unknown trajectories. From this model and after some analysis, he obtains a system of six delay-differential equations for the evolution of the states, the velocities and the time delays.

To describe his model, we denote by $x_i(t)$ $(i = 1, 2)$ the positions of the two point charges on the axis in a given inertial system at time t, the time of an observer in that system. Let $v_i(t) = x_i'(t)$ $(i = 1, 2)$ be the velocities of the charges. As mentioned above, we omit radiation reaction but allow an external electric field, $E_{\text{ext}}(t, x)$, in the x-direction, that is assumed to be continuous over some open set D in the (t, x)-plane. Then the equation of motion of charge i is

$$\frac{m_i v_i'(t)}{[1 - v_i^2(t)/c^2]^{3/2}} = q_i E_j\bigl(t, x_i(t)\bigr) + q_i E_{\text{ext}}\bigl(t, x_i(t)\bigr), \quad i, j \in \{1, 2\},\ j \neq i, \tag{2.1.1}$$

where m_i is the rest mass and q_i is the magnitude of charge i, c is the speed of light, and $E_j(t, x)$ is the electric field at (t, x) due to other charge $j \neq i$. The magnetic field of charge j is not involved in this one-dimensional case.

The field at time t and at the point $x_i(t)$ produced by charge j is assumed to be that computed from the Liénard–Wiechert potentials. The expression for this field involves a time lag, $t-\tau_{ji}$, representing the instant at which a light signal would have to leave charge j in order to arrive at $x_i(t)$ at the instant t. Therefore, the delay $\tau_{ji}(t)$ must be a solution of the functional equation

$$\tau_{ji}(t)=\bigl|x_i(t)-x_j\bigl(t-\tau_{ji}(t)\bigr)\bigr|/c. \tag{2.1.2}$$

Clearly, $\tau_{ji}(t)$ cannot be written explicitly.

Because of the occurrence of time delays in the model equation (2.1.1), one needs to specify initial trajectories of the two charges over some appropriate interval $[\alpha, t_0]$. Consider now those initial trajectories and their extensions $(x_1(t), x_2(t))$ defined on some interval $[\alpha, \beta)$, where $\beta > t_0$, such that

(a) each $x_i'(t)$ is continuous and $|x_i'(t)| < c$ for all $t \in [\alpha, \beta)$;
(b) $x_2(t) > x_1(t)$ and $(t, x_i(t)) \in D$ for all $t \in [t_0, \beta)$;
(c) the two functional equations $\tau_{ji}^0 = |x_i(t_0) - x_j(t_0 - \tau_{ji}^0)|/c$ have solutions τ_{ji}^0, $i \neq j$, $i, j \in \{1, 2\}$.

Then Driver proves that $(x_1(t), x_2(t))$ is a solution of (2.1.1)–(2.1.2) if and only if it satisfies the following system of six delay differential equations for $t \in (t_0, \beta)$:

$$\begin{cases} x_i'(t)=v_i(t), \\ \tau_{ji}'(t)=\dfrac{(-1)^i v_i(t)-(-1)^i v_j(t-\tau_{ji}(t))}{c-(-1)^i v_j(t-\tau_{ji}(t))}, \\ \dfrac{v_i'(t)}{[1-v_i^2(t)/c^2]^{3/2}}=\dfrac{(-1)^i a_i c}{\tau_{ji}^2(t)}\cdot\dfrac{c+(-1)^i v_j(t-\tau_{ji}(t))}{c-(-1)^i v_j(t-\tau_{ji}(t))}+q_i E_{\text{ext}}\bigl(t, x_i(t)\bigr)/m_i, \end{cases} \tag{2.1.3}$$

where $\tau_{ji}(t_0) = \tau_{ji}^0$, $a_i = q_1 q_2/(4\pi\epsilon_0 m_i c^3)$ (a constant, and in particular, ϵ_0 is the dielectric constant of free space), and $(i, j) = (1, 2)$ or $(2, 1)$.

It is shown in Driver [58] that if given initial trajectories satisfy condition (a) for $\alpha \leqslant t \leqslant t_0$, condition (b) at t_0, and condition (c), and if $E_{\text{ext}}(t, x)$ is Lipschitz continuous with respect to x in each compact subset of D and if the initial velocity of each particle is Lipschitz continuous, then a unique solution does exist. This solution can be continued as long as the charges do not collide ($\lim x_1(t) = \lim x_2(t)$ as t approaches the right endpoint of the maximal interval for existence) and neither $(t, x_1(t))$ nor $(t, x_2(t))$ approaches the boundary D.

We remark here that in Driver and Norris [64], the above Lipschitz continuity for the initial velocities is relaxed to the integrability of the initial velocity on $[\alpha, t_0]$. In Driver [61], one special case was given where the positions and velocities of the particles *at some instant* will determine the state of the system. More precisely, in this example of electrodynamic equations of motion, instantaneous values of positions and velocities of the particles will determine their trajectories, if the solutions are defined for all future time. This property was frequently conjectured, asserted, or implicitly assumed, as in Newtonian mechanics and as indicated by the long list of related references in Driver [61], but this property should not be expected for general electrodynamic equations.

In the case where $E_{\text{ext}}(t,x)=0$ for all $(t,x)\in\mathbb{R}^2$ and if $q_1q_2>0$ (two point charges of like sign), then $\lim_{t\to\infty}[x_2(t)-x_1(t)]=\infty$ and $|v_i(t)|\leqslant\bar{c}<c$ for all $t\geqslant\alpha$. This is a quite interesting result as it indicates that the delay $\tau_{ji}(t)$ may become unbounded, as such, one obtains a system of functional differential equation with unbounded state-dependent delays.

It is noted that if three-dimensional motions are considered, then one obtains a functional differential system of neutral type where the delays are dependent on the states, and the change rate of v_i at the current time also depends on its historical value $v_j'(t-\tau_{ji})$. More precisely, if we introduce a unit vector

$$u_i=\frac{x_i-x_j(t-\tau_{ji})}{c\tau_{ji}}$$

and a scalar quantity

$$\gamma_{ij}=1-\frac{1}{c}v_j(t-\tau_{ji})\cdot u_i$$

as Driver [62] does, where $\cdot$ indicates the dot or scalar product in $\mathbb{R}^3$ (note, of course, x_1, x_2 are now vectors in $\mathbb{R}^3$), then the Lorentz force law yields

$$v_i'(t)=\frac{q_i(1-|v_i|^2/c^2)^{1/2}}{m_i}\big[E_j+(v_i/c\cdot E_j)(u_i-v_i/c)-(v_i/c\cdot u_i)E_j\big], \tag{2.1.4}$$

where E_j is the retarded (vector-valued) electric field arriving at x_i at the instant t from particle j. This field, in $\mathbb{R}^3$, can be found from the Liénard–Weichert potentials as

$$\begin{aligned}E_j={}&\frac{kcq_j}{\tau_{ji}^2\gamma_{ij}^3}\big[u_i-v_i(t-\tau_{ji})/c\big]\big[1-|v_j|^2(t-\tau_{ji})\big]\\&+\frac{kq_j}{\tau_{ji}\gamma_{ij}^3}u_i\times\big(\big[u_i-v_j(t-\tau_{ji})/c\big]\times v_j'(t-\tau_{ji})\big),\end{aligned} \tag{2.1.5}$$

where $k>0$ is a constant depending on the units, and $\times$ indicates the vector cross product in $\mathbb{R}^3$. The dynamical adaptation for τ_{ji} is given by

$$\tau_{ji}'(t)=\frac{u_i\cdot[v_i-v_j(t-\tau_{ji})]}{c\gamma_{ij}}. \tag{2.1.6}$$

In the above discussions, the motion of each particle is influenced by the electromagnetic fields of the others, and due to the finite speed of the propagation of these fields, the model equations describing the motion of charged particles via action at a distance will involve time delays which depends on the state of the whole system. In Driver [63] and in Hoag and Driver [111], it is noted that if one considers that the basic laws of physics are symmetric with respect to time reversal, then the existence of these delays implies that there should also be advanced terms in the equations, and thus one is led to a system of functional differential equations with mixed arguments (Hoag and Driver [111]), and of neutral type (Driver [63]).

In summary and in conclusion, despite the fact that much of the work by Driver and his collaborators on electrodynamics was published nearly 40 years ago, many interesting questions related to the fundamental issues of electrodynamics remain unsolved mathematically and Driver's models remain as a source of inspiration for the theoretical development and a testing tool for new results.

2.2. *Position control*

State-dependent delays arise naturally in automatic position control if in the feedback loop running times of signals between the object of study and a reference point are taken into account.

In [205], the following simple and idealized situation is considered: An object moves along a line and regulates its position relative to an obstacle by means of signals which are reflected by the obstacle. Let $x(t)$ denote the position of the object at time $t \in \mathbb{R}$. The aim of control is that the object should not collide with the obstacle located at $-w < 0$, and that the object should be close to a preferred position at distance w from the obstacle, i.e., at $0 \in \mathbb{R}$. The difficulty for the control is that measurement of the position via signal running times takes time during which the object is moving. More precisely, assume that signals travel from the object to the obstacle at a speed $c > 0$ and are reflected. The object senses the reflected signals and measures the signal running time $s(t)$ between the emission and detection at time t:

$$cs = \left|x\bigl(t - s(t)\bigr) + w\right| + \left|x(t) + w\right|.$$

Then it uses s to compute a distance d from the obstacle according to

$$d = \frac{c}{2}s.$$

This seems to be reasonable since it gives the true distance at time t at least if at times $t - s(t)$ and t the object is in the same position. We must however emphasize that in general d is only a *computed* length and not the true distance, and that we consider a situation where there is no direct, immediate access to the true position.

Depending on the computed distance $d - w$ from the preferred the object adjusts its speed in size and direction, with a reaction time lag $r > 0$ which is assumed to be constant. This negative feedback mechanism is then described by the differential equation

$$x'(t) = v\bigl(d(t - r) - w\bigr)$$

where $v : \mathbb{R} \to \mathbb{R}$ represents negative feedback with respect to the preferred position $0 \in \mathbb{R}$ in the sense that

$$\delta v(\delta) < 0 \quad \text{for all } \delta \neq 0$$

holds. Therefore, we are led to the system

$$x'(t) = v\left(\frac{c}{2}s(t-r) - w\right), \tag{2.2.1}$$

$$cs(t) = \left|x\big(t - s(t)\big) + w\right| + \left|x(t) + w\right| \tag{2.2.2}$$

for positive parameters c, w, r and a negative feedback nonlinearity v. Notice that for motion $x : \mathbb{R} \to \mathbb{R}$ with speed $|x'|$ bounded by a constant $b < c$, that is, slower than the signal speed, the equation

$$s = \frac{1}{c}\big(\left|x(t-s) + w\right| + \left|x(t) + w\right|\big)$$

equivalent to (2.2.2) has a unique solution $s = \sigma(x|_{(-\infty,t]})$ because for given x and t the right-hand side of the equation defines a contraction $[0, \infty) \to [0, \infty)$. Then one can take the right-hand side of Eq. (2.2.2) with $t - r - \sigma(x|_{(-\infty,t-r]})$ and $t - r$ instead of $t - s$ and t, respectively, and replace $cs(t - r)$ in Eq. (2.2.1), which yields a single delay differential equation with state-dependent delay. Essentially the same reasoning shows that for Lipschitz continuous solutions $x : [-h, t_e) \to \mathbb{R}$, $h > 0$ and $t_e > 0$, which have Lipschitz constants strictly less than c and satisfy suitable boundedness conditions, the system (2.2.1)–(2.2.2) can be rewritten as a single delay differential equation of the form (1.0.1).

A closely related model, with an explicit expression for a fraction of the signal running time instead of Eq. (2.2.2) for the total signal running time, was mentioned earlier by Nussbaum [177]. A similar model was also proposed by Messer [161].

It is perhaps more realistic to replace the first order differential equation (2.2.1) with the constant reaction lag $r > 0$ by Newton's law

$$x'' = A$$

with an instantaneous restoring force A which depends on the computed distance d. Also friction might be taken into account. Such a model was studied in [208], by means of the fundamental theory presented in the next section.

We describe the main results from [205,208], namely existence of stable periodic orbits, in Section 7.3.

In [32,33] Büger and Martin study a case of velocity control which involves signal running times. They consider an object travelling along a line which, ideally, should have a certain prescribed constant velocity v_0 throughout its whole journey. The object regulates its velocity $v = x'$ by adjusting its acceleration $a = v'$ according to a negative feedback relation

$$a \cdot (v - v_0) < 0 \quad \text{for all } v \in \mathbb{R} \setminus \{v_0\},$$

that is, the object speeds up if $v < v_0$ and slows down when $v > v_0$. This strategy is remotely controlled by a base located at $x = 0$ to which the object transfers its current velocity v with a certain transmission velocity $c_1 > 0$, and the negative feedback information is then transmitted back from the base to the object with a transmission velocity $c > 0$. The total signal running time of the signals emitted from the object in the past and transmitted back from the base to the object as to arrive at time t is given by

$$s = s_1 + s_2$$

where

$$s_1 = \frac{1}{c_1}\bigl|x(t-s)\bigr| \quad \text{and} \quad s_2 = \frac{1}{c}\bigl|x(t)\bigr|.$$

If an additional constant reaction time $r > 0$ of the base is taken into account then the total transmission delay is

$$r + s(t)$$

where now

$$s(t) = \frac{1}{c_1}\bigl|x\bigl(t - r - s(t)\bigr)\bigr| + \frac{1}{c}\bigl|x(t)\bigr|.$$

The preceding equation and the second order differential equation

$$x''(t) = A\bigl(x'\bigl(t - r - s(t)\bigr) - v_0\bigr)$$

with a negative feedback nonlinearity $A : \mathbb{R} \to \mathbb{R}$ constitute the model. Büger and Martin investigated a simplification which neglects the running time of the signal from the object to the base, in which case the model is reduced to the single equation

$$x''(t) = A\left(x'\left(t - r - \frac{1}{c}\bigl|x(t)\bigr|\right) - v_0\right) \tag{2.2.3}$$

with an explicitly given state-dependent delay. It is shown in [32,33] that close to segments of constant velocity solutions $t \mapsto v_0 t + c$ there exist segments of solutions of Eq. (2.2.3) for which $t - r - \frac{1}{c}|x(t)|$ remains bounded from above by some T. In this case the object reacts only to velocities achieved before $t = T$, which may not be adequate for larger velocities reached later. The phenomenon occurs for solutions whose speed $|v| = |x'|$ approaches the signal speed c or grows beyond. Büger and Martin call it the *escaping disaster*. In [33] they design another control mechanism which overcomes this kind of instability of the constant velocity solutions.

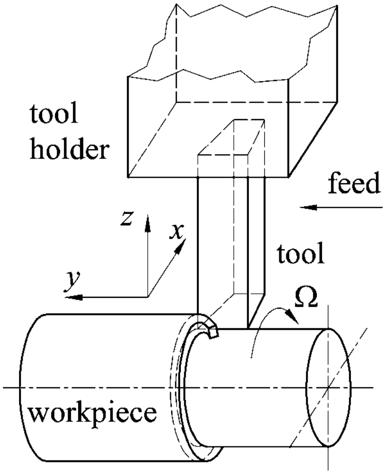


Fig. 1. A cutting process.

2.3. *Mechanical models*

State-dependent models have been proposed and investigated in mechanics, as well. Johnson [123] studied a steel rolling mill control system in 1972 where state-dependent delays have already been encountered. Nevertheless, state-dependent delay models have not frequently used in mechanics, since the required mathematical methods, like linearization techniques, were only recently developed (see Sections 3.4 and 3.6).

Insperger, Stépán and Turi [117] proposed a two degree of freedom model for turning process. This models a machine tool where a workpiece is rotating, the tool cuts the surface that was formed in the previous cut, see Fig. 1. The chip thickness is determined by the current and a previous position of the tool and the workpiece. In standard models the time delay between two succeeding cuts is considered to be equal to the period of the workpiece rotation. More realistic models given in the machine tool literature include the feed motion and the consequent trochoidal path of the cutter tooth. In this case the time delay between the succeeding cuts is not constant, it changes periodically in time. Time periodic delays also arise in the model of varying spindle speed machining. If the regeneration process is modeled more accurately, then the vibration of the tool is also included in the model, and this results a model with state-dependent delay. The system can be modeled as a two degree of freedom oscillator that is excited by the cutting force, so the governing equations are

$$m\ddot{x}(t) + c_x\dot{x}(t) + k_x x(t) = F_x, \tag{2.3.1}$$

$$m\ddot{y}(t) + c_y\dot{y}(t) + k_y y(t) = -F_y, \tag{2.3.2}$$

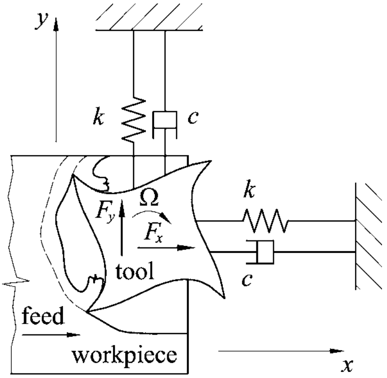


Fig. 2. A milling process.

where the associated model parameters are mass m, damping c_x, c_y and stiffness k_x, k_y. The x and y components of the cutting force are, respectively,

$$F_x = K_x w h^q, \quad F_y = K_y w h^q,$$

where K_x and K_y are the cutting coefficients in the x and y components, w is the depth of cut, h is the chip thickness, and the exponent $0 < q < 1$ is constant. The time delay between the present and the previous cuts is determined by the equation

$$R\Omega\tau(x_t) = 2R\pi + x(t) - x\bigl(t - \tau(x_t)\bigr), \tag{2.3.3}$$

where Ω is the spindle speed given in [rad/s] and R is the radius of the workpiece. This is an implicit equation for the time delay, and τ depends on the solution, as well. The chip thickness satisfies

$$h = v\tau(x_t) + y(t) - y\bigl(t - \tau(x_t)\bigr),$$

where v is the feed speed. Therefore the model equations are

$$m\ddot{x}(t) + c_x\dot{x}(t) + k_x x(t) = K_x w\bigl(v\tau(x_t) + y(t) - y\bigl(t - \tau(x_t)\bigr)\bigr)^q,$$

$$m\ddot{y}(t) + c_y\dot{y}(t) + k_y y(t) = -K_y w\bigl(v\tau(x_t) + y(t) - y\bigl(t - \tau(x_t)\bigr)\bigr)^q,$$

where the delay function is defined by (2.3.3). In [117] the linearized stability of an equilibrium solution was also studied using the method of [101] and [108]. It was shown that the linearized equation is almost equal to the standard constant delay machine tool vibration equation, the difference is a term with a small coefficient.

A related problem was studied in [116], where a two degree freedom model of milling process is considered. In this machine a tool with equally spaced teeth rotating with constant spindle speed, and cuts the surface that was formed in the previous cut, see Fig. 2. In this case the system again can be described by (2.3.1)–(2.3.2), but now the damping and stiffness parameters are equal in both x and y directions: $c_x = c_y = c$, $k_x = k_y = k$, and it was shown that the model equations have the form

$$m\ddot{x}(t) + c\dot{x}(t) + kx(t) = \sum_{j=1}^{N} \alpha_{x,j}(t)\Big(R\big(1 - \cos(\Omega\tau_j - \vartheta)\big) + \big(v\tau_j + x(t-\tau_j) - x(t)\big)\sin\varphi_j(t) + \big(y(t-\tau_j) - y(t)\big)\cos\varphi_j(t)\Big)^q, \tag{2.3.4}$$

$$m\ddot{y}(t) + c\dot{y}(t) + ky(t) = \sum_{j=1}^{N} \alpha_{y,j}(t)\Big(R\big(1 - \cos(\Omega\tau_j - \vartheta)\big) + \big(v\tau_j + x(t-\tau_j) - x(t)\big)\sin\varphi_j(t) + \big(y(t-\tau_j) - y(t)\big)\cos\varphi_j(t)\Big)^q. \tag{2.3.5}$$

Here N is the number of teeth,

$$\alpha_{x,j}(t) = wg\big(\varphi_j(t)\big)\big(K_t\cos\big(\varphi_j(t)\big) + K_n\sin\big(\varphi_j(t)\big)\big),$$
$$\alpha_{y,j}(t) = wg\big(\varphi_j(t)\big)\big(K_n\cos\big(\varphi_j(t)\big) - K_t\sin\big(\varphi_j(t)\big)\big),$$

K_t and K_n are tangential and normal cutting coefficients, g is a screen function, it is equal to 1 if the jth tooth is active, and it is 0 if not, $\phi_j(t) = -\Omega t + (j-1)\vartheta$, Ω is the spindle speed, $\vartheta = 2\pi/N$ is the pitch angle, and the time delays $\tau_j = \tau_j(t, x_t, y_t)$ are defined by the implicit relations

$$\begin{aligned}&\big(v\tau_j + x(t-\tau_j) - x(t)\big)\cos\phi_j(t) - \big(y(t-\tau_j) - y(t)\big)\sin\phi_j(t)\\ &\quad = R\sin(\Omega\tau_j - \vartheta)\end{aligned} \tag{2.3.6}$$

for $j = 1, \ldots, N$. It is easy to check that the functions on the right-hand sides of (2.3.4) and (2.3.5) are periodic in time with period $\bar{\tau} = 2\pi/(N\Omega)$, and the time delay functions τ_j are periodic in time with period $T = N\bar{\tau}$. Moreover, $\tau_j(t + \bar{\tau}, x_t, y_t) = \tau_{j-1}(t, x_t, y_t)$ gives the connection between time delays associated to two succeeding cuts. Note that if the vibration of the tool is not included in the delay model, i.e., $x(t) = 0$ and $y(t) = 0$, then Eq. (2.3.6) is simplified to

$$v\tau_j\cos\phi_j(t) = R\sin(\Omega\tau_j - \vartheta),$$

which yields that the delay depends only on time. This case of time periodic delay was investigated earlier in the literature. If the feed is negligible relatively to the diameter, i.e., $v\tau_j \ll R$, then if we substitute $v\tau_j = 0$ into the previous equation we get $\sin(\Omega\tau_j - \vartheta) = 0$.

This gives the constant delays $\tau_j = \bar{\tau} = \vartheta/\Omega$ that was usually used in standard milling models. Linearized stability of model (2.3.4)–(2.3.5) was also studied in [116] using the results of [101] and was compared to the stability of the standard time-periodic time delay models.

2.4. *Delay adaptation in neural networks and distributed systems*

A synaptic connection between two neurons is referred as to a delay line if the signal transmission is delayed. In neural circuits, time delays arise because interneural distances and axonal conduction times are finite [5,141–143,147,163,219]. In several sensory systems, delay lines are essential for coordinating activities. Examples include the auditory system of barn owls, echo location in bats, and the lateral line system of weakly electric fish. See [41] and the survey article [40]. Delay lines are also important for managing distributed systems, for such systems a fundamental problem concerns how the flow of information from distinct, independent components can be best regulated to optimize a prescribed performance of the network. For example, in parallel computing machines the asynchronous output of independent processors must be integrated to yield well-defined results.

Several possible biophysical mechanisms can be envisioned by which adjustable delays could be achieved, and recent development in the physiology of synapses and dendrites suggests that not only synaptic weights, but also synaptic delays vary [1,215] and changing synaptic delays have significant impact on the neural signal processing.

Time delays have important influence on learning algorithms. As noted in [19], not only delays affect the learning of other parameters such as gains, time constants or synaptic weights but also *delays themselves may be part of the adjustable parameters of a neural system so as to increase the range of its dynamics*. There exist numerous examples of finely tuned delays and delay lines, and *certainly many delays are subject to variations, for instance during the growth of an organism.*

Much of the existing work related to delay adaptation in neural networks have been concentrating on the fine tuning of a selected set of parameters in architectures already endowed of a certain degree of structure. In these applications, the delays are arranged in orderly arrays of delay lines, these delay lines are essentially part of a feedforward network for which the learning task is much more simple and the delays adjust on a slow time scale. In [199], the successive parts of a spoken word are delayed differentially in order to arrive simultaneously onto a unit assigned to the recognition of that particular word. Baldi and Atiya [19] viewed this as a "time warping" technique for optimal matching in the sense that for a given input $I(t)$, the output of the ith delay line is given by the convolution

$$O_i(t) = \int_0^\infty K(i,s) I(t-s)\,\mathrm{d}s,$$

where $K(i,s)$ is a Gaussian delay kernel

$$K(i,s) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-(s-iT)^2/(2\sigma^2)}$$

and T is a parameter used to discretize the possible delays. In this application, the adjustment of delays on a slow time scale may take place across several speakers during the basic development of a speech recognizer and thus the adjustment is independent of the dynamics of the network. However, it is also desirable that delays adjust on a fast scale to adapt in real time to one particular speaker during normal functioning. The above work of Baldi and Atiya, as well as those in [51,166], develops models for delay adaptation with the help of a global teacher signal. It is shown in [51] that a network with adaptable delays can achieve smaller errors than a network with fixed delays, if they both have the same number of neurons and connections. Other rules and algorithms have been developed by which excitatory postsynaptic potentials from different synapses can be gradually pulled into coincidence, see, for example, [113].

Examples of self-organized delay adaptation can be found in [73]. For example, time delays in the optic nerve are equalized [192], signals in visual callosal axons arrive simultaneously at all axonal endings [115], and neurons in vitro can inhibit the formation of a myelin sheet by firing at a low frequency [193].

Two different mechanisms, delay shift and delay selection, are investigated in [73,74] for the self-organized adaptation of transmission delays in the nervous system.

To formulate the model for a network consisting of a large number of presynaptic neurons and one postsynaptic neuron which receives its input lines, we consider the idealized situation where there is a continuous set of input connections described by two functions, $\rho(\tau,t)$ and $\omega(\tau,t)$, for the delays and weights, respectively: $\rho(\tau,t)\,\mathrm{d}\tau$ gives the fraction of connections with delays in $[\tau,\tau+\mathrm{d}\tau]$, and $\omega(\tau,t)$ is the average weight of connections with delay τ. We assume that delays in the system adapt on a developmental time scale, and thus the model formulated below does not involve the internal dynamics of the network. In the continuous description, the input density $J(\tau,t)$ provided by the synapses at delay τ after presentation of a pattern has the simple form

$$J(\tau,t)=\rho(\tau,t)\omega(\tau,t).$$

The dynamics of the input are governed by two simultaneous equations: a balance equation for the input density

$$\frac{\partial}{\partial t}J(\tau,t)=-\frac{\partial}{\partial\tau}\bigl[J(\tau,t)v(\tau,t)\bigr]+Q(\tau,t),\tag{2.4.1}$$

and a continuity equation for $\rho(\tau,t)$, indicating the conservation of the number of neural connections

$$\frac{\partial}{\partial}\rho(\tau,t)=-\frac{\partial}{\partial\tau}\bigl[\rho(\tau,t)v(\tau,t)\bigr].\tag{2.4.2}$$

The drift velocity, $v(\tau,t)$, and the source term, $Q(\tau,t)$, are defined according to the Hebbian principles. In particular, in the case of delay shifts, the weights are not modified and the source term vanishes. Therefore, the dynamics is completely governed by (2.4.2) where the velocity, $v = \mathrm{d}\tau/\mathrm{d}t$, of the delays realizes the Hebbian adaptation

$$v(\tau, t) = \gamma_\tau \int_{-\infty}^{+\infty} W_\tau(\tau - s) P(s, t)\, \mathrm{d}s, \tag{2.4.3}$$

and γ_τ denotes the learning rate, $W_\tau(x)$ denotes a learning window for delay adaptation. $W_\tau(x)$ should be positive when the presynaptic contribution precedes the postsynaptic spike, and negative in the other case, and this rule will adjust the delays such that their effects will align in time at the soma [72]. The distribution of spike times, $P(s, t)$, of a neuron depends on the input and its statistics, and this is assumed to be of the form $P(s, t) = \beta J(s, t)$ in the work of [73,74] and it is justified if the input is sufficiently high and if there is some random background activity. Note that delay shift mechanism assumes that the transmission delays themselves are altered. This mechanism is possible because transmission velocities in the nervous system can be altered, for example, by changing the length and thickness of dendrites and axons, the extent of myelination of axons, or the density and type of ion channels. See [74,75].

In the case of delay selection, the drift velocity of the delays vanishes and the total input of the postsynaptic neuron is not conserved. Equations (2.4.1) and (2.4.2) result in

$$\rho(\tau, t) \frac{\partial}{\partial t} \omega(\tau, t) = Q(\tau, t). \tag{2.4.4}$$

Again, the source term is derived from the Hebbian rule by

$$Q(\tau, t) = \gamma_\omega \omega(\tau, t) \rho(\tau, t) \int_{-\infty}^{+\infty} W_\omega(\tau - s) P(s, t)\, \mathrm{d}s,$$

with γ_ω denoting the corresponding learning rate, and W_ω representing a learning window that is maximal just before the time of spiking—leading to a selection of delay lines for which the effects align at soma.

In the aforementioned delay-adaptation models, it is assumed that delays in the system adapt on a developmental time scale, and thus the temporal development of ρ and ω is determined by an average over an ensemble of presynaptic input patterns. In the recent work [114], the self-organized adaptation of transmission delays is incorporated into the projective adaptive resonance theory developed in [38,39], and this self-organized adaptation of delay is driven by the dissimilarity between input patterns and stored patterns in a neural network designed for pattern recognition from data sets in high dimensional spaces. This adaptation can be regarded as a consequence of the Hebbian learning law, and the dynamic adaptation can be modeled by a nonlinear differential equation and hence a system of delay differential equations with adaptive delay is used.

We now describe the model for such a network that consists of two layers of neurons. Denote the nodes in F_1 layer (Comparison/Input Processing layer) by P_i, $i \in \Lambda_p := \{1, \ldots, m\}$; nodes in F_2 layer (Clustering layer) by C_j, $j \in \Lambda_c := \{m+1, \ldots, m+n\}$; the activation of F_1 node P_i by x_i, the activation of F_2 node C_j by y_j; the bottom-up weight from P_i to C_j by z_{ij}, the top-down weight from C_j to P_i by w_{ji}.

The STM (short-term memory) equations for neurons in F_1 layer are given by

$$\epsilon_p \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = -x_i(t) + I_i, \quad t \geqslant -1,\ i \in \Lambda_p, \tag{2.4.5}$$

where $0 < \epsilon_p \ll 1$, I_i is the constant input imposed on P_i.

The change of the STM for a F_2 neuron depends on the internal decay, the excitation from self-feedback, the inhibition from other F_2 neurons and the excitation by the bottom-up filter inputs from F_1 neurons. Namely, we have the STM equations for the committed neurons in F_2 layer:

$$\begin{aligned}\epsilon_c \frac{\mathrm{d}y_j(t)}{\mathrm{d}t} &= -y_j(t) + \bigl[1 - Ay_j(t)\bigr]\bigl[f_c\bigl(y_j(t)\bigr) + T_j(t)\bigr] \\ &\quad - \bigl[B + Cy_j(t)\bigr] \sum_{k \in \Lambda_c \setminus \{j\}} f_c\bigl(y_k(t)\bigr), \quad t \geqslant 0,\ j \in \Lambda_c,\end{aligned} \tag{2.4.6}$$

where $0 < \epsilon_c \ll 1$, $f_c : \mathbb{R} \to \mathbb{R}$ is a signal function, A, B, and C are non-negative constants, and the bottom-up filter input T_j is given by

$$T_j(t) = D \sum_{i \in \Lambda_p} z_{ij}(t) f_p\bigl(x_i\bigl(t - \tau_{ij}(t)\bigr)\bigr)\mathrm{e}^{-\alpha\tau_{ij}(t)}, \quad t \geqslant 0, \tag{2.4.7}$$

where D is a scaling constant, $f_p : \mathbb{R} \to \mathbb{R}$ is the signal function of the input layer. It is assumed here that the signal transmissions between two layers are not instantaneous and the signal decays exponentially at a rate $\alpha > 0$. This assumption that signal strength decays if the transmission is delayed can be replaced by the mechanism of *delay selection* by replacing (2.4.7) by

$$\begin{aligned}T_j(t) = D_f &\sum_{i \in \Lambda_p, \tau_{ij}(t)=0} z_{ij}(t) f_p\bigl(x_i(t)\bigr) \\ &+ D_d \sum_{i \in \Lambda_p, \tau_{ij}(t)>0} z_{ij}(t) f_p\bigl(x_i\bigl(t - \tau_{ij}(t)\bigr)\bigr)\end{aligned}$$

with two different weight factors $0 < D_d \ll D_f$.

The term τ_{ij} is the signal transmission delay between the cluster neuron C_j and the input neuron P_i. It is assumed that this delay is driven by the dissimilarity in the sense that the signal processing from the input neuron P_i to the cluster neuron C_j is faster when the output from P_i is similar to the corresponding component of w_{ji} of the feature vector $w_j = (w_{ji})_{i \in \Lambda_p}$ of the cluster neuron C_j. Therefore, we have

$$\beta \frac{\mathrm{d}\tau_{ij}(t)}{\mathrm{d}t} = -\tau_{ij}(t) + E\bigl[1 - h_{ij}(t)\bigr], \quad t \geqslant 0,\ i \in \Lambda_p,\ j \in \Lambda_c, \tag{2.4.8}$$

where $\beta > 0$, $E \in (0, 1)$ are constants and

$$h_{ij}(t) = S\bigl(d\bigl(f_p\bigl(x_i(t), w_{ji}(t)\bigr)\bigr), z_{ij}(t)\bigr)$$

is the similarity measure between the output signal $f_p(x_i(t))$ and the corresponding component $w_{ji}(t)$ of the feature vector of the cluster neuron C_j, with respect to the significance factor of the bottom-up synaptic weight $z_{ij}(t)$, here d is the usual absolute value function and $S:\mathbb{R}^+ \times [0,1] \to [0,1]$ is a given function, non-increasing with respect to the first argument and nondecreasing with respect to the second argument. Moreover, $S(0,1) = 1$ (the similarity measure is 1 with complete similarity and maximal synaptic bottom-up weight) and $S(+\infty, z) = S(x, 0) = 0$ for all $z \in [0,1]$ and $x \in \mathbb{R}^+$ (the similarity measure is 0 with complete dissimilarity or minimal bottom-up synaptic weight). Therefore, if $\tau_{ij}(0) = 0$ then from (2.4.8) it follows that $0 \leqslant \tau_{ij}(t) \leqslant E$ for all $t \in \mathbb{R}^+$, and moreover, if $h_{ij}(t) = 1$ on an interval $[0, b)$ for a given $b > 0$ then $\tau_{ij}(t) = 0$ for all $t \in [0, b)$.

The equation governing the change of the weights follows from the usual synaptic conservation rule and only connections to activated neurons are modified. The top-down weights are modified so that the template will point to the direction of the delayed and exponentially decayed outputs from F_1 layer. Therefore, we have

$$\gamma \frac{\mathrm{d}w_{ji}(t)}{\mathrm{d}t} = f_c\big(y_j(t)\big)\big[-w_{ji}(t) + f_p\big(x_i\big(t - \tau_{ij}(t)\big)\mathrm{e}^{-\alpha\tau_{ij}(t)}\big)\big] \tag{2.4.9}$$

for $t \geqslant 0$, $i \in \Lambda_p$, $j \in \Lambda_c$, where $\gamma > 0$ is a given constant.

The bottom-up weights are changed according to the competitive learning law and Weber law that says that LTM size should vary inversely with input pattern scale. Thus the LTM equations for committed neurons C_j in F_2 layer are

$$\begin{aligned}\delta \frac{\mathrm{d}z_{ij}(t)}{\mathrm{d}t} = f_c\big(y_j(t)\big)\Big[&\big(1 - z_{ij}(t)\big)h_{ij}(t)L - z_{ij}(t)\big(1 - h_{ij}(t)\big) \\ &- z_{ij}(t) \sum_{k \in \Lambda_p \setminus \{i\}} h_{kj}(t)\Big], \quad t \geqslant 0,\ i \in \Lambda_p,\ j \in \Lambda_c,\end{aligned} \tag{2.4.10}$$

where $0 < \delta \ll \gamma = \mathrm{O}(1)$ and $L > 0$ is a given constant.

Equations (2.4.5)–(2.4.10) give a system of functional differential equations where the dynamics of the delay $\tau_{ij}(t)$ is adaptive and is described by the nonlinear equation (2.4.8). The dynamics is investigated in [114] in the case where the signal functions f_p and f_c are step functions, though much remains to be done in the general case.

2.5. *Disease transmission and threshold phenomena*

Delay differential equations with state-dependent delay arise naturally from the modeling of infection disease transmission, the modeling of immune response systems and the modeling of respiration, where the delay is due to the time required to accumulate an appropriate dosage of infection or antigen concentration.

Following [201], we consider a particular infectious disease in an isolated population that is divided into several disjoint classes (compartments) of individuals: the susceptible class (those individuals who are not infective but are capable of contracting the disease and become infective), the exposed class (those who are exposed but not yet infective), and the infective class (those individuals who are capable of transmitting the disease to others), and the removed class (those who have had the disease and are dead, or have recovered and are permanently immune, or are isolated until recovery and permanent immunity occur). Let $S(t)$, $E(t)$, $I(t)$ and $R(t)$ be the size of each class at time t, and assume the following:

(i) the rate of exposure of susceptibles to infectives at time t is given by $-rS(t)I(t)$;

(ii) an individual who becomes infective at time t recovers from the infection (is thus removed from the population) at time $t+\sigma$, where σ is a positive constant;

(iii) an individual who is first exposed at time τ becomes infective at time t if

$$\int_{\tau}^{t} \bigl[\rho_1(x)+\rho_2(x)I(x)\bigr]\,\mathrm{d}x = m,$$

where ρ_1, ρ_2 are nonnegative functions and m is a given positive constant;

(iv) the population size remains to be a constant N.

The motivation for assumption (iii), the basis for a threshold model, is that human body can often control a small exposure to an infection, that is, there is a tolerance level below which the body's immune system can combat exposure to infection. When too large an exposure results, the individual contracts the disease. The amount of exposure received depends on the duration of the exposure and the amount of infectivity around the individual, that is assumed to be proportional to the number of infective individuals in the population. Thus, during the time interval $[t, t+h]$ an exposure of $\int_t^{t+h} \rho_2(x)I(x)\,\mathrm{d}x$ is accumulated where ρ_2 is a proportionality function which is a measure of the amount of infection communicated per infective (virulence). When the total exposure reaches the threshold m, the individual moves from class (E) to class (I). The term ρ_1 is the rate of accumulation of exposure independent of the number of infectives (such as constant input of virus from the external environment). In what follows, we consider the simple case where $\rho_1 = 0$, $\rho_2 = \rho$.

We assume the initial distribution of the infectives is given by a monotone function $I_0 : [-\sigma, 0] \to [0, \infty)$ such that $I_0(-\sigma) = 0$, and $I_0(0) > 0$ infective individuals are inserted in the population at $t = 0$. It is convenient to extend I_0 to the whole real line

$$I_0^{\mathrm{e}}(t) = \begin{cases} 0, & |t| > \sigma; \\ I_0(t), & -\sigma < t \leqslant 0; \\ I_0(0) - I_0(t-\sigma), & 0 \leqslant t \leqslant \sigma, \end{cases}$$

so that the extension $I_0^{\mathrm{e}}(t)$ describes the number of initial infectives who are still present as infectives at time $t \in [-\sigma, \infty)$.

From (i) and (ii), it follows that the number of new infectives introduced into the population at time t is given by $-\int_{t-\sigma}^{t} \frac{\mathrm{d}}{\mathrm{d}x} S(\tau(x))H(x)\,\mathrm{d}x$, where $H(x) = 0, x < 0$ and $H(x) = 1$, $x > 0$. Therefore,

$$I(t) = I_0^{\mathrm{e}}(t) - \int_{t-\sigma}^{t} \frac{\mathrm{d}}{\mathrm{d}x} S\bigl(\tau(x)\bigr)H(x)\,\mathrm{d}x, \quad t \geqslant 0.$$

For the infection to spread, some of the initial susceptible population must become infective before time σ. Thus, we assume there is $t_0 < \sigma$ so that the following "admissibility" condition,

$$\int_0^{t_0} \rho(x) I_0^{\mathrm{e}}(x)\,\mathrm{d}s = m,$$

must be met.

Then the model equations become

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} S(t) = -rS(t)I(t), \\ I(t) = I_0^{\mathrm{e}}(t) - \displaystyle\int_{t-\sigma}^{t} \frac{\mathrm{d}}{\mathrm{d}x} S\big(\tau(x)\big) H(x)\,\mathrm{d}x, \\ E(t) = S\big(\tau(t)\big) - S(t), \\ R(t) = N - S(t) - I(t) - E(t), \\ \displaystyle\int_{\tau(t)}^{t} \rho(x) I(x)\,\mathrm{d}x = m. \end{cases}$$

If we further adopt the convention that $\tau(t) = 0$ for $t \leqslant t_0$, then we obtain a state-dependent delay differential equation for $S(t)$:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} S(t) = -rS(t)I(t), \\ I(t) = I_0(t) + S\big(\tau(t-\sigma)\big) - S\big(\tau(t)\big), \\ \displaystyle\int_{\tau(t)}^{t} \rho(x) I(x)\,\mathrm{d}x = m, \quad t > t_0, \\ \tau(t) = 0, \quad t \leqslant t_0. \end{cases}$$

Other related models involving threshold conditions that determine the state-dependence of delay can be found in Gatica and Waltman [87–89], Hoppensteadt and Waltman [112], Smith [189], and Waltman [201]. Relatively complete references can be found in the work of Kuang and Smith [136,137], where the prototype equation takes the form

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} x(t) = -\nu x(t) - \mathrm{e}^{-\eta\tau} f\big(x(t-\tau)\big), \\ \displaystyle\int_{t-\tau}^{t} k\big(x(t), x(s)\big)\,\mathrm{d}s = m, \end{cases} \tag{2.5.1}$$

with nonnegative constants ν, η and m, and a positive function k, as well as a nonlinearity $f : \mathbb{R} \to \mathbb{R}$. Again, in the contents of epidemiological modeling, $x(t)$ may represent the proportion of a population which is infective at time t and the second equation in system (2.5.1) may reflect that an individual who is first exposed to the disease at time $t - \tau$ becomes infectious at time t if, during the interval from $t - \tau$ to t, a threshold level of exposure is accumulated where per unit time exposure depends on the infective fraction $x(s)$ via $k(x(t), x(s))$.

2.6. *Population models and size-dependent interaction*

Recent efforts in modeling state-dependent phenomena in population dynamics involve structured models, and state-dependent delay normally arises from a certain threshold condition. For example, in the work [170], Nisbet and Gurney considered insect populations which have several life stages (instars). After constructing a mathematical model consisting of an equation for the mass density function of the population, and under the homogeneity assumption for the population at each life stage, they reduced the model to a system of delay differential equations for the size of the population in each life stage. The threshold delays then appear due to the assumption that the insect must spend an amount of time in the larval stage sufficient to accumulate a threshold amount of food. See also [3].

This idea was adopted in the work of Arino, Hbid and Bravo de la Parra [8] for the growth of a population of fish where they introduce an additional stage between the eggs and the mobile larvae, the so-called (S1) larval stage. The state variable for this stage is $n_1(a,t)$, and the passage through (S1) is described with the help of another variable $q_1(a,t)$, the amount of food eaten up to time t by an individual entered in (S1) a units of time earlier. The introduction of this variable makes it possible to formulate the condition for any individual to have eaten a certain amount of food Q_1 (threshold) during the whole duration (bounded above by T_1) an individual can spend in (S1). The variation of ingested food is governed by the standard structured model (see [162]) subject to zero boundary and initial conditions that can be solved by integration along the characteristics to give

$$q_1(a,t) = \int_{t-a}^{t} \frac{K_1}{N_1(\sigma)+C_1}\,\mathrm{d}\sigma, \qquad t > a,$$

where K_1 is the quantity of food flowing into the species habitat per unit of volume, per unit of time, C_1 is the food(converted into a number of individuals) taken per unit of volume by consumers other than (S1) stage, and $N_1(t) = \int_0^{T_1} n_1(a,t)\,\mathrm{d}a$ is the population in stage (S1) which is susceptible to enter the next stage at time t per unit of volume. Hence an individual moves out of the (S1) stage exactly at time t if it entered in the (S1) stage $a_1(t)$ units of time earlier, where $a_1(t)$ is given by the threshold condition

$$\int_{t-a_1(t)}^{t} \frac{K_1}{N_1(\sigma)+C_1}\,\mathrm{d}\sigma = Q_1,$$

from which it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}a_1(t) = -\frac{N_1(t-a_1(t)) - N_1(t)}{N_1(t)+C_1}. \tag{2.6.1}$$

The introduction of $a_1(t)$ through the above threshold condition ties the change of individual states to the dynamics of the population at the population level. From the definition of $a_1(t)$, one naturally has $n_1(a,t) = 0$ for all $a > a_1(t)$. The dynamics for $n_1(a,t)$ with $0 < a < a_1(t)$ is given by

$$\begin{cases} \dfrac{\partial}{\partial a} n_1(a,t) + \dfrac{\partial}{\partial t} n_1(a,t) = -f(a) n_1(a,t), \quad 0 < a < a_1(t),\ t > 0, \\ n_1(a,0) = 0, \qquad n_1(0,t) = B(t), \end{cases}$$

where the function f is related to the individual resistance to fluctuation of food capacities, and $B(t)$ is the density of eggs laid per unit of volume at time t. In the special case where the eggs of a given year are determined directly in terms of passive larvae that survived some years earlier, we have

$$B(t) = kN_1(t-r)$$

for some positive constants r and k. This yields

$$\begin{cases} N_1(t) = k \displaystyle\int_{t-a_1(t)}^{t} \exp\left(-\int_0^{t-a} f(\sigma)\,\mathrm{d}\sigma\right) N_1(a-r)\,\mathrm{d}a, \\ \displaystyle\int_{t-a_1(t)}^{t} \frac{K_1}{N_1(\sigma)+C_1}\,\mathrm{d}\sigma = Q_1, \end{cases} \tag{2.6.2}$$

or, equivalently,

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} N_1(t) = kN_1(t-r) - \big(1 - a_1'(t)\big) \exp\left(-\displaystyle\int_0^{a_1(t)} f(\sigma)\,\mathrm{d}\sigma\right) kN_1(t-r) \\ \qquad\qquad - k \displaystyle\int_{t-a_1(t)}^{t} f(t-a) \exp\left(-\int_0^{t-a} f(\sigma)\,\mathrm{d}\sigma\right) N_1(t-a)\,\mathrm{d}a, \\ \displaystyle\int_{t-a_1(t)}^{t} \frac{K_1}{N_1(\sigma)+C_1}\,\mathrm{d}\sigma = Q_1. \end{cases} \tag{2.6.3}$$

Note that in the above system of FDEs, there are two components of the delay: a constant delay r and a state-dependent delay.

As the second equation can be written as (2.6.1) by differentiation, it is natural that Arino, Hadeler and Hbid [7] and Magal and Arino [148] considered the system with adaptive delays

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} x(t) = -f\big(x\big(t-\tau(t)\big)\big), \\ \dfrac{\mathrm{d}}{\mathrm{d}t} \tau(t) = h\big(x(t), \tau(t)\big). \end{cases}$$

The existence of periodic solutions for the above system with adaptive delays is considered, and their results will be described in Section 7.2.

In the context of population dynamics, the delay arises frequently as the maturation time from birth to adulthood and this time is in some cases a function of the total population [11]. In [52], the consequence of size-dependent competition among the individuals is investigated by using a system of delay differential equations with state-dependent delays for a population consisting of only two distinct size classes, juveniles and adults, under the assumption that individuals are born at a size $s = s_b$ and remain juvenile as long as $s < s_m$, and individuals mature on reaching the maturation size threshold $s = s_m$. Therefore, if the density of juvenile and adult individuals at time t are denoted by $J(t)$ and $A(t)$, respectively, and if juveniles and adults feed on a shared resource, denoted by $F(t)$, then

$$s_m - s_b = \int_{t-\tau(t)}^{t} \epsilon_g a F(x)\,\mathrm{d}x, \tag{2.6.4}$$

where it is assumed that juvenile individuals feed at a rate aF and use all ingested food for growth in size with conversion efficiency ϵ_g, and $\tau(t)$ is the juvenile delay at time t. Assuming further that adult individuals feed at a rate qaF and use all ingested food for reproduction with conversion efficiency ϵ_b (here q is the ratio between the adult and juvenile feeding rate), and that the (instantaneous) mortality is inversely proportional to food intake with proportionality constant μ, de Roos and Persson obtain the following set of equations for the dynamics of $(J(t), A(t), F(t))$:

$$\begin{cases} \dfrac{\mathrm{d}F(t)}{\mathrm{d}t} = D - aFA - qaFA, \\ \dfrac{\mathrm{d}J(t)}{\mathrm{d}t} = R(t) - R\bigl(t-\tau(t)\bigr)P(t)\dfrac{F(t)}{F(t-\tau(t)))} - \dfrac{\mu}{aF(t)}J(t), \\ \dfrac{\mathrm{d}A(t)}{\mathrm{d}t} = R\bigl(t-\tau(t)\bigr)P(t)\dfrac{F(t)}{F(t-\tau(t))} - \dfrac{\mu}{qaF(t)}A(t), \end{cases} \tag{2.6.5}$$

where

$$R(t) = \epsilon_b qaF(t)A(t)$$

is the total population birth rate at time t and

$$P(t) = \exp\left(-\int_{t-\tau(t)}^{t} \frac{\mu}{aF(x)}\,\mathrm{d}x\right)$$

denotes the probability that an individual which should mature at time t has survived its juvenile period. The ratio $F(t)/F(t-\tau(t))$ counts for the change in maturation rate due to a change in the juvenile delay $\tau(t)$, and this can be derived by considering the model formulation in terms of a hyperbolic partial differential equation for the size distribution of the consumer population $n(t, s)$. The corresponding boundary conditions reflect the fact that the flow rate at time t across the boundary of the size domain $s = s_b$ into the juvenile class equals to total population birth rate $R(t)$ at that time, and that the flow rate at time t across the boundary $s = s_m$ into the adult class equals $\epsilon_g aF(t)n(t, s_m)$.

Note that differentiation of (2.6.4) then yields an equation to govern the evolution of the delay

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = 1 - \frac{F(t)}{F(t-\tau)}.$$

Similarly, in [23], it is showed that size dependent birth processes where the lifespan of individuals is a function of the current population size lead to a certain integral equation,

$$x(t) = \int_{t-L(x(t))}^{t} b\big(x(s)\big)\,\mathrm{d}s, \tag{2.6.6}$$

differentiation then leads to the delay differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \frac{b(x(t)) - b(x(t-L(x(t))))}{1 - L'(x(t))b(x(t-L(x(t))))}. \tag{2.6.7}$$

One should emphasize that, despite a close correspondence between solutions of (2.6.6) and (2.6.7), caution must be employed in using one equation to investigate the other. For example, any constant is a solution of (2.6.7), whereas equation (2.6.6) only admits constant solutions whose values satisfy $x = b(x)L(x)$.

We note that early development of state-dependent delay models in economics and population biology was motivated by phenomenological considerations. For example, in [25, 146] the dynamics of price adjustment in a single commodity market that involves time delays due to production lags and storage policies is considered. The importance of the incorporation of a variable production delay is pointed out as certain commodities, once produced, may be stored for a variable period of time until market prices are deemed advantageous by the producer. In the field of population dynamics, motivated by the observation in Gambell [85] that for antarctic whale and seal populations, the length of time to maturity is a function of the amount of food (mostly krill) available, Aiello, Freedman and Wu [2] propose a stage-structured model of population growth, where the time to maturity is itself state dependent and the special form is suggested by the work [6] that describes how the duration of larval development of flies is a nonlinear increasing function of larval density. This model is further analyzed in [224]. An alternative version, which is designed to address the drawback in the Aiello–Freedman–Wu model that the maturation time for any newborn depends on the existing population size at the same time, is proposed in [76] based on the assumption that the maturation time depends on the size of the population which existed at the time of birth. In particular, if $r(t)$ is the date of birth of individuals who become mature at time $t \geqslant 0$, then the age length up to maturity at time t will be given by the function $\tau(z(r(t)))$. Consequently, $r(t)$ can be solved implicitly by

$$r(t) = t - \tau\big(z\big(r(t)\big)\big), \quad t \geqslant 0.$$

This modification in defining the density-delay dependent term results in the following system for the population with two stages: immature and mature, whose densities are denoted by $z(t) - x_m(t)$ and $x_m(t)$:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} z(t) = -\gamma z(t) + (\alpha + \gamma) x_m(t) - f\big(x_m(t)\big) x_m(t), \\ \dfrac{\mathrm{d}}{\mathrm{d}t} x_m(t) = \alpha x_m\big(r(t)\big) r'(t) \exp\big[-\gamma \tau\big(z\big(r(t)\big)\big)\big] - f\big(x_m(t)\big) x_m(t), \end{cases}$$

where $\gamma > 0$ and $f(x_m)$ are the mortality rates during the immature and mature stages. A detailed derivation of the above model and a careful analysis of conditions which assure existence, uniqueness, positiveness and boundedness of solutions can be found in [76], and the additional analysis such as the existence of steady-state solutions and how the delay affects stability can be found in [9,11].

Not much progress has been made for population dynamics involving both spatial dispersal and state-dependent delay, which would naturally involve certain types of partial functional differential equations. In [184], a new class of non-local partial functional differential equations is proposed for the evolution of a single species population that involves delayed feedback, where the delay such as the time length for reproduction, is selective and the selection depends on the status of the system. The abstract model in the work [184] is the following non-local partial differential equation with state-dependent selective delay:

$$\frac{\partial}{\partial t} u(t,x) + A u(t,x) + d u(t,x) = \big(F(u_t)\big)(x), \quad x \in \Omega, \tag{2.6.8}$$

where

$$\big(F(u_t)\big)(x) := \int_{-r}^{0} \left\{ \int_{\Omega} b\big(u(t+\theta, y)\big) f(x-y)\, \mathrm{d}y \right\} \xi\big(\theta, \|u(t)\|\big)\, \mathrm{d}\theta,$$

A is a densely-defined self-adjoint positive linear operator with domain $D(A) \subset L^2(\Omega)$ and with compact resolvent, Ω is a smooth bounded domain in $\mathbb{R}^n$, $f : (\Omega - \Omega) \to \mathbb{R}$ is a certain bounded function, $b : \mathbb{R} \to \mathbb{R}$ is a locally Lipschitz map and satisfies $|b(w)| \leqslant C_1 |w| + C_2$ with $C_1 \geqslant 0$ and $C_2 \geqslant 0$, d is a positive constant. The function $u(\cdot,\cdot) : [-r, +\infty) \times \Omega \to \mathbb{R}$ is given such that for any t the function $u(t) \equiv u(t,\cdot) \in L^2(\Omega)$, $\|\cdot\|$ is the $L^2(\Omega)$-norm. The function $\xi : [-r, 0] \times \mathbb{R} \to \mathbb{R}$ represents the state-selective delay, for example, if

$$\xi(\theta, s) = \mathrm{e}^{-\beta(g(s)-\theta)^2}, \quad \theta \in [-r, 0], \ s \in \mathbb{R},$$

then the function g gives the coordinate of the maximum of ξ. Thus the system selects the maximal historical impact on the current change rate according to the system's current state.

2.7. *Cell biology and hematological disorders*

The monograph [162] contains a brief discussion of a modified structured model for the control of the bone marrow stem cell population which supplies the circulating red blood cell population. For this case, the maturing stem cell population is structured by a maturation variable and the rate of maturation is assumed to depend only on the total mature red blood cell population. Since a threshold level of maturation is required in order for an immature cell to enter the population, a threshold delay differential equation arises naturally.

Age-structured models have also been used in the study of hematological disorders, and these models can be reduced via the method of characteristics to systems of threshold type differential delay equations, see [24,26,149,185]. Here we describe the work [149].

We first note that the precursor cells begin from a pool of burst-forming units of the erythroid line (BFU) that have differentiated into a self-sustaining population which eventually leads to the production of mature erythrocytes. At some point, the BFUs further differentiate and start down a proliferative path that can ultimately produce erythrocytes. Early in this proliferative phase of development the hormone Epo, alone with other hormones, affects the number of BFU that become erythrocytes. Increase in the concentration of Epo may increase the number of BFU recruited to mature into erythrocytes. Alternatively, there may be a relatively constant supply of committed BFU, but only the cells tagged with sufficient Epo survive the rapidly proliferating colony forming units (CFU) phase to complete maturation.

Let $p(t,\mu)$ denote the population of precursor cells at time t with age μ, and let $V(E)$ be the velocity of maturation, which may depend on the hormone concentration, E. The maturity level μ for erythropoiesis can represent the accumulation of hemoglobin in the precursor cells. If $S_0(E)$ is the number of cells recruited into the proliferating precursor population, then the entry of new precursor cells into the age-structured model satisfies the boundary condition

$$V(E)p(t,0) = S_0(E). \tag{2.7.1}$$

Let μ_F be the maximum age for a cell reaching maturity, then the dynamics of the precursor cells is governed by the age-structured model

$$\left(\frac{\partial}{\partial t} + V(E)\frac{\partial}{\partial \mu}\right)p(t,\mu) = \bigl[\beta(\mu,E) - H(\mu)\bigr]p(t,\mu), \quad t > 0,\ 0 < \mu < \mu_F, \tag{2.7.2}$$

where $\beta(\mu,E)$ is the net birth rate for proliferating precursor cells, and $H(\mu)$ is the disappearance rate given by

$$H(\mu) = \frac{h(\mu - \bar{\mu})}{\int_{\mu}^{\mu_F} h(s - \bar{\mu})\,\mathrm{d}s}$$

and $h(\mu - \bar{\mu})$ is the distribution of maturity levels of the cells when released into the circulating blood, with $\bar{\mu}$ being the mean age of mature precursor cells and the normalization $\int_0^{\mu_F} h(\mu - \bar{\mu})\,\mathrm{d}\mu = 1$.

Let $m(t, \nu)$ be the population of mature non-proliferating cells at time t and age ν, and assume the mature cells age at a constant rate W. The boundary condition for cells entering the maturation population is given by

$$Wm(t, 0) = V(E) \int_0^{\mu_F} h(\mu - \bar{\mu}) p(t, \mu)\,\mathrm{d}\mu. \tag{2.7.3}$$

The complete feedback involves the Epo level E, its growth is regulated by the total population of mature cells. Therefore, we need to determine the maximum age $\nu_F(t)$ of erythrocytes, which varies in t as the destruction of erythrocytes occurs by active removal of the oldest cells. We assume a constant erythrocyte removal rate Q, which can be justified by either assuming a constant supply of markers or a constant number of phagocytes that become satiated in their destruction of the oldest erythrocytes, then consideration of the moving boundary condition at $\nu = \nu_F(t)$ yields

$$\bigl(W - \nu_F'(t)\bigr)m\bigl(t, \nu_F(t)\bigr) = Q, \tag{2.7.4}$$

or equivalently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\nu_F(t) = W - \frac{Q}{m(t, \nu_F(t))}. \tag{2.7.5}$$

To obtain this unusual moving boundary condition, we notice that when erythrocytes age, their cell membrane breaks down and marcophages destroy the least pliable cells. On the other hand, since we assume the macrophages are in constant supply and are saturated in their consumption of erythrocytes, the age of destruction of erythrocytes then varies. In particular, during the time interval $[t, t + \Delta t]$, one can use the Mean Value Theorem to find $\chi, \eta \in [t, t + \Delta t]$ so that

$$Q\Delta t + \bigl[\nu_F(t + \Delta t) - \nu_F(t)\bigr]m\bigl(\eta, \nu_F(\eta)\bigr) = W\Delta t m\bigl(\chi, \nu_F(\chi)\bigr)$$

from which (2.7.4) follows.

The dynamics of matured cells before its maximum age is governed by the age-structured model

$$\left(\frac{\partial}{\partial t} + W\frac{\partial}{\partial \nu}\right)m(t, \nu) = -W\gamma(\nu)m(t, \nu), \quad t > 0,\ 0 < \nu < \nu_F(t), \tag{2.7.6}$$

where $\gamma(\nu)$ is the age-dependent death rate of mature cells.

To complete the feedback cycle, we note that the Epo level E is governed by a negative feedback of the total population of mature cells

$$M(t) = \int_0^{\nu_F(t)} m(t, \nu)\,\mathrm{d}\nu, \tag{2.7.7}$$

via

$$\frac{\mathrm{d}}{\mathrm{d}t}E(t) = F\big(M(t)\big) - kE(t), \tag{2.7.8}$$

where k is the decay constant for the hormone and $F(M)$ is a monotonically decreasing function of M.

In the simple case, that we consider in the remaining part of this subsection, where $V(E) = W = 1$ and $\beta(\mu, E) = \beta(\mu)$ is independent of E, if the initial condition

$$p(0, \mu) = \phi(\mu), \qquad m(0, \nu) = \psi(\nu)$$

is given, then the method of characteristics yields that

$$p(t, \mu) = \begin{cases} \phi(\mu - t)\exp\left[\int_0^t F(\mu - t + w)\,\mathrm{d}w\right], & t < \mu; \\ S_0\big(E(t-\mu)\big)\exp\left[\int_{t-\mu}^t F(w - t + \mu)\,\mathrm{d}w\right], & t > \mu, \end{cases} \tag{2.7.9}$$

with $F(\mu) = \beta(\mu) - H(\mu)$, and

$$m(t, \nu) = \begin{cases} \psi(\nu - t)\exp\left[-\int_0^t (\nu + \sigma - t)\,\mathrm{d}\sigma\right], & t < \nu; \\ \int_0^{\mu_F} h(\mu - \bar{\mu})p(t - \nu, \mu)\,\mathrm{d}\mu \exp\left[-\int_0^\nu \gamma(\sigma)\,\mathrm{d}\sigma\right], & t > \nu. \end{cases} \tag{2.7.10}$$

If t is sufficiently large, then we get

$$M(t) = \int_0^{\nu_F(t)} \int_0^{\mu_F} h(\mu - \bar{\mu})p(t - \nu, \mu)\,\mathrm{d}\mu \exp\left[-\int_0^\nu \gamma(\sigma)\,\mathrm{d}\sigma\right]\mathrm{d}\nu. \tag{2.7.11}$$

Thus, (2.7.5), (2.7.8) and (2.7.11), with p given in (2.7.9) form a complete system of integro-differential equations with the delay ν_F being adaptive.

3. A framework for the initial value problem

3.1. *Preliminaries*

For differential delay equations with state-dependent delays it is less obvious than in case of time-invariant delays on which state space IVPs are well-posed. For initial data in the familiar space $C = C([-h, 0]; \mathbb{R}^n)$ solutions to equations with state-dependent delay are in general not unique in cases where for similar equations with constant delay the IVP is well-posed. Winston [216] gave the following example: The functions defined by

$$x(t) = t + 1 \quad \text{and} \quad x(t) = t + 1 - t^{3/2}$$

for small $t > 0$ both are solutions of the equation

$$x'(t) = -x\bigl(t - \bigl|x(t)\bigr|\bigr)$$

with initial values

$$x(t) = \begin{cases} -1, & \text{if } t \leqslant -1; \\ \frac{3}{2}(t+1)^{1/3} - 1, & \text{if } -1 < t \leqslant -\frac{7}{8}; \\ \frac{10}{7}t + 1, & \text{if } -\frac{7}{8} < t \leqslant 0. \end{cases}$$

Early results on existence, uniqueness, and continuous dependence for solutions of IVPs with state-dependent delays are due to Driver [57–62], who studied cases of the two-body problem of electrodynamics; see also work by Driver and Norris [64], Travis [198], and Hoag and Driver [111]. The latter investigated equations with both delayed and advanced state-dependent shifted arguments. Winston [216,218] studied uniqueness for special scalar equations, among others. Well-posed initial value problems are the basis for work on periodic solutions, notably by Nussbaum [174], Alt [4], Mallet-Paret and Nussbaum [151, 152,155], Kuang and Smith [136,137,190], Mallet-Paret, Nussbaum and Paraskevopoulos [156], Arino, Hadeler and Hbid [7], Krisztin and Arino [132], Magal and Arino [148], and in [205]. Further existence and uniqueness results are due to Gatica and Waltman [88,89] and Jackiewicz [119,121], and to Ito and Kappel [118] in an approach to more general IVPs. The delay differential and integral equations addressed in these results belong to special classes where the state-dependent delay appears explicitly or is defined implicitly by an additional equation. Typically, the IVP is uniquely solved for initial and other data which satisfy suitable Lipschitz conditions. Manitius [160] and Brokate and Colonius [29] deal with differentiability of operators given by the right-hand side of differential equations with state-dependent delay, in the context of control theory. Section 6 below reports about work of Hartung and Turi [106] who proved differentiability of solutions with respect to parameters—including Lipschitz continuous initial data—in Sobolev spaces and related quasi-normed spaces. Louihi, Hbid, and Arino [144] consider a class of equations with state-dependent delay in the framework of nonlinear semigroup theory. Their results are used in an approach of Ouifki and Hbid [178] to periodic solutions which is described in Section 8.3 below.

Let us now see where the difficulty with uniqueness arises if a given differential equation with state-dependent delay is written in the general form (1.0.1). Our discussion of the uniqueness problem will naturally lead us to a manifold on which Eq. (1.0.1) defines a continuous semiflow with continuously differentiable solution operators. We shall obtain continuously differentiable local stable and unstable manifolds of the semiflow at stationary points, also center manifolds for the solution operators, and a convenient Principle of Linearized Stability, among others.

We begin with an equation

$$x'(t) = g\big(x\big(t - r(x_t)\big)\big), \tag{3.1.1}$$

with a given map $g : \mathbb{R}^n \to \mathbb{R}^n$ and a given delay functional $r : U \to [0, h]$, for $h > 0$, $n \in \mathbb{N}$, and $U \subset (\mathbb{R}^n)^{[-h,0]}$. Equation (3.1.1) has the form (1.0.1) for

$$f = g \circ ev \circ \big(id \times (-r)\big)$$

where

$$ev : \big(\mathbb{R}^n\big)^{[-h,0]} \times [-h, 0] \to \mathbb{R}^n$$

is the evaluation map defined by

$$ev(\phi, s) = \phi(s).$$

Notice that the restriction of ev to $C \times [-h, 0]$ is not locally Lipschitz continuous: Lipschitz continuity would imply Lipschitz continuity of elements $\phi \in C$. Therefore in general f is not locally Lipschitz continuous on open subsets of C, and the familiar results on existence, uniqueness, and dependence of solutions on initial data and parameters for RFDEs from, say, [54,98] fail.

The difficulty just mentioned disappears if C is replaced by the smaller Banach space $C^1 = C^1([-h, 0]; \mathbb{R}^n)$ since the restricted evaluation map

$$Ev : C^1 \times [-h, 0] \ni (\phi, s) \mapsto \phi(s) \in \mathbb{R}^n$$

is continuously differentiable, with

$$\mathrm{D}_1 Ev(\phi, s)\chi = Ev(\chi, s) \quad \text{and} \quad \mathrm{D}_2 Ev(\phi, s)1 = \phi'(s).$$

So, for $g : \mathbb{R}^n \to \mathbb{R}^n$ and $r : U \to [0, h]$, $U \subset C^1$ open, both continuously differentiable, the resulting functional

$$f = g \circ Ev \circ \big(id \times (-r)\big)$$

is continuously differentiable from U to $\mathbb{R}^n$, with

$$\begin{aligned} \mathrm{D}f(\phi)\chi &= \mathrm{D}g\big(\phi\big(-r(\phi)\big)\big)\big[\mathrm{D}_1 Ev\big(\phi, -r(\phi)\big)\chi - \mathrm{D}_2 Ev\big(\phi, -r(\phi)\big)\mathrm{D}r(\phi)\chi\big] \\ &= \mathrm{D}g\big(\phi\big(-r(\phi)\big)\big)\big[\chi\big(-r(\phi)\big) - \mathrm{D}r(\phi)\chi\phi'\big(-r(\phi)\big)\big] \end{aligned}$$

for $\phi \in U$ and $\chi \in C^1$.

However, yet another obstacle is in the way. Suppose $U \subset C^1$ is open, $f : U \to \mathbb{R}^n$ is continuously differentiable, and the IVP

$$x'(t) = f(x_t) \quad \text{for } t > 0, \quad x_0 = \phi, \tag{3.1.2}$$

is well-posed for $\phi \in U$. A solution $x : [-h, t_e) \to \mathbb{R}^n$, $0 < t_e \leqslant \infty$, has segments $x_t \in U \subset C^1$, $0 \leqslant t < t_e$. Therefore x is continuously differentiable, and the curve $[0, t_e) \ni t \mapsto x_t \in C^1$ is continuous. Continuity at $t = 0$ yields

$$\phi'(0) = x'(0) = f(x_0) = f(\phi),$$

a necessary condition on initial data which may or may not be satisfied.

In any case, the last equation suggests to consider the IVP (3.1.2) for initial data only in the closed subset

$$X_f = \{\phi \in U\colon\ \phi'(0) = f(\phi)\}$$

of $U \subset C^1$.

From now on, a solution is understood to be a continuously differentiable function $x : [-h, t_*) \to \mathbb{R}^n$, $0 < t_* \leqslant \infty$, which satisfies $x_t \in U$ for $0 \leqslant t < t_*$, $x_0 = \phi$, and $x'(t) = f(x_t)$ for $0 < t < t_*$.

Incidentally, notice that X_f is a nonlinear analogue of the domain

$$\{\phi \in C^1\colon\ \phi'(0) = L\phi\}$$

of the generator of the semigroup given by the solutions to the linear IVP

$$y'(t) = Ly_t, \quad y_0 = \phi \in C,$$

for $L : C \to \mathbb{R}^n$ linear continuous [54,98].

3.2. *The semiflow on the solution manifold*

There is a mild smoothness condition (S) on the functional $f : U \to \mathbb{R}^n$, $U \subset C^1$ open, which is often satisfied if f represents an equation with state-dependent delay, and which implies all the desired results, namely

(S1) f is continuously differentiable,
(S2) each derivative $\mathrm{D}f(\phi)$, $\phi \in U$, extends to a linear map $\mathrm{D}_e f(\phi) : C \to \mathbb{R}^n$, and
(S3) the map

$$U \times C \ni (\phi, \chi) \mapsto \mathrm{D}_e f(\phi)\chi \in \mathbb{R}^n$$

is continuous.

Let us see what condition (S) means for the example

$$f = g \circ Ev \circ \big(id \times (-r)\big)$$

above, with $r : U \to [0, h]$, $U \subset C^1$ open, and $g : \mathbb{R}^n \to \mathbb{R}^n$ continuously differentiable. Suppose the delay functional r satisfies condition (S) (with $n = 1$). If then $\mathrm{D}_e f$ is defined by the formula above for $\mathrm{D}f$, with $\mathrm{D}r$ replaced by $\mathrm{D}_e r$, one sees that condition (S) holds for f. Notice that r satisfies (S) provided it is the restriction of a continuously differentiable map $V \to [0,h]$, $V \subset C$ open, to $U = V \cap C^1$.

THEOREM 3.2.1. *Suppose $U \subset C^1$ is open, $f : U \to \mathbb{R}^n$ has property* (S), *and $X_f \neq \emptyset$. Then X_f is a continuously differentiable submanifold of U with codimension n, and each $\phi \in X_f$ uniquely defines a noncontinuable solution $x^\phi : [-h, t_+(\phi)) \to \mathbb{R}^n$ of the IVP* (3.1.2). *All segments x_t^ϕ, $0 \leqslant t < t_+(\phi)$ and $\phi \in X_f$, belong to X_f, and the relations*

$$F(t,\phi) = x_t^\phi, \quad \phi \in X_f,\ 0 \leqslant t < t_+(\phi)$$

define a domain $\Omega \subset \mathbb{R} \times X_f$ and a continuous semiflow $F : \Omega \to X_f$. Each map

$$F(t,\cdot) : \{\phi \in X_f\colon\ (t,\phi) \in \Omega\} \to X_f$$

is continuously differentiable, and for all $(t,\phi) \in \Omega$ and $\chi \in T_\phi X_f$ we have

$$\mathrm{D}_2 F(t,\phi)\chi = v_t^{\phi,\chi}$$

with the solution $v^{\phi,\chi} : [-h, t_+(\phi)) \to \mathbb{R}^n$ of the linear IVP

$$v'(t) = \mathrm{D}f\bigl(F(t,\phi)\bigr)v_t, \quad v_0 = \chi. \tag{3.2.1}$$

At each (t,ϕ) with $\phi \in X_f$ and $h < t < t_+(\phi)$, the partial derivative $\mathrm{D}_1 F(t,\phi)$ *exists, and*

$$\mathrm{D}_1 F(t,\phi)1 = \bigl(x_t^\phi\bigr)'.$$

The restriction of F to the submanifold $\{(t,\phi) \in \Omega\colon\ h < t\}$ of $\mathbb{R} \times X_f$ is continuously differentiable.

Notice that the tangent spaces of the manifold X_f are given by

$$T_\phi X_f = \{\chi \in C^1\colon\ \chi'(0) = \mathrm{D}f(\phi)\chi\}.$$

Theorem 3.2.1 is proved in [206,207]. The construction of the semiflow is also sketched in [209].

The first part (S1) of the hypothesis and continuity of each extension $\mathrm{D}_e f(\phi)$ suffice for X_f to be a continuously differentiable submanifold. The proof of this is simple: For

$$p : C^1 \ni \phi \mapsto \phi'(0) \in \mathbb{R}^n,$$

$$X_f = (p - f)^{-1}(0),$$

and the Implicit Function Theorem yields local representations of X_f as graphs over the kernels of the linear maps $\mathrm{D}(p - f)(\phi) = p - \mathrm{D}f(\phi))$, $\phi \in X_f$, provided these linear maps are surjective. In case $n = 1$ surjectivity follows since using the continuity of $\mathrm{D}_e f(\phi)$ at $0 \in C$ one finds $\chi \in C^1 \subset C$ with $\chi'(0) = 1$ and $\mathrm{D}f(\phi)\chi = \mathrm{D}_e f(\phi)\chi < 1 = p\chi$, which gives $(p - \mathrm{D}f(\phi))C^1 = \mathbb{R}$. In case $n > 1$ similar arguments yield a basis of $\mathbb{R}^n$ in $(p - \mathrm{D}f(\phi))C^1$.

Condition (S3) implies that the map $D_e f$ is locally bounded. From this one derives easily the following local Lipschitz property:

(L) For every $\phi \in U$ there are a neighbourhood V and $L \geqslant 0$ with $|f(\psi) - f(\chi)| \leqslant L\|\psi - \chi\|_C$ for all $\psi \in V$, $\chi \in V$,

see Corollary 1 in [207]. Notice that the last norm is the norm on C and not the larger norm on the smaller space C^1. So (L) is not a consequence of (S1). Property (L) (together with (S1), (S2) and continuity of each $\mathrm{D}_e f(\phi)$) yields existence and continuity of the semiflow as well as the properties of the maps $F(t,\cdot)$ stated in Theorem 3.2.1. Only for the two last statements of Theorem 3.2.1, on smoothness, the full hypothesis (S) is needed.

It is worth noting that continuity of the map

$$U \ni \phi \mapsto \mathrm{D}_e f(\phi) \in L\big(C, \mathbb{R}^n\big),$$

which seems only slightly stronger than property (S3) above, does in general *not* hold for functionals f which represent differential equations with state-dependent delay.

A key issue in the proof of Theorem 3.2.1 is how the local Lipschitz property (L) is used for the construction of local solutions to the IVP. Let us briefly explain this.

In order to solve the integrated version

$$x(t) = \phi(0) + \int_0^t f(x_s)\,\mathrm{d}s, \quad x_0 = \phi \in X_f,$$

of the IVP for $0 \leqslant t \leqslant T$ by a continuously differentiable map

$$x : [-h, T] \to \mathbb{R}^n,$$

which is continuously differentiable with respect to ϕ, the desired solution is first written as the sum of a linear, continuously differentiable continuation $\hat{\phi}$ of ϕ and of a function u which is zero on $[-h, 0]$. The fixed point problem for u is

$$u(t) = \int_0^t \big(f\big(u_s + \hat{\phi}_s\big) - f(\phi)\big)\,\mathrm{d}s, \quad 0 \leqslant t \leqslant T.$$

The advantage of this formulation is that for the operator $u \mapsto A(\phi, u)$ given by the right hand side dependence on ϕ is more explicit than in the equation for x. For $A(\phi,\cdot)$ to become a contraction with respect to the norm $\|u\|_{T,1} = \max_{0\leqslant s\leqslant T}|u(s)| + \max_{0\leqslant s\leqslant T}|u'(s)|$ we use the following estimate for $v = A(\phi, u)$, $\overline{v} = A(\phi, \overline{u})$, and $0 \leqslant t \leqslant T$:

$$\big|v'(t) - \overline{v}'(t)\big| = \big|f\big(u_t + \hat{\phi}_t\big) - f\big(\overline{u}_t + \hat{\phi}_t\big)\big| \leqslant L\|u_t - \overline{u}_t\|_C$$

(due to the local Lipschitz property, for T and u and $\overline{u}$ small)

$$\leqslant L \max_{0\leqslant s\leqslant T} \bigl|u(s) - \overline{u}(s)\bigr| = L \max_{0\leqslant s\leqslant T} \left| \int_0^s \bigl(u'(r) - \overline{u}'(r)\bigr)\,\mathrm{d}r \right| \leqslant LT\|u - \overline{u}\|_{T,1}.$$

The proof that the semiflow F is continuously differentiable on the manifold given by $t > h$ is based on growth estimates and smoothness properties of solutions to the IVP

$$v'(t) = \mathrm{D}_e f\bigl(F(t,\phi)\bigr)v_t, \quad v_0 = \chi \tag{3.2.2}$$

for initial data $\chi \in C$.

The previous result is optimal in the sense that typically the semiflow F has no partial derivatives with respect to the first variable for $0 \leqslant t < h$.

The framework presented up to here is instrumental in the proof in [208] that for a certain differential system with state-dependent delay, which models position control by echo, hyperbolic stable periodic orbits exist.

With regard to results for solutions $x : [-h, t_+) \to \mathbb{R}^n$ to equations of the form (3.1.1) and generalizations thereof in the sense that x is continuous but differentiable only for $0 < t < t_+$, notice that for $h \leqslant t < t_+$ we have

$$x_t \in X_f$$

(with $f = g \circ ev \circ (id \times (-r))$ in case of Eq. (3.1.1)), so all dynamical properties like structure and stability of invariant sets (stationary points, periodic orbits, unstable manifolds, global attractors, ...) are determined by the semiflow F on X_f—provided the mild smoothness condition (S) holds.

Some notions of the approach to well-posedness and smoothness which we described here are related to ideas from earlier work. The Lipschitz property (L) from the proof of Theorem 3.2.1 was used before in [205] and is analogous to the notion of being *locally almost Lipschitzian* from [156]. The condition that (S1) holds and each $\mathrm{D}_e f(\phi)$ is continuous was introduced in [156] as *almost Fréchet differentiability*. It also is a special case of a smoothness condition used in [129]. Sets analogous to X_f were considered in [134] as state space for neutral functional differential equations, and by Louihi, Hbid and Arino [144]. In [144] the IVP for a certain class of equations of the form (3.1.1) defines a semigroup of locally Lipschitz continuous solution operators on the space $C^{0,1}$. The positively invariant subset $E \subset C^{0,1}$ of data $\phi \in C^1$ with $\phi'(0) = g(\phi(-r(\phi)))$ is identified as the domain of strong continuity of the semigroup. In the present notation, $E = X_f$ for $f = g \circ Ev \circ (id \times (-r))$. It is stated in [144] that E is a Lipschitz manifold.

3.3. *Compactness*

Recall that a map from a subset of a Banach space into a Banach space is called compact if images of bounded sets have compact closure. A simple compactness result for the semiflow of Theorem 3.2.1 is the following. Notice in which way the hypotheses strengthen property (S).

PROPOSITION 3.3.1. *Suppose $f : U \to \mathbb{R}^n$, $U \subset C^1$ open, is bounded, satisfies condition* (S), *and*

(Lb) *for every bounded set $B \subset U$ there exists $L_B \geqslant 0$ with $|f(\phi) - f(\psi)| \leqslant L_B \|\phi - \psi\|_C$ for all ϕ, ψ in B.*

Then all maps $F(t, \cdot)$, $t \geqslant h$, are compact.

PROOF. Let $t \geqslant h$ and let a bounded set $B \subset \{\phi \in X_f\colon (t, \phi) \in \Omega\}$ be given. In order that $\overline{F(t, B)} \subset C^1$ be compact it is enough to show that every sequence in $F(t, B)$ has a convergent subsequence, which follows by means of the Ascoli–Arzelà Theorem provided both sets $M = \{x_t^\phi \in C^1\colon \phi \in B\}$ and $M' = \{(x_t^\phi)' \in C\colon \phi \in B\}$ are bounded with respect to the norm $\|\cdot\|_C$ and equicontinuous. The boundedness of B and f in combination with Eq. (1.0.1) shows that

$$b = \sup_{\phi \in B, -h \leqslant s \leqslant t} \left|\left(x^\phi\right)'(s)\right| < \infty.$$

Then the boundedness of B and integration yield

$$\sup_{\phi \in B, -h \leqslant s \leqslant t} \left|x^\phi(s)\right| < \infty.$$

In particular, M' and M are bounded with respect to $\|\cdot\|_C$. The boundedness of M' yields the equicontinuity of M. We also have that the set

$$Y = \left\{F(s, \phi)\colon \phi \in B, 0 \leqslant s \leqslant t\right\} \subset X_f \subset U$$

is bounded with respect to $\|\cdot\|_{C^1}$. Now equicontinuity of M' follows from the estimate

$$\begin{aligned}
\left|\left(x_t^\phi\right)'(s) - \left(x_t^\phi\right)'(u)\right| &= \left|\left(x^\phi\right)'(t+s) - \left(x^\phi\right)'(t+u)\right| \\
&= \left|f\left(F(t+s, \phi)\right) - f\left(F(t+u, \phi)\right)\right| \\
&\leqslant L_Y \left\|F(t+s, \phi) - F(t+u, \phi)\right\|_C \\
&= L_Y \max_{-h \leqslant v \leqslant 0} \left|x^\phi(t+s+v) - x^\phi(t+u+v)\right| \\
&= L_Y \max_{-h \leqslant v \leqslant 0} \left|\int_{t+u+v}^{t+s+v} (x^\phi)'(w)\,\mathrm{d}w\right| \leqslant L_Y b|s-u|
\end{aligned}$$

for all $\phi \in B$ and all s, u in $[-h, 0]$. □

In case of the example $f = g \circ Ev \circ (id \times (-r))$ with $r : U \to [0, h]$, $U \subset C^1$ open, and $g : \mathbb{R}^n \to \mathbb{R}^n$ continuously differentiable the hypotheses of Proposition 3.3.1 are satisfied if in addition U is convex, r has property (S), g is bounded, and $\mathrm{D}_e r$ and $\mathrm{D}g$ map bounded sets onto bounded sets: The boundedness of f is obvious. The formula which computes $\mathrm{D}_e f$ from $\mathrm{D}g$ and $\mathrm{D}_e r$ shows that $\mathrm{D}_e f$ maps bounded sets onto bounded subsets. For a given bounded set $B \subset U$ there is a larger convex bounded subset $B^* \subset U$, and for ϕ, ψ in B^* we have

$$\begin{aligned}\left|f(\phi) - f(\psi)\right| &= \left|\int_0^1 \mathrm{D}f\big(\psi + t(\phi - \psi)\big)(\phi - \psi)\,\mathrm{d}t\right| \\ &= \left|\int_0^1 \mathrm{D}_e f\big(\psi + t(\phi - \psi)\big)(\phi - \psi)\,\mathrm{d}t\right| \\ &\leqslant \int_0^1 \left|\mathrm{D}_e f\big(\psi + t(\phi - \psi)\big)(\phi - \psi)\right|\mathrm{d}t \\ &\leqslant \sup_{\chi \in B^*} \left\|\mathrm{D}_e f(\chi)\right\| \|\phi - \psi\|_C,\end{aligned}$$

which yields property (Lb).

For another result on compactness, for solution operators on the space $C^{0,1}$, see [144].

3.4. *Linearization at equilibria*

Theorem 3.2.1 reveals in particular how to linearize semiflows defined by differential equations with state-dependent delay, an issue which had been mysterious before. Consider $f : U \to \mathbb{R}^n$, $U \subset C^1$ open, with property (S). Let a stationary point $\phi_0 \in X_f$ of the semiflow F from Theorem 3.2.1 be given. The linearization of F at ϕ_0 is the semigroup T of the linear continuous operators $T(t) = \mathrm{D}_2 F(t, \phi_0)$, $t \geqslant 0$, on the Banach space $T_{\phi_0} X_f$ with the norm $\|\cdot\|_{C^1}$. T is strongly continuous since the solutions $v^{\phi_0,\chi} : [-h, \infty) \to \mathbb{R}^n$ of the IVP (3.2.1) with $\chi \in T_{\phi_0} X_f$ are continuously differentiable.

Before the present approach was available a heuristic technique had been developed in order to circumvent the linearization problem in studies of local stability and instability properties of equilibria. This technique associates to the given nonlinear equation an auxiliary linear equation in the following way: First the delay is frozen at equilibrium, then the resulting nonlinear equation with constant delay is linearized. Of course, this makes sense only for equations where the delay appears explicitly, like, e.g., in Eq. (3.1.1). Let us use Eq. (3.1.1) to show that the auxiliary equation found by the heuristic technique is

$$v'(t) = \mathrm{D}_e f(\phi_0) v_t$$

in our framework. Suppose for simplicity that $n = 1$, that $g : \mathbb{R} \to \mathbb{R}$ satisfies $g(0) = 0$ and that g and the delay functional $r : U \to [0, h]$, $U \subset C^1$ open, are continuously differentiable. Freezing the delay in Eq. (3.1.1) at $\phi_0 = 0$ and linearizing the resulting RFDE

$$x'(t) = g\big(x\big(t - r(0)\big)\big)$$

at the zero solution yields

$$v'(t) = g'(0) v\big(t - r(0)\big).$$

On the other hand, for the functional $f_{g,r} = g \circ Ev \circ (id \times (-r))$ and $\chi \in C^1$ the computation of derivatives above yields

$$\mathrm{D} f_{g,r}(0)\chi = \mathrm{D}g(0)\chi\big(-r(0)\big) = g'(0)\chi\big(-r(0)\big).$$

Obviously the right-hand side of this equation defines a continuous linear extension $\mathrm{D}_e f_{g,r}(0) : C \to \mathbb{R}$ of $\mathrm{D} f_{g,r}(0)$, which verifies our previous statement. In other words, the auxiliary equation found by the heuristic method yields the true linear variational equation by restriction to the tangent space $T_0 X_f$.

In the sequel we clarify the relation between the spectral properties of the linearization T of the semiflow F of Theorem 3.2.1 at a stationary point $\phi_0 \in X_f$ and the strongly continuous semigroup T_e on the space C which is defined by the solutions of the IVP (3.2.2). Recall that the generator $G_e : dom \to C$ of T_e is given by $dom = \{\phi \in C^1 \colon\ \phi'(0) = D_e f(\phi_0)\phi\}$ and $G_e\phi = \phi'$.

We have $T_{\phi_0} X_f = dom$, $T(t)\phi = T_e(t)\phi$ for $t \geqslant 0$ and $\phi \in dom$, and the norm $\|\cdot\|_{C^1}$ coincides with the graph norm $\|\cdot\|_e = \|\cdot\|_C + \|G_e \cdot\|_C$ of the operator G_e.

It is a simple general fact that for a strongly continuous semigroup S of linear continuous operators on a Banach space B, with generator $A : D_A \to B$, the induced linear operators

$$D_A \ni x \mapsto S(t)x \in D_A$$

are continuous with respect to the graph norm $\|\cdot\|_A = \|\cdot\| + \|A\cdot\|$ and form a strongly continuous semigroup on the Banach space $(D_A, \|\cdot\|_A)$, with the generator A_d defined on the domain

$$D_d = \{x \in D_A \colon\ Ax \in D_A\}$$

of A^2 and given by $A_d x = Ax$. Proofs are immediate from $S(t)D_A \subset D_A$ and

$$S(t)Ax = AS(t)x \quad \text{on } D_A$$

for all $t \geqslant 0$.

Our semigroups and their generators are precisely in the relation just described, so we have

$$\begin{aligned} D = (D_d =) &\left\{\chi \in C^1 \colon\ \chi'(0) = \mathrm{D}f(\phi_0)\chi,\, \chi' \in C^1,\, \chi''(0) = \mathrm{D}f(\phi_0)\chi'\right\} \\ &= \left\{\chi \in C^2 \colon\ \chi'(0) = \mathrm{D}f(\phi_0)\chi,\, \chi''(0) = \mathrm{D}f(\phi_0)\chi'\right\} \end{aligned}$$

and $G\chi = \chi'$ on D.

We return to the general case. Below, ρ and σ denote resolvent sets and spectra, respectively.

PROPOSITION 3.4.1. *Suppose B is a Banach space over $\mathbb{C}$, S is a strongly continuous semigroup on B with generator $A : D_A \to B$, and A_d is the generator of the semigroup induced on the Banach space $(D_A, \|\cdot\|_A)$. Then the following holds.*

(i) *$\rho(A) \subset \rho(A_d) \subset \rho(A) \cup \{\lambda \in \mathbb{C}\colon\ A - \lambda I$ injective and $(A - \lambda I)D_A \neq B\}$, and $\sigma(A_d) \subset \sigma(A)$. For all $\lambda \in \rho(A)$, $(A_d - \lambda I)^{-1}x = (A - \lambda I)^{-1}x$ on D_A.*

(ii) *Let γ be a simple closed curve in $\rho(A)$ and let $\sigma_{\gamma,d} = int(\gamma) \cap \sigma(A_d)$, $\sigma_\gamma = int(\gamma) \cap \sigma(A)$. Then the spectral projections P_d, P and generalized eigenspaces $\mathcal{G}_d = P_d D_A$, $\mathcal{G} = PB$ associated with the sets $\sigma_{\gamma,d} \subset \sigma(A_d)$ and $\sigma_\gamma \subset \sigma(A)$, respectively, satisfy*

$$P_d x = Px \quad \text{on } D_A, \quad \mathcal{G}_d \subset \mathcal{G}, \quad \text{and} \quad \overline{\mathcal{G}_d}^B = \mathcal{G}.$$

PROOF. 1. Proof of (i).

1.1. Let $\lambda \in \rho(A)$ be given. The continuity of the inverse with respect to the norm on B and the definition of the graph norm combined show that for all $y \in D_A$ we have

$$\begin{aligned}\left\|(A - \lambda I)^{-1}y\right\|_A &= \left\|(A - \lambda I)^{-1}y\right\| + \left\|A(A - \lambda I)^{-1}y\right\| \\ &\leqslant \left\|(A - \lambda I)^{-1}y\right\| + \left\|(A - \lambda I)(A - \lambda I)^{-1}y\right\| \\ &\quad + \left\|\lambda(A - \lambda I)^{-1}y\right\| \\ &\leqslant \big((1 + |\lambda|)\left\|(A - \lambda I)^{-1}\right\| + 1\big)\|y\|.\end{aligned}$$

It follows that $(A - \lambda I)^{-1}$ defines a linear continuous map R from $(D_A, \|\cdot\|_A)$ into itself. The range of R contains the domain D_d of the generator A_d since for any $x \in D_A$ with $Ax \in D_A$ we have that $y = (A_d - \lambda I)x = (A - \lambda I)x \in D_A$, hence $x = (A - \lambda I)^{-1}y \in (A - \lambda I)^{-1}D_A = RD_A$.

Conversely, for $x = Ry = (A - \lambda I)^{-1}y$ with $y \in D_A$ we have $x \in D_A$ and $(A - \lambda I)x = y \in D_A$, hence $Ax = \lambda x + y \in D_A$, or $x \in D_d$.

So R maps D_A injectively onto the domain D_d of A_d. We have $R(A_d - \lambda I)y = R(A - \lambda I)y = y$ on D_d, $(A_d - \lambda I)Rx = (A - \lambda I)Rx = x$ on D_A, and it follows that $\lambda \in \rho(A_d)$ and $(A_d - \lambda I)^{-1}y = (A - \lambda I)^{-1}y$ on D_A.

The shown inclusion of resolvent sets yields the asserted result for spectra.

1.2. Proof of the remaining inclusion. Let $\lambda \in \rho(A_d)$ be given. $A - \lambda I$ is injective since $(A - \lambda I)x = 0$ implies $Ax \in D_A$, or $x \in D_d$, and therefore $(A_d - \lambda I)x = (A - \lambda I)x = 0$, which in turn gives $x = 0$, by $\lambda \in \rho(A_d)$.

Consider the injective continuous map

$$A_c\colon \big(D_A, \|\cdot\|_A\big) \ni x \mapsto (A - \lambda I)x \in B.$$

In case $(A - \lambda I)D_A = B$ the open mapping theorem shows that the inverse of A_c is continuous, which implies that

$$B \ni y \mapsto (A - \lambda I)^{-1}y \in B$$

is continuous. So in this case, $\lambda \in \rho(A)$, and the assertion becomes obvious.

2. Proof of (ii). The representation of $P_d x$ and Px, $x \in D_A$, as contour integrals in $(D_A, \| \cdot \|_A)$ and B, respectively, along the curve γ with integrands given by

$$(A_d - \lambda I)^{-1} x = (A - \lambda I)^{-1} x, \quad \lambda \in |\gamma|,$$

shows that $P_d x = Px$ on D_A. Consequently,

$$\mathcal{G}_d = P_d D_A \subset P D_A \subset PB = \mathcal{G},$$

and finally

$$\overline{\mathcal{G}_d}^B = \overline{P_d D_A}^B = \overline{P D_A}^B = PB$$

(since $\overline{D_A} = B$ and since P has closed range)

$$= \mathcal{G}. \qquad \square$$

It is easy to see that for a given semiflow S on a Banach space B over $\mathbb{R}$ and its generator $A : D_A \to B$, the complexification $A_{\mathbb{C}}$ of A coincides with the generator of the complexified semigroup $S_{\mathbb{C}} : t \mapsto S(t)_{\mathbb{C}}$, and that the semigroup induced by $S_{\mathbb{C}}$ on $(D_A)_{\mathbb{C}}$ is generated by the complexification of the generator $A_d : D_d \to (D_A, \| \cdot \|_A)$.

Returning to our case of semigroups T_e on C, T on $T_{\phi_0} X_f$ and their generators G_e, G, respectively, we recall that the embedding $(dom, \| \cdot \|_e) \to C$ is compact, by the Ascoli–Arzelà theorem. This yields that all resolvents of $(G_e)_{\mathbb{C}}$, which define continuous maps from $C_{\mathbb{C}}$ onto the complexification of the Banach space $(dom, \| \cdot \|_e)$, are compact. Therefore $\sigma(G_e) := \sigma((G_e)_{\mathbb{C}})$ is discrete and consists of eigenvalues with finite-dimensional generalized eigenspaces. Using Proposition 3.4.1 and the remarks following it we infer

$$\rho(G) = \rho(G_e), \qquad \sigma(G) = \sigma(G_e),$$

and for the spectral projections and generalized eigenspaces of $G_{\mathbb{C}}$ and $(G_e)_{\mathbb{C}}$ which are associated with $\lambda \in \sigma(G) = \sigma(G_e)$,

$$P(\lambda)\chi = P_e(\lambda)\chi \quad \text{on } T_{\phi_0} X_f = dom,$$

$$\mathcal{G}(\lambda) = \mathcal{G}_e(\lambda).$$

Recall that to the right of any line parallel to the imaginary axis there are at most a finite number of eigenvalues of G_e. Let C_u and C_c denote the unstable and center spaces of G_e, i.e., the finite-dimensional realified generalized eigenspaces given by the eigenvalues with positive real part and on the imaginary axis, respectively. The stable space C_s of G_e is the realified generalized eigenspace given by the spectrum with negative real part. More precisely, C_s is the realification of the space $(id - P_{\geqslant 0}) C_{\mathbb{C}}$ where $P_{\geqslant 0} : C_{\mathbb{C}} \to \mathbb{C}_{\mathbb{C}}$ denotes the spectral projection associated with the finite spectral set of all $\lambda \in \sigma(G_e)$ with $Re\lambda \geqslant 0$. We have

$$C_u \oplus C_c \subset D \subset dom = T_{\phi_0} X_f$$

while the infinite-dimensional, complementary space C_s is not contained in *dom*. Using the previous relations between spectral projections and generalized eigenspaces of the generators G and G_e it follows easily that the unstable and center spaces of G coincide with C_u and C_c, respectively, and that the stable space of G is

$$C_s \cap dom.$$

For each map $F(t,\cdot)$, $t > 0$, we have

$$\sigma\big(\mathrm{D}F(t,\cdot)(\phi_0)\big) = \sigma\big(T(t)\big) \subset \{0\} \cup \big\{\mathrm{e}^{zt}\colon\ z \in \sigma(G)\big\}, \tag{3.4.1}$$

and the linear unstable, center, and stable spaces for the derivative $\mathrm{D}F(t,\cdot)(\phi_0) = \mathrm{D}_2F(t,\phi_0) = T(t)$, which are defined by the eigenvalues outside, on, and inside the unit circle, are

$$C_u,\ C_c,\ \text{and } C_s \cap T_{\phi_0}X_f,\ \text{respectively.}$$

3.5. *Local invariant manifolds at stationary points*

Consider $f: U \to \mathbb{R}^n$, $U \subset C^1$ open, with property (S) and the associated semiflow $F: \Omega \to X_f$ of Theorem 3.2.1. It is convenient to set

$$\Omega_t = \big\{\phi \in X_f\colon\ (t,\phi) \in \Omega\big\} \quad \text{and} \quad F_t = F(t,\cdot),$$

for each $t \geqslant 0$. A stationary point $\phi_0 \in X_f$ of the semiflow F is a fixed point for each map $F_t: \Omega_t \to X_f$, $t > 0$. Results about local invariant manifolds for continuously differentiable maps on open subsets of Banach spaces [97,44,171,133] yield continuously differentiable local stable, center, and unstable manifolds

$$W_s \subset X_f, \quad W_c \subset X_f, \quad W_u \subset X_f$$

of F_t $t > 0$, at the fixed point ϕ_0, with tangent spaces at ϕ_0 given by

$$T_{\phi_0}W_s = C_s \cap T_{\phi_0}X_f, \qquad T_{\phi_0}W_c = C_c, \qquad T_{\phi_0}W_u = C_u.$$

In the sequel we recall details of this and use them to show that local stable manifolds W_s of any map F_a, $a > 0$, provide local stable manifolds of the semiflow F at the given stationary point. The last fact is familiar for flows of continuously differentiable vector fields in finite dimensions, but here we are concerned with semiflows on a Banach manifold of infinite dimension, and some care should be taken which properties of the underlying delay differential equation (1.0.1) ensure the result.

For local unstable manifolds an analogous result holds, and the proof is similar. Existence of smooth local center manifolds for the semiflow is more difficult, and the relationship between center manifolds of its time-t-maps and center manifolds for the semiflow is more subtle. This will be discussed in the next section.

So let a stationary point $\phi_0 \in X_f$ of the semiflow F be given. Recall from the preceding subsection the linear stable, center, and unstable spaces

$$C_s \cap T_{\phi_0} X_f, C_c, C_u$$

of the generator G of the semigroup T on $T_{\phi_0} X_f$. Choose $\beta > 0$ so that

$$-\beta > \max\left\{Re\, z\colon\ z \in \sigma(G), Re\, z < 0\right\}.$$

Let $a > 0$. In order to introduce the local stable manifold of the map $F_a : \Omega_a \to X_f$ at the fixed point ϕ_0 we use a manifold chart of X_f at ϕ_0. Let $Y = T_{\phi_0} X_f$. There is a subspace $E \subset C^1$ of dimension n which is a complement of Y in C^1. Let $P : C^1 \to C^1$ denote the projection along E onto Y. Then the equation $K(\phi) = P(\phi - \phi_0)$ defines a manifold chart on an open neighbourhood V of ϕ_0 in $\Omega_a \subset X_f$, with $Y_0 = K(V)$ an open neighbourhood of $0 = K(\phi_0)$ in the Banach space Y (with the norm given by $\|\cdot\|_{C^1}$). The inverse of K is given by a continuously differentiable map $R : Y_0 \to C^1$. Both derivatives $\mathrm{D}K(\phi_0)$ and $\mathrm{D}R(0)$ are given by the identity on Y. We may assume that there is a Lipschitz constant $L_R \geqslant 0$ so that

$$\left\|R(\chi) - R(\psi)\right\|_{C^1} \leqslant L_R \|\chi - \psi\|_{C^1} \quad \text{for all } \chi \in Y_0, \psi \in Y_0.$$

Choose an open neighbourhood $Y_1 \subset Y_0$ of 0 in Y with $F_a(R(Y_1)) \subset V$. In local coordinates the map F_a is represented by the continuously differentiable map

$$H : Y_1 \ni \chi \mapsto K\bigl(F_a\bigl(R(\chi)\bigr)\bigr) \in Y.$$

Obviously, $H(0) = 0$, $\mathrm{D}H(0) = \mathrm{D}F_a(\phi_0) = T(a)$, and $H(Y_1) \subset Y_0$. Using the last statement of the preceding subsection we infer that the linear stable, center, and unstable spaces of H at its fixed point $0 \in Y_1$ are $C_s \cap Y$, C_c, and C_u, respectively. Set $\lambda = \mathrm{e}^{-a\beta}$. Then (3.4.1) gives

$$\max\left\{|\zeta|\colon\ \zeta \in \sigma\bigl(\mathrm{D}H(0)\bigr), |\zeta| < 1\right\} < \lambda < 1.$$

The Stable Manifold Theorem (see Theorem I.2 in [133]) yields the following result.

PROPOSITION 3.5.1. *There exist* $\alpha \in (0, \lambda)$, *convex open neighbourhoods* $C_{s,2}$ *of* 0 *in* $C_s \cap Y$ *and* $C_{cu,2}$ *of* 0 *in* $C_c \oplus C_u$ *with* $N = C_{s,2} + C_{cu,2} \subset Y_2$, *a continuously differentiable map* $w : C_{s,2} \to C_{cu,2}$ *with* $w(0) = 0$ *and* $\mathrm{D}w(0) = 0$, *and an equivalent norm* $\|\cdot\|_H$ *on* Y *such that the following holds.*

(i) *The graph* $W = \{\chi + w(\chi)\colon\ \chi \in C_{s,2}\}$ *is equal to the set of all initial points* $\psi = \psi_0$ *of trajectories* $(\psi_j)_0^\infty$ *of* H *which satisfy* $\lambda^{-j}\psi_j \in N$ *for all* $j \in \mathbb{N}_0$ *and* $\lambda^{-j}\psi_j \to 0$ *as* $j \to \infty$.

(ii) $H(W) \subset W$.

(iii) $\|H(\phi) - H(\psi)\|_H \leqslant \alpha\|\phi - \psi\|_H$ *for all* $\psi \in W$, $\phi \in W$.

(iv) *For every trajectory* $(\psi_j)_0^\infty$ *of* H *with* $\lambda^{-j}\psi_j \in N$ *for all* $j \in \mathbb{N}_0$,

$$\psi_0 \in W.$$

Here, trajectories are defined by the equations $\psi_{j+1} = H(\psi_j)$ for all integers $j \geqslant 0$. The local stable manifold of F_a at ϕ_0 is the continuously differentiable submanifold

$$W_s = R(W)$$

of X_f. Obviously, $W_s \subset V$, $\phi_0 \in W_s$, and $T_{\phi_0}W_s = C_s \cap Y = C_s \cap T_{\phi_0}X_f$.

COROLLARY 3.5.2.

(i) $F_a(W_s) \subset W_s$, *and each neighbourhood of* ϕ_0 *in* W_s *contains a neighbourhood* $W_{s,1}$ *of* ϕ_0 *in* W_s *with* $F_a(W_{s,1}) \subset W_{s,1}$.

(ii) *There exists* $c_s \geqslant 0$ *so that for every trajectory* $(\psi_j)_0^\infty$ *of* F_a *in* W_s *and for all integers* $j \geqslant 0$,

$$\|\psi_j - \phi_0\|_{C^1} \leqslant c_s\alpha^j\|\psi_0 - \phi_0\|_{C^1}.$$

PROOF. 1. The first inclusion in assertion (i) follows from

$$K\bigl(F_a(W_s)\bigr) = K\bigl(F_a\bigl(R(W)\bigr)\bigr) = H(W) \subset W = K\bigl(R(W)\bigr)$$

by application of R. Proof of the second part of (i): For $\epsilon > 0$, set

$$Y_{H,\epsilon} = \bigl\{\psi \in Y\colon\ \|\psi\|_H < \epsilon\bigr\}.$$

Any given neighbourhood of ϕ_0 in $V \subset X_f$ contains $V_\epsilon = R(Y_{H,\epsilon})$ for some $\epsilon > 0$, and $R(W \cap Y_{H,\epsilon}) = R(W) \cap R(Y_{H,\epsilon}) = W_s \cap V_\epsilon$. Part (iii) of Proposition 3.5.1 yields $F_a(W_s \cap V_\epsilon) = R(K(F_a(R(W \cap Y_{H,\epsilon})))) = R(H(W \cap Y_{H,\epsilon})) \subset R(W \cap Y_{H,\epsilon}) = W_s \cap V_\epsilon$.

2. Proof of assertion (ii). There are positive constants $c_1 \leqslant c_2$ with

$$c_1\|\chi\|_{C^1} \leqslant \|\chi\|_H \leqslant c_2\|\chi\|_{C^1} \quad \text{for all } \chi \in Y.$$

Let a trajectory $(\psi_j)_0^\infty$ of F_a in W_s be given. The points $\chi_j = K(\psi_j) \in W$ form a trajectory of H since

$$\chi_{j+1} = K(\psi_{j+1}) = K\bigl(F_a(\psi_j)\bigr) = K\bigl(F_a\bigl(R(\chi_j)\bigr)\bigr) = H(\chi_j)$$

for each integer $j \geqslant 0$. Hence

$$\begin{aligned}\|\psi_j-\phi_0\|_{C^1} &= \big\|R(\chi_j)-R(0)\big\|_{C^1}\\ &\leqslant L_R\|\chi_j\|_{C^1}\\ &\leqslant L_R c_1^{-1}\|\chi_j\|_H\\ &\leqslant L_R c_1^{-1}\alpha^j\|\chi_0\|_H\\ &\leqslant L_R\frac{c_2}{c_1}\alpha^j\|\chi_0\|_{C^1}\\ &= L_R\frac{c_2}{c_1}\alpha^j\|P\|_{C^1}\|\psi_0-\phi_0\|_{C^1}.\end{aligned}$$

□

We want to show that the semiflow F maps a piece W^s of W_s close to ϕ_0 into W_s, in the sense that $F([0,\infty)\times W^s)\subset W_s$. The proof requires a quantitative version of continuous dependence on initial conditions. Notice that the proof of the latter employs the Lipschitz property (L) of the functional f.

PROPOSITION 3.5.3. *There exist an open neighbourhood* $X_{f,a}$ *of* ϕ_0 *in* X_f *and a constant* $c_a \geqslant 0$ *so that* $[0,a]\times X_{f,a}\subset\Omega$ *and*

$$\big\|F(t,\phi)-\phi_0\big\|_{C^1}\leqslant c_a\|\phi-\phi_0\|_{C^1}\quad \textit{for all } (t,\phi)\in[0,a]\times X_{f,a}.$$

PROOF. 1. Using continuity of the semiflow and compactness of the interval $[0,a]$ we find an open neighbourhood $X_{f,a}$ of the stationary point ϕ_0 in X_f so that $[0,a]\times X_{f,a}\subset\Omega$ and $F([0,a]\times X_{f,a})$ is contained in a neighbourhood of ϕ_0 in the domain U of f on which the Lipschitz estimate (L) holds.

2. Let $\xi=\phi_0(0)$, and let $\phi\in X_{f,a}$ be given. Set $x=x^\phi$. For $0\leqslant t\leqslant a$, we have

$$\begin{aligned}\big|x(t)-\xi\big| &= \left|x(0)-\xi+\int_0^t x'(s)\,\mathrm{d}s\right| = \left|x(0)-\xi+\int_0^t f(x_s)\,\mathrm{d}s\right|\\ &\leqslant \|x_0-\phi_0\|_C+L\int_0^t\|x_s-\phi_0\|_C\,\mathrm{d}s,\end{aligned}$$

as $f(\phi_0)=0$. For any $t\in[0,a]$ there exists $t_0\in[t-h,t]$ with $\|x_t-\phi_0\|_C=|x(t_0)-\xi|$. In case $t_0<0$ we obtain

$$\|x_t-\phi_0\|_C\leqslant\|x_0-\phi_0\|_C$$

while in case $t_0\geqslant 0$,

$$\begin{aligned}\|x_t-\phi_0\|_C &\leqslant \|x_0-\phi_0\|_C+L\int_0^{t_0}\|x_s-\phi_0\|_C\,\mathrm{d}s\\ &\leqslant \|x_0-\phi_0\|_C+L\int_0^t\|x_s-\phi_0\|_C\,\mathrm{d}s.\end{aligned}$$

In both cases,

$$\|x_t - \phi_0\|_C \leqslant \|x_0 - \phi_0\|_C + L\int_0^t \|x_s - \phi_0\|_C \, \mathrm{d}s.$$

Gronwall's lemma yields

$$\|F(t,\phi) - \phi_0\|_C \leqslant \|\phi - \phi_0\|_C \, \mathrm{e}^{Lt} \quad \text{on } [0,a] \times X_{f,a}.$$

Consequently, for all $(t,\phi) \in [0,a] \times X_{f,a}$,

$$\left|\left(x^\phi\right)'(t)\right| = \left|f\left(x_t^\phi\right) - f(\phi_0)\right| \leqslant L\left\|x_t^\phi - \phi_0\right\|_C \leqslant L\,\mathrm{e}^{Lt}\|\phi - \phi_0\|_C.$$

It follows that

$$\left\|\left(x_t^\phi\right)' - (\phi_0)'\right\|_C \leqslant L\,\mathrm{e}^{Lt}\|\phi - \phi_0\|_{C^1},$$

and finally

$$\left\|F(t,\phi) - \phi_0\right\|_{C^1} \leqslant (L+1)\mathrm{e}^{La}\|\phi - \phi_0\|_{C^1}$$

on $[0,a] \times X_{f,a}$. □

Now we can prove the desired invariance property of W_s with respect to the semiflow F, and an exponential estimate. Set

$$\gamma = -\frac{\log \alpha}{a};$$

then $0 < \beta < \gamma$.

PROPOSITION 3.5.4. *There exists an open neighbourhood W^s of ϕ_0 in W_s so that $[0,\infty) \times W^s \subset \Omega$, $F([0,\infty) \times W^s) \subset W_s$, and there is a constant $c_w \geqslant 0$ so that*

$$\left\|F(t,\psi) - \phi_0\right\|_{C^1} \leqslant c_w\,\mathrm{e}^{-\gamma t}\|\psi - \phi_0\|_{C^1} \quad \text{for all } \psi \in W^s,\ t \geqslant 0. \tag{3.5.1}$$

PROOF. 1. Set $V_N = R(N)$. Choose c_s according to Corollary 3.5.2(ii) and $X_{f,a}$ and c_a according to Proposition 3.5.3. It follows that there is an open neighbourhood W^s of ϕ_0 in $W_s \cap X_{f,a} \subset V_N \cap X_{f,a}$ so that

$$F_a\left(W^s\right) \subset W^s, \tag{3.5.2}$$

$$F\left([0,a] \times W^s\right) \subset V_N, \tag{3.5.3}$$

and

$$\left\{\chi \in Y\colon \|\chi\|_{C^1} \leqslant \|P\|_{C^1}\, c_a\, c_s \sup_{\eta \in W^s} \|\eta - \phi_0\|_{C^1}\right\} \subset N. \tag{3.5.4}$$

Using (3.5.2) and (3.5.3) and properties of the semiflow we get $[0,\infty) \times W^s \subset \Omega$ and $F([0,\infty) \times W^s) \subset V_N$.

2. Proof of $F([0,\infty) \times W^s) \subset W_s$. Let $t \geqslant 0$ and $\psi \in W^s$ be given. The assertion $\rho = F(t,\psi) \in W_s$ is equivalent to

$$K(\rho) \in K(W_s) = W. \tag{3.5.5}$$

By the remarks in part 1 the point ρ defines a trajectory $(\rho_j)_0^\infty$ of F_a in V_N, with $\rho_0 = \rho$, and the point $\psi \in W^s$ defines a trajectory $(\psi_j)_0^\infty$ of F_a in $W^s \subset V_N$, with $\psi_0 = \psi$. The points $\chi_j = K(\rho_j)$ form a trajectory of H in N since

$$\chi_{j+1} = K(\rho_{j+1}) = K\bigl(F_a(\rho_j)\bigr) = K\bigl(F_a\bigl(R(\chi_j)\bigr)\bigr) = H(\chi_j)$$

for all integers $j \geqslant 0$. Proposition 3.5.1(iv) shows that (3.5.5) follows from

$$\lambda^{-j}\chi_j \in N \quad \text{for all integers } j \geqslant 0. \tag{3.5.6}$$

Proof of (3.5.6). Let $j \in \mathbb{N}_0$, $k \in \mathbb{N}_0$, $ka \leqslant t < (k+1)a$. Then

$$\begin{aligned}
\|\chi_j\|_{C^1} &= \bigl\|K(\rho_j)\bigr\|_{C^1} \\
&= \bigl\|P(\rho_j - \phi_0)\bigr\|_{C^1} \\
&\leqslant \|P\|_{C^1}\|\rho_j - \phi_0\|_{C^1} \\
&= \|P\|_{C^1}\bigl\|F(ja,\rho) - \phi_0\bigr\|_{C^1} \\
&= \|P\|_{C^1}\bigl\|F\bigl(t-ka, F\bigl((j+k)a,\psi\bigr)\bigr) - \phi_0\bigr\|_{C^1} \\
&\leqslant \|P\|_{C^1} c_a \|\psi_{j+k} - \phi_0\|_{C^1} \\
&\leqslant \|P\|_{C^1} c_a c_s \alpha^{j+k}\|\psi - \phi_0\|_{C^1},
\end{aligned}$$

hence

$$\bigl\|\lambda^{-j}\chi_j\bigr\|_{C^1} \leqslant \|P\|_{C^1} c_a c_s \cdot 1 \cdot \sup_{\eta \in W^s} \|\eta - \phi_0\|_{C^1},$$

which yields

$$\lambda^{-j}\chi_j \in N,$$

according to (3.5.4).

3. Proof of (3.5.1). Let $\psi \in W^s$, $t \geqslant 0$, $j \in \mathbb{N}_0$, $ja \leqslant t < (j+1)a$. Then

$$\begin{aligned}
\bigl\|F(t,\psi) - \phi_0\bigr\|_{C^1} &= \bigl\|F\bigl(t - ja, F(ja,\psi)\bigr) - \phi_0\bigr\|_{C^1} \\
&\leqslant c_a\bigl\|F(ja,\psi) - \phi_0\bigr\|_{C^1}
\end{aligned}$$

$$
\begin{aligned}
&\leqslant c_a c_s \alpha^j \|\psi - \phi_0\|_{C^1} \\
&= c_a c_s \mathrm{e}^{j \log \alpha} \|\psi - \phi_0\|_{C^1} \\
&= c_a c_s \mathrm{e}^{(t \log \alpha)/a} \mathrm{e}^{(j - t/a) \log \alpha} \|\psi - \phi_0\|_{C^1} \\
&\leqslant c_a c_s \mathrm{e}^{(t \log \alpha)/a} \alpha^{-1} \|\psi - \phi_0\|_{C^1}. \qquad \square
\end{aligned}
$$

The C^1-submanifold W^s of X_f is the local stable manifold of F at ϕ_0. It is locally positively invariant under F, with tangent space

$$T_{\phi_0} W^s = C_s \cap T_{\phi_0} X_f,$$

and it has the following uniqueness property: There exists a constant $c > 0$ so that all initial data $\psi \in X_f$ with $[0, \infty) \times \{\psi\} \subset \Omega$ and

$$\mathrm{e}^{\beta t} \left\| F(t, \psi) - \phi_0 \right\|_{C^1} < c \quad \text{for all } t \geqslant 0$$

belong to W^s. This property is easily established by means of estimates as in the preceding proofs.

We said already that analogously to the approach to local *stable* manifolds just presented one obtains continuously differentiable local *unstable* manifolds for the semiflow from local unstable manifolds of the maps F_a, $a > 0$.

Local *unstable* manifolds for certain classes of differential equations with state-dependent delay were also obtained in earlier work, by Krishnan [126,127] and more generally in [129]. The proofs in [126,127] and [129] proceed without knowledge of a semiflow and use the heuristic approach with the auxiliary linear equation mentioned in Section 3.4. It is remarkable that in [129] higher order differentiability of local unstable manifolds is achieved.

Related to work on local invariant manifolds is a result of Arino and Sanchez [10] about saddle point behaviour of solutions close to a stationary point which is hyperbolic, i.e., there is no spectrum of the generator G_e on the imaginary axis. Also in [10] no semiflow is available, and the auxiliary linear equation is used.

3.6. *The principle of linearized stability*

For $f : U \to \mathbb{R}^n$, $U \subset C^1$ open, with property (S) and the associated semiflow F of Theorem 3.2.1 the results of Section 3.4 and Proposition 3.5.4 yield the following Principle of Linearized Stability.

THEOREM 3.6.1. *If all eigenvalues of G_e have negative real part then ϕ_0 is exponentially asymptotically stable for the semiflow F.*

For applications, recall that the eigenvalues of G_e and their multiplicities are given by the familiar transcendental characteristic equation which is obtained from the Ansatz $v(t) = \mathrm{e}^{\lambda t} c$ for a solution to the equation

$$v'(t) = \mathrm{D}_e f(\phi_0) v_t.$$

Earlier Principles of Linearized Stability, for certain classes of differential equations with state-dependent delay, are due to Cooke and Huang [48] and to Hartung and Turi [105, 108]. The proofs employ the heuristic approach described in Section 3.4. Related is work by Hartung [101] on exponential stability of periodic solutions to nonautonomous, periodic differential equations with state-dependent delay, and a result by Győri and Hartung [92] who derive exponential stability of the zero solution of the nonautonomous equation

$$x'(t) = a(t) x\big(t - r\big(t, x(t)\big)\big)$$

from exponential stability of the zero solution to the linear nonautonomous RFDE

$$y'(t) = a(t) y\big(t - r(t, 0)\big).$$

4. Center manifolds

4.1. *Preliminaries*

Assume $f : U \to \mathbb{R}^n$, $U \subset C^1$ open, with property (S) of Section 3, and define X_f as in Section 3. By Theorem 3.2.1, in case $X_f \neq \emptyset$ the solutions of the equation

$$x'(t) = f(x_t) \tag{1.0.1}$$

with initial function $x_0 = \phi \in X_f$ define a semiflow $F : \Omega \to X_f$. Assume that $0 \in U$ and 0 is a stationary point of F. Define $L = \mathrm{D} f(0)$, $L_e = \mathrm{D}_e f(0)$ and

$$r : U \ni \phi \mapsto f(\phi) - L\phi \in \mathbb{R}^n.$$

Clearly, r also satisfies (S), and $r(0) = 0$, $\mathrm{D}r(0) = 0$. Then (1.0.1) is equivalent to

$$x'(t) = L x_t + r(x_t). \tag{4.1.1}$$

The solutions of the IVP

$$y'(t) = L_e y_t, \quad y_0 = \phi \in C$$

define the strongly continuous semigroup $(T_e(t))_{t \geqslant 0}$ on C with generator $G_e : \mathrm{dom}(G_e) \to C$, $\mathrm{dom}(G_e) = \{\phi \in C^1 \colon \phi'(0) = L_e(\phi)\}$, $G_e \phi = \phi'$. The realified generalized eigenspaces of G_e given by the eigenvalues with negative, zero and positive real part are the stable C_s, center C_c and unstable C_u spaces, respectively. We have the decomposition

$$C = C_s \oplus C_c \oplus C_u,$$

C_s is infinite dimensional, C_c and C_u are finite dimensional, $C_c \subset \mathrm{dom}(G_e)$, $C_u \subset \mathrm{dom}(G_e)$. The set $C_s^1 = C_s \cap C^1$ is a closed subset of C^1. Hence we get the decomposition

$$C^1 = C_s^1 \oplus C_c \oplus C_u \tag{4.1.2}$$

of C^1.

The derivatives $D_2F(t, 0)$, $t \geqslant 0$, form the strongly continuous semigroup $T(t)$, $t \geqslant 0$, on $T_0X_f = \mathrm{dom}(G_e)$ with generator G. The stable, center and unstable subspaces of G are $C_s \cap \mathrm{dom}(G_e) = C_s^1 \cap \mathrm{dom}(G_e)$, C_c and C_u, respectively, and

$$T_0X_f = \mathrm{dom}(G_e) = \bigl(C_s^1 \cap \mathrm{dom}(G_e)\bigr) \oplus C_c \oplus C_u.$$

In the sequel we assume

$$\dim C_c \geqslant 1.$$

The main result of this section guarantees the existence of a Lipschitz smooth (local) center manifold of F at the stationary point 0.

THEOREM 4.1.1. *There exist open neighbourhoods $C_{c,0}$ of* 0 *in C_c and $C^1_{su,0}$ in $C_s^1 \oplus C_u$ with $N = C_{c,0} + C^1_{su,0} \subset U$, a Lipschitz continuous map $w_c : C_{c,0} \to C^1_{su,0}$ such that* $w_c(0) = 0$ *and for the graph*

$$W_c = \{\phi + w_c(\phi) \colon \phi \in C_{c,0}\}$$

of w_c the following holds.

(i) *$W_c \subset X_f$, and W_c is a* $\dim C_c$*-dimensional Lipschitz smooth submanifold of X_f.*

(ii) *If $x : \mathbb{R} \to \mathbb{R}^n$ is a continuously differentiable solution of* (1.0.1) *on* $\mathbb{R}$ *with $x_t \in N$ for all $t \in \mathbb{R}$, then $x_t \in W_c$ for all $t \in \mathbb{R}$.*

(iii) *W_c is locally positively invariant with respect to the semiflow F, i.e., if $\phi \in W_c$ and $\alpha > 0$ such that $F(t, \phi)$ is defined for all $t \in [0, \alpha)$, and $F(t, \phi) \in N$ for all $t \in [0, \alpha)$, then $F(t, \phi) \in W_c$ for all $t \in [0, \alpha)$.*

The proof will be given in Section 4.3. We follow the approach of [54] by applying the Lyapunov–Perron method. The variation-of-constants formula is from [54] which requires dual spaces and adjoint operators. Other forms of the variation-of-constants formula, obtained, e.g., via integrated semigroups or extrapolation theory, could also be used. The existence of a global center manifold for a modified version of Eq. (4.1.1) is formulated as a fixed point problem with a parameter. However, as the right-hand side of (4.1.1) has smoothness properties only in the space C^1, the space, where we look for fixed points, should contain smoother functions than the corresponding space in [54]. This is why the proof is not a straightforward application of the technique of [54]. The C^1-smoothness of W_c also holds under the hypotheses of Theorem 4.1.1. A proof can be found in [131].

4.2. *The linear inhomogeneous equation*

Let $|\cdot|$ be a norm in $\mathbb{R}^n$. The spaces C, C^1 and their norms are defined as in Section 1. Let $L^\infty(-h,0;\mathbb{R}^n)$ denote the Banach space of measurable and essentially bounded functions from $(-h,0)$ into $\mathbb{R}^n$ equipped with the essential least upper bound norm $\|\cdot\|_\infty$.

We denote dual spaces and adjoint operators by an asterisk $*$ in the sequel. The elements $\phi^\odot$ of C^* for which the curve

$$[0,\infty)\ni t\mapsto T_e^*(t)\phi^\odot\in C^*$$

is continuous form a closed subspace $C^\odot$ (of C^*) which is positively invariant under $T_e^*(t)$, $t\geqslant 0$. The operators

$$T_e^\odot(t): C^\odot\ni\phi^\odot\mapsto T_e^*(t)\phi^\odot\in C^\odot,\quad t\geqslant 0,$$

constitute a strongly continuous semigroup on $C^\odot$. Similarly, we can introduce the dual space $C^{\odot*}$ and the semigroup of adjoint operators $T_e^{\odot*}(t)$, $t\geqslant 0$, which is strongly continuous on $C^{\odot\odot}$. There is an isometric isomorphism between $\mathbb{R}^n\times L^\infty(-h,0;\mathbb{R}^n)$ equipped with the norm $\|(\alpha,\phi)\|=\max\{|\alpha|,\|\phi\|_\infty\}$ and $C^{\odot*}$. We will identify $C^{\odot*}$ with $\mathbb{R}^n\times L^\infty(-h,0;\mathbb{R}^n)$ and omit the isomorphism. The original state space is sun-reflexive in the sense that, for the norm-preserving linear map $j: C\to C^{\odot*}$ given by $j(\phi)=(\phi(0),\phi)$, we have $j(C)=C^{\odot\odot}$. We also omit the embedding operator j and identify C and $C^{\odot\odot}$. All of these results as well as the decomposition of $C^{\odot*}$ and the variation-of-constants formula can be found in [54].

Let $Y^{\odot*}$ denote the subspace $\mathbb{R}^n\times\{0\}$ of $C^{\odot*}$. For the kth unit vector e_k in $\mathbb{R}^n$ set $r_k^{\odot*}=(e_k,0)\in Y^{\odot*}$. Let $l:\mathbb{R}^n\to Y^{\odot*}$ be the linear map given by $l(e_k)=r_k^{\odot*}$, $k\in\{1,2,\dots,n\}$. Then l has an inverse l^{-1}, and $\|l\|=\|l^{-1}\|=1$.

Let $G_e^{\odot*}$ denote the generator of $T_e^{\odot*}$. For the spectra $\sigma(G_e)$ and $\sigma(G_e^{\odot*})$ we have $\sigma(G_e)=\sigma(G_e^{\odot*})$. Recall that we assumed

$$\sigma(G_e)\cap\mathrm{i}\mathbb{R}\neq\emptyset.$$

Then $C^{\odot*}$ can be decomposed as

$$C^{\odot*}=C_s^{\odot*}\oplus C_c\oplus C_u,\tag{4.2.1}$$

where $C_s^{\odot*}$, C_c, C_u are closed subspaces of $C^{\odot*}$, C_c and C_u are contained in C^1, $1\leqslant\dim C_c<\infty$, $\dim C_u<\infty$. The subspaces $C_s^{\odot*}$, C_c and C_u are invariant under $T_e^{\odot*}(t)$,

$t \geqslant 0$, and $T_e(t)$ can be extended to a one-parameter group on both C_c and C_u. There exist real numbers $K \geqslant 1$, $a < 0$, $b > 0$ and $\epsilon > 0$ with $\epsilon < \min\{-a, b\}$ such that

$$\begin{aligned} \|T_e(t)\phi\| &\leqslant K\mathrm{e}^{bt}\|\phi\|, && t \leqslant 0,\ \phi \in C_u, \\ \|T_e(t)\phi\| &\leqslant K\mathrm{e}^{\epsilon|t|}\|\phi\|, && t \in \mathbb{R},\ \phi \in C_c, \\ \|T_e^{\odot *}(t)\phi\| &\leqslant K\mathrm{e}^{at}\|\phi\|, && t \geqslant 0,\ \phi \in C_s^{\odot *}. \end{aligned} \tag{4.2.2}$$

Using the identification of C and $C^{\odot\odot}$, we obtain $C_s^1 = C^1 \cap C_s^{\odot *}$. The decompositions (4.1.2) and (4.2.1) define the projection operators P_s, P_c, P_u and $P_s^{\odot *}$, $P_c^{\odot *}$, $P_u^{\odot *}$ with ranges C_s^1, C_c, C_u and $C_s^{\odot *}$, C_c, C_u, respectively.

We need a variation-of-constants formula for solutions of

$$x'(t) = L_e x_t + q(t) \tag{4.2.3}$$

with a continuous function $q : \mathbb{R} \to \mathbb{R}^n$.

If c, d are reals with $c \leqslant d$, and $w : [c, d] \to C^{\odot *}$ is continuous, then the weak-star integral

$$\int_c^d T_e^{\odot *}(d - \tau)w(\tau)\,\mathrm{d}\tau \in C^{\odot *}$$

is defined by

$$\left(\int_c^d T_e^{\odot *}(d - \tau)w(\tau)\,\mathrm{d}\tau\right)(\phi^{\odot}) = \int_c^d T_e^{\odot *}(d - \tau)w(\tau)(\phi^{\odot})\,\mathrm{d}\tau$$

for all $\phi^{\odot} \in C^{\odot}$.

If $I \subset \mathbb{R}$ is an interval, $q : I \to \mathbb{R}^n$ is continuous and $x : I + [-h, 0] \to \mathbb{R}^n$ is a solution of (4.2.3) on I, then the curve $u : I \ni t \mapsto x_t \in C$ satisfies the integral equation

$$u(t) = T_e(t - s)u(s) + \int_s^t T_e^{\odot *}(t - \tau)Q(\tau)\,\mathrm{d}\tau, \quad t, s \in I,\ s \leqslant t, \tag{4.2.4}$$

with $Q(t) = l(q(t))$, $t \in I$. Moreover, if $Q : I \to Y^{\odot *}$ is continuous, and $u : I \to C$ satisfies (4.2.4), then there is a continuous function $x : I + [-h, 0] \to \mathbb{R}^n$ such that $x_t = u(t)$ for all $t \in I$, and x satisfies (4.2.3) with $q(t) = l^{-1}(Q(t))$, $t \in I$. So, there is a one-to-one correspondence between the solutions of (4.2.3) and (4.2.4).

For a Banach space B with norm $\|\cdot\|$ and a real $\eta \geqslant 0$, we define the Banach space

$$C_\eta(\mathbb{R}, B) = \left\{b \in C(\mathbb{R}, B) \colon \sup_{t \in \mathbb{R}} \mathrm{e}^{-\eta|t|}\|b(t)\| < \infty\right\}$$

with norm

$$\|b\|_{C_\eta^0(\mathbb{R}, B)} = \sup_{t \in \mathbb{R}} \mathrm{e}^{-\eta|t|}\|b(t)\|.$$

For $\eta \geqslant 0$, we introduce the notation

$$Y_\eta = C_\eta\big(\mathbb{R}, Y^{\odot *}\big), \quad C_\eta^0 = C_\eta(\mathbb{R}, C), \quad C_\eta^1 = C_\eta\big(\mathbb{R}, C^1\big).$$

We need the following smoothing property of Eq. (4.2.4).

PROPOSITION 4.2.1. *Let $\eta \geqslant 0$, $Q \in Y_\eta$, $u \in C_\eta^0$, and assume that u satisfies*

$$u(t) = T_e(t-s)u(s) + \int_s^t T_e^{\odot *}(t-\tau)Q(\tau)\,\mathrm{d}\tau, \quad -\infty < s \leqslant t < \infty.$$

Then $u \in C_\eta^1$ and

$$\|u\|_{C_\eta^1} \leqslant \big(1 + \mathrm{e}^{\eta h}\|L_e\|\big)\|u\|_{C_\eta^0} + \mathrm{e}^{\eta h}\|Q\|_{Y_\eta}.$$

PROOF. Define $q : \mathbb{R} \to \mathbb{R}^n$ by $q(t) = l^{-1}(Q(t))$, $t \in \mathbb{R}$. Then $q \in C_\eta(\mathbb{R}, \mathbb{R}^n)$, and

$$\|q\|_{C_\eta(\mathbb{R},\mathbb{R}^n)} = \|Q\|_{Y_\eta}.$$

The function $x : \mathbb{R} \to \mathbb{R}^n$, given by $x(t) = u(t)(0)$, satisfies $x_t = u(t)$, $t \in \mathbb{R}$, and Eq. (4.2.3) holds for all $t \in \mathbb{R}$. Then x is C^1-smooth, $x_t \in C^1$ for all $t \in \mathbb{R}$, and the mapping $\mathbb{R} \ni t \mapsto x_t \in C^1$ is continuous. Moreover, for all $t \in \mathbb{R}$,

$$\begin{aligned}
\big|x'(t)\big| &\leqslant \|L_e\|\|x_t\|_C + \big|q(t)\big| \\
&= \|L_e\|\big\|u(t)\big\|_C + \big\|Q(t)\big\|_{Y^{\odot *}} \\
&\leqslant \mathrm{e}^{\eta|t|}\big(\|L_e\|\|u\|_{C_\eta^0} + \|Q\|_{Y_\eta}\big).
\end{aligned}$$

Hence

$$\begin{aligned}
\sup_{t\in\mathbb{R}} \mathrm{e}^{-\eta|t|}\big\|x_t'\big\|_C &= \sup_{t\in\mathbb{R}} \mathrm{e}^{-\eta|t|} \sup_{-h\leqslant s\leqslant 0} \big|x'(t+s)\big| \\
&\leqslant \big(\|L_e\|\|u\|_{C_\eta^0} + \|Q\|_{Y_\eta}\big) \sup_{t\in\mathbb{R}} \mathrm{e}^{-\eta|t|} \sup_{-h\leqslant s\leqslant 0} \mathrm{e}^{\eta|t+s|} \\
&\leqslant \mathrm{e}^{\eta h}\big(\|L_e\|\|u\|_{C_\eta^0} + \|Q\|_{Y_\eta}\big).
\end{aligned}$$

Therefore, $u \in C_\eta^1$, and

$$\begin{aligned}
\|u\|_{C_\eta^1} &= \sup_{t\in\mathbb{R}} \mathrm{e}^{-\eta|t|}\|x_t\|_{C^1} = \sup_{t\in\mathbb{R}} \mathrm{e}^{-\eta|t|}\big(\|x_t\|_C + \big\|x_t'\big\|_C\big) \\
&\leqslant \|u\|_{C_\eta^0} + \mathrm{e}^{\eta h}\|L_e\|\|u\|_{C_\eta^0} + \mathrm{e}^{\eta h}\|Q\|_{Y_\eta}.
\end{aligned}$$

□

For a given $Q:\mathbb{R}\to Y^{\odot *}$ we (formally) define

$$(\mathcal{K}Q)(t)=\int_0^t T_e^{\odot *}(t-\tau)P_c^{\odot *}Q(\tau)\,\mathrm{d}\tau+\int_\infty^t T_e^{\odot *}(t-\tau)P_u^{\odot *}Q(\tau)\,\mathrm{d}\tau$$
$$+\int_{-\infty}^t T_e^{\odot *}(t-\tau)P_s^{\odot *}Q(\tau)\,\mathrm{d}\tau.$$

PROPOSITION 4.2.2. *Assume* $\eta\in(\epsilon,\min\{-a,b\})$. *Then the mapping*

$$\mathcal{K}_\eta : Y_\eta \ni Q\mapsto \mathcal{K}Q\in C_\eta^1$$

is linear bounded with norm

$$\|\mathcal{K}_\eta\|\leqslant K\bigl(1+\mathrm{e}^{\eta h}\|L_e\|\bigr)\left(\frac{1}{\eta-\epsilon}+\frac{1}{-a-\eta}+\frac{1}{b-\eta}\right)+\mathrm{e}^{\eta h}.$$

If $Q\in Y_\eta$ *then* $u=\mathcal{K}Q$ *is the unique solution of*

$$u(t)=T_e(t-s)u(s)+\int_s^t T_e^{\odot *}(t-\tau)Q(\tau)\,\mathrm{d}\tau,\quad -\infty<s\leqslant t<\infty, \tag{4.2.5}$$

in C_η^1 *with* $P_c^{\odot *}u(0)=0$.

PROOF. Lemma IX.3.2 of [54] shows that $\mathcal{K}$ as a mapping from Y_η into C_η^0 is linear bounded such that its norm is bounded by

$$K\left(\frac{1}{\eta-\epsilon}+\frac{1}{-a-\eta}+\frac{1}{b-\eta}\right),$$

moreover $u=\mathcal{K}Q$ with $Q\in Y_\eta$ is the unique solution of (4.2.5) with $P_c^{\odot *}u(0)=0$. Hence Proposition 4.2.1 yields the boundedness of $\mathcal{K}_\eta$ with the stated bound for the norm. □

4.3. *Construction of a center manifold*

Now we prove Theorem 4.1.1.

As $\dim C_c<\infty$, there is a norm $|\cdot|_c$ on C_c which is C^∞-smooth on $C_c\setminus\{0\}$. Then

$$|\phi|_1=\max\bigl\{|P_c\phi|_c,\ \bigl\|(\mathrm{id}_{C^1}-P_c)\phi\bigr\|_{C^1}\bigr\},\quad \phi\in C^1,$$

defines the new norm $|\cdot|_1$ on C^1 which is equivalent to $\|\cdot\|_{C^1}$.

Let $\rho:\mathbb{R}\to\mathbb{R}$ be a C^∞-smooth function so that $\rho(t)=1$ for $t\leqslant 1$, $\rho(t)=0$ for $t\geqslant 2$, and $\rho(t)\in(0,1)$ for $t\in(1,2)$.

Define

$$\hat{r}(\phi) = \begin{cases} r(\phi), & \text{if } \phi \in U; \\ 0, & \text{if } \phi \notin U. \end{cases}$$

For any $\delta > 0$, let

$$r_\delta(\phi) = \hat{r}(\phi)\rho\left(\frac{|P_c\phi|_c}{\delta}\right)\rho\left(\frac{|(\mathrm{id}_{C^1} - P_c)\phi|_1}{\delta}\right), \quad \phi \in C^1.$$

For $\gamma > 0$ set $B_\gamma(C^1) = \{\phi \in C^1\colon\ |\phi|_1 < \gamma\}$.

Choose $\delta_0 > 0$ so that

$$B_{2\delta_0}\bigl(C^1\bigr) \subset U,$$

and $r|_{B_{2\delta_0}(C^1)}$, $\mathrm{D}r|_{B_{2\delta_0}(C^1)}$ are bounded. Then, for any $\delta \in (0, \delta_0)$

$$r_\delta|_{\{\phi\in C^1\colon\ |(\mathrm{id}_{C^1} - P_c)\phi|_1 < \delta\}}(\phi) = \hat{r}(\phi)\rho\left(\frac{|P_c\phi|_c}{\delta}\right), \quad \phi \in C^1,$$

and $r_\delta|_{\{\phi\in C^1\colon\ |(\mathrm{id}_{C^1} - P_c)\phi|_1 < \delta\}}$ is a bounded and C^1-smooth function with bounded derivative.

There exist $\delta_1 \in (0, \delta_0)$ and a nondecreasing function $\mu\colon [0, \delta_1] \to [0, 1]$ such that μ is continuous at 0, $\mu(0) = 0$, and for all $\delta \in (0, \delta_1]$ and for all $\phi, \psi \in C^1$

$$\begin{aligned} &\bigl|r_\delta(\phi)\bigr| \leqslant \delta\mu(\delta), \\ &\bigl|r_\delta(\phi) - r_\delta(\psi)\bigr| \leqslant \mu(\delta)\|\phi - \psi\|_{C^1}. \end{aligned} \tag{4.3.1}$$

For a proof of completely analogous estimates see, e.g., Proposition II.2 in [133].

For $\delta \in (0, \delta_1]$ we consider the modified equations

$$x'(t) = Lx_t + r_\delta(x_t), \quad t \in \mathbb{R}, \tag{4.3.2}$$

and

$$u(t) = T_e(t - s) + \int_s^t T_e^{\odot *}(t - \tau)l\bigl(r_\delta\bigl(u(\tau)\bigr)\bigr)\,\mathrm{d}\tau, \quad -\infty < s \leqslant t < \infty. \tag{4.3.3}$$

These equations are equivalent in the following sense: If $x\colon \mathbb{R} \to \mathbb{R}^n$ is C^1-smooth and is a solution of Eq. (4.3.2), then $u\colon \mathbb{R} \ni t \mapsto x_t \in C^1$ is a solution of Eq. (4.3.3), and conversely, a continuous $u\colon \mathbb{R} \to C^1$ satisfying (4.3.3) defines a C^1-smooth solution of (4.3.2) by $x(t) = u(t)(0)$, $t \in \mathbb{R}$.

Now we fix the reals $\eta \in (\epsilon, \min\{-a, b\})$ and $\delta \in (0, \delta_1)$ such that

$$\|\mathcal{K}_\eta\| \mu(\delta) < \frac{1}{2}. \tag{4.3.4}$$

Let the substitution operator

$$R : (C^1)^{\mathbb{R}} \to (Y^{\odot *})^{\mathbb{R}}$$

of the map $C^1 \ni \phi \mapsto l(r_\delta(\phi)) \in Y^{\odot *}$ be given by

$$R(u)(t) = l\big(r_\delta\big(u(t)\big)\big).$$

Inequalities (4.3.1) and $\|l\| = 1$ yield that

$$R\big(C^1_\eta\big) \subset Y_\eta,$$

and for the induced map

$$R_{\delta\eta} : C^1_\eta \to Y_\eta,$$

the inequalities

$$\big\| R_{\delta\eta}(u) \big\|_{Y_\eta} \leqslant \delta\mu(\delta), \quad u \in C^1_\eta,$$

and

$$\big\| R_{\delta\eta}(u) - R_{\delta\eta}(v) \big\|_{Y_\eta} \leqslant \mu(\delta) \|u - v\|_{C^1_\eta}, \quad u, v \in C^1_\eta, \tag{4.3.5}$$

hold.

Let the mapping $S : C_c \to C^1_\eta$ be given by $(S\phi)(t) = T_e(t)\phi$, $\phi \in C_c$, $t \in \mathbb{R}$. For all $\phi \in C_c$ we have $\|T_e(t)\phi\|_{C^1} = \|T_e(t)\phi\|_C + \|\frac{\mathrm{d}}{\mathrm{d}t}(T_e(t)\phi)\|_C$ and $\frac{\mathrm{d}}{\mathrm{d}t}(T_e(t)\phi) = T_e(t)G_e\phi = T_e(t)\phi'$. Therefore, by applying the second inequality in (4.2.2) and $\eta > \epsilon$, we find

$$\begin{aligned} \|S\phi\|_{C^1_\eta} &= \sup_{t \in \mathbb{R}} \mathrm{e}^{-\eta|t|} \big(\big\| T_e(t)\phi \big\|_C + \big\| T_e(t)\phi' \big\|_C \big) \\ &\leqslant K \big(\|\phi\|_C + \big\|\phi'\big\|_C \big) \\ &\leqslant K \|\phi\|_{C^1}. \end{aligned} \tag{4.3.6}$$

Define the mapping

$$\mathcal{G} : C^1_\eta \times C_c \to C^1_\eta$$

by

$$\mathcal{G}(u, \phi) = S\phi + \mathcal{K}_\eta \circ R_{\delta\eta}(u), \quad u \in C^1_\eta,\ \phi \in C_c.$$

For all $u, v \in C^1_\eta$ and $\phi \in C_c$, (4.3.4) and (4.3.5) yield

$$\begin{aligned}\left\|\mathcal{G}(u,\phi) - \mathcal{G}(v,\phi)\right\|_{C^1_\eta} &\leqslant \|\mathcal{K}_\eta\| \left\|R_{\delta\eta}(u) - R_{\delta\eta}(v)\right\|_{Y_\eta} \\ &\leqslant \|\mathcal{K}_\eta\| \mu(\delta) \|u - v\|_{C^1_\eta} \\ &\leqslant \frac{1}{2} \|u - v\|_{C^1_\eta}.\end{aligned}$$

If $\gamma > 0$ and $\phi \in C_c$ with $\|\phi\|_{C^1} \leqslant \gamma/(2K)$, and $u \in \overline{B_\gamma(C^1_\eta)}$, then, by using (4.3.4), (4.3.5) and (4.3.6),

$$\begin{aligned}\left\|\mathcal{G}(u,\phi)\right\|_{C^1_\eta} &\leqslant K\|\phi\|_{C^1} + \|\mathcal{K}_\eta\| \mu(\delta) \|u\|_{C^1_\eta} \\ &\leqslant \frac{\gamma}{2} + \frac{\gamma}{2} = \gamma.\end{aligned}$$

Therefore, $\mathcal{G}(\cdot,\phi)$ maps $\overline{B_\gamma(C^1_\eta)}$ into itself provided $\gamma \geqslant 2K\|\phi\|_{C^1}$. In addition, $\mathcal{G}(\cdot,\phi)$ is Lipschitz continuous with Lipschitz constant $1/2$.

Consequently, there is a map

$$u^* : C_c \to C^1_\eta$$

such that, for each $\phi \in C_c$, $u = u^*(\phi)$ is the unique solution in C^1_η of the equation

$$u = \mathcal{G}(u,\phi).$$

The mapping u^* is globally Lipschitz continuous since

$$\left\|u^*(\phi) - u^*(\psi)\right\|_{C^1_\eta} \leqslant K\|\phi - \psi\|_{C^1} + \frac{1}{2}\left\|u^*(\phi) - u^*(\psi)\right\|_{C^1_\eta},$$

yielding

$$\left\|u^*(\phi) - u^*(\psi)\right\|_{C^1_\eta} \leqslant 2K\|\phi - \psi\|_{C^1}$$

for all $\phi, \psi \in C_c$.

The set

$$W = \left\{u^*(\phi)(0) \colon \phi \in C_c\right\}$$

is called the global center manifold of Eq. (4.3.2) at the stationary point 0.

Setting

$$w : C_c \ni \phi \mapsto (\mathrm{id}_{C^1} - P_c)u^*(\phi)(0) \in C^1_s \oplus C_u,$$

we get the graph representation

$$W = \{\phi + w(\phi) \colon \phi \in C_c\}$$

for W.

For all $\phi \in C_c$ we have

$$\begin{aligned} |w(\phi)|_1 &= \|w(\phi)\|_{C^1} = \|(\mathrm{id}_{C^1} - P_c)u^*(\phi)(0)\|_{C^1} \\ &= \|\mathcal{K}_\eta(R_{\delta\eta}(u^*(\phi)))(0)\|_{C^1} \leqslant \|\mathcal{K}_\eta(R_{\delta\eta}(u^*(\phi)))\|_{C^1_\eta} \\ &\leqslant \|\mathcal{K}_\eta\| \|R_{\delta\eta}(u^*(\phi))\|_{Y_\eta} \\ &\leqslant \|\mathcal{K}_\eta\| \delta\mu(\delta) < \delta. \end{aligned}$$

An important consequence is that

$$W \subset \{\phi \in C^1 \colon |(\mathrm{id}_{C^1} - P_c)\phi|_1 < \delta\},$$

that is, W is contained in the δ-neighbourhood of C_c, where r_δ is C^1-smooth with bounded derivative. This fact is essential in the proof of the C^1-smoothness of the center manifold.

Setting

$$\begin{aligned} &C_{c,0} = \{\phi \in C_c \colon |\phi|_1 < \delta\}, \\ &C^1_{su,0} = \{\phi \in C^1_s \oplus C_u \colon |\phi|_1 < \delta\}, \\ &N = C_{c,0} + C^1_{su,0} = \{\phi \in C^1 \colon |\phi|_1 < \delta\}, \\ &w_c = w|_{C_{c,0}}, \\ &W_c = \{\phi + w_c(\phi) \colon \phi \in C_{c,0}\}, \end{aligned}$$

we obtain that $w_c(C_{c,0}) \subset C^1_{su,0}$ and w_c is Lipschitz continuous. As $\mathcal{G}(0,0) = 0$, it follows that $u^*(0) = 0$, and consequently $w_c(0) = 0$.

Let $v \in C^1_\eta$ be a solution of Eq. (4.3.3). Define $z : \mathbb{R} \to C^1$ by

$$z(t) = v(t) - T_e(t)P_c v(0), \quad t \in \mathbb{R}.$$

Obviously, $v \in C^1_\eta$ implies $z \in C^1_\eta$. Moreover,

$$z(t) = T_e(t-s)z(s) + \int_s^t T_e^{\odot *}(t-\tau) l(r_\delta(v(\tau)))\, \mathrm{d}\tau, \quad -\infty < s \leqslant t < \infty.$$

As $P_c z(0) = 0$ and $v \in C^1_\eta$, Proposition 4.2.2 yields

$$z = \mathcal{K}_\eta(R_{\delta\eta}(v)).$$

Therefore

$$v = S\big(P_c v(0)\big) + \mathcal{K}_\eta\big(R_{\delta\eta}(v)\big),$$

and

$$v(0) = u^*\big(P_c v(0)\big) \in W^c.$$

For any $t \in \mathbb{R}$ and $\hat{v} : \mathbb{R} \ni s \mapsto v(t+s) \in C^1$, it is clear that $\hat{v} \in C^1_\eta$, and $\hat{v}$ is also a solution of Eq. (4.3.3). Therefore, $v(t) = \hat{v}(0) \in W$ follows for all $t \in \mathbb{R}$. Consequently, for each $v \in C^1_\eta$ satisfying Eq. (4.3.3), $v(t) \in W$ holds for all $t \in \mathbb{R}$.

If $x : \mathbb{R} \to \mathbb{R}^n$ is a solution of Eq. (1.0.1) with $x_t \in N$ for all $t \in \mathbb{R}$, then (4.3.2) also holds, and $u(t) = x_t$, $t \in \mathbb{R}$, satisfies Eq. (4.3.3) since $r|_N = r_\delta|_N$, and $u \in C^1_\eta$. Thus, $x_t \in W$, $t \in \mathbb{R}$. This proves (ii) of Theorem 4.1.1.

In order to show (iii) in Theorem 4.1.1, let $\phi \in W_c$. Then $u^*(P_c\phi) \in C^1_\eta$, and $u^*(P_c\phi)(t) \in W$ for all $t \in \mathbb{R}$. Let $\beta \in (0, \infty]$ be maximal so that

$$u^*(P_c\phi)(t) \in N \quad \text{for all } t \in [0, \beta),$$

that is,

$$u^*(P_c\phi)(t) \in W_c \quad \text{for all } t \in [0, \beta).$$

Then there exists a C^1-smooth function $y : [-h, \beta) \to \mathbb{R}^n$ so that $y_t = u^*(P_c\phi)(t)$, $t \in [0, \beta)$, and

$$\begin{cases} y'(t) = Ly_t + r(y_t), & 0 < t < \beta, \\ y_0 = \phi. \end{cases}$$

If $x^\phi : [-h, \alpha)$ is also a solution of the above IVP with $x^\phi_t \in N$, $t \in [0, \alpha)$, then the result on unique continuation of solutions in Section 3 yields $\alpha \leqslant \beta$ and $x^\phi(t) = y(t)$, $t \in [-h, \alpha)$.

For any $\phi \in W_c$, the function $x : \mathbb{R} \to \mathbb{R}^n$ defined by $x_t = u^*(P_c\phi)(t)$, $t \in \mathbb{R}$, is continuously differentiable and satisfies Eq. (4.3.2). Consequently, $x'(t) = Lx_t + r_\delta(x_t)$, $t \in \mathbb{R}$. In particular, $\phi'(0) = L\phi + r_\delta(\phi)$. Using $\phi \in W_c \subset N$ and $r_\delta|_N = r|_N$, $\phi \in X_f$ follows. Thus, $W_c \subset X_f$.

Recall from Section 3.5 that there is an n-dimensional subspace $E \subset C^1$ which is a complement of $Y = T_0 X_f$ in C^1. If $e_1, \dots, e_n$ is a basis for E, then using the decomposition $C^1 = C^1_s \oplus C_c \oplus C_u$ of C^1, for each $i \in \{1, \dots, n\}$ we have

$$e_i = s_i + c_i + u_i$$

for some $s_i \in C^1_s$, $c_i \in C_c$, $u_i \in C_u$. As $C_c \oplus C_u \subset Y$, we have $s_i \notin Y$. Then the subspace $\hat{E}$ spanned by the vectors

$$\hat{e}_i = e_i - c_i - u_i, \quad i \in \{1, \dots, n\},$$

is also an n-dimensional complementary subspace of Y in C^1, and in addition $\hat{E} \subset C_s^1$. Therefore, without loss of generality, we may assume $E \subset C_s^1$. Then

$$C_s^1 = E \oplus \left(C_s^1 \cap Y\right),$$

$$Y = \left(C_s^1 \cap Y\right) \oplus C_c \oplus C_u,$$

and

$$C^1 = E \oplus \left(C_s^1 \cap Y\right) \oplus C_c \oplus C_u = E \oplus Y.$$

Let $P : C^1 \to C^1$ denote the projection along E onto Y. There is an open neighbourhood V of 0 in X_f so that $P : V \to Y$ is a manifold chart of X_f. Set $Y_0 = P(V)$. The inverse of $P : V \to Y_0$ is C^1-smooth. If $\delta > 0$ is sufficiently small, then $W_c \subset V$ and $PW_c \subset Y_0$. In order to complete the proof of (i) in Theorem 4.1.1, it is enough to show that PW_c is a $\dim C_c$-dimensional Lipschitz submanifold of Y. Indeed,

$$PW_c = \left\{P\left(\phi + w_c(\phi)\right) \colon \phi \in C_{c,0}\right\} = \left\{\phi + Pw_c(\phi) \colon \phi \in C_{c,0}\right\}.$$

As $w_c(\phi) \in C_s^1 \oplus C_u$, we have

$$Pw_c(\phi) \in \left(C_s^1 \cap Y\right) \oplus C_u.$$

Thus, PW_c is the graph of the Lipschitz continuous map

$$\left\{\phi \in C_c \colon |\phi|_1 < \delta\right\} \ni \chi \mapsto Pw_c(\chi) \in \left(C_s^1 \cap Y\right) \oplus C_u.$$

This completes the proof of Theorem 4.1.1.

4.4. *Discussion*

It is also true that the local center manifold W_c given in Theorem 4.1.1 is a C^1-submanifold of X_f. The proof will appear in [131]. It is based on the fact that for the global center manifold W of the modified Eq. (4.3.1)

$$W \subset \left\{\phi \in C^1 \colon \left|(\mathrm{id}_{C^1} - P_c)\phi\right|_1 < \delta\right\},$$

and r_δ is C^1-smooth on the subset $\{\phi \in C^1 \colon |(\mathrm{id}_{C^1} - P_c)\phi|_1 < \delta\}$ of C^1 with bounded derivative. The techniques of [54] or [133] can be modified to our situation in order to show that the map $w : C_c \to C_s^1 \oplus C_u$ is C^1-smooth.

Another way to obtain C^1-smooth center manifolds is to consider, for some $a > 0$, the time-a map

$$F_a : \Omega_a \to X_f,$$

and construct a local center manifold of F_a at its fixed point 0. However, as a 2-dimensional ordinary differential equation example shows in [130], the obtained center manifold of F_a is not necessarily a locally invariant center manifold of the semiflow F. There is a standard technique to overcome this difficulty (see, e.g., [130]). The idea is that the modification of the map F_a should be done through the modification of the semiflow F. This requires a modification and extension of F from a small neighbourhood of $[0, a] \times \{0\}$ in Ω to a certain global semiflow. This is a nontrivial task. It is an open problem to work out the complete proof by using this approach.

Local bifurcation results for functional differential equations with state-dependent delay through the center manifold reduction would require C^k-smooth local center manifolds also with $k > 1$. As far as we know such results are not available at the moment.

A first step towards the proof of a C^k-smooth center manifold could be a C^k-smooth version of the results of Section 3. Then a C^k-smooth time-a map could be the basis to construct a C^k-smooth center manifold as suggested above for $k = 1$.

Another possible way is the extension of the approach explained in this section. Notice that it does not require the existence of a smooth semiflow. We remark that this idea worked for a construction of C^k-smooth unstable manifolds under natural conditions on f which are satisfied by equations with state-dependent delay [129].

5. Hopf bifurcation

Hopf bifurcation is the phenomenon that under certain conditions small periodic orbits appear close to a stationary point when in the underlying differential equation a parameter is varied and passes a critical value. In this section we state a Hopf bifurcation theorem for differential equations with state-dependent delay which has recently been proved by M. Eichmann [65]. The equation considered is a parametrized version of Eq. (1.0.1), namely

$$x'(t) = g(\alpha, x_t). \tag{5.0.1}$$

The map $g : J \times U \to \mathbb{R}^n$ in Eq. (5.0.1) is defined on the product of an open interval $J \subset \mathbb{R}$ and an open subset $U \subset C^1 = C^1([-h, 0], \mathbb{R}^n)$, with $h > 0$ and $n \in \mathbb{N}$. Let $C = C([-h, 0], \mathbb{R}^n)$ and let C^2 denote the Banach space of twice continuously differentiable functions $\phi : [-h, 0] \to \mathbb{R}^n$, with the norm given by $\|\phi\|_{C^2} = \|\phi\|_C + \|\phi'\|_C + \|\phi''\|_C$. The set

$$U^* = U \cap C^2$$

is an open subset of C^2. The following hypotheses on smoothness are assumed.

- (H1) The mapping $g : J \times U \to \mathbb{R}^n$ is continuously differentiable.
- (H2) For each $(\alpha, \phi) \in J \times U$ the partial derivative $\mathrm{D}_2 g(\alpha, \phi)$ of g with respect to ϕ extends to a continuous linear map

$$\mathrm{D}_{2,e} g(\alpha, \phi) : C \to \mathbb{R}^n.$$

(H3) The mapping

$$J \times U \times C \ni (\alpha, \phi, \chi) \mapsto \mathrm{D}_{2,e} g(\alpha, \phi)\chi \in \mathbb{R}^n$$

is continuous.

(H4) The restriction $g^* = g|J \times U^*$ is twice continuously differentiable.

(H5) For each $(\alpha, \phi) \in J \times U^*$ the second order partial derivative $D_2^2 g^*(\alpha, \phi): C^2 \times C^2 \to \mathbb{R}^n$ of g^* with respect to ϕ has a continuous bilinear extension

$$\mathrm{D}_{2,e}^2 g^*(\alpha, \phi): C^1 \times C^1 \to \mathbb{R}^n.$$

(H6) The mappings

$$J \times U^* \times C^1 \times C^1 \ni (\alpha, \phi, \chi_1, \chi_2) \mapsto \mathrm{D}_{2,e}^2 g^*(\alpha, \phi)(\chi_1, \chi_2) \in \mathbb{R}^n$$

and

$$J \times U^* \times C^1 \ni (\alpha, \phi, \chi) \mapsto \mathrm{D}_{2,e}^2 g^*(\alpha, \phi)(\chi, \cdot) \in L\big(C^2, \mathbb{R}^n\big)$$

are continuous.

The hypotheses (H1)–(H3) imply that each map $g(\alpha, \cdot): U \to \mathbb{R}^n$, $\alpha \in J$, satisfies the hypotheses (S1)–(S3) of Theorem 3.2.1. Notice that condition (H3) is weaker than continuity of the map

$$J \times U \ni (\alpha, \phi) \mapsto \mathrm{D}_{2,e} g(\alpha, \phi) \in L\big(C, \mathbb{R}^n\big).$$

In (H4), differentiability refers to the norm given by $\|(\alpha, \phi)\| = |\alpha| + \|\phi|_{C^2}$ on $\mathbb{R} \times C^2$. Notice that condition (H6) is weaker than continuity of the map

$$J \times U^* \ni (\alpha, \phi) \mapsto \mathrm{D}_{2,e}^2 g^*(\alpha, \phi) \in L^2\big(C^1, \mathbb{R}^n\big),$$

where $L^2(C^1, \mathbb{R}^n)$ denotes the Banach space of continuous bilinear maps $C^1 \times C^1 \to \mathbb{R}^n$, with the appropriate norm.

Suppose $\phi^* \in U^* \subset C^2$ satisfies

$$g^*\big(\alpha, \phi^*\big) = 0 \quad \text{for all } \alpha \in J,$$

so that ϕ^* is a stationary point for all $\alpha \in J$.

For $\alpha \in J$, set $L(\alpha) = \mathrm{D}_2 g(\alpha, \phi^*)$, and let $A(\alpha)$ denote the generator of the strongly continuous semigroup on C given by the IVP

$$y'(t) = \mathrm{D}_{2,e} g(\alpha, \phi^*) y_t, \quad y_0 = \chi \in C.$$

The spectral assumptions for the Hopf bifurcation theorem from [65] are the following.

(L1) There is a continuously differentiable map $\lambda : I \to \mathbb{C}$, $I \subset J$ an open interval, such that each $\lambda(\alpha)$, $\alpha \in I$, is a simple eigenvalue of $A(\alpha)$.

(L2) For some $\alpha_0 \in I$, $Re\,\lambda(\alpha_0) = 0$ and $\omega_0 = Im\,\lambda(\alpha_0) > 0$ and

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}(Re\,\lambda)(\alpha_0) \neq 0.$$

(L3) For every integer $k \in \mathbb{Z} \setminus \{-1, 1\}$, $\mathrm{i}k\omega_0$ is not an eigenvalue of $A(\alpha_0)$.

The following local Hopf bifurcation theorem is obtained in [65]:

THEOREM 5.0.1. *Suppose* (H1)–(H6) *and* (L1)–(L3) *hold. Then there are an interval $M \subset \mathbb{R}$ with $0 \in M$ and continuously differentiable mappings $u^* : M \to C^1$, $\omega^* : M \to \mathbb{R}$ and $\alpha^* : M \to I$ with*

$$u^*(0) = \phi^*, \quad \alpha^*(0) = \alpha_0, \quad \omega^*(0) = \omega_0$$

such that for each $a \in M$ there is a periodic solution $x : \mathbb{R} \to \mathbb{R}^n$ of the equation

$$x'(t) = g\big(\alpha^*(a), x_t\big)$$

with $x_0 = u^*(a)$ and with period $\frac{\omega^*(a)}{2\pi}$.

To our knowledge, Theorem 5.0.1 is the first Hopf bifurcation result for differential equations with state-dependent delay. A related earlier result is due to H.L. Smith [189] who proved bifurcation of periodic solutions from a stationary point for a system of integral equations with state-dependent delay, by reduction to a Hopf bifurcation theorem of Hale and de Oliveira [95] for equations with time-invariant but parameter-dependent delay.

6. Differentiability of solutions with respect to parameters

6.1. *Preliminaries*

This section deals with nonautonomous parametrized state-dependent delay systems of the form

$$x'(t) = g\big(t, x(t), x\big(t - \tau(t, x_t, \sigma)\big), \theta\big), \quad t \in [0, T], \tag{6.1.1}$$

with initial condition

$$x(t) = \phi(t), \quad t \in [-h, 0]. \tag{6.1.2}$$

Here σ and θ are parameters in the delay function τ and in g belonging to normed linear spaces Σ and Θ, respectively. In the sequel we consider also the initial function ϕ in the IVP (6.1.1)–(6.1.2) as parameters, and denote the corresponding solution by $x(\cdot; \phi, \sigma, \theta)$, and its segment function at t by $x(\cdot; \phi, \sigma, \theta)_t$.

Suppose, e.g., a system of the form

$$y'(s) = \tilde{g}\big(s, y(s), y\big(s - \tilde{\tau}(s, y_t)\big)\big), \quad s \in [t_0, t_0 + T]$$

is given. Then introducing $t = s - t_0$ and $x(t) = y(t + t_0)$ we can transform the equation into the form (6.1.1) with $g(t, \psi, u, \theta) = \tilde{g}(t + \theta, \psi, u)$ with $\theta = t_0$, and $\tau(s, \psi, \sigma) = \tilde{\tau}(s + \sigma, \psi)$ with $\sigma = t_0$. In this case $\Sigma = \Theta = \mathbb{R}$. Of course, (6.1.1) contains more general cases as well, e.g., σ and θ can be coefficient functions in the delay function τ and g, respectively. In this case Σ and Θ will be infinite dimensional function spaces.

As we have seen in Section 3, in general the IVP (6.1.1)–(6.1.2) has a unique solution only if the initial function ϕ is Lipschitz continuous, or equivalently, if ϕ belongs to the Banach space $W^{1,\infty}$ of absolutely continuous functions $\phi : [-h, 0] \to \mathbb{R}^n$ with essentially bounded derivatives, with the norm defined by

$$|\phi|_{W^{1,\infty}} = \max\big\{|\phi|_C, \operatorname{ess\,sup}\big\{\big|\phi'(s)\big| \colon s \in [-h, 0]\big\}\big\}.$$

We define the parameter space for the IVP (6.1.1)–(6.1.2) as

$$\Gamma = W^{1,\infty} \times \Sigma \times \Theta,$$

and the norm on Γ by

$$|\gamma|_\Gamma = \big|(\phi, \sigma, \theta)\big|_\Gamma = |\phi|_{W^{1,\infty}} + |\sigma|_\Sigma + |\theta|_\Theta.$$

We assume throughout this section that $\Omega_1 \subset \mathbb{R}^n$, $\Omega_2 \subset \mathbb{R}^n$, $\Omega_3 \subset \Theta$, $\Omega_4 \subset C$ and $\Omega_5 \subset \Sigma$ are open subsets of the respective spaces, $T > 0$ is finite or $T = \infty$ (in the latter case $[0, T]$ means $[0, \infty)$), and

- (D1) (i) $g : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \Theta \supset [0, T] \times \Omega_1 \times \Omega_2 \times \Omega_3 \to \mathbb{R}^n$ is continuous; and
  - (ii) g is locally Lipschitz continuous with respect to its second, third and fourth variables in the following sense: For every $\alpha \in (0, T]$, for every compact subsets $M_i \subset \Omega_i$ $(i = 1, 2)$ of $\mathbb{R}^n$, and for every closed and bounded subset $M_3 \subset \Omega_3$ of Θ there exists $L_1 = L_1(\alpha, M_1, M_2, M_3)$ such that $|f(t, v, w, \theta) - f(t, \bar{v}, \bar{w}, \bar{\theta})| \leqslant L_1(|v - \bar{v}| + |w - \bar{w}| + |\theta - \bar{\theta}|_\Theta)$, for $t \in [0, \alpha]$, $v, \bar{v} \in M_1$, $w, \bar{w} \in M_2$, and $\theta, \bar{\theta} \in M_3$;
- (D2) (i) $\tau : \mathbb{R} \times C \times \Sigma \supset [0, T] \times \Omega_4 \times \Omega_5 \to \mathbb{R}$ is continuous, $0 \leqslant \tau(t, \psi, \sigma) \leqslant h$ for $t \in [0, T]$, $\psi \in \Omega_4$, and $\sigma \in \Omega_5$; and
  - (ii) $\tau(t, \psi, \sigma)$ is locally Lipschitz-continuous in ψ and σ in the following sense: For every $\alpha \in (0, T]$, for every compact subset $M_4 \subset \Omega_4$ of C and for every closed, bounded subset $M_5 \subset \Omega_5$ of Σ there exists a constant $L_2 = L_2(\alpha, M_4, M_5)$ such that $|\tau(t, \psi, \sigma) - \tau(t, \bar{\psi}, \bar{\sigma})| \leqslant L_2(|\psi - \bar{\psi}|_C + |\sigma - \bar{\sigma}|_\Sigma)$ for $t \in [0, \alpha]$, $\psi, \bar{\psi} \in M_4$, and $\sigma, \bar{\sigma} \in M_5$.

We, of course, assume that a parameter $\bar{\gamma} = (\bar{\phi}, \bar{\sigma}, \bar{\theta}) \in \Gamma$ satisfies the compatibility condition

$$\bar{\phi}(0) \in \Omega_1, \quad \bar{\phi}\big(-\tau\big(0, \bar{\phi}, \bar{\sigma}\big)\big) \in \Omega_2, \quad \bar{\theta} \in \Omega_3, \quad \bar{\phi} \in \Omega_4, \quad \text{and} \quad \bar{\sigma} \in \Omega_5. \tag{6.1.3}$$

It is known [57] that the IVP (6.1.1)–(6.1.2) has a unique solution for each parameter $(\bar{\phi}, \bar{\sigma}, \bar{\theta}) \in \Gamma$, moreover, the solution is Lipschitz continuous with respect to the parameters [99]. We denote the open ball with radius δ centered at $\bar{\gamma}$ in Γ by $G_\Gamma(\bar{\gamma}, \delta)$, i.e., $G_\Gamma(\bar{\gamma}, \delta) = \{\gamma \in \Gamma\colon\ |\gamma - \bar{\gamma}|_\Gamma < \delta\}$. The next result is proved, e.g., in [99].

THEOREM 6.1.1. *Suppose* (D1)(i)–(ii) *and* (D2)(i)–(ii). *For any* $\bar{\gamma} = (\bar{\phi}, \bar{\sigma}, \bar{\theta}) \in \Gamma$ *satisfying* (6.1.3) *there exist* $\alpha > 0$, $\delta > 0$ *and* $L = L(\alpha, \bar{\gamma}, \delta)$ *such that the IVP* (6.1.1)–(6.1.2) *has a unique solution on* $[-h, \alpha]$ *for any* $\gamma \in G_\Gamma(\bar{\gamma}, \delta)$, *and*

$$\big|x(\cdot;\gamma)_t - x(\cdot;\bar{\gamma})_t\big|_{W^{1,\infty}} \leqslant L|\gamma - \bar{\gamma}|_\Gamma \quad \textit{for } t \in [0, \alpha], \ \gamma \in G_\Gamma(\bar{\gamma}, \delta).$$

6.2. *Pointwise differentiability with respect to parameters*

In this subsection we study differentiability of the function $\gamma \mapsto x(t;\gamma)$ where t is fixed. In addition to (D1)(i)–(ii) and (D2)(i)–(ii) we need

(D1) (iii) $g\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \Theta \supset [0, T] \times \Omega_1 \times \Omega_2 \times \Omega_3 \to \mathbb{R}^n$ is continuously differentiable with respect to its second, third and fourth variables;

(D2) (iii) $\tau\colon \mathbb{R} \times C \times \Sigma \supset [0, T] \times \Omega_4 \times \Omega_5 \to [0, \infty)$ is continuously differentiable with respect to its second and third variables.

As it was shown in Section 3.2, solutions corresponding to a parameter value from the parameter set

$$\Pi = \big\{(\phi, \sigma, \theta) \in \Gamma\colon\ \phi'(0) = g\big(0, \phi(0), \phi\big(-\tau(0, \phi, \sigma)\big), \theta\big),\ \phi \in C^1\big\}$$

are continuously differentiable on $[-h, \alpha]$. The key point of the proof of the following result is the same as that of Theorem 3.2.1, namely, the differentiability of the evaluation map $Ev\colon C^1 \times [-h, 0] \to \mathbb{R}^n$.

THEOREM 6.2.1. *Suppose* (D1)(i)–(iii) *and* (D2)(i)–(iii), $(\bar{\phi}, \bar{\sigma}, \bar{\theta}) \in \Pi$, *and let* $\delta > 0$ *and* $\alpha > 0$ *be such that the IVP* (6.1.1)–(6.1.2) *has a unique solution on* $[-h, \alpha]$ *for any* $\gamma \in G_\Gamma(\bar{\gamma}, \delta)$. *Then for any* $t \in [0, \alpha]$ *the function*

$$x(t;\cdot)\colon \Gamma \supset G_\Gamma(\bar{\gamma}, \delta) \to \mathbb{R}^n$$

is differentiable at $\bar{\gamma}$, *and its derivative is given by*

$$\mathrm{D}_2 x(t;\bar{\gamma})u = z(t;\bar{\gamma}, u), \qquad u \in \Gamma,$$

where $z(\cdot;\bar{\gamma}, u)$ *is the solution of the time-dependent delay system*

$$z'(t;\bar{\gamma}, u) = \mathrm{D}_2 g\big(t, \bar{x}(t), \bar{x}\big(t - \bar{r}(t)\big), \bar{\theta}\big) z(t;\bar{\gamma}, u)$$

$$
\begin{aligned}
&\quad + \mathrm{D}_3 g\big(t, \bar{x}(t), \bar{x}\big(t - \bar{r}(t)\big), \bar{\theta}\big)\Big(-\bar{x}'\big(t - \bar{r}(t)\big)\mathrm{D}_2\tau(t, \bar{x}_t, \bar{\sigma}) \\
&\quad \times z(\cdot; \bar{\gamma}, u)_t + z\big(t - \bar{r}(t); \bar{\gamma}, u\big) - \bar{x}'\big(t - \bar{r}(t)\big)\mathrm{D}_3\tau\big(t, \bar{x}_t, \bar{\sigma}\big)u^{\sigma}\Big) \\
&\quad + \mathrm{D}_4 g\big(t, \bar{x}(t), \bar{x}\big(t - \bar{r}(t)\big), \bar{\theta}\big)u^{\theta}, \quad t \in [0, \alpha], \\
z(t; \bar{\gamma}, u) &= u^{\phi}(t), \quad t \in [-r, 0],
\end{aligned}
$$

and where $\bar{x}(t) = x(t; \bar{\gamma})$, $\bar{r}(t) = \tau(t, \bar{x}_t, \bar{\sigma})$, *and* $u = (u^{\phi}, u^{\sigma}, u^{\theta}) \in \Gamma$.

We refer to [99] for the proof of Theorem 6.2.1. Here we just make some remarks on the choice of the norm on Γ. It is clear that at some point in the proof of Theorem 6.2.1 it is necessary to be able to differentiate the composite function $F(t, \psi) = \psi(-\tau(t, \psi))$. Suppose $\psi \in C^1$ is fixed, and consider

$$
\begin{aligned}
F(t, \psi + u) - F(t, \psi) &= \psi\big(-\tau\big(t, \psi + u\big)\big) + u\big(-\tau(t, \psi + u)\big) - \psi\big(-\tau(t, \psi)\big) \\
&= \psi'\big(-\tau(t, \psi)\big)\big(-\tau(t, \psi + u) + \tau(t, \psi)\big) \\
&\quad + u\big(-\tau(t, \psi)\big) + \omega(t, \psi, u) + u\big(-\tau(t, \psi + u)\big) \\
&\quad - u\big(-\tau(t, \psi)\big),
\end{aligned}
$$

where

$$
\begin{aligned}
\omega(t, \psi, u) = \psi\big(-\tau(t, \psi + u)\big) - \psi\big(-\tau(t, \psi)\big) \\
- \psi'\big(-\tau(t, \psi)\big)\big(-\tau(t, \psi + u) + \tau(t, \psi)\big).
\end{aligned}
$$

Therefore we expect that

$$
\mathrm{D}_2 F(t, \psi)u = -\psi'\big(-\tau(t, \psi)\big)\mathrm{D}_2\tau(t, \psi)u + u\big(-\tau(t, \psi)\big)
$$

using an appropriate norm on the domain. The assumptions $\psi \in C^1$, (D2)(iii) and the chain rule combined imply immediately that $|\omega(t, \psi, u)|/|u|_C \to 0$ as $|u|_C \to 0$. But to control the term $u(-\tau(t, \psi + u)) - u(-\tau(t, \psi))$ the C-norm is not suitable, we need the stronger $W^{1,\infty}$-norm: The Mean Value Theorem yields

$$
\frac{|u(-\tau(t, \psi + u)) - u(-\tau(t, \psi))|}{|u|_{W^{1,\infty}}} \leqslant \big|\tau(t, \psi + u) - \tau(t, \psi)\big|,
$$

and the right-hand side of the preceding inequality tends to 0 as $|u|_{W^{1,\infty}} \to 0$, due to the continuity of τ. Therefore $\mathrm{D}_2 F$ defined above is, in fact, the derivative of the function $F(t, \cdot)\colon W^{1,\infty} \to \mathbb{R}^n$ at $\psi \in C^1$ for any t.

6.3. *Differentiability with respect to parameters in norm*

Since the condition $\bar{\gamma} \in \Pi$ in the previous subsection may be inconvenient for certain applications, we explore different spaces to study differentiability in it for the case when the initial function and the solution segments are only $W^{1,\infty}$ functions.

First we introduce some notation and definitions. $W_\alpha^{1,p}$ $(1 \leqslant p < \infty)$ denotes the Banach space of absolutely continuous functions $\psi : [-h, \alpha] \to \mathbb{R}^n$ of finite norm

$$|\psi|_{W_\alpha^{1,p}} = \left(\int_{-h}^{\alpha} \left|\psi(s)\right|^p + \left|\psi'(s)\right|^p \,\mathrm{d}s \right)^{1/p}.$$

Similarly, $W_\alpha^{1,\infty}$ denotes the Banach space $W_\alpha^{1,\infty}([-h,\alpha],\mathbb{R}^n)$.

Consider a linear space Y and let $|\cdot|$ and $\|\cdot\|$ be two norms defined on Y. We say that $(Y, |\cdot|)$ is a quasi-Banach space with respect to the norm $\|\cdot\|$ if for any $r > 0$ the set $\{y \in Y\colon\ \|y\| \leqslant r\}$ is complete in the norm $|\cdot|$. See [96].

In addition to (D1)(i)–(iii) and (D2)(i)–(iii) we use in this subsection the following condition

(D2) (iv) $\tau : \mathbb{R} \times C \times \Sigma \supset [0,T] \times \Omega_4 \times \Omega_5 \to [0,\infty)$ is continuously differentiable with respect to its first variable; and

(v) $\mathrm{D}_1\tau$, $\mathrm{D}_2\tau$ and $\mathrm{D}_3\tau$ are locally Lipschitz continuous in the following sense: For every $\alpha \in (0,T]$, for every compact subset $M_4 \subset \Omega_4$ of C and for every closed, bounded subset $M_5 \subset \Omega_5$ of Σ there exists a constant $L_3 = L_3(\alpha, M_4, M_5)$ such that

$$\left|\mathrm{D}_1\tau(t,\psi,\sigma) - \mathrm{D}_1\tau\left(t,\bar{\psi},\bar{\sigma}\right)\right| \leqslant L_3\left(\left|\psi-\bar{\psi}\right|_C + |\sigma-\bar{\sigma}|_\Sigma\right),$$

$$\left\|\mathrm{D}_2\tau(t,\psi,\sigma) - \mathrm{D}_2\tau\left(t,\bar{\psi},\bar{\sigma}\right)\right\|_{\mathcal{L}(C,\mathbb{R})} \leqslant L_3\left(\left|\psi-\bar{\psi}\right|_C + |\sigma-\bar{\sigma}|_\Sigma\right),$$

$$\left\|\mathrm{D}_3\tau(t,\psi,\sigma) - \mathrm{D}_3\tau\left(t,\bar{\psi},\bar{\sigma}\right)\right\|_{\mathcal{L}(\Sigma,\mathbb{R})} \leqslant L_3\left(\left|\psi-\bar{\psi}\right|_C + |\sigma-\bar{\sigma}|_\Sigma\right)$$

for $t \in [0,\alpha]$, $\psi, \bar{\psi} \in M_4$, and $\sigma, \bar{\sigma} \in M_5$.

Hale and Ladeira [96] investigated differentiability of solutions to the constant delay equation

$$x'(t) = f\left(x(t), x(t-\tau)\right)$$

with respect to the delay, τ. They showed by means of an extension of the Uniform Contraction Principle to quasi-Banach spaces that the map

$$[0,h] \to W_\alpha^{1,1}, \quad \tau \mapsto x(\cdot;\tau)$$

is differentiable. Note that in their proof the integral norm of $W_\alpha^{1,1}$ played a crucial role; it can be replaced by a more general $W_\alpha^{1,p}$-norm, but not by the stronger $W_\alpha^{1,\infty}$-norm. This result suggests that the set $W_\alpha^{1,\infty}$ equipped with the norm $|\cdot|_{W_\alpha^{1,p}}$ could possibly be used as the state space for solutions. It might be a reasonable choice since (see, e.g., [99]) the parameter map $(\phi,\theta,\sigma)\mapsto x(\cdot;\phi,\theta,\sigma)_t$ is Lipschitz continuous in both the $|\cdot|_{W^{1,\infty}}$ and $|\cdot|_{W^{1,p}}$-norms while the time map $t\mapsto x(\cdot;\phi,\theta,\sigma)_t$ is continuous only in the $|\cdot|_{W^{1,p}}$, but not in the $|\cdot|_{W^{1,\infty}}$-norm. This indicates that the set $W^{1,\infty}$ equipped with the $|\cdot|_{W^{1,p}}$-norm (which is not a Banach space, it is only a quasi-Banach space with respect to the $|\cdot|_{W^{1,\infty}}$-norm) could be considered as a "natural" state space for state-dependent delay equations.

We follow the usual procedure to study differentiability with respect to parameters: Introducing $y(t)=x(t)-\tilde\phi(t)$ where $\tilde\phi$ is the extension of ϕ to $[-h,\alpha]$ by $\tilde\phi(t)=\phi(0)$ for $0<t\leqslant\alpha$, we rewrite the IVP (6.1.1)–(6.1.2) as a fixed point equation $S(y,\phi,\theta,\sigma)=y$, with the operator S given by

$$S(y,\phi,\theta,\sigma)(t)=\begin{cases}0, & t\in[-h,0],\\ \displaystyle\int_0^t g\big(u,y(u)+\tilde\phi(u),\Lambda\big(u,y_u+\tilde\phi_u,\sigma\big),\theta\big)\,\mathrm{d}u, & t\in[0,\alpha],\end{cases}$$

and with $\Lambda(t,\psi,\sigma)=\psi(-\tau(t,\psi,\sigma))$. In order to apply the Uniform Contraction Principle in this setting we need continuous differentiability of S with respect to y, ϕ, θ and σ in the $W^{1,p}_\alpha$ norm. It turns out that instead of the pointwise differentiability of Λ with respect to ψ and σ studied in the previous subsection it is enough to have the differentiability of the composite function $t\mapsto\Lambda(t,x_t,\sigma)$ with respect to x and σ in a norm of "L^p-type", for $x\in W^{1,\infty}_\alpha$.

Brokate and Colonius [29] studied equations of the form

$$x'(t)=f\big(t,x\big(r\big(t,x(t)\big)\big)\big),\quad t\in[a,b]$$

and investigated differentiability of the composition operator

$$A: W^{1,\infty}\big([a,b];\mathbb{R}\big)\supset\bar X\to L^p\big([a,b];\mathbb{R}\big),\quad A(x)(t)=x\big(r\big(t,x(t)\big)\big).$$

They assumed that r is twice continuously differentiable satisfying $a\leqslant r(t,v)\leqslant b$ for all $t\in[a,b]$ and $v\in\mathbb{R}$, and considered as domain of A the set

$$\bar X=\Big\{x\in W^{1,\infty}\big([a,b];\mathbb{R}\big)\colon \text{there exists } \epsilon>0 \text{ s.t. } \frac{\mathrm{d}}{\mathrm{d}t}\big(r\big(t,x(t)\big)\big)\geqslant\epsilon \\ \text{for a.e. } t\in[a,b]\Big\}.$$

It was shown in [29] that under these assumptions A is continuously differentiable with the derivative given by

$$\big(\mathrm{D}A(x)u\big)(t)=x'\big(r\big(t,x(t)\big)\big)\mathrm{D}_2r\big(t,x(t)\big)u(t)+u\big(r\big(t,x(t)\big)\big)$$

for $u\in W^{1,\infty}([a,b],\mathbb{R})$.

Both the strong $W^{1,\infty}$ norm on the domain and the weak L^p norm on the range, together with the choice of the domain seemed to be necessary to obtain the results in [29]. Note that Manitius in [160] used a similar domain and norm when he studied linearization for a class of state-dependent delay systems.

Clearly, to apply the Uniform Contraction Principle to the operator S we have to use the same norm on the domain and range of S. It turns out that the following "product norm" preserves the essential properties of the different norms used in [96] and [29]: Let $x \in W^{1,\infty}_\alpha$, and decompose x as $x = y + \tilde{\phi}$, (where $\phi(t) = x(t)$ for $t \in [-r, 0]$, and $\tilde{\phi}$ is the extension of ϕ to $[-r, \alpha]$ by $\tilde{\phi}(t) = \phi(0)$), and define the norm of x by

$$|x|_{X^{1,p}_\alpha} = \left(\int_0^\alpha \left|y'(u)\right|^p \mathrm{d}u\right)^{1/p} + |\phi|_{W^{1,\infty}},$$

and consider the normed linear space $X^{1,p}_\alpha \equiv (W^{1,\infty}_\alpha, |\cdot|_{X^{1,p}_\alpha})$. The norm $|\cdot|_{X^{1,p}_\alpha}$ is weaker than the $|\cdot|_{W^{1,\infty}_\alpha}$ norm, but stronger than the $|\cdot|_{W^{1,p}_\alpha}$ norm (see [106]).

This norm is "strong enough" that the methods of [29], with minor modifications, provide differentiability of the composition map

$$B : X^{1,p}_\alpha \times \Sigma \supset U_1 \times U_2 \to L^p\big([0, \alpha]; \mathbb{R}^n\big), \qquad B(x, \sigma)(t) = \Lambda(t, x_t, \sigma),$$

on a suitable domain $U_1 \times U_2$. On the other hand, $|\cdot|_{X^{1,p}_\alpha}$ is "weak enough" that using the differentiability of the operator B above we can obtain differentiability of the operator $S : X^{1,p}_\alpha \times W^{1,\infty} \times \Theta \times \Sigma \supset V_1 \times V_2 \times V_3 \times V_4 \to X^{1,p}_\alpha$ with respect to y, ϕ, θ and σ. Moreover it is possible to use a modification of the Uniform Contraction Principle to get differentiability of the fixed point (the solution of the IVP) with respect to the parameters ϕ, θ and σ in the $|\cdot|_{X^{1,p}_\alpha}$ norm. Since this product norm is stronger than the $|\cdot|_{W^{1,p}_\alpha}$ norm, the result implies the differentiability of solutions in the latter norm as well. For more details and the proof of the next result we refer to [106].

THEOREM 6.3.1. *Suppose* (D1)(i)–(iii) *and* (D2)(i)–(v), *and let* $\bar{\delta} > 0$ *and* $\alpha > 0$ *be such that the IVP* (6.1.1)–(6.1.2) *has a unique solution on* $[-h, \alpha]$ *for any* $\gamma \in G_\Gamma(\bar{\gamma}, \bar{\delta})$, *and suppose there exists* $\epsilon > 0$ *such that the solution* $\bar{x} = x(\cdot; \bar{\gamma})$ *satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(t - \tau(t, \bar{x}_t, \bar{\sigma})\big) \geqslant \epsilon \quad a.e.\ t \in [0, \alpha].$$

Then there exists $\delta > 0$ *such that the functions*

$$\Gamma \supset G_\Gamma(\bar{\gamma}, \delta) \to X^{1,p}_\alpha, \qquad \gamma \mapsto x(\cdot; \gamma)$$

and

$$\Gamma \supset G_\Gamma(\bar{\gamma}, \delta) \to W^{1,p}_\alpha, \qquad \gamma \mapsto x(\cdot; \gamma)$$

are continuously differentiable.

An application of differentiability of solutions with respect to parameters was given in [100], where estimation of unknown parameters in state-dependent delay equations was studied. The goal of the work was to find a parameter value which minimizes a least square cost function $P(\gamma) = \sum_{k=1}^{N}(x(t_i;\gamma) - y_i)^2$, where y_i $(i = 1, \dots, N)$ are measurements of the solution at time points t_i $(i = 1, \dots, N)$. The so-called method of quasilinearization was adopted and numerically tested for state-dependent delay equations for cases where the parameters were infinite dimensional, e.g., the initial function. This algorithm is based on Newton's method and uses the derivative of P, hence also $\mathrm{D}_2x(t;\gamma)$. The convergence of the estimation method was observed also for those cases where $\mathrm{D}_2x(t;\gamma)$ did not exist in the pointwise sense of Theorem 6.2.1 but only in a norm, as stated in Theorem 6.3.1. For example, when the initial function is approximated by piecewise linear splines, then Theorem 6.2.1 is not applicable to solutions corresponding to such parameters since they belong to $W^{1,\infty}$ but not to C^1.

7. Periodic solutions via fixed point theory

Over the last 40 years the most general results on existence of periodic solutions to autonomous delay differential equations, with constant and also with state-dependent delay, have been obtained using topological fixed point theorems and the fixed point index. The first step in applying these tools is the construction of a *return map*: For initial data in a suitably chosen set K one follows the solution segments until they return to K. Fixed points of the return map define periodic solutions. The search for a suitable set K requires a priori knowledge about the desired periodic solutions and about their role in the global dynamics generated by the delay equation. Hypotheses of fixed point theorems must also be satisfied. This indicates that in general the search for K may be a nontrivial task. The finer (more restrictive) a structure a domain K of a return map has, the more qualitative information about the periodic solution is provided.

It is not uncommon that domains of return maps or their closures contain a known fixed point which also is a stationary point of the semiflow. Therefore, to obtain a *nonconstant* periodic solution, one needs to find another fixed point. While impossible in case the trivial fixed point is globally attracting, there is hope in case the trivial fixed point is unstable. A weak topological notion of instability, which proved very useful, is Browder's concept of *ejectivity* [30]. A fixed point x of a map $f : M \to N$, N a topological space and $M \subset N$, is called ejective if there exists a neighbourhood V of x in M so that for each $y \in V \setminus x$ there is $j \in \mathbb{N}$ with $f^j(y) \notin V$. Ejectivity was deeply explored and first applied by Nussbaum (see, e.g., [174,175]), who also proved the first result on existence of periodic solutions for differential equations with state-dependent delay [174]. In the next subsection, Section 7.1, we describe a very general result of Mallet-Paret, Nussbaum, and Paraskevopoulos [156] on existence of periodic solutions and its proof, which is based on ejectivity.

The Section 7.2 deals with a model from Section 2.6 where the delay is governed by a differential equation which involves the state. This example was studied by Arino, Hadeler and Hbid [7] and Magal and Arino [148] and shows some of the specific difficulties caused by state-dependent delays in the search for periodic solutions.

Ejectivity is used also in the existence proofs in [4,136,137,151].

Of course, ejectivity does not adequately reflect the unstable behaviour of solutions to decent differential equations. Close to unstable manifolds of equilibria, or in cones around such unstable manifolds, solutions to delay and other differential equations behave much more regular than expressed by ejectivity. Accordingly one can prove existence of periodic solutions and global bifurcation from stationary points also without recourse to ejectivity. Schauder's fixed point theorem and simple calculations of the fixed point index suffice if only unstable solution behaviour close to equilibria is exploited to a larger extent. This was done in [202,203] and in Chapter XV of [54] for a class of RFDEs with constant delay. For equations with state-dependent delay, a proof of existence of periodic solutions along these lines has not yet been carried out.

Beyond existence and outside the scope of purely topological tools, uniqueness and stability of periodic orbits are of interest. In Section 7.3 we present results from [205, 208] where return maps are contractions or have a locally attracting fixed point, with the associated periodic orbit nontrivial, stable and hyperbolic. The approach applies to single equations and systems with state-dependent delay where the nonlinearities are given by functions which do not vary much on long intervals.

7.1. *A general result by continuation*

In [156] Mallet-Paret, Nussbaum and Paraskevopoulos prove existence of periodic solutions for a rather general class of scalar RFDEs which include equations with state-dependent delay. They consider Eq. (1.0.1) with $f : C \to \mathbb{R}$ continuous, $C = C([-h, 0], \mathbb{R})$, and assume that there are $\tau_0 \in (0, h]$ and a locally Lipschitz continuous function $g : \mathbb{R} \to \mathbb{R}$ satisfying the negative feedback condition

$$\xi g(\xi) < 0 \quad \text{for all } \xi \neq 0$$

so that

$$f(\phi) = g\big(\phi(-\tau_0)\big)$$

for ϕ in the closed hyperplane H given by $\phi(0) = 0$. In particular, $f(0) = 0$, τ_0 is the delay on H, and the constant function zero is a solution to Eq. (1.0.1). In Section 3 we mentioned the property (alL) from [156] of being *almost locally Lipschitzian*, to which condition (L) in Section 3 is closely related. Property (alL) requires local Lipschitz estimates for f which involve the norm on C but only arguments of f which are Lipschitz continuous. For f with property (alL) and at most linear growth Lipschitz continuous initial data $\phi \in C$ uniquely determine solutions $x : [-h, \infty) \to \mathbb{R}$ of Eq. (1.0.1) with $x_0 = \phi$. Also bounds for solutions and continuous dependence on initial data are established. Under a more restrictive negative feedback condition, now for the functional f and involving also data $\phi \in C \setminus H$, it is shown that segments x_t, $t \geqslant 0$, of solutions which start from $x_0 = \phi$ in a closed bounded convex set

$$G^+ \subset \big\{\phi \in H \colon \phi(t) \geqslant 0 \text{ for all } t \in [-h, 0]\big\}$$

return to G^+ at a certain well-defined zero $z_2 = z_2(x_0) > 0$ provided the solution has at least one sign change on $(0, \infty)$ and a zero thereafter. On a subset $U^+ \subset G^+$ which is open with respect to the topology on G^+ induced by C the previous result yields a continuous return map $\Gamma_0 : U^+ \to G^+$. Actually, more is achieved here for later use: $\Gamma_0 = \Gamma(\cdot, 0)$, for a homotopy $\Gamma : U^+ \times [0, 1] \to G^+$ of modified return maps $\Gamma(\cdot, \alpha)$ associated with the members of a one-parameter family of RFDEs; the equation at $\alpha = 1$ has the property that the values of initial data on $[-h, -\tau_0)$ have no influence on the solution.

The extremal point $0 \in G^+$ of G^+ does not belong to U^+, and each fixed point of the return map Γ_0 defines a non-constant periodic solution of Eq. (1.0.1). A fixed point exists if the fixed point index $i_{G^+}(\Gamma_0, U^+)$ of Γ_0 is defined and non-zero.

The remaining steps towards

$$i_{G^+}\bigl(\Gamma_0, U^+\bigr) \neq 0 \tag{7.1.1}$$

require some sort of linearization of the RFDE at the zero solution, in order to describe and exploit conditions for instability of the zero solution. Here the hypothesis (aFd) that f is *almost Fréchet differentiable* at 0 comes into play. It requires that the restriction of f to the space $C^{0,1} = C^{0,1}([-h, 0], \mathbb{R})$ of Lipschitz continuous data has a derivative at $0 \in C^1 \subset C^{0,1}$. The properties (alL) and (aFd) combined imply that this derivative $\mathrm{D}(f|_{C^{0,1}})(0)$ extends to a continuous linear map $\mathrm{D}_e(f|_{C^{0,1}})(0) : C \to \mathbb{R}$, like in conditions (S2) and (S3) from Section 3. A look back at the hypotheses on f shows that the recipe *freeze the delay at equilibrium, then linearize* mentioned in Section 3 is not sufficient to compute this extended derivative. The desired linear equation has the form

$$\begin{aligned} v'(t) &= \mathrm{D}_e(f|_{C^{0,1}})(0)v_t \\ &= -\beta\, v(t) - \gamma v(t - \tau_0) \end{aligned} \tag{7.1.2}$$

with constants $\beta \geqslant 0$, $\gamma \geqslant 0$, and is considered for initial data in the Banach space $C_{\tau_0} = C([-\tau_0, 0], \mathbb{R})$. On the cone

$$K = \bigl\{\phi \in C_{\tau_0} \colon\ \phi(t) \geqslant 0 \text{ for all } t \in [-\tau_0, 0],\ \phi(0) = 0\bigr\}$$

there is a return map $S_{\beta,\gamma} : K \to K$ similar to Γ_0, given by the solutions to Eq. (7.2.2), with $S_{\beta,\gamma}(0) = 0$. In case the zero solution of Eq. (7.2.2) is unstable the trivial fixed point $0 \in K$ has index zero;

$$i_K(S_{\beta,\gamma}, 0) = 0. \tag{7.1.3}$$

To see this one can use ejectivity as in [174,175], or follow [176], or proceed as in [203,54].

The major steps in the proof of (7.2.1) are a reduction to

$$i_{G^+}\bigl(\Gamma_1, U^+\bigr) \neq 0$$

by homotopy invariance, and the deduction of the preceding inequality from (7.2.3). In this last step all basic properties normalization, additivity, homotopy invariance, and commutativity of the fixed point index are instrumental.

The skillful proof in [156] overcomes more obstacles than this brief outline indicates. The main theorem of [156] yields a wide variety of results on existence of periodic solutions for explicitly given RFDEs, in particular also for equations with multiple state-dependent delays of the form

$$x'(t) = G\big(x(t), x\big(t - r_1\big(x(t)\big)\big), \ldots, x\big(t - r_m\big(x(t)\big)\big)\big).$$

7.2. *Periodic solutions when delay is described by a differential equation*

As mentioned above, Nussbaum [174] used the ejective fixed point theorem to prove the existence of periodic solutions to the equation

$$x'(t) = -\alpha x\big(t - 1 - \big|x(t)\big|\big)\big(1 - x^2(t)\big).$$

Alt [4] and Kuang and Smith [136,137] obtained periodic solutions for equations where the delay is given by a threshold condition.

While the major steps towards existence of so-called slowly oscillating periodic solutions remain essentially the same as for delay differential equations with constant delays, the technical details become more involved when delays are state-dependent. This refers both to the construction of the domain of a return map and to the verification that a trivial fixed point is ejective. To illustrate this we describe in the sequel work of Arino, Hadeler and Hbid [7] and of Magal and Arino [148] for the system

$$\begin{cases} x'(t) = -f\big(x\big(t - r(t)\big)\big), \\ r'(t) = q\big(x(t), r(t)\big). \end{cases} \tag{7.2.1}$$

Here the variation of the delay is determined by an ordinary differential equation. Standing assumptions are that the functions $f : \mathbb{R} \to [-M, M]$, $M > 0$, and $q : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are continuously differentiable, with

$$xf(x) > 0 \quad \text{for all } x \neq 0 \text{ and } f \text{ nondecreasing,}$$

and that for some fixed reals $r_2 > r_1 > 0$ and for all $x \in \mathbb{R}$,

$$q(x, r_2) < 0 < q(x, r_1).$$

Let $C = C([-r_2, 0], \mathbb{R})$. The hypotheses ensure that the IVP given by the system (7.2.1) and initial data $(\phi, r_0) \in C \times [r_1, r_2]$ with ϕ Lipschitz continuous has a unique solution $(x, r) = (x^{\phi, r_0}, r^{\phi, r_0})$, with x defined on $[-r_2, \infty)$ and r defined on $[0, \infty)$. The solution component x is Lipschitz continuous with $|x'(t)| \leqslant M$ for all $t > 0$, and the hypothesis on q yields $r_1 < r(t) < r_2$ for all $t > 0$.

Incidentally, note that the system (7.2.1) can be reformulated as a special case of (1.0.1) with $n = 2$, $h = r_2$, $U = (\mathbb{R} \times [0, h])^{[-h,0]}$, and the corresponding functional being given by $(-f(\phi(-r_0)), q(\phi(0), r_0))$ for $(\phi, r_0) \in U$.

It also follows that the function $q(0,\cdot)$ has zeros in $[r_1, r_2]$. The additional hypothesis

$$\frac{\partial q}{\partial r}(0,r) < 0 \quad \text{on } [r_1, r_2]$$

implies that there is exactly one zero r_0^* of $q(0,\cdot)$ in $[r_1, r_2]$, and that the constant solution $r^*: t \mapsto r_0^*$ of the autonomous equation

$$r' = q(0,r)$$

is asymptotically stable. Below we shall see that this fact causes complications in view of ejectivity of a fixed point corresponding to the constant solution of system (7.2.1) given by $x(t) = 0$ and r^*.

In order to obtain that $t \mapsto t - r(t)$ is strictly increasing it is furthermore assumed that

$$q(x,r) < 1 \quad \text{on } \mathbb{R} \times [r_1, r_2].$$

The notion of a slowly oscillating solution, which is familiar from work on equations with constant delay $h > 0$ like, e.g.,

$$x'(t) = -f\big(x(t-h)\big)$$

and means that zeros are isolated and spaced at distances larger than the delay h, is modified for the x-components of solutions to the system (7.2.1) according to the following definition: A function $x : [t_0, \infty) \to \mathbb{R}$, $t_0 \in \mathbb{R}$, is called slowly oscillating if its zeros form a disjoint union of closed intervals Z whose left endpoints have no accumulation point and are spaced at distances not less than r_2, with

$$\lim_{t \nearrow a} \operatorname{sign}\big(x(t)\big) = -\lim_{s \searrow b} \operatorname{sign}\big(x(s)\big)$$

at each compact interval $Z = [a,b]$ with $t_0 < a$. Notice that the definition allows nonconstant functions which are zero on some unbounded interval $[t_1, \infty)$, $t_1 > t_0$, a phenomenon which occurs among the first components of solutions to the system (7.2.1).

The first step towards a return map is to show that certain initial data define solutions with slowly oscillating first component. In [7] it is shown that the first components of solutions with initial value component ϕ in the cone

$$\Gamma = \big\{\phi \in C([-r_2, 0]; \mathbb{R})\colon \text{ there exists } \theta \in [-r_2, -r_1] \text{ so that } \phi(\theta) = 0$$
$$\text{and } \phi(s) < 0 \text{ for } s < \theta;\, 0 \leqslant \phi(\theta) \leqslant \phi(0) \text{ for } \theta \leqslant s \leqslant 0\big\}$$

return to Γ at a sequence of times whose distances are not less than r_2. More precisely, if (x, r) is a solution with initial value (ϕ, r_0) and $\pm\phi \in \Gamma$ then there are two sequences, possibly finite, of points t_i^* and zeros t_i of x in $[0, \infty)$ such that

$$t_0^* \geqslant 0, \quad t_i^* + (r_2 - r_1) \leqslant t_{i+1} \leqslant t_{i+1}^* - r_1$$

and $\pm(-1)^{i+1}x$ is nondecreasing on the interval $[t_i^*, t_{i+1}^*]$. In particular, $\pm(-1)^i x_{t_i^*} \in \Gamma$ for each index $i > 0$. To reach the conclusion, it is required that the delay variation satisfies the smallness condition $(r_2 - r_1)|f(x)| < |x|$ for $x \neq 0$.

Uniqueness with respect to initial data is needed, and the domain for a return map should be convex. Therefore Γ must be modified. The smaller set

$$\Gamma_1 = \{\phi \in \Gamma\colon \phi \text{ is Lipschitzian and nondecreasing}\}$$

is convex, and closed in $C^{0,1}([-r_2, 0], \mathbb{R})$. Let $E = \Gamma_1 \times [r_1, r_2]$. For each integer $j \geqslant 1$, it is now natural to introduce operators P_j and P_j^+ on the space E by

$$P_j(\phi, r_0) = \bigl(x_{t_j^*}, r\bigl(t_j^*\bigr)\bigr),$$

$$P_j^+(\phi, r_0) = \bigl((-1)^j x_{t_j^*}, r\bigl(t_j^*\bigr)\bigr).$$

Obviously, the existence of $P_j(\phi, r_0)$ is subject to the condition that $t_1^*, t_2^*, \dots, t_{j-1}^*$ exist. If this condition is satisfied but the solution starting from $(x_{t_{j-1}^*}, r(t_{j-1}^*))$ does not cross zero, then it is proved in [7] that $x(t) \to 0$ and $r(t) \to r^*$ as $t \to \infty$. In this case, one defines $P_j(\phi, r_0) = (0, r^*)$.

P_1 sends bounded sets of E into bounded sets of the product space $C([-r_2, 0], \mathbb{R}) \times [r_1, r_2]$, and P_1^+ is a compact and continuous operator from E into itself.

It is important to know when a first positive zero $t_1 = t_1(\phi, r_0)$ exists: This property is ensured for every $(\phi, r_0) \in (\Gamma_1 \cup (-\Gamma_1)) \times [r_1, r_2]$ if there exist $M > 0$ and $R_0 > 0$ with $M(r_2 - r_1) \leqslant R_0$, $|f(x)| \leqslant M$ for $x \in \mathbb{R}$ and $(2r_1 - r_2)|f(x) \geqslant |x|$ for $|x| \leqslant R_0$.

Unfortunately, the fixed point $(0, r^*)$ of any return map defined in a set containing $\{0\} \times [r_1, r_2]$ is not ejective. This follows immediately from positive invariance of the set $\{0\} \times [r_1, r_2]$ and asymptotic stability of the solution r^* to the equation $r' = q(0, r)$.

In order to obtain ejectivity smaller convex subsets

$$E_K = \bigl\{(\phi, r_0) \in E\colon \bigl|r_0 - r_0^*\bigr| \leqslant K\|\phi\|_C, \|\phi\|_C = \bigl|\phi(0)\bigr|\bigr\}$$

for $K > 0$ are introduced which contain from the obstacle for ejectivity $\{0\} \times [r_1, r_2]$ only the point $(0, r_0^*)$. For $R > 0$ and $K > 0$ let

$$E_{R,K} = \bigl\{(\phi, r_0) \in E\colon \|\phi\|_C \leqslant R, \bigl|r_0 - r_0^*\bigr| \leqslant K\|\phi\|_C\bigr\},$$

and define

$$\Gamma_2 = \Bigl\{\phi \in \Gamma_1\colon \phi(0) = \sup_{-r_2 \leqslant s \leqslant 0} \bigl|\phi(s)\bigr|\Bigr\}.$$

In order for P_1^+ to map $(\Gamma_2 \times [r_1, r_2]) \cap E_{R,K}$ into E_K, a restriction on r_1 is needed. In [7] it is shown that for each $R > 0$ there exist $\tilde{r}_1 > 0$ and $K > 0$ such that for each $r_1 > \tilde{r}_1$

the operator P_1^+ maps $(\Gamma_2 \times [r_1, r_2]) \cap E_{R,K}$ into E_K. The key to the proof of the above statement is that for each $R > 0$ there exists $C(r_1) \in (0, 1)$ such that

$$\left\| x_{t_1^*}^{\phi, r_0} \right\|_C \geqslant C(r_1) \|\phi\|_C \tag{7.2.2}$$

for each $(\phi, r_0) \in \Gamma_2 \times [r_1, r_2]$ with $0 < \|\phi\|_C \leqslant R$. Estimates of this type exclude super-exponential decay of slowly oscillating solutions and play an important role in work on the global dynamics of equations with constant delay.

The map $(\phi, r_0) \mapsto -x_{t_1^*}^{\phi, r_0}$ sends $E = \Gamma_1 \times [r_1, r_2]$ into Γ_2. The fact that $t_1 = t_1(\phi, r_0)$ is a zero of $x = x^{\phi, r_0}$ and integration of the first equation in (7.2.1) yield the estimate

$$\left\| x_{t_1^*}^{\phi, r_0} \right\|_C \leqslant M r_2,$$

for each $(\phi, r_0) \in E$. With this preparation, we can define $R = M r_2$ and introduce the set

$$X = \left\{ \phi \in \Gamma_2 \colon \ \|\phi\|_C \leqslant R, \ \operatorname*{ess\,sup}_{-r_2 \leqslant s \leqslant 0} \left| \phi'(s) \right| \leqslant M \right\}.$$

Choose $K > 0$ and $\tilde{r}_1 > 0$ as above, assume $r_1 > \tilde{r}_1$, and set

$$Y = \bigl(X \times [r_1, r_2] \bigr) \cap E_{R,K}.$$

Y is closed and convex, and the iterate P_1^2 defines a map $P : Y \to Y$.

It remains to show that $(0, r_0^*)$ is an ejective fixed point of P. Ejectivity of $(0, r_0^*)$ is a consequence of instability of the constant solution $t \mapsto (0, r_0^*)$ to the system (7.2.1), which in turn follows from instability for the linearized system. In [7] the technique *freeze the delay at equilibrium, then linearize* mentioned in Section 3.4 is applied to the first equation in (7.2.1) and yields

$$y'(t) = -y\bigl(t - r_0^*\bigr), \tag{7.2.3}$$

which is unstable for $r_0^* > \frac{\pi}{2}$ when f normalized so that $f'(0) = 1$. More precisely, the eigenvalues of the generator of the semigroup generated by Eq. (7.2.2) on the space

$$C\bigl(\bigl[-r_0^*, 0\bigr], \mathbb{R}\bigr)$$

with largest real part are a complex conjugate pair $u \pm i v$ in the open right halfplane, and the associated realified generalized eigenspace U is 2-dimensional and consists of segments of solutions

$$\mathbb{R} \ni t \mapsto \mathrm{e}^{ut} \bigl(a \cos(vt) + b \sin(vt) \bigr)$$

to Eq. (7.2.3). For $(a, b) \neq (0, 0)$ these solutions are slowly oscillating, due to $|v| < \frac{\pi}{r_0^*}$. The problem is now to transfer such unstable solution behaviour to the slowly oscillating solution components x of solutions (x, r) to the nonlinear system (7.2.1) which start from small initial data $(\phi, \rho) \in Y$ in the other state space $C([-r_2, 0], \mathbb{R}) \times [r_1, r_2]$. The proof in [7] proceeds by contradiction. Ejectivity means that there exists $\epsilon > 0$ so that for each $(\phi, r_0) \in Y$ with $0 < \|\phi\|_C \leqslant \epsilon$ there is an integer $j > 0$ such that $(x, r) = (x^{\phi,r_0}, r^{\phi,r_0})$ and $t_j^* = t_j^*(\phi, r_0)$ satisfy $\|x_{t_j^*}\|_C \geqslant \epsilon$ or $|r(t_j^*) - r_0^*| \geqslant \epsilon$. It can be shown that in case the previous statement on ejectivity is not true, then there exists a constant $d > 0$ so that for all $(\phi, r_0) \in Y$ with $\|\phi\|_C \leqslant \epsilon$ and for all $t \geqslant 0$ we have $\|x_t^{\phi,r_0}\|_C \leqslant d\epsilon$ and $|r(t) - r_0^*| \leqslant d\epsilon$.

The next step is to show that the spectral projection Π_U onto the realified generalized eigenspace U associated with $u \pm v$ in $C([-r_0^*, 0], \mathbb{R})$ satisfies

$$\gamma := \inf_{\|\phi\|=1, \phi\in\tilde{\Gamma}_2} \|\Pi_U \phi\| > 0, \tag{7.2.4}$$

under a further restriction about the variation of the delay

$$r_2 - r_1 < \delta \tag{7.2.5}$$

for some constant $\delta > 0$ (explicitly given in [7]) that depends on r_0^* only. Here and in what follows, $\tilde{\Gamma}_2$ is defined in a similar fashion as Γ_2, except we replace Γ_1 and Γ by $\tilde{\Gamma}_1$ and $\tilde{\Gamma}$ respectively, where the domain of ϕ is $[-r_0^*, 0]$ rather than $[-r_2, 0]$.

Let $\sup_{j\in\mathbb{N}} \|x_{t_j^*}\phi, r_0\|_C = \tilde{\epsilon} \leqslant \epsilon$. We can choose an integer $j_0 > 0$ so that $\|x_{t_{j_0}^*}\phi, r_0\|_C \geqslant C\tilde{\epsilon}$ and thus, $|y(t_{j_0}^*)| \geqslant C\tilde{\epsilon}\gamma$, where $y(t) = \Pi_U x_t$, C and γ are given in (7.2.2) and (7.2.4). Note that the first equation of the system (7.2.1) can be written as $x'(t) = -x(t - r^*) + o(\tilde{\epsilon})$ for large t, by using the linearization and the fact that $|x'(t)| \leqslant M$ and $|r(t) - r_0^*| \leqslant d\tilde{\epsilon}$ for all $t \geqslant 0$.

Therefore,

$$y'(t) = A_U y(t) + \mathrm{o}(\tilde{\epsilon})$$

with $A_U = \left(\begin{smallmatrix} \alpha & -\beta \\ \beta & \alpha \end{smallmatrix}\right)$. It then follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left|y(t)\right| = \alpha\left|y(t)\right| + \mathrm{o}(\tilde{\epsilon}).$$

Therefore,

$$\left|y(t)\right| = \mathrm{e}^{\alpha(t-\sigma)}\left[\left|y(\sigma)\right| + \int_\sigma^t \mathrm{e}^{\alpha(\sigma-s)}\mathrm{o}(\tilde{\epsilon})\,\mathrm{d}s\right],$$

which implies

$$\left|y(t)\right| \geqslant \mathrm{e}^{\alpha(t-t_{j_0}^*)}\left[C\tilde{\epsilon}\gamma - \frac{\mathrm{o}(\tilde{\epsilon})}{\alpha}\right] \geqslant \mathrm{e}^{\alpha(t-t_{j_0}^*)}\frac{C\gamma}{2}\tilde{\epsilon}$$

if $\tilde{\epsilon}$ is small. This leads to a contradiction since $y(t) = \Pi_U x_t$ should be bounded by a constant multiple of $\tilde{\epsilon}$ for all $t \geqslant 0$.

In summary, under a few technical conditions including the negative feedback condition on the state variable x and the delay r, the smallness of the delay variation and an instability condition for an associated linear equation, Arino, Hadeler and Hbid obtain in [7] the existence of periodic solution with slowly oscillating first component.

In [148] Magal and Arino obtain such periodic solutions under weaker conditions on the delay variation—(7.2.5) is no longer needed—by means of a different argument which uses a modification of ejectivity and employs other cones of initial data. Magal and Arino consider a cone that was already used by Kuang and Smith in [136,137], namely,

$$\begin{aligned}E_{\mathrm{KS}} = \big\{&(\phi, r_0) \in C^{0,1}\big([-r_2, 0], \mathbb{R}\big) \times [r_1, r_2]\colon \phi(-r_0) = 0\\ &\text{and } \phi \text{ nonincreasing on } [-r_0, 0]\big\}.\end{aligned}$$

Using arguments similar to those in [7] one can show that for $\epsilon = \pm 1$ and $(\epsilon\phi, r_0) \in E_{KS}$ there are reals $t_i^* = t_i^*(\phi, r_0)$ and zeros $t_i = t_i(\phi, r_0)$ of $x = x^{\phi,r_0}$ such that $t_0 = -r_0$, $t_0^* = 0$, $t_i^* \leqslant t_{i+1}$, $t_i = t_i^* - r(t_i^*)$ for integers $i \geqslant 0$, and $\epsilon(-1)^{i+1}x(t)$ is nonincreasing on $[t_i^*, t_{i+1}^*]$, $x(t_i) = 0$, and $x(t_i^*) \neq 0$ if $\phi(0) \neq 0$. So, $(\epsilon(-1)^{i+1}x_{t_i^*}, r(t_i^*)) \in E_{\mathrm{KS}}$. Let $X_0 = C^1([-r_2, 0], \mathbb{R}) \times [r_1, r_2]$, introduce

$$E_0 = \big\{(\phi, r_0) \in X_0\colon \phi'(s) \geqslant 0 \text{ for } s \in [-r_0, 0], \phi(-r_0) = 0, \phi'(0) = 0\big\},$$

and define the return maps P_j and P_j^+ on E_0 exactly as in [7]. For each integer $j > 0$ one finds $P_{2j}(E_0) \subset E_0$.

The fact that for initial data $(\phi, r_0) \in C^1([-r_2, 0], \mathbb{R}) \times [r_1, r_2]$ the map $t \mapsto t - r^{\phi,r_0}(t)$ is increasing implies that on $[0, \infty)$ the solution $(x^{\phi,r_0}, r^{\phi,r_0})$ does not depend on the restriction of ϕ to $[-r_2, -r_0)$. The preceding observation suggests to consider initial data $(\tilde{\phi}, r_0)$ with $\tilde{\phi}$ defined only on $[-r_0, 0]$, and secondly, to modify the fixed point problem by transforming the initial delay r_0 to 1. Let

$$X_1 = C^1\big([-1, 0], \mathbb{R}\big) \times [r_1, r_2].$$

Magal and Arino introduce maps

$$L : X_0 \to X_1 \quad \text{and} \quad Q : X_1 \to X_0$$

by

$$L(\phi, r_0) = (\psi, r_0), \quad \psi(s) = \phi(s\, r_0)$$

and

$$Q(\psi, r_0) = (\phi, r_0), \quad \phi(s) = \psi(s/r_0) \quad \text{on } [-r_0, 0],$$

$$\phi(s) = \frac{\psi'(-1)}{r_0}(s - r_0) \quad \text{on } [-r_2, -r_0).$$

Let E_1 denote the analogue of E_0 where X_0 and r_0 are replaced by X_1 and 1, respectively. E_1 is closed and convex, and $Q(E_1) \subset E_0$, $L(E_0) \subset E_1$. The maps $F_{2j} = L \circ P_{2j} \circ (Q|E_1)$ send E_1 into E_1. We have $F_2(0, r_0^*) = (0, r_0^*)$, and fixed points (ϕ, r_0) of F_2 with $\phi \neq 0$ yield periodic solutions of the system (7.2.1) with slowly oscillating first component. However, the second component of $F_2 : E_1 \to C^1([-1,0], \mathbb{R}) \times [r_1, r_2]$ is not continuous at points $(0, r_0)$ with $r_0 \neq r_0^*$, which requires a further modification. The map $\tilde{F}_2 : E_1 \to E_1$ resulting from this is continuous and compact (with respect to the topology on $C^1([-1,0], \mathbb{R})$) and retains the property that nontrivial fixed points define periodic solutions to (7.2.1) with slowly oscillating first component. It is shown in [148] that

(i) $\tilde{F}_2(0, r_0^*) = (0, r_0^*)$,

(ii) $\tilde{F}_2(\{0\} \times [r_1, r_2]) \subset \{0\} \times [r_1, r_2]$, and

(iii) for every $\epsilon > 0$ there exist $c > 0$ and $\gamma \in [0, 1)$ with

$$\left|\left(\tilde{F}_1^{2j}\right)_2(\phi, r_0) - r_0^*\right| \leqslant \gamma \left|r_0 - r_0^*\right|$$

for all $(\phi, r_0) \in E_1$ with

$$\|\phi\|_{C^1([-1,0],\mathbb{R})} + \left|r_0 - r_0^*\right| \leqslant \epsilon$$

and

$$\|\phi\|_{C^1([-1,0],\mathbb{R})} \leqslant c\left|r_0 - r_0^*\right|,$$

and for all integers $j \geqslant 1$.

Here the index 2 denotes the second component of the map $\tilde{F}_2^{2j}$. Properties (ii) and (iii) combined exclude that the fixed point $(0, r^*)$ of $\tilde{F}_2$ is ejective. Now the modification of ejectivity comes into play. Let X be a Banach space, $A \subset Y \subset X$, and assume that $g : Y \to Y$ has a fixed point $x_0 \in \partial_Y A$. Then x_0 is called semi-ejective on $Y \setminus A$ if there is a neighbourhood V of x_0 in Y so that for each $y \in V \setminus A$ there is an integer $m \geqslant 1$ with

$$g^m(y) \in Y \setminus V.$$

Arguments similar to those in [7] which exploit the instability of Eq. (7.2.3) yield in [148] that the fixed point $(0, r_0^*)$ is semi-ejective on $E_1 \setminus (\{0\} \times [r_1, r_2])$. Finally, an extension of the ejective fixed point theorem in [148] to the case of semi-ejective fixed points guarantees that $\tilde{F}_2$ has a fixed point (ϕ, r_0) with $\phi \neq 0$, which defines the desired periodic solution.

7.3. *Attracting periodic orbits*

In [205] the system (2.2.1)–(2.2.2)

$$x'(t) = v\left(\frac{c}{2}s(t - r) - w\right),$$

$$cs = \left|x(t-s) + w\right| + \left|x(t) + w\right|,$$

with positive parameters c, r, w, is studied. The function $v : \mathbb{R} \to \mathbb{R}$ is assumed to satisfy a negative feedback condition

$$\delta v(\delta) < 0 \quad \text{for all } \delta \neq 0.$$

For continuous and bounded v the system defines a continuous semiflow on a set O which is open in a compact set $M \subset C = C([-h,0],\mathbb{R})$, with $h = \frac{4w}{c} + r$. The set M consists of Lipschitz continuous functions. The set O contains a closed subset K of initial data which define solutions whose segments x_t return to K, after an excursion into the ambient space O. This yields a continuous return map, on a domain without stationary points. If v is Lipschitz continuous and close to constants in $(-\infty, -\beta]$ and $[\beta, \infty)$, respectively, with $\beta > 0$ sufficiently small, then one can estimate the Lipschitz constant of the return map in terms of Lipschitz constants for v and for its restrictions $v|_{(-\infty,-\beta]}$ and $v|_{[\beta,\infty)}$; under suitable further assumptions the return map becomes a contraction. The unique fixed point of the contracting return map belongs to a periodic orbit of the system (2.2.1)–(2.2.2) which is stable and exponentially attracting with asymptotic phase.

The observation which led to the method used in [205] and in earlier work on equations with constant delays is the following: If the function $g : \mathbb{R} \to \mathbb{R}$ in the equation

$$y'(t) = g\big(y(t-1)\big) \tag{7.3.1}$$

is constant on some interval I and if y remains long enough in I, say, for $t_0 - 1 \leqslant t \leqslant t_0$, then for $t \geqslant t_0$ the solution y depends only on $y(t_0)$ and $g(I)$. This can be used to design simple-looking nonlinearities g, representing negative feedback, for which periodic solutions of Eq. (7.3.1) can be computed explicitly, see, e.g., Chapter XV in [54]. Moreover, solutions which start from initial data close to the periodic orbit eventually merge into it. This is an extremely strong kind of orbital stability, giving hope that also for nonlinearities which are only close to constants on some intervals attracting periodic orbits may exist. If instead of the scalar equation (7.3.1) more generally systems are considered then suitable nonlinearities which are constant on nontrivial intervals yield low-dimensional subsets of the state space which are positively invariant under the semiflow and absorb flowlines from a neighbourhood.

In [208] the system

$$u' = v, \tag{7.3.2}$$

$$v' = -r\,v + A(p), \tag{7.3.3}$$

$$p = \frac{c}{2}s - w, \tag{7.3.4}$$

$$cs = u(t-s) + u(t) + 2w \tag{7.3.5}$$

is studied. In contrast to the system (2.2.1)–(2.2.2) the model (7.3.2)–(7.3.5) for position control by echo is now based on Newton's law. Instead of the constant time lag $r > 0$ in Eq. (2.2.1) there is now a friction term $-rv$ in Eq. (7.3.3).

For suitable positive values of the parameters w, c, r and for certain functions $A : \mathbb{R} \to \mathbb{R}$ which represent negative feedback and are constant on $(-\infty, -\beta]$ and on $[\beta, \infty)$, with $\beta > 0$ small, the existence of a hyperbolic stable periodic orbit is established. The proof begins with a reformulation of the system as an equation of the form (1.0.1), with a continuously differentiable functional f defined on an open subset U of the space $C^1 = C^1([-h, 0], \mathbb{R}^2)$, for suitable $h > 0$. Several steps of the proof rely on the smoothness properties of the semiflow F on the solution manifold $X_f \subset U$ which are provided by Theorem 3.2.1. In X_f a thin, infinite-dimensional set I of initial data ϕ is found to which the flowlines $F(\cdot, \phi)$ return, after a journey through the ambient part of the manifold. The associated return map is not necessarily compact, which precludes an immediate application of Schauder's theorem in order to find a fixed point—not to speak of an attracting fixed point. But the return map is semiconjugate to an interval map which is differentiable. Estimates of derivatives of the interval map yield a unique, attracting fixed point of the latter, which can be lifted to the return map. The proof that the resulting periodic orbit of the system (7.3.2)–(7.3.5) is hyperbolic and stable involves a continuously differentiable Poincaré return map, on a hyperplane transversal to the periodic orbit, in addition to the previous return map on the thin set $I \subset X_f$, and a discussion of derivatives of iterates.

8. Attractors, singular perturbation, small delay, generic convergence, stability and oscillation

This section deals with limiting behaviour. Section 8.1 is concerned with long term dynamics and reports about the structure of a global attractor [132]. Section 8.2 describes work of Mallet-Paret and Nussbaum [151–155] about the asymptotic shape of periodic solutions when a parameter becomes large, Section 8.3 sketches an approach of Ouifki and Hbid [178] to existence of periodic solutions when delays are small, Section 8.4 reports about work of Bartha [21] on generic convergence of solutions, and Section 8.5 comments on further results about stability and oscillatory solution behaviour.

8.1. *An attracting disk*

The paper [132] studies the equation

$$x'(t) = -\mu x(t) + f\bigl(x\bigl(t - r\bigl(x(t)\bigr)\bigr)\bigr) \tag{8.1.1}$$

with $\mu > 0$, $f \in C^2(\mathbb{R}, \mathbb{R})$, $f(0) = 0$, $f'(u) < 0$ for all $u \in \mathbb{R}$, $r \in C^1(\mathbb{R}, \mathbb{R})$, $r(0) = 1$, and $\sup_{u \in \mathbb{R}} f(u) < \infty$ provided $r(u) \geqslant 0$ for all $u \in \mathbb{R}$. The case $\mu = 0$ can also be handled with a slight modification. Then the delayed logistic equation (or Wright's equation)

$$y'(t) = -\alpha y\bigl(t - r\bigl(y(t)\bigr)\bigr)\bigl[1 + y(t)\bigr]$$

with state-dependent delay and solutions satisfying $y(t) > -1$ is a particular case. Indeed, after the transformation $x = \log(1+y)$ we obtain

$$x'(t) = -\alpha\bigl[\mathrm{e}^{x(t-r(\mathrm{e}^{x(t)}-1))} - 1\bigr].$$

The aim is to describe the asymptotic behaviour of the slowly oscillatory solutions of Eq. (8.1.1). Here a solution x of (8.1.1) is called slowly oscillatory if $|z' - z| > r(0) = 1$ for every pair of zeros z', z of x. The results are in part analogous to those of Walther [204] for the constant delay case $r \equiv 1$.

Let I_r denote the largest subinterval of $\mathbb{R}$ with $0 \in I_r$ and $r(u) \geqslant 0$ for all $u \in I_r$. First it is shown that for every element ϕ of the space $BC((-\infty, 0], I_r)$ of bounded continuous functions from $(-\infty, 0]$ into I_r, there is a solution $x : \mathbb{R} \to \mathbb{R}$ of Eq. (8.1.1) through ϕ, i.e., x is continuous on $\mathbb{R}$, continuously differentiable on $(0, \infty)$, (8.1.1) holds for all $t > 0$, and $x|_{(-\infty,0]} = \phi$. If ϕ is Lipschitz continuous then x is unique.

In the next step four positive constants A, B, R, K are constructed such that

$$r\bigl([-B, A]\bigr) \subset (0, R],$$
$$\max_{(u,v)\in[-B,A]\times[-B,A]} \bigl|-\mu u + f(v)\bigr| \leqslant K,$$

moreover, for any solution $x : \mathbb{R} \to \mathbb{R}$ of (8.1.1) with $x|_{(-\infty,0]} \in BC((-\infty, 0], I_r)$ there exists $T \geqslant 0$ such that

$$x(t) \in [-B, A] \quad \text{for all } t \geqslant T.$$

Therefore, from the point of view of the asymptotic ($t \to \infty$) behaviour, only those solutions are interesting which have values in $[-B, A]$.

Let C_R denote the space $C([-R, 0], \mathbb{R})$ equipped with the supremum norm. The set

$$L_K = \left\{\phi \in C_R\colon\ \phi\bigl([-R, 0]\bigr) \subset [-B, A],\quad \sup_{-R\leqslant s<t\leqslant 0} \frac{|\phi(t) - \phi(s)|}{t - s} \leqslant K\right\}$$

is a compact convex subset of C_R. For every $\phi \in L_K$, Eq. (8.1.1) has a unique solution $x^\phi : [-R, \infty) \to \mathbb{R}$ with $x|_{[-R,0]} = \phi$ and $x(t) \in [-B, A]$ for all $t \geqslant 0$. Then the relations

$$F(t, \phi) = x_t^\phi, \quad t \geqslant 0, \qquad x_t^\phi(s) = x^\phi(t + s), \quad -R \leqslant s \leqslant 0,$$

define a continuous semiflow F on L_K. In the sequel, only those solutions of (8.1.1) are considered whose segments are in L_K.

Define the compact subset

$$S = \bigl\{\phi \in L_K\colon\ \mathrm{sch}\bigl(\phi, [t - 1, t]\bigr) \leqslant 1 \text{ for all } t \in [-R + 1, 0]\bigr\}$$

of L_K, where $\mathrm{sch}(\phi, [t-1, t])$ denotes the number of sign changes of ϕ on the interval $[t - 1, t]$. If x is a slowly oscillatory solution of (8.1.1), then all segments x_t belong to S.

The set S is positively invariant under the semiflow F. The restriction of F to $[0,\infty)\times S$ defines the continuous semiflow F_S on the compact metric space S. The global attractor $\mathcal{A}$ of F_S has the following properties:

(i) $\mathcal{A}$ is a compact connected subset of $S \subset L_K$.

(ii) For each $\phi \in \mathcal{A}$ there is a unique solution of (8.1.1) through ϕ on $\mathbb{R}$, which is also denoted by x^ϕ. The map $F_{\mathcal{A}} : \mathbb{R}\times\mathcal{A} \ni (t,\phi) \mapsto x_t^\phi \in \mathcal{A}$ is a continuous flow.

(iii) $\mathcal{A}$ is the union of $0 \in C_R$ and the segments x_t of the globally defined slowly oscillating solutions $x:\mathbb{R}\to[-B,A]$ of Eq. (8.1.1).

The first main result is that a Poincaré–Bendixson type theorem holds on $\mathcal{A}$: The α- and ω-limit sets of phase curves in $\mathcal{A}$ are either $\{0\}$ or periodic orbits given by slowly oscillating periodic solutions. The second main result is that in case $\mathcal{A}\neq\{0\}$, the set $\mathcal{A}$ is homeomorphic to the 2-dimensional closed unit disk so that the unit circle corresponds to a periodic orbit given by a slowly oscillating periodic solution.

Below we list some of the technical tools used in the proofs.

There is an additional assumption on the delay function r: either

$$\left|r'(u)\right| < \frac{1}{K}, \quad u \in [-B, A],$$

or

$$r \in C^2\big([-B,A],\mathbb{R}\big), \text{ and there is } a \in (0,1) \text{ with } r''(u) \leqslant a\mu\big[r'(u)\big]^2, \\ u \in [-B,A].$$

This assumption and the fact that the dependence of the delay on the state is of the simple form $r(x(t))$ seem to be crucial in several parts of the proof.

An important consequence of the above hypothesis on r is that the function

$$t \mapsto t - r\big(x(t)\big)$$

is strictly increasing for solutions of (8.1.1). Another important fact is that for a suitable weighted difference $v:\mathbb{R}\to\mathbb{R}$ of two solutions x and y on $\mathbb{R}$, an equation of the form

$$v'(t) = \alpha(t)v\big(t - r\big(x(t)\big)\big) \tag{8.1.2}$$

holds on $\mathbb{R}$ with a negative, bounded and continuous α. The backward uniqueness of solutions is a corollary.

A modified version of the well-known discrete Lyapunov functional of Mallet-Paret and Sell [150,157,158] is also introduced. Instead of on intervals with fixed length, the sign changes are counted for a solution x on intervals of the form $[t - r(x(t)), t]$. The properties are completely analogous to those of the constant delay case. In particular, this functional can be used to exclude the existence of solutions decaying to zero at ∞ or $-\infty$ faster than any exponential. By applying the discrete Lyapunov functional to a weighted difference of two solutions satisfying Eq. (8.1.2), the number of sign changes of the difference can be controlled.

A return map P on the compact convex set

$$U = \{\phi \in L_K\colon\ \phi(0) = 0,\ \phi(s) \geqslant 0 \text{ for all } s \in [-1, 0]\}$$

is defined by $P(\phi) = x^{\phi}_{z_2}$, where at z_2 the second sign change of x^{ϕ} in $(0, \infty)$ occurs, and $P(\phi) = 0$ if there is at most one sign change in $(0, \infty)$. P is not necessarily continuous on U. However, $P|_{\mathcal{A}\cap U}$ is continuous. In addition, P restricted to $\{\phi \in \mathcal{A} \cap U\colon\ P(\phi) \neq 0\}$ is a homeomorphism onto $\mathcal{A} \cap U \setminus \{0\}$. It is also an essential step that $\mathcal{A} \cap U$ is connected. The elements of $\mathcal{A} \cap U \setminus \{0\}$ are exactly those segments x_s of globally defined slowly oscillating solutions $x : \mathbb{R} \to [-B, A]$ for which $x(s) = 0$ and $x'(s) < 0$.

An asymptotic expansion for slowly oscillating solutions converging to zero as $t \to -\infty$ is also proved. It relates solutions of Eq. (8.1.1) to solutions of the associated linear equation

$$y'(t) = -\mu y(t) + f'(0)y(t-1).$$

This result is used to verify that for any two elements ϕ, ψ of $\mathcal{A}$, the difference of the solutions $x^{\phi} - x^{\psi}$ has at most one sign change in all intervals $[t - r(x^{\phi}(t)), t]$ and $[t - r(x^{\psi}(t)), t]$, $t \in \mathbb{R}$. This fact is important in the proof of the injectivity of the map

$$\Pi : \mathcal{A} \ni \phi \mapsto \begin{pmatrix} \phi(0) \\ \phi(-r(\phi(0))) \end{pmatrix} \in \mathbb{R}^2.$$

The paper [22] considers Eq. (8.1.1) in the positive feedback case, i.e., $f' > 0$, and proves certain results which are analogous to the constant delay case in [133].

We remark that in the constant delay case, it is also known that the attractor of the slowly oscillating solutions is a C^1-smooth submanifold of the phase space [210]. Another remarkable result is that the domain of attraction is an open dense subset of the phase space [159]. Whether these remain true for the state-dependent delay case are open problems.

8.2. *Limiting profiles for a singular perturbation problem*

In their series of papers [151,152,155] Mallet-Paret and Nussbaum determine the asymptotic shape, or *limiting profile*, of periodic solutions to equations of the form

$$\epsilon x'(t) = f\bigl(x(t), x\bigl(t - r\bigl(x(t)\bigr)\bigr)\bigr), \tag{8.2.1}$$

for $\epsilon \to 0$. A limiting profile is a subset Ω of the plane $\mathbb{R}^2$ which arises as limit of a sequence of solutions

$$x^k = \{\bigl(t, x^k(t)\bigr) \in \mathbb{R}^2\colon\ t \in \mathbb{R}\}, \quad k \in \mathbb{N},$$

to Eq. (8.2.1) with $\epsilon = \epsilon_k$, in case $\lim_{k\to\infty} \epsilon_k = 0$. Convergence of a sequence of closed subsets of the plane means that intersections with given compact subsets converge in the Hausdorff metric.

The standing hypotheses in [155] are the following. f is a Lipschitz continuous real-valued map defined on a square $I \times I$, $I = [-D, C]$ with $C > 0$ and $D > 0$, and $r : I \to [0, \infty)$ is Lipschitz continuous with

$$r(0) = 1 \quad \text{and} \quad r(\xi) > 0 \quad \text{for } -D < \xi < C.$$

The zero set $f^{-1}(0) \subset I \times I$ is a strictly decreasing continuous function $g : I \to I$ with $g(0) = 0$, and f is positive below its zero set and negative above it. Moreover,

$$\left|g^2(\xi)\right| < |\xi| \quad \text{on } (-D, C) \setminus \{0\}.$$

In case $r(C) > 0$ it is assumed that $g(C) = -D$ while in case $r(-D) > 0$, $g(-D) = C$. Finally, f is differentiable at $(0, 0)$ with $\mathrm{D}_2 f(0, 0)1 < D_1 f(0, 0)1$.

Notice that the distribution of the signs of f in $I \times I$ generalizes the negative feedback inequality for functions of a single variable.

The periodic solutions considered are slowly oscillating in the sense that their zeros are spaced at distances larger than $r(0) = 1$, which is the delay at equilibrium. Their minimal periods are given by 3 consecutive zeros, and they are sine-like in the sense that the period interval between 3 successive zeros consists of 3 adjacent subintervals on each of which the periodic solution is monotone. Existence of sine-like slowly oscillating periodic solutions for sufficiently small $\epsilon > 0$ is proved in [152].

Limiting profiles exist, due to a result from [152] that for every sequence of parameters $\epsilon_k > 0$ with $\lim_{k\to\infty} \epsilon_k = 0$ and for every sequence of sine-like slowly oscillating periodic solutions $x^k : \mathbb{R} \to I$ of Eq. (8.2.1) with $\epsilon = \epsilon_k$ there is a subsequence (k_j) for which the graphs x^{k_j} converge.

It may happen that the limiting profile is simply the abscissa $\mathbb{R} \times \{0\}$. Theorem 5.1 in [152] provides sufficient conditions on r which exclude this case, like for example $r'(0) \neq 0$.

The first step towards the description of nontrivial limiting profiles is an appropriate interpretation of the formal limit of Eq. (8.2.1) for $\epsilon \to 0$, which reads

$$0 = f\big(x(t), x\big(t - r\big(x(t)\big)\big)\big) \tag{8.2.2}$$

and can be considered as a difference equation in implicit form for functions on the real line. But this is too narrow, as limiting profiles may contain vertical line segments. Notice that for any point $(t, x(t))$ on a solution $x : \mathbb{R} \to I$ of Eq. (8.2.2) there is another point $(s, x(s))$ on x with

$$s = t - r\big(x(t)\big) \leqslant t \quad \text{and} \quad x(s) = g\big(x(t)\big)$$

since Eq. (8.2.2) is solved for the second argument by g. This suggests to consider the *backdating map* $\Phi : \mathbb{R} \times I \to \mathbb{R} \times I$ given by

$$\Phi(\tau, \xi) = \big(\tau - r(\xi), g(\xi)\big)$$

and its trajectories (τ_n, ξ_n), which satisfy the system

$$\begin{aligned} 0 &= f(\xi_n, \xi_{n-1}), \\ \tau_{n-1} &= \tau_n - r(\xi_n). \end{aligned}$$

Properties of the backdating map are the key to the description of limiting profiles.

Theorem 1.3 in [155] establishes that the minimal periods $p^k > 2$ of sine-like slowly oscillating periodic solutions $x^k : \mathbb{R} \to I$, $k \in \mathbb{N}$, of Eq. (8.2.1) with $\epsilon = \epsilon_k$, $\lim_{k\to\infty} \epsilon_k = 0$, are bounded.

Each nontrivial limiting profile $\Omega \subset \mathbb{R} \times I$ is periodic, i.e.,

$$\Omega = \Omega + (p, 0)$$

with the (existing) limit $p \geqslant 2$ of the minimal periods of the approximating sequence of periodic solutions, and the intersection of Ω with a suitable vertical strip of width p can be written as the union of a horizontal line segment below the abscissa, an ascending part, a horizontal line segment above the abscissa, and a descending part. The horizontal parts may be singletons, and the ascending and descending parts may contain both horizontal and vertical line segments.

The first main result, Theorem A in [155], describes a nontrivial limiting profile Ω in the following way, using the (existing) limits $\mu > 0$ of the maxima and $-\nu < 0$ of the minima of the approximating periodic solutions: There is a sequence of continuous functions $\psi_n : [-\nu, \mu] \setminus \{0\} \to \mathbb{R}$, $n \in \mathbb{Z}$, with right and left limits at 0, so that for each integer n the function $(-1)^n \psi_n$ is increasing, $\psi_n \leqslant \psi_{n+1}$, $\psi_{n+2} = \psi_n + p$, and

$$\Omega = \left(\bigcup_n \psi_n^* \cup \big(A_n \times \{0\}\big) \right) \cup \left(\bigcup_n B_n \times \{\lambda_n\} \right)$$

where

$$\psi_n^* = \big\{ \big(\psi_n(\xi), \xi\big) \colon\ 0 \neq \xi \in [-\nu, \mu] \big\},$$

$$A_n = \begin{cases} \big[\psi_n(0-), \psi_n(0+)\big] & \text{if } n \text{ is even}, \\ \big[\psi_n(0+), \psi_n(0-)\big] & \text{if } n \text{ is odd}, \end{cases}$$

and

$$B_n = \big[\psi_n(\lambda_n), \psi_{n+1}(\lambda_n)\big]$$

with

$$\lambda_n = \begin{cases} \mu & \text{if } n \text{ is even}, \\ -\nu & \text{if } n \text{ is odd}. \end{cases}$$

Moreover, there exist $\delta_0 > 0$ and $\delta_1 > 0$ so that for $0 \neq \xi \in [-\nu, \delta_0]$,

$$\psi_{2m}(\xi) = \max_{-\nu \leqslant s \leqslant \xi} \big(r(s) + \psi_{2m-1}\big(g(s)\big)\big) \tag{8.2.3}$$

while for $0 \neq \xi \in [-\delta_1, \mu]$,

$$\psi_{2m+1}(\xi) = \max_{\xi \leqslant s \leqslant \mu} \big(r(s) + \psi_{2m}\big(g(s)\big)\big). \tag{8.2.4}$$

The nonlocal *max-plus operators* given by the right-hand sides of Eqs. (8.2.3)–(8.2.4) bear analogies with linear Fredholm integral operators. To see this, replace addition in (8.2.3)–(8.2.4) by multiplication and maximization by integration. In [153,154] Mallet-Paret and Nussbaum study max-plus operators and associated eigenvalue problems; the theory is applied in [155].

Theorem B in [155] deals with monotone delay functions r and provides more detailed information about limiting profiles, in terms of f, g, r and $h = r + r \circ g$. Here the functions ψ_n for n odd are solutions to an eigenvalue problem

$$p + \psi_n(\xi) = \max_{\xi \leqslant s \leqslant \mu} \big(h(s) + \psi_n\big(g^2(s)\big)\big)$$

for a max-plus operator, with the period p as an *additive eigenvalue*. The functions ψ_n for n even are computed from ψ_n for n odd.

Theorem C in [155] establishes uniqueness of limiting profiles, under further conditions on the data f and r.

A simple-looking example for which there is a unique limiting profile is the equation

$$\epsilon x'(t) = -x(t) - kx\big(t - 1 - cx(t)\big)$$

with $k > 1$ and $c > 0$. In this case, $p = 1 + k$ and

$$\Omega \cap \big((-1, k) \times \mathbb{R}\big) = \left\{\left(\tau, \frac{1}{c}\tau\right) \colon -1 < \tau < k\right\},$$

$$\Omega \cap \big(\{k\} \times \mathbb{R}\big) = \left\{(k, \xi) \colon -\frac{1}{c} \leqslant \xi \leqslant \frac{k}{c}\right\}.$$

So Ω is uniquely determined and looks like sawteeth.

8.3. *Small delay*

In [178] Ouifki and Hbid obtain existence of periodic solutions for a system of the form

$$x'(t) = g\big(x\big(t - r(x_t)\big)\big) \tag{8.3.1}$$

with a map $g:\mathbb{R}^2 \to \mathbb{R}^2$ which satisfies $g(0)=0$, is smooth of class C^4, and has Jacobian $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ at $0 \in \mathbb{R}^2$. Also the delay functional $r:C \to [0,h] \subset \mathbb{R}$, $C = C([-h,0];\mathbb{R}^2)$, is assumed to be smooth of class C^4, and several smallness conditions are imposed. The approach in [178] is based on a decomposition of the functional $f:C \to \mathbb{R}^2$ corresponding to the right hand side of Eq. (8.3.1) into a smooth map $f_r : C \to \mathbb{R}^2$ and a remainder term, both depending on r. The decomposition holds for arguments in a closed subset E_1 of the space $C^{0,1} = C^{0,1}([-h,0];\mathbb{R}^2)$. The set E_1 is contained in the analogue $E \subset C^1 = C^1([-h,0];\mathbb{R}^2) \subset C^{0,1}$ of the solution manifold X_f from Section 3; in [144] it was shown that a nonlinear semigroup on $C^{0,1}$ generated by equations like (8.3.1) becomes strongly continuous if restricted to E. Smallness assumptions on r yield that the set E_1 is positively invariant. Each truncated equation

$$y'(t) = f_r(y_t) \tag{8.3.2}$$

defines a semiflow on the space C, as f_r is sufficiently smooth. The delay $\mu = r(0)$ at equilibrium is then considered as a parameter; Eq. (8.3.2) is rewritten as

$$y'(t) = f_{\tilde{r}}(\mu, z_t), \tag{8.3.3}$$

with $\tilde{r}(\phi) = r(\phi) - r(0)$. Under assumptions which guarantee certain stability properties of the stationary point 0 of Eq. (8.3.3) with $\mu = 0$ and $\tilde{r}$ small, a combination of center manifold theory with a Hopf bifurcation theorem yields attracting periodic orbits $o(\mu,\tilde{r})$ of the truncated equation (8.3.2) for small $\mu > 0$ and small $\tilde{r}$. Upon that a return map is constructed following solutions of the original equation (8.3.1) which start from initial data in E_1 close to a chosen point on $o(\mu,\tilde{r})$. This requires further smallness properties of $\tilde{r}$; closeness refers to the topology of $C^{0,1}$. With respect to this topology the return map is continuous and compact, Schauder's theorem is applied, and resulting fixed points define periodic solutions of Eq. (8.3.1).

8.4. *Generic convergence*

In [21] Bartha considers the scalar equation

$$x'(t) = -\mu x(t) + f\big(x\big(t - r\big(x(t)\big)\big)\big) \tag{8.4.1}$$

assuming that $\mu > 0$, $f \in C^1(\mathbb{R},\mathbb{R})$, $f(0)=0$, $f'(u) > 0$ for all $u \in \mathbb{R}$, $r \in C^1(\mathbb{R},\mathbb{R})$, and there is $A > 0$ with

$$\big|f(u)\big| < \mu|u| \quad \text{for all } u \in \mathbb{R} \setminus (-A, A).$$

In addition, it is also required that

$$r(u) > 0 \quad \text{for all } u \in [-A, A].$$

Setting $R = \max_{u\in[-A,A]} |r(u)|$, the metric space X is defined as the space of Lipschitz continuous functions $\phi : [-R, 0] \to [-A, A]$ equipped with the metric

$$d(\phi, \psi) = \max_{s\in[-R,0]} \bigl|\phi(s) - \psi(s)\bigr|.$$

Then, for every $\phi \in X$, Eq. (8.4.1) has a unique solution $x^\phi : [-R, \infty) \to [-A, A]$ with $x^\phi|_{[-R,0]} = \phi$, and the relations

$$F(t, \phi) = x_t^\phi, \quad t \geqslant 0, \quad x_t^\phi(s) = x^\phi(t + s), \quad -R \leqslant s \leqslant 0,$$

define a continuous semiflow on X.

Using the standard ordering

$$\phi \leqslant \psi \quad \text{iff} \quad \phi(s) \leqslant \psi(s), \quad -R \leqslant s \leqslant 0,$$

it is relatively straightforward to show that the semiflow F is monotone, i.e., $F(t, \phi) \leqslant F(t, \psi)$ whenever $\phi \in X$, $\psi \in X$, $\phi \leqslant \psi$ and $t \geqslant 0$. However, the strongly order preserving property (SOP), which is a crucial hypothesis in the generic convergence theorem of Smith and Thieme in [191], does not hold in general for F. Recall that SOP of F means the monotonicity of F, and that in case $\phi \in X$, $\psi \in X$, $\phi \leqslant \psi$, $\phi \neq \psi$ there exist $t_0 > 0$ and open subsets U, V of X with $\phi \in U$ and $\psi \in V$ such that $F(t_0, U) \leqslant F(t_0, V)$. Here, for subsets S, T of X we write $S \leqslant T$ if $\phi \leqslant \psi$ holds for all $\phi \in S$ and $\psi \in T$. The main reason of the failure of the SOP property for F is that for different elements ϕ, ψ of X with $\phi \leqslant \psi$, $x^\phi(t) = x^\psi(t)$ may happen for all $t \geqslant 0$.

In [21] the SOP property is replaced by the weaker mildly order preserving property (MOP). Introduce

$$\phi <_F \psi$$

for elements ϕ, ψ of X if $\phi \leqslant \psi$, $\phi \neq \psi$, and $F(t, \phi) \neq F(t, \psi)$ for all $t \geqslant 0$. Then F is said to be MOP if it is monotone, and for every ϕ, ψ in X with $\phi <_F \psi$, there exist $t_0 > 0$ and open subsets U, V of X with $\phi \in U$ and $\psi \in V$ such that $F(t_0, U) \leqslant F(t_0, V)$.

[21] proves that F has the MOP property. An important step toward this result is that for two globally defined solutions $x : \mathbb{R} \to [-A, A]$ and $y : \mathbb{R} \to [-A, A]$ with $x_0 = y_0$, it is true that

$$x(t) = y(t) \quad \text{for all } t \in \mathbb{R}.$$

The abstract generic convergence result of Smith and Thieme from [191] is modified in [21] so that SOP is replaced by MOP. Then the main result of [21] is that there is an open dense subset Y of X such that

$$\lim_{t\to\infty} x^\phi(t) \text{ exists}$$

for all $\phi \in Y$.

8.5. *Stability and oscillation*

Several results in the literature which deal with limiting behaviour of solutions to nonautonomous differential equations with non-constant delay are also valid for equations with state-dependent delays, see, e.g., [128,222]. Most of these papers concentrate on the behaviour of given solutions, not on questions of existence, uniqueness, and continuous dependence.

Here we list a few papers where the presence of the state-dependent delay is emphasized since it causes new technical difficulties.

Kuang [135] considers the scalar nonautonomous state-dependent delay differential equation

$$x'(t) = -g\big(t, x(t)\big) - \mathrm{e}^{-\eta\tau(x_t)} f\big(t, x\big(t - \tau(x_t)\big)\big).$$

Sharp conditions for the boundedness of solutions, global and uniform stability of the trivial solution are presented.

Cooke and Huang [47] study the scalar equation

$$x'(t) = x(t)\left(a - bx(t) - \sum_{i=1}^{L} b_i x(t - r_i) - cx\big(t - \tau(x_t)\big)\right),$$

where a, b_i, c are positive constants, τ is a functional of the history of $x(\cdot)$ over all times before t. They obtain results on convergence of positive solutions, periodic and oscillatory behaviour which extend work of G. Seifert for the constant delay case.

Győri and Hartung [93] consider the linear delay differential systems

$$x'(t) = A_i(t)x\big(t - \sigma_i(t)\big) \tag{8.5.1}$$

with continuous coefficient functions $A_i : [0, \infty) \to \mathbb{R}^{n\times n}$ and continuous delay functions $\sigma_i : [0, \infty) \to \mathbb{R}$, $i \in \{1, 2\}$, such that, for some $r > 0$, $0 \leqslant \sigma_i(t) \leqslant t + r$, $t \geqslant 0$, and $\liminf_{t\to\infty}[t - \sigma_i(t)] > 0$. Assuming that the zero solution of Eq. (8.5.1) with $i = 1$ is exponentially stable, explicit neighbourhoods of A_1 and σ_1 are constructed so that if A_2 and σ_2 belong to the corresponding neighbourhoods of A_1 and σ_1, then the zero solution of Eq. (8.5.1) with $i = 2$ is also exponentially stable. As an application, among others, sufficient conditions are given to guarantee the exponential stability of the zero solution of the scalar equation

$$x'(t) = a(t)x\big(t - \tau(t, x_t)\big)$$

with delay functional τ defined by the threshold relation

$$\int_{t-\tau(t,x_t)}^{t} f(t, s, x_t)\,\mathrm{d}s = m$$

provided that the zero solution of

$$x'(t) = a(t)x\big(t - \tau(t, 0)\big)$$

is exponentially stable.

Cao, Fan and Gard [37] study the two-stage population model of Aiello, Freedman and Wu [2] with density-dependent delay. They show that no Hopf bifurcation can occur in the sense that the characteristic equation, associated with linearization at any strictly positive equilibrium, never has imaginary roots. Instability can arise together with the creation of multiple equilibria. The attractivity regions of the equilibrium points are also estimated.

Bélair [24] considers an age-structured model, and reduces it to a system of delay differential equations with state-dependent delay. Assuming that a center manifold reduction is valid, a supercritical Hopf bifurcation is established.

Rai and Robertson [182,183] study stage-structured population models with delay where the delay is a function of the total population density, and they prove positivity, boundedness and stability of the solutions.

Bartha [20] addresses the stability and convergence of solutions for a class of neutral functional differential equations with state-dependent delay. The equation is transformed into a retarded differential equation with infinite delay. The state-dependent delay causes that the transformation depends on each particular solution. For the retarded equation with infinite delay a result of Krisztin [128] can be applied to get sharp sufficient conditions for the stability of the zero solution. The second part of [20] contains attractivity results.

Pinto [180] gives conditions assuring asymptotic expansions of the form

$$y(t) = \exp\left(\int_{t_0}^{t} a(s)\,\mathrm{d}s\right)\left[\xi + \mathrm{O}\left(\int_{t}^{\infty} \lambda(s)\,\mathrm{d}s\right)\right] \quad \text{as } t \to \infty$$

for the solutions of the scalar state-dependent differential equation

$$y'(t) = a(t)y\big(t - r\big(t, y(t)\big)\big)$$

with certain $\xi \in \mathbb{R}$, $\lambda \in L_1[0, \infty)$ constructed by means of the coefficient function a and the delay function r. These results are extended to systems in [83].

Asymptotic solution behaviour for various classes of autonomous equations with state-dependent delays is investigated also in [45,46,53,104].

Gatica and Rivero [86] obtain sufficient conditions for the oscillation of all nontrivial solutions of a scalar equation with a state-dependent delay given by a threshold condition.

Domoshnitsky, Drakhlin and Litsyn [56] study the equation

$$x'(t) + \sum_{i=1}^{m} A_i(t)x\big(t - (H_i x)(t)\big) = f(t)$$

in $\mathbb{R}^n$ with measurable and essentially bounded functions A_i and f, and measurable delay functional H_i. Sufficient conditions are obtained for the boundedness, oscillation and nonoscillation of the solutions by using an associated linear equation.

Additional stability and oscillation results can be found in [36,46,84,92,172,173,194, 195,200,217,220,223].

9. Numerical methods

9.1. *Preliminaries*

The study of numerical approximation for state-dependent delay equations goes back at least to the mid sixties of the last century [28,77], and since then it is an intensively investigated research area [12–14,17,18,27,31,35,49,68–70,78,79,90,125,140,145,168,169,196, 197,211].

In this section we concentrate on numerical methods for state-dependent delay equations (SD-DDEs) of the form

$$x'(t) = f\big(t, x(t), x\big(t - \tau\big(t, x(t)\big)\big)\big), \quad t \in [t_0, t_N] \tag{9.1.1}$$

with an associated initial condition

$$x(t) = \phi(t), \quad t \in [t_0 - h, t_0]. \tag{9.1.2}$$

For simplicity we assume that $f : [t_0, t_N] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous and Lipschitz continuous in its second and third variables, $\tau : [t_0, t_N] \times \mathbb{R} \to [0, h]$ is continuous, and $\phi \in C^{0,1}([t_0, t_N]; \mathbb{R})$, therefore the IVP (9.1.1)–(9.1.2) has a unique solution. The results we present can usually be easily extended to the system or multiple delays case.

We mention that there are a large number of publications dealing with other types of state-dependent differential equations including neutral SD-DDEs of the form

$$x'(t) = f\big(t, x(t), x\big(t - \tau\big(t, x(t)\big)\big), x'\big(t - \sigma\big(t, x(t)\big)\big)\big)$$

(see, e.g., [16,27,34,42,69,119,122,169,179]), the so-called "implicit" neutral SD-DDEs of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(x(t) - g\big(t, x\big(t - \sigma\big(t, x(t)\big)\big)\big)\big) = f\big(t, x(t), x\big(t - \tau\big(t, x(t)\big)\big)\big)$$

[102,140], Volterra differential equations with state-dependent delays [13,17,31,35,120, 196], and differential-algebraic equations with state-dependent delays [109].

9.2. *Continuous Runge–Kutta methods for ODEs*

Before we discuss approximation of state-dependent equations, first we recall some basic notations and definitions for numerical approximation of ODEs. Consider the IVP

$$x'(t) = g\big(t, x(t)\big), \quad t \in [t_0, t_N], \tag{9.2.1}$$

$$x(t_0) = x_0, \tag{9.2.2}$$

and suppose mesh points $\Delta = \{t_0 < t_1 < \cdots < t_N\}$ are given. *Discrete* one- or multistep methods associated to the IVP (9.2.1)–(9.2.2) produce a sequence $y_0, \ldots, y_N$ to approximate the solution x at the mesh points Δ. In this paper for simplicity we restrict our discussion to a popular class of one-step methods, the Runge–Kutta (RK) methods. A discrete RK method is defined by $y_0 = x_0$, and

$$y_{n+1} = y_n + h_n \sum_{i=1}^{s} b_i k_{n,i}, \quad n = 0, 1, \ldots, N-1, \tag{9.2.3}$$

where the stage values $k_{n,1}, \ldots, k_{n,s}$ are determined by the system of algebraic equations

$$k_{n,i} = g\left(t_n + h_n c_i, y_n + h_n \sum_{j=1}^{s} a_{ij} k_{n,j}\right), \quad i = 1, 2, \ldots, s, \tag{9.2.4}$$

and $h_n = t_{n+1} - t_n$. s is called the number of *stages*, the b_i's are the *weights*, the c_i's are the *abscissae* of the method satisfying $c_i \in [0, 1]$. The coefficients a_{ij} are collected in a matrix $\mathbf{A}$, the weights and abscissae in the vectors $\mathbf{b}$ and $\mathbf{c}$, and the parameters are usually listed in the Butcher tableau $\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{b}^{\mathrm{T}} \end{array}$. If the matrix $\mathbf{A}$ is lower triangular with zero diagonal entries then the RK method is explicit, otherwise (9.2.3) and (9.2.4) implicitly define the sequence $y_0, \ldots, y_N$.

In the mid eighties of the last century and at the beginning of the nineties the interest in the study and application of *continuous extensions* of numerical methods has been increased, since if a *dense output* is required by an ODE solver, e.g., when plotting the graph of the numerical solution, then a discrete solver is not efficient enough. Continuous methods are especially important for the approximation of delay equations where the evaluation of the approximate solution is needed in between mesh points.

One possible way to derive a continuous extension of the RK method (called CRK method) (9.2.3)–(9.2.4) is the following: Define

$$u(t_n + \theta h_n) = y_n + h_n \sum_{i=1}^{s} w_i(\theta) k_{n,i}, \quad \theta \in [0, 1],\ n = 0, 1, \ldots, N-1, \tag{9.2.5}$$

where w_i are polynomials of degree less than or equal to δ satisfying

$$w_i(0) = 0, \quad w_i(1) = b_i, \quad i = 1, \ldots, s.$$

This formula is called *interpolant of the first class* of the discrete RK method, δ is the *degree of interpolant*. The Butcher tableau of a CRK method has the form $\begin{array}{c|c} \mathbf{c} & \mathbf{A} \\ \hline & \mathbf{w}^{\mathrm{T}}(\theta) \end{array}$.

There are several methods to define the polynomial interpolation w_i in (9.2.5). A typical way is to use a cubic Hermite interpolation, or when a higher order interpolant is required, a so-called fully Hermite interpolation [71,110]. Another approach is to consider (9.2.4)–(9.2.5) as a discrete RK method with coefficients a_{ij}/θ, weights $w_i(\theta)/\theta$ and step size θh_n, and apply order conditions known for the discrete RK method, see, e.g., [27]. We mention that there are more general extensions of discrete RK methods than that of the form (9.2.5) (see, e.g., [27,110]), but we do not go into details here.

Let z be the solution of the local problem

$$z'(t) = g\bigl(t, z(t)\bigr), \quad t \in [t_n, t_{n+1}], \tag{9.2.6}$$

$$z(t_n) = z^*. \tag{9.2.7}$$

We say that the discrete RK method (9.2.3)–(9.2.4) has *order p* if $p \geqslant 1$ is the largest integer such that for all C^p-functions g in (9.2.1) and for all meshes Δ we have

$$\bigl|z(t_n) - y_n\bigr| = \mathrm{O}\bigl(h_n^{p+1}\bigr)$$

uniformly with respect to z^* in any bounded subset of $\mathbb{R}$ and $n = 0, \ldots, N-1$, where z is the solution of (9.2.6)–(9.2.7). Similarly, the CRK method (9.2.4)–(9.2.5) has *uniform order q* if $q \geqslant 1$ is the largest integer such that for all C^q-functions g in (9.2.1) and for all meshes Δ we have

$$\max\bigl\{\bigl|z(t) - u(t)\bigr|\colon t_n \leqslant t \leqslant t_{n+1}\bigr\} = \mathrm{O}\bigl(h_n^{q+1}\bigr)$$

uniformly with respect to z^* in any bounded subset of $\mathbb{R}$ and $n = 0, \ldots, N-1$, where z is the solution of (9.2.6)–(9.2.7).

We recall the following result from [27]:

THEOREM 9.2.1. *If the discrete RK method* (9.2.3)–(9.2.4) *has order p and if g is a C^p-function*, *then the method is convergent of global order p on any bounded interval* $[t_0, t_N]$, *i.e.*,

$$\max\bigl\{\bigl|x(t_n) - y_n\bigr|\colon n = 1, \ldots, N\bigr\} = \mathrm{O}\bigl(h^p\bigr),$$

where $h = \max\{h_0, \ldots, h_{N-1}\}$.

Moreover, *if the CRK method* (9.2.4)–(9.2.5) *has uniform order q*, *then it has uniform global order* $q' = \min\{p, q+1\}$, *i.e.*,

$$\max\bigl\{\bigl|x(t) - u(t)\bigr|\colon t_0 \leqslant t \leqslant t_N\bigr\} = \mathrm{O}\bigl(h^{q'}\bigr).$$

It can be checked (see [27]) that, in order to get the uniform order q, the interpolant must be of degree $\delta \geqslant q$, and if a discrete RK method has a continuous extension $u(t)$ of uniform order q with degree $\delta > q$, then it has a continuous extension $\tilde{u}(t)$ of uniform order q with degree $\delta = q$, as well.

As an example we present the Butcher tableau of a continuous extension of the classical four-stage discrete RK method which has uniform order 3:

$$\begin{array}{c|cccc} 0 & 0 & & & \\ \frac{1}{2} & \frac{1}{2} & 0 & & \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \\ 1 & 0 & 0 & 1 & 0 \\ \hline & w_1(\theta) & w_2(\theta) & w_3(\theta) & w_4(\theta) \end{array} \quad \text{where} \quad \begin{aligned} w_1(\theta) &= \tfrac{2}{3}\theta^3 - \tfrac{3}{2}\theta^2 + \theta, \\ w_2(\theta) &= -\tfrac{2}{3}\theta^3 + \theta^2, \\ w_3(\theta) &= -\tfrac{2}{3}\theta^3 + \theta^2, \\ w_4(\theta) &= \tfrac{2}{3}\theta^3 - \tfrac{1}{2}\theta^2. \end{aligned}$$

One-step collocation methods can be considered as CRK methods: Pick distinct abscissae $c_1, \dots, c_s \in [0, 1]$, and define

$$\begin{aligned} \ell_i(v) &= \prod_{k=1,\ k\neq i}^{s} \frac{v - c_k}{c_i - c_k}, && i = 1, \dots, s, \\ a_{ij} &= \int_0^{c_i} \ell_j(v)\,\mathrm{d}v, && i, j = 1, \dots, s, \\ w_i(\theta) &= \int_0^{\theta} \ell_i(v)\,\mathrm{d}v, && i = 1, \dots, s. \end{aligned}$$

Then the corresponding CRK method (9.2.4)–(9.2.5) defines a polynomial u of degree $\leqslant s$ satisfying the collocation equations

$$u'(t_n + h_n c_i) = g\big(t_n + h_n c_i, u(t_n + h_n c_i)\big), \quad i = 1, \dots, s, \quad u(t_n) = y_n.$$

It is known (see, e.g., [27]) that the uniform global order of this CRK method is $q' = s$ or $s + 1$.

Nowadays an efficient differential equation solver uses a higher order continuous method, or usually a pair of higher order methods (to estimate local errors in the step size selection). For other examples of CRK methods we refer to [27,49,70,71].

9.3. *The standard approach to approximation of SD-DDEs*

A typical approach (called “the standard approach” in [27]) to obtain a numerical approximation to the solution x of the state-dependent IVP (9.1.1)–(9.1.2) is the numerical analogue of the method of steps well-known for computing exact solutions of constant delay equations. In this subsection we describe this approach, but for simplicity, we formulate it using the class of one-step CRK methods. Clearly, it can be adopted using many other types of ODE discretization techniques.

Pick mesh points $\Delta = \{t_0 < t_1 < \cdots < t_N\}$, let $y_0 = \phi(t_0)$, and consider the sequence of local IVPs for $n = 0, 1, \dots, N - 1$:

$$z_n'(t) = f\big(t, z_n(t), y\big(t - \tau\big(t, z_n(t)\big)\big)\big), \quad t \in [t_n, t_{n+1}], \tag{9.3.1}$$

$$z_n(t_n) = y_n, \tag{9.3.2}$$

where

$$y(s) = \begin{cases} \phi(s), & s \in [t_0 - h, t_0], \\ u(s), & s \in [t_0, t_n], \\ z_n(s), & s \in [t_n, t_{n+1}]. \end{cases}$$

Here u denotes the function $u : [t_0, t_n] \to \mathbb{R}$ whose restriction to $[t_i, t_{i+1}]$ $(i = 0, \dots, n-1)$ is the numerical solution, i.e., a CRK interpolant of the solution of the ith IVP. Then u is already defined in the previous steps. Now we solve this IVP on $[t_n, t_{n+1}]$ using the CRK method of the form (9.2.5) corresponding to $g(t, x) = f(t, x, u(t - \tau(t, x)))$, i.e., where $k_{n,i}$'s are defined by

$$\begin{cases} t_{n,i} = t_n + h_n c_i, \\ y_{n,i} = y_n + h_n \sum_{j=1}^{s} a_{ij} k_{n,j}, \\ k_{n,i} = f\big(t_{n,i}, y_{n,i}, y\big(t_{n,i} - \tau(t_{n,i}, y_{n,i})\big)\big). \end{cases} \tag{9.3.3}$$

Then we extend u from $[t_0, t_n]$ to $[t_0, t_{n+1}]$ by this CRK interpolant, define $y_{n+1} = u(t_{n+1})$, and continue with the next local IVP in the same manner.

If τ is bounded below by a positive constant $\bar{\tau}$, and the step size is chosen so that $h_n \leqslant \bar{\tau}$, then (9.3.1) is an ODE, since y in (9.3.1) never takes an argument from $[t_n, t_{n+1}]$, so it is explicitly defined. Therefore if the discrete RK method is explicit, i.e., $\mathbf{A}$ is lower triangular with zeros in the diagonal, then its continuous extension (9.2.5)–(9.3.3) is also explicit.

The numerical difficulty arises in the *vanishing delay case*, i.e., when τ can be arbitrary small. Then h_n can be such that $t - \tau(t, z(t)) > t_n$ for some $t \in [t_n, t_{n+1}]$, so the evaluation of y in (9.3.3) depends also on the unknown interpolant u on $[t_n, t_{n+1}]$. (This is called *overlapping*.) In this case the method is implicit, even if the original discrete RK method was explicit. Therefore the existence of the numerical approximation, i.e., the solvability of the algebraic equations (9.2.4)–(9.3.3) for the stage values $k_{n,1}, \dots, k_{n,s}$ is not trivial. It can be shown [27] that the above problem has a positive answer: Suppose the functions f, τ and ϕ are Lipschitz continuous, then in the overlapping case there always exist a sufficiently small h_n and a suitable degree of interpolants so that the implicit relations (9.2.4)–(9.3.3) have a unique solution for the stage values $k_{n,1}, \dots, k_{n,s}$; therefore u has a unique extension from $[t_0, t_n]$ to $[t_n, t_{n+1}]$. The proof of this result shows that this extension can be determined as a limit of a fixed point iteration, and therefore, it is common to estimate it by iteration using a predictor–corrector method. Next we show a possible algorithm to compute the nth step of the approximation in the case when the original discrete RK method is explicit, i.e., $a_{ij} = 0$ for $i \leqslant j$ and $c_1 = 0$. In this algorithm we first (Step 1) predict a starting value of the stage values using the last computed approximate solution value in the overlapping case instead of interpolating values. If overlapping occurs, then in an iteration (Step 2) we correct the stage values. It can be done using a fixed number of steps (m in the algorithm), or by testing the numerical convergence in a loop. Finally (Step 3) we update the interpolant.

Step 1: Prediction—computation of initial stage values

```
$t_{n,1} = t_n$
$z_{n,1}^{(0)} = y_n$
$k_{n,1}^{(0)} = f\big(t_n, z_{n,1}^{(0)}, u\big(t_n - \tau\big(t_n, z_{n,1}^{(0)}\big)\big)\big)$
for $i = 2, \dots, s$ do
    $t_{n,i} = t_n + h_n c_i$
    $z_{n,i}^{(0)} = y_n + h_n \sum_{j=1}^{i-1} a_{ij} k_{n,j}^{(0)}$
    $d_{n,i}^{(0)} = t_{n,i} - \tau\big(t_{n,i}, z_{n,i}^{(0)}\big)$
    if $d_{n,i}^{(0)} \leqslant t_n$ then
        $k_{n,i}^{(0)} = f\big(t_{n,i}, z_{n,i}^{(0)}, u\big(d_{n,i}^{(0)}\big)\big)$
    else
        $k_{n,i}^{(0)} = f\big(t_{n,i}, z_{n,i}^{(0)}, z_{n,i}^{(0)}\big)$
    end if
end for
```

Step 2: Correction by iteration is needed if $d_{n,i}^{(0)} > t_n$ was for any $i \geqslant 2$ in Step 1

```
for $r = 1, \dots, m$ do
    for $i = 1, \dots, s$ do
        $z_{n,i}^{(r)} = y_n + h_n \sum_{j=1}^{i-1} a_{ij} k_{n,j}^{(r-1)}$
        $d_{n,i}^{(r)} = t_{n,i} - \tau\big(t_{n,i}, z_{n,i}^{(r-1)}\big)$
        if $d_{n,i}^{(r)} \leqslant t_n$ then
            $k_{n,i}^{(r)} = f\big(t_{n,i}, z_{n,i}^{(r)}, u\big(d_{n,i}^{(r)}\big)\big)$
        else
            $\theta_i = \dfrac{d_{n,i}^{(r)} - t_n}{h_n}$
            $\hat{u}_i = y_n + h_n \sum_{j=1}^{i-1} w_j(\theta_i) k_{n,j}^{(r)} + h_n \sum_{j=i}^{s} w_j(\theta_i) k_{n,j}^{(r-1)}$
            $k_{n,i}^{(r)} = f\big(t_{n,i}, z_{n,i}^{(r)}, \hat{u}_i\big)$
        end if
    end for
end for
```

Step 3: Computation of the extension of u to $[t_n, t_{n+1}]$

$$u(t_n + \theta h_n) = y_n + h_n \sum_{i=1}^{s} w_i(\theta) k_{n,i}^{(m)}, \quad \theta \in [0, 1].$$

Most modern differential equation solvers use non-uniform mesh size, therefore the selection of the step size in each integration step (the so-called *primary step size selection*) is an important practical issue in such softwares. Concerning this topic we refer to [27,110,125] for more details.

The next result says that in order a method be of order p it is necessary that the solution be at least C^p on each interval $[t_n, t_{n+1}]$. We say that the function x has *discontinuity of order* p at ξ if $x^{(p-1)}$ exists and is continuous at ξ, and $x^{(p)}$ has jump discontinuity at ξ, i.e., $x^{(p)}(\xi+) \neq x^{(p)}(\xi-)$.

For the proof of the next result see [27]:

THEOREM 9.3.1. *Suppose f, τ and ϕ are C^p functions. Moreover,*

(1) *the mesh $\Delta = \{t_0 < t_1 < \cdots < t_N\}$ includes all discontinuity points $\xi_1 < \xi_2 < \cdots < \xi_m$ of the solution of order $\leqslant p$;*

(2) *the discrete RK method* (9.2.1)–(9.2.3) *is consistent of order p and the CRK method* (9.2.4)–(9.2.5) *is consistent of uniform order q.*

Then the method (9.2.4)–(9.3.3) *for solving the IVP* (9.1.1)–(9.1.2) *is convergent of uniform global order $q' = \min\{p, q+1\}$, i.e.*

$$\max\big\{\big|x(t) - u(t)\big|\colon t_0 \leqslant t \leqslant t_N\big\} = \mathrm{O}\big(h^{q'}\big),$$

where $h = \max\{h_0, \ldots, h_{N-1}\}$.

We note that there are many papers which follow the basic method of steps described in this subsection, i.e., approximate the solution of an SD-DDE by that of an ODE, but defined by different discrete or continuous one- or multistep ODE solvers or with different definitions of the associated ODE [15,49,68,70,77,91,124,125,169,179].

Rewriting a constant delay equation as an equivalent abstract Cauchy problem and using Trotter–Kato-type approximations is another popular approach especially in control applications. This technique was extended to SD-DDEs in [118].

9.4. *Tracking of derivative discontinuities*

Theorem 9.3.1 indicates that a high order method must locate all the discontinuity points of the solutions up to order p, and add them to the mesh. For constant or time-dependent delay equations this is a relatively simple task, but in the state-dependent delay case it leads to significant difficulties. Indeed, the location of the discontinuity points can not be computed *a priori* because they depend on the solution, and, on the other hand, only approximate locations can be computed.

We assume throughout this subsection that f, τ and ϕ are all C^p-functions. One can show that in case $\phi'(t_0-) \neq f(t_0, \phi(t_0), \phi(-\tau(t_0)))$ the corresponding solution x of the IVP (9.1.1)–(9.1.2) has discontinuity of order 1 at $t = t_0$, but for $t > t_0$ the solution is $C1$-smooth. If $\xi > t_0$ is such that $\xi - \tau(\xi, x(\xi)) = t_0$, it is easy to check that $x''(\xi-) \neq x''(\xi+)$, so x has a discontinuity of order 2 at $t = \xi$. Similarly, if we set $\xi_0 = t_0$ and define a (finite or infinite) sequence by the relation $\xi_{i+1} - \tau(\xi_{i+1}, x(\xi_{i+1})) = \xi_i$ for $i = 0, 1, \ldots,$ the sequence $\xi_0, \xi_1, \ldots$ consists of discontinuity points of increasing order: ξ_i has order $i + 1$. If a solution x is such that the time lag function $t \mapsto t - \tau(t, x(t))$ is strictly monotone increasing (which is satisfied in many applications), the above sequence will contain all discontinuity points of the solution.

On the other hand, if the above time lag function is not strictly monotone increasing, then the equation $\xi - \tau(\xi, x(\xi)) = \xi_i$ may have many solutions ξ. It can be checked (see [27,78,79,168]) that at a solution ξ of the above equation the function x has a discontinuity of order $\geqslant 1$, if and only if the graph of $t \mapsto t - \tau(t, x(t))$ crosses the level ξ_i, i.e., the root ξ has odd multiplicity. Therefore in this case the discontinuity points can be naturally stored in a tree: $\xi_{0,1} = t_0$ is the root of the tree, level 1 of the tree contains the solutions $\xi_{1,1}, \ldots, \xi_{1,\ell_1}$ of

$$\xi - \tau\big(\xi, x(\xi)\big) = \xi_{i,j} \tag{9.4.1}$$

with odd multiplicity for $i = 0$ and $j = 1$. Then for $i = 1$ and any $j = 1, \ldots, \ell_1$ the solutions $\xi_{2,1}, \ldots, \xi_{2,\ell_2}$ of (9.4.1), if exist, are placed in the second level of the tree, as the descendents of the respective $\xi_{1,j}$, etc.

It was shown in [168] that there are only finitely many computationally important points (i.e., of order less or equal to p) in the discontinuity propagation tree, so they can be ordered and relabeled as an increasing sequence $t_0 = \xi_0 < \xi_1 < \xi_2 < \cdots < \xi_\ell \leqslant t_N$. In the sequel we use this notation.

Let ξ_j be a descendent of ξ_i in the discontinuity tree, i.e.,

$$\xi_j - \tau\big(\xi_j, x(\xi_j)\big) = \xi_i. \tag{9.4.2}$$

It is easy to see that as discontinuity propagates from ξ_i to ξ_j, smoothing occurs. More precisely, we recall the next result from [168].

THEOREM 9.4.1. *Suppose ξ_i and ξ_j satisfying* (9.4.2) *are discontinuity points of order k_i and k_j, respectively, and ξ_j has odd multiplicity m_j. Then $m_j k_i + 1 \leqslant k_j$.*

Since the exact discontinuity points $\xi_1, \dots, \xi_\ell$ depend on the solution, suitable approximations are required. Let u denote the continuous interpolant (in the previous subsection it was a CRK interpolant) of the numerical approximation of x using certain mesh and approximate values $y_0 = \phi(t_0), y_1, \dots, y_N$. Let $\tilde{\xi}_0 = t_0$, and define the approximate discontinuity points as solutions of

$$\xi - \tau\bigl(\xi, u(\xi)\bigr) = \tilde{\xi}_i \tag{9.4.3}$$

where u is the approximate solution of the IVP satisfying $|x(\tilde{\xi}_j) - u(\tilde{\xi}_j)| = \mathrm{O}(h^p)$, and $h = \max\{h_0, \dots, h_{N-1}\}$. The following result is cited from [78]:

THEOREM 9.4.2. *Let u be a pth order approximation of x at $\tilde{\xi}_j$, i.e., $|x(\tilde{\xi}_j) - u(\tilde{\xi}_j)| = \mathrm{O}(h^p)$, $h = \max\{h_0, \dots, h_{N-1}\}$, m_j be the multiplicity of ξ_j in* (9.4.2), *and $|\xi_i - \tilde{\xi}_i| = \mathrm{O}(h^{r_i})$. Then $|\xi_j - \tilde{\xi}_j| = \mathrm{O}(h^{r_j})$ where $r_j = \min\{p, r_i\}/m_j$.*

Feldstein and Neves [78] suggested the following *secondary step size control* to select the step size of the numerical integration method so that the approximate discontinuity points are collected to the mesh to keep the global order of the method high: Suppose at the nth step $y_0, \dots, y_n$ are defined, and the approximate discontinuity points found so far are $t_0 = \tilde{\xi}_0 < \tilde{\xi}_1 < \tilde{\xi}_2 < \cdots < \tilde{\xi}_{\ell_n} \leqslant t_N$.

Step 1: Predict the next approximate value of the integration method by $y_{n+1} = u(t_n + h_n)$ using the continuous method u of order p and the step size h_n selected by the primary step size control method of u.

Step 2: For $i = 1, \dots, \ell_n$ find the first i such that

$$\bigl(t_n - \tau(t_n, y_n) - \tilde{\xi}_i\bigr)\bigl(t_{n+1} - \tau(t_{n+1}, y_{n+1}) - \tilde{\xi}_i\bigr) < 0,$$

i.e., (9.4.3) corresponding to the right-hand side $\tilde{\xi}_i$ has a solution $\tilde{\xi}$. If such i exists then proceed with Step 3, otherwise we accept h_n and y_{n+1} and finish this algorithm, i.e, go to the next iterate of computing y_{n+2}.

Step 3: Using a root-finding method (e.g., bisection) combined with the definition of u find an approximate solution $\tilde{\xi}$ with the above property (if more solutions are found, the least one is used).

Step 4: Include $\tilde{\xi}$ to the new mesh $\hat{t}_0 = t_0, \dots, \hat{t}_n = t_n, \hat{t}_{n+1} = \tilde{\xi}, \hat{t}_{n+2} = t_{n+1}$, redefine the approximate solution values $\hat{y}_0 = y_0, \dots, \hat{y}_n = y_n, \hat{y}_{n+1} = u(\tilde{\xi})$. Restart the numerical integration from $\hat{t}_{n+1} = \tilde{\xi}$ to $\hat{t}_{n+2}$. Let $\hat{u}$ be the new interpolant on $[\hat{t}_{n+1}, t_{n+2}]$, and define $\hat{y}_{n+2} = \hat{u}(t_{n+2})$.

Step 5: Continue with the next iterate.

See [78,79] for more details. Concerning the global order of the above method combined with a pth order continuous numerical approximation method u Feldstein and Neves [78] showed:

THEOREM 9.4.3. *Suppose the order of discontinuity of ξ_j is k_j $(j = 1, \dots, \ell)$, $\tilde{\xi}_j$ is the corresponding approximate discontinuity point, i.e., the solution of* (9.4.3), $|\xi_j - \tilde{\xi}_j| =$

$\mathrm{O}(h^{r_j})$, $\hat{\xi}_j$ is a numerical approximation of $\tilde{\xi}_j$ with order $|\hat{\xi}_j - \tilde{\xi}_j| = \mathrm{O}(h^{s_j})$, and the uniform order of u is p. Then

$$\max\{|x(t) - u(t)|\colon t \in [t_0, t_N]\} = \mathrm{O}(h^s) \quad \textit{where } s = \min_{j=1,\dots,\ell}\{p, k_j r_j, k_j s_j\}.$$

Moreover, if $s_j \geqslant p/k_j$ for $j = 1, \dots, \ell$, then

$$\max\{|x(t) - u(t)|\colon t \in [t_0, t_N]\} = \mathrm{O}(h^p).$$

In this subsection we assumed that the parameters of the IVP, f, τ and ϕ are all C^p-functions. If they are only piecewise C^p-functions, i.e., there are points where any of the tree parameters has smaller smoothness, then starting from such a point we can build a discontinuity propagation tree similar to what we described in this subsection. Such points are called *secondary discontinuity points*. Similarly, multiple delays can also be handled (see, e.g., [27]).

We note that the extension of the notion of the discontinuity tree from the scalar case to the system case is far from being obvious, and was investigated by Willé and Baker [212, 213]. They associated a so called dependency network of oriented graphs to the discontinuity points, see also [27]. For more discussions on numerical problems related to tracking discontinuity points we refer to [211,214].

Tracking discontinuities can be computationally expensive, especially when the number of discontinuity points is large. Another typical approach to handle the loss of numerical accuracy due to the presence of derivative discontinuities is the method of *defect control* developed by Enright and Hayashi [67,69] (see also [16,79]). In this method the size of the defect, i.e., $\max\{|u'(t) - f(t, u(t), u(t - \tau(t, u(t))))|\colon t \in [t_n, t_{n+1}]\}$ is monitored and its size is controlled at each step of the integration.

9.5. *Concluding remarks*

There are several software packages available for solving SD-DDEs. Without completeness we list some of them: ARCHI (Paul [179]), DDE-STRIDEL (Butcher [34]), DDVERK (Enright and Hayashi [68]), DMRODE (Neves [167]), DKLAG6 (Corwin, Sarafyan and Thompson [49]), RADAR5 (Guglielmi and Hairer [90]), SNDDELM (Jackiewicz and Lo [122]), SYSDEL (Karouri and Vaillancourt [125]). We refer to [27,79,197] for further discussion and comparison of available solvers for state-dependent delay equations.

In this section we discussed some problems related to the design and analysis of numerical approximation schemes for state-dependent delay equations. Other important qualitative issues, like stability of numerical methods are not discussed here, we mention [12, 27,140] for studies in this directions. We also mention that there are many topics beyond the scope of this survey, e.g., numerical bifurcation analysis (see [66,145]), boundary value problems (see [18]) or parameter estimation (see [14,100,103,107,164]) for state-dependent delay equations.

Acknowledgements

The work of F. Hartung and T. Krisztin was supported by the Hungarian Scientific Research Fund, grant No. T 046929 (FH)/T 049516 (TK); the work of J. Wu was supported by Natural Sciences and Engineering Research Council of Canada and by Canada Research Chairs Program.

References

[1] H. Agmon-Snir and I. Segev, *Signal delay and input synchronization in passive dendritic structures*, J. Neurophysiology **70** (1993), 2066–2085.

[2] W.G. Aiello, H.I. Freedman and J. Wu, *Analysis of a model representing state-structured population growth with state-dependent time delay*, SIAM J. Appl. Math. **52** (1992), 855–869.

[3] J.F.M. Al-Omari and S.A. Gourley, *Dynamics of a stage-structured population model incorporating a state-dependent maturation delay*. Nonlinear Analysis: Real World Applications **6** (2005), 13–33.

[4] W. Alt, *Periodic solutions of some autonomous differential equations with variable time delay*, Functional Differential Equations and Approximation of Fixed Points, Bonn, 1978, H.O. Peitgen and H.O. Walther, eds., Lecture Notes in Math., Vol. 730, Springer, Berlin (1979), 16–31.

[5] U. an der Heiden, *Analysis of Neural Networks*, Springer, Heidelberg (1979).

[6] H.G. Andrewartha and L.C. Birch, *The Distribution and Abundance of Animals*, University of Chicago Press, Chicago (1954).

[7] O. Arino, K.P. Hadeler and M.L. Hbid, *Existence of periodic solutions for delay differential equations with state-dependent delay*, J. Differential Equations **144** (1998), 263–301.

[8] O. Arino, M.L. Hbid and R. Bravo de la Parra, *A mathematical model of growth of population of fish in the larval stage: Density-dependence effects*, Math. Biosci. **150** (1998), 1–20.

[9] O. Arino and E. Sanchez, *Delays induced in population dynamics*, Mathematical Modelling of Population Dynamics, Bȩdlewo, 2002, Banach Center Publications, Vol. 63, Polish Academy of Sciences, Warsaw (2004), 9–46.

[10] O. Arino and E. Sanchez, *A saddle point theorem for functional state-dependent delay equations*, Discrete Continuous Dynamical Systems **12** (2005), 687–722.

[11] O. Arino, E. Sanchez and A. Fathallah, *State-dependent delay differential equations in population dynamics: modeling and analysis*, Topics in Functional Differential and Difference Equations, Lisbon, 1999, T. Faria and P. Freitas, eds., Fields Inst. Communications, Vol. 29, Amer. Math. Soc., Providence, RI (2001), 19–36.

[12] C.T.H. Baker, *Retarded differential equations*, J. Comput. Appl. Math. **125** (2000), 309–335.

[13] C.T.H. Baker, *Numerical analysis of Volterra functional and integral equations*, The State of the Art in Numerical Analysis, I.S. Duff and G.A. Watson, eds., Clarendon Press, Oxford (1996).

[14] C.T.H. Baker, G.A. Bocharov and F.A. Rihan, *A report on the use of delay differential in numerical modelling in the biosciences*, NA Report 343, Department of Mathematics, University of Manchester (1999).

[15] C.T.H. Baker and C.A.H. Paul, *A global convergence theorem for a class of parallel continuous explicit Runge–Kutta methods and vanishing lag delay differential equations*, SIAM J. Numer. Anal. **33** (1996), 1559–1576.

[16] C.T.H. Baker, C.A.H. Paul and D.R. Willé, *Issues in the numerical solutions of evolutionary delay differential equations*, Adv. Comput. Math. (1995), 171–196.

[17] C.T.H. Baker and D.R. Willé, *On the propagation of derivative discontinuities in Volterra retarded integro-differential equations*, New Zealand J. Math. **29** (2000), 103–113.

[18] V.L. Bakke and Z. Jackiewicz, *The numerical solution of boundary-value problems for differential equations with state dependent deviating arguments*, Appl. Math. **34** (1989), 1–17.

[19] P. Baldi and A.F. Atiya, *How delays affect neural dynamics and learning*, IEEE Trans. Neural Networks **5** (1994), 612–621.

[20] M. Bartha, *On stability properties for neutral differential equations with state-dependent delay*, Differential Equations Dynamical Systems **7** (1999), 197–220.

[21] M. Bartha, *Convergence of solutions of an equation with state-dependent delay*, J. Math. Anal. Appl. **254** (2001), 410–432.

[22] M. Bartha, *Periodic solutions for differential equations with state-dependent delay and positive feedback*, Nonlinear Anal. TMA **53** (2003), 839–857.

[23] J. Bélair, *Population models with state-dependent delays*, Mathematical Population Dynamics, New Brunswick (NJ), 1991, O. Arino, D.E. Axelrod and M. Kimmel, eds., Lecture Notes in Pure and Appl. Math., Vol. 131, M. Dekker, New York (1991) 165–176.

[24] J. Bélair, *Stability analysis of an age-structured model with a state-dependent delay*, Canadian Appl. Math. Quarterly **6** (1998), 305–319.

[25] J. Bélair and M.C. Mackey, *Consumer memory and price fluctuations on commodity markets: an integrodifferential model*, J. Dynamics Differential Equations **1** (1989), 299–325.

[26] J. Bélair and J. Mahaffy, *Variable maturation velocity and parameter sensitivity in a model of haematopoiesis*, IMA J. Math. Appl. Med. Biol. **18** (2001), 193–211.

[27] A. Bellen and M. Zennaro, *Numerical Methods for Delay Differential Equations*, Oxford Science Publications, Clarendon Press, Oxford (2003).

[28] R. Bellman and K.L. Cooke, *On the computational solution of a class of functional differential equations*, J. Math. Anal. Appl. **12** (1965), 495–500.

[29] M. Brokate and F. Colonius, *Linearizing equations with state-dependent delay*, Appl. Math. Optim. **21** (1990), 45–52.

[30] F.E. Browder, *A further generalization of the Schauder fixed point theorem*, Duke Math. J. **32** (1965), 575–578.

[31] H. Brunner, *Collocation Methods for Volterra Integral and Related Functional Differential Equations*, Cambridge Monographs on Applied and Computational Mathematics, Vol. 15, Cambridge University Press, Cambridge (2004).

[32] M. Büger and M.R.W. Martin, *Stabilizing control for an unbounded state-dependent delay differential equation*, Dynamical Systems and Differential Equations, Kennesaw, GA, 2000, Discrete and Continuous Dynamical Systems (Added Volume) (2001), 56–65.

[33] M. Büger and M.R.W. Martin, *The escaping disaster: A problem related to state-dependent delays*, Z. Angew. Math. Phys. **55** (2004), 547–574.

[34] J.C. Butcher, *The adaptation of STRIDE to delay differential equations*, Appl. Numer. Math. **9** (1992), 415–425.

[35] B. Cahlon, *Numerical solutions for functional integral equations with state-dependent delay*, Appl. Numer. Math. **9** (1992), 291–305.

[36] B. Cahlon, *Oscillatory solutions of Volterra integral equations with state-dependent delay*, Dynamical Systems Appl. **2** (1993), 461–469.

[37] Y. Cao, J. Fan and T.C. Gard, *The effects of state-dependent time delay on a stage-structured population growth model*, Nonlinear Anal. TMA **19** (1992), 95–105.

[38] Y. Cao and J. Wu, *Projective ART for clustering data sets in high dimensional spaces*, Neural Networks **15** (2002), 105–120.

[39] Y. Cao and J. Wu, *Dynamics of projective adaptive resonance theory model: The foundation of PART algorithm*, IEEE Trans. Neural Networks **15** (2004), 245–260.

[40] C.E. Carr, *Processing of temporal information in the brain*, Ann. Rev. Neuroscience **16** (1993), 223–243.

[41] C.E. Carr and M. Konishi, *A circuit for detection of interaural time differences in the brain stem of the barn owl*, J. Neuroscience **10** (1990), 3227–3246.

[42] R.N. Castleton and L.J. Grimm, *A first order method for differential equations of neutral type*, Math. Comput. **27** (1973), 571–577.

[43] F. Chen and J. Shi, *Periodicity in a nonlinear predator-prey system with state dependent delays*, Acta Math. Appl. Sinica (English Ser.) **21** (2005), 49–60.

[44] S.N. Chow and K. Lu, C^k *centre unstable manifolds*, Proc. Roy. Soc. Edinburgh Ser. A **108** (1988), 303–320.

[45] K.L. Cooke, *Asymptotic theory for the delay-differential equation* $u'(t) = -au(t - r(u(t)))$, J. Math. Anal. Appl. **19** (1967), 160–173.

[46] K.L. Cooke, *Asymptotic equivalence of an ordinary and a functional differential equation*, J. Math. Anal. Appl. **51** (1975), 187–207.

[47] K.L. Cooke and W. Huang, *A theorem of George Seifert and an equation with state-dependent delay*, Delay and Differential Equations, A.M. Fink et al., eds., World Scientific, Singapore (1992), 65–77.

[48] K. Cooke and W. Huang, *On the problem of linearization for state-dependent delay differential equations*, Proc. Amer. Math. Soc. **124** (1996), 1417–1426.

[49] S.P. Corwin, D. Sarafyan and S. Thompson, *DKLAG6: a code based on continuously imbedded sixth-order Runge–Kutta methods for the solution of state-dependent functional differential equations*, Appl. Numer. Math. **24** (1997), 319–330.

[50] El Hadi Ait Dads and K. Ezzinbi, *Boundedness and almost periodicity for some state-dependent delay differential equations*, Electronic J. Differential Equations **67** (2002), 13 pp.

[51] S.P. Day and M.R. Davenport, *Continuous-time temporal back-propagation with adaptable time delays*, IEEE Trans. Neural Networks **4** (1993), 348–354.

[52] A.M. de Roos and L. Persson, *Competition in size-structured populations: mechanisms inducing cohort formation and population cycles*, Theor. Population Biology **63** (2003), 1–16.

[53] W. Desch and J. Turi, *Asymptotic theory for a class of functional-differential equations with state-dependent delays*, J. Math. Anal. Appl. **199** (1996), 75–87.

[54] O. Diekmann, S.A. van Gils, S.M. Verduyn Lunel and H.O. Walther, *Delay Equations: Functional-, Complex- and Nonlinear Analysis*, Springer, New York (1995).

[55] A. Domoshnitsky and M. Drakhlin, *Periodic solutions of differential equations with delay depending on solution*, Nonlinear Anal. TMA **30** (1997), 2665–2672.

[56] A. Domoshnitsky, M. Drakhlin and E. Litsyn, *On equations with delay depending on solution*. Nonlinear Anal. TMA **49** (2002), 689–701.

[57] R.D. Driver, *Existence theory for a delay-differential system*, Contributions to Differential Equations **1** (1963), 317–336.

[58] R.D. Driver, *A two-body problem of classical electrodynamics: The one-dimensional case*, Ann. Phys. **21** (1963), 122–142.

[59] R.D. Driver, *A functional-differential system of neutral type arising in a two-body problem of classical electrodynamics*, International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics, J. LaSalle and S. Lefschetz, eds., Academic Press, New York (1963), 474–484.

[60] R.D. Driver, *The "backwards" problem for a delay-differential system arising in a two-body problem of classical electrodynamics*, Proceedings of the Fifth International Conference on Nonlinear Oscillations, Vol. 2, Yu.A. Mitropol'skii and A.N. Sharkovskii, eds., Izdanie Inst. Mat. Akad. Nauk Ukrain. SSR, Kiev (1969), 137–143.

[61] R.D. Driver, *A "backwards" two-body problem of classical relativistic electrodynamics*, Phys. Rev. **178** (1969), 2051–2057.

[62] R.D. Driver, *A neutral system with state-dependent delay*, J. Differential Equations **54** (1984), 73–86.

[63] R.D. Driver, *A mixed neutral system*, Nonlinear Anal. TMA **8** (1984), 155–158.

[64] R.D. Driver and M. Norris, *Note on uniqueness for a one-dimensional two-body problem of classical electrodynamics*, Ann. Phys. **42** (1967), 347–351.

[65] M. Eichmann, *A local Hopf bifurcation theorem for differential equations with state dependent delay*, Ph.D. thesis, Universität Gießen, Gießen (2006).

[66] K. Engelborghs, T. Luzyanina and G. Samaey, *DDE-BIFTOOL v. 2.00: A Matlab package for bifurcation analysis of delay differential equations*, Report TW 330, Katholieke Universiteit Leuven, Belgium (October 2001).

[67] W.H. Enright, *Continuous numerical methods for ODEs with defect control*, J. Comput. Appl. Math. **125** (2000), 159–170.

[68] W.H. Enright and H. Hayashi, *A delay differential equation solver based on a continuous Runge–Kutta method with defect control*, Numer. Algorithms **16** (1997), 349–364.

[69] W.H. Enright and H. Hayashi, *Convergence analysis of the solution of retarded and neutral delay differential equations by continuous numerical methods*, SIAM J. Numer. Anal. **35** (1998), 572–585.

[70] W.H. Enright and M. Hu, *Interpolating Runge–Kutta methods for vanishing delay differential equations*, Computing **55** (1995), 223–236.

[71] W.H. Enright, K.R. Jackson, S.P. Nørsett and P.G. Thomsen, *Interpolants for Runge–Kutta formulas*, ACM Trans. Math. Software **12** (1986), 193–218.

[72] C.W. Eurich, J.D. Cowan and J.G. Milton, *Hebbian delay adaptation in a network of integrate-and-fire neurons*, Artificial Neural Networks, W. Gerstner, A. Germond, M. Hasler and J.-D. Nicoud, eds., Springer, Berlin (1997), 157–162.

[73] C.W. Eurich, U. Ernst, K. Pawelzik, J.D. Cowan and J.G. Milton, *Dynamics of self-organized delay adaptation*, Phys. Rev. Lett. **82** (1999), 1594–1597.

[74] C.W. Eurich, K. Pawelzik, K. Ernst, U. Thiel, J.D. Cowan and J.G. Milton, *Delay adaptation in the nervous system*, Neurocomputing **32–33** (2000), 741–748.

[75] C.W. Eurich, M.C. Mackey and H.C. Schwegler, *Recurrent inhibitory dynamics: the role of state dependent distributions of conduction delay times*, J. Theoret. Biol. **216** (2002), 31–50.

[76] A. Fathallah and O. Arino, *On a model representing stage-structured population growth with state-dependent time delay*, J. Biol. Systems **5** (1997), 469–488.

[77] A. Feldstein, *Discretization methods for retarded ordinary differential equations*, Doctoral thesis and Technical Report, Department of Mathematics, University of California, Los Angeles (1964).

[78] A. Feldstein and K.W. Neves, *High order methods for state-dependent delay differential equations with nonsmooth solutions*, SIAM J. Numer. Anal. **21** (1984), 844–863.

[79] A. Feldstein, K.W. Neves and S. Thompson, *Sharpness results for state-dependent differential equations: An overview*, Appl. Numer. Math., to appear.

[80] R. Gabasov and S.V. Churakova, *On the existence of optimal controllers in systems with time lags*, Differentsial'nye Uravneniya **3** (1967), 1074–1080.

[81] R. Gabasov and S.V. Churakova, *Necessary optimality conditions in time lag systems*, Automat. Remote Control **1** (1968), 37–54.

[82] R. Gabasov and S.V. Churakova, *Sufficient conditions for optimality in systems with a delay*, Automat. Remote Control **2** (1968), 193–209.

[83] J. Gallardo and M. Pinto, *Asymptotic integration of nonautonomous delay-differential systems*, J. Math. Anal. Appl. **199** (1996), 654–675.

[84] J. Gallardo and M. Pinto, *Asymptotic constancy of solutions of delay-differential equations of implicit type*, Rocky Mountain J. Math. **28** (1998), 487–504.

[85] R. Gambell, *Birds and mammals—Antarctic whales*, Antarctica, W.N. Bonner and D.W.H. Walton, eds., Pergamon Press, New York (1985), 223–241.

[86] J.G. Gatica and J. Rivero, *Qualitative behavior of solutions of some state-dependent delay equations*, Delay and Differential Equations, A.M. Fink, et al., eds., World Scientific, Singapore (1992), 36–56.

[87] J.G. Gatica and P. Waltman, *A threshold model of antigen antibody dynamics with fading memory*, Nonlinear Phenomena in Mathematical Sciences, V. Lakshmikantham, ed., Academic Press, New York (1982), 425–439.

[88] J.G. Gatica and P. Waltman, *Existence and uniqueness of solutions of a functional differential equation modeling thresholds*, Nonlinear Anal. TMA **8** (1984), 1215–1222.

[89] J.G. Gatica and P. Waltman, *A system of functional differential equations modeling threshold phenomena*, Appl. Anal. **28** (1988), 39–50.

[90] N. Guglielmi and E. Hairer, *Implementing Radau IIA methods for stiff delay differential equations*, Computing **67** (2001), 1–12.

[91] I. Győri and F. Hartung, *Numerical approximations for a class of differential equations with time- and state-dependent delays*, Appl. Math. Lett. **8** (1995), 19–24.

[92] I. Győri and F. Hartung, *On the exponential stability of a state-dependent delay equation*, Acta Sci. Math. (Szeged) **66** (2000), 71–84.

[93] I. Győri and F. Hartung, *On equi-stability with respect to parameters in functional differential equations*, Nonlinear Funct. Anal. Appl. **7** (2002), 329–351.

[94] H. Halanay and J.A. Yorke, *Some new results and problems in the theory of differential-delay equations*, SIAM Rev. **13** (1971), 55–80.

[95] J.K. Hale and J.F.C. de Oliveira, *Hopf bifurcation for functional equations*, J. Math. Anal. Appl. **74** (1980), 41–59.

[96] J.K. Hale and L.A.C. Ladeira, *Differentiability with respect to delays*, J. Differential Equations **92** (1991), 14–26.

[97] J.K. Hale and X.B. Lin, *Symbolic dynamics and nonlinear semiflows*, Ann. Mat. Pura Appl. **144** (1986), 229–259.

[98] J.K. Hale and S.M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Springer, New York (1993).

[99] F. Hartung, *On differentiability of solutions with respect to parameters in a class of functional differential equations*, Funct. Differential Equations **4** (1997), 65–79.

[100] F. Hartung, *Parameter estimation by quasilinearization in functional differential equations with state-dependent delays: a numerical study*, J. Nonlinear Anal. TMA **47** (2001), 4557–4566.

[101] F. Hartung, *Linearized stability in periodic functional differential equations with state-dependent delays*, J. Comput. Appl. Math. **174** (2005), 201–211.

[102] F. Hartung, T.L. Herdman and J. Turi, *On existence, uniqueness and numerical approximation for neutral equations with state-dependent delays*, Appl. Numer. Math. **24** (1997), 393–409.

[103] F. Hartung, T.L. Herdman and J. Turi, *Parameter identifications in classes of neutral differential equations with state-dependent delays*, Nonlinear Anal. TMA **39** (2000), 305–325.

[104] F. Hartung and J. Turi, *On the asymptotic behavior of the solutions of a state-dependent delay equation*, Differential and Integral Equations **8** (1995), 1867–1872.

[105] F. Hartung and J. Turi, *Stability in a class of functional differential equations with state-dependent delays*, Qualitative Problems for Differential Equations and Control Theory, C. Corduneanu, ed., World Scientific, Singapore (1995), 15–31.

[106] F. Hartung and J. Turi, *On differentiability of solutions with respect to parameters in state-dependent delay equations*, J. Differential Equations **135** (1997), 192–237.

[107] F. Hartung and J. Turi, *Identification of parameters in delay equations with state-dependent delays*, Nonlinear Anal. TMA **29** (1997), 1303–1318.

[108] F. Hartung and J. Turi, *Linearized stability in functional differential equations with state-dependent delays*, Dynamical Systems and Differential Delay Equations, Kennesaw, GA, 2000, Discrete and Continuous Dynamical Systems (Added Volume) (2001), 416–425.

[109] R. Hauber, *Numerical treatment of retarded differential-algebraic equations by collocation methods*, Adv. Comput. Math. **7** (1997), 573–592.

[110] D.J. Higham, *Highly continuous RungeKutta interpolants*, ACM Trans. Math. Software **17** (1991), 368–386.

[111] J.T. Hoag and R.D. Driver, *A delayed-advanced model for the electrodynamics two-body problem*, Nonlinear Anal. TMA **15** (1990), 165–184.

[112] F.C. Hoppensteadt and P. Waltman, *A flow mediated control model or respiration*, Lectures Math. Life Sci., Vol. 12, Amer. Math. Soc., Providence, RI (1979), 211–218.

[113] H. Hüning, H. Glünder and G. Palm, *Synaptic delay learning in pulse-coupled neurons*, Neural Comput. **10** (1998), 555–565.

[114] J. Hunter, J.G. Milton and J. Wu, *Adaptive delay in projective clustering*, in progress (2005).

[115] G.M. Innocenti, P. Lehmann and J.-C. Nouzel, *Computational structure of visual callosal axons*, European J. Neuroscience **6** (1994), 918–935.

[116] T. Insperger, G. Stépán, F. Hartung and J. Turi, *State dependent regenerative delay in milling processes*, Proceedings of the ASME International Design Engineering Technical Conferences, Long Beach, CA (2005), paper no. DETC2005-85282 (CD-ROM).

[117] T. Insperger, G. Stépán and J. Turi, *State-dependent delay model for regenerative cutting processes*, Proceedings of the Fifth EUROMECH Nonlinear Dynamics Conference, Eindhoven, The Netherlands (2005), 1124–1129.

[118] K. Ito and F. Kappel, *Approximation of semilinear Cauchy problems*, Nonlinear Anal. TMA **24** (1995), 51–80.

[119] Z. Jackiewicz, *Existence and uniqueness of solutions of neutral delay-differential equations with state-dependent delays*, Funkcial. Ekvac. **30** (1987), 9–17.

[120] J. Jackiewicz, *One-step methods for neutral delay-differential equations with state dependent delays*, Zastos. Math. **20** (1990), 445–463.

[121] Z. Jackiewicz, *A note on existence and uniqueness of solutions of neutral functional-differential equations with state-dependent delay*. Comment. Math. Univ. Carolin. **36** (1995), 15–17.

[122] Z. Jackiewicz and E. Lo, *The numerical integration of neutral functional differential equations by Adams predictor-corrector methods*, Appl. Numer. Math. **8** (1991) 477–491.

[123] R.A. Johnson, *Functional equations, approximations, and dynamic response of systems with variable time-delay*, IEEE Trans. Automatic Control **17** (1972), 398–401.

[124] A. Karoui and R. Vaillancourt, *An adapted Runge–Kutta pair for state-dependent delay differential equations*, C.R. Math. Rep. Acad. Sci. Canada **15** (1993), 193–198.

[125] A. Karoui and R. Vaillancourt, *Computer solutions of state-dependent delay differential equations*, Comput. Math. Appl. **27** (1994), 37–51.

[126] H.P. Krishnan, *An analysis of singularly perturbed delay-differential equations and equations with state-dependent delays*, Ph.D. thesis, Brown University, Providence, RI (1998).

[127] H.P. Krishnan, *Existence of unstable manifolds for a certain class of delay differential equations*, Electronic J. Differential Equations **32** (2002), 1–13.

[128] T. Krisztin, *On stability properties for one-dimensional functional differential equations*, Funkcial. Ekvac. **34** (1991), 241–256.

[129] T. Krisztin, *A local unstable manifold for differential equations with state-dependent delay*, Discrete Continuous Dynamical Systems **9** (2003), 993–1028.

[130] T. Krisztin, *Invariance and noninvariance of center manifolds of time-t maps with respect to the semiflow*, SIAM J. Math. Anal. **36** (2004), 717–739.

[131] T. Krisztin, *C^1-smoothness of center manifolds for differential equations with state-dependent delays*, Nonlinear Dynamics and Evolution Equations, Fields Institute Communications **48** (2006), 213–226.

[132] T. Krisztin and O. Arino, *The 2-dimensional attractor of a differential equation with state-dependent delay*, J. Dynamics Differential Equations **13** (2001), 453–522.

[133] T. Krisztin, H.O. Walther and J. Wu, *Shape, Smoothness, and Invariant Stratification of an Attracting Set for Delayed Monotone Positive Feedback*, Fields Institute Monograph Series, Vol. 11, Amer. Math. Soc., Providence, RI (1999).

[134] T. Krisztin and J. Wu, *Monotone semiflows generated by neutral equations with different delays in neutral and retarded parts*, Acta Math. Univ. Comenian. **63** (1994), 207–220.

[135] Y. Kuang, $3/2$ *stability results for nonautonomous state-dependent delay differential equations*, Differential Equations to Biology and to Industry, M. Martelli et al., eds., World Scientific, Singapore (1996), 261–269.

[136] Y. Kuang and H.L. Smith, *Periodic solutions of differential delay equations with threshold-type delays*, Oscillation and Dynamics in Delay Equations, J.R. Graef and J.K. Hale, eds., Contemp. Math., Vol. 120, Amer. Math. Soc., Providence, RI (1992), 153–176.

[137] Y. Kuang and H.L. Smith, *Slowly oscillating periodic solutions of autonomous state-dependent delay differential equations*, Nonlinear Anal. TMA **19** (1992), 855–872.

[138] Y. Li and Y. Kuang, *Periodic solutions in periodic state-dependent delay equations and population models*, Proc. Amer. Math. Soc. **130** (2002) 1345–1353.

[139] W. Liu, *Positive periodic solutions to state-dependent delay differential equations*, Appl. Math. Ser. A (Chinese Ed.) **17** (2002), 22–28.

[140] Y. Liu, *Numerical solutions of implicit neutral functional differential equations*, SIAM J. Numer. Anal. **36** (1999), 516–528.

[141] A. Longtin and J. Milton, *Complex oscillations in the human pupil light reflex with mixed and delayed feedback*, Math. Biosci. **90** (1988), 183–199.

[142] A. Longtin and J. Milton, *Modelling autonomous oscillations in the human pupil light reflex using nonlinear delay-differential equations*, Bull. Math. Biol. **51** (1989), 605–624.

[143] A. Longtin and J. Milton, *Insight into the transfer function, gain and oscillation onset for the pupil light reflex using nonlinear delay-differential equations*, Biol. Cybern. **61** (1989), 51–58.

[144] M. Louihi, M.L. Hbid and O. Arino, *Semigroup properties and the Crandall–Liggett approximation for a class of differential equations with state-dependent delays*, J. Differential Equations **181** (2002), 1–30.

[145] T. Luzyanina, K. Engelborghs and D. Roose, *Numerical bifurcation analysis of differential equations with state-dependent delays*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **11** (2001), 737–753.

[146] M.C. Mackey, *Commodity price fluctuations: Price-dependent delays and nonlinearities as explanatory factors*, J. Econom. Theory **48** (1989), 497–509.

[147] M.C. Mackey and J. Milton, *Feedback delays and the origin of blood cell dynamics*, Comment. Theoret. Biol. **1** (1990), 299–327.

[148] P. Magal and O. Arino, *Existence of periodic solutions for a state-dependent delay differential equation*, J. Differential Equations **165** (2000), 61–95.

[149] J. Mahaffy, J. Bélair and M. Mackey, *Hematopoietic model with moving boundary condition and state-dependent delay: Applications in erythropoiesis*, J. Theoret. Biol. **190** (1998), 135–146.

[150] J. Mallet-Paret, *Morse decompositions for differential delay equations*, J. Differential Equations **72** (1988), 270–315.

[151] J. Mallet-Paret and R.D. Nussbaum, *Boundary layer phenomena for differential-delay equations with state-dependent time-lags: I*, Arch. Rational Mech. Anal. **120** (1992), 99–146.

[152] J. Mallet-Paret and R.D. Nussbaum, *Boundary layer phenomena for differential-delay equations with state-dependent time-lags: II*, J. Reine Angew. Math. **477** (1996), 129–197.

[153] J. Mallet-Paret and R.D. Nussbaum, *Eigenvalues for a class of homogeneous cone maps arising from max-plus operators*, Discrete and Continuous Dynamical Systems **8** (2002), 519–562.

[154] J. Mallet-Paret and R.D. Nussbaum, *A basis theorem for a class of max-plus eigenproblems*, J. Differential Equations **189** (2003), 616–639.

[155] J. Mallet-Paret and R.D. Nussbaum, *Boundary layer phenomena for differential-delay equations with state-dependent time-lags: III*, J. Differential Equations **189** (2003), 640–692.

[156] J. Mallet-Paret, R.D. Nussbaum and P. Paraskevopoulos, *Periodic solutions for functional differential equations with multiple state-dependent time lags*, Topol. Methods Nonlinear Anal. **3** (1994), 101–162.

[157] J. Mallet-Paret and G. Sell, *Systems of differential delay equations: Floquet multipliers and discrete Lyapunov functions*, J. Differential Equations **125** (1996), 385–440.

[158] J. Mallet-Paret and G. Sell, *The Poincaré–Bendixson theorem for monotone cyclic feedback systems with delay*, J. Differential Equations **125** (1996), 441–489.

[159] J. Mallet-Paret and H.O. Walther, *Rapid oscillations are rare in scalar systems governed by monotone negative feedback with a time lag*, Preprint (1994).

[160] A. Manitius, *On the optimal control of systems with a delay depending on state, control, and time*, Séminaires IRIA, Analyse et Contrôle de Systèmes, IRIA, France (1975), 149–198.

[161] J.A. Messer, *Steuerung eines Roboterarms durch eine Differentialgleichung mit Zeitverzögerung*, Preliminary report (2002).

[162] J.A.J. Metz and O. Diekmann, *The Dynamics of Physiologically Structured Populations*, Lecture Notes in Biomath., Vol. 68, Springer, Berlin (1986).

[163] J. Milton, *Dynamics of Small Neural Networks*, Amer. Math. Soc., Providence, RI (1996).

[164] K. Murphy, *Estimation of time- and state-dependent delays and other parameters in functional-differential equations*, SIAM J. Appl. Math. **50** (1990), 972–1000.

[165] A.D. Myshkis, *On certain problems in the theory of differential equations with deviating argument*, Uspekhi Mat. Nauk **32** (1977), 173–202. English translation in: Russian Math. Surveys **32** (1977), 181–213.

[166] H. Napp-Zinn, M. Jansen and R. Eckmiller, *Recognition and tracking of impulse patterns with delay adaptation in biology-inspired pulse-processing neural net (BNP) hardware*, Biol. Cybern. **74** (1996), 449–453.

[167] K.W. Neves, *Automatic integration of functional differential equations: an approach*, ACM Trans. Math. Software **1** (1975), 357–368.

[168] K.W. Neves and A. Feldstein, *Characterization of jump discontinuities for state dependent delay differential equations*, J. Math. Anal. Appl. **56** (1976), 689–707.

[169] K.W. Neves and S. Thompson, *Software for the numerical solution of systems of functional differential equations with state-dependent delays*, Appl. Numer. Math. **9** (1992), 385–401.

[170] R.M. Nisbet and W.S.C. Gurney, *The systematic formulation of population models for insects with dynamically varying instar duration*, Theoret. Population Biol. **23** (1983), 114–135.

[171] A. Neugebauer, *Invariante Mannigfaltigkeiten und Neigungslemmata für Abbildungen in Banachräumen*, Diploma thesis, Ludwig-Maximilians-Universität München, München (1988).

[172] K. Niri, *Oscillations in differential equations with state-dependent delays*, Nonlinear Oscillations **6** (2003), 252–259.

[173] S. Nishihira and T. Yoneyama, *Uniform stability of the functional-differential equation* $x'(t) = A(t)x(t - r(t, x_t))$, Structure of Functional Equations and Mathematical Methods (Kyoto, 1996), Surikaisekikenkyusho Kokyuroku **984** (1997), 71–74.

[174] R.D. Nussbaum, *Periodic solutions of some nonlinear autonomous functional differential equations*, Ann. Mat. Pura Appl. (4) **101** (1974), 263–306.

[175] R.D. Nussbaum, *A global bifurcation theorem with applications to functional differential equations*, J. Funct. Anal. **19** (1975), 319–339.

[176] R.D. Nussbaum, *Periodic solutions of some integral equations from the theory of epidemics*, Nonlinear Systems and Applications, V. Lakshmikantham, ed., Academic Press, New York (1977), 235–257.

[177] R.D. Nussbaum, Personal communication.

[178] R. Ouifki and M.L. Hbid, *Periodic solutions for a class of functional differential equations with state-dependent delay close to 0*, Math. Models Methods Appl. Sci. **13** (2003), 807–841.

[179] C.A.H. Paul, *A user guide to ARCHI—an explicit (Runge–Kutta) code for solving delay and neutral differential equations*, NA Report 283, Department of Mathematics, University of Manchester (1995).

[180] M. Pinto, *Asymptotic integration of the functional-differential equation* $y'(t) = a(t)y(t - r(t, y))$, J. Math. Anal. Appl. **175** (1993), 46–52.

[181] S.D. Poisson, *Sur les équations aux différences melées*, J. de l'Ecole Polytechnique Paris (1) 6, cahier 13 (1806), 126–147.

[182] S. Rai and R.L. Robertson, *Analysis of a two-stage population model with space limitations and state-dependent delay*, Canad. Appl. Math. Quart. **8** (2000), 263–279.

[183] S. Rai and R.L. Robertson, *A stage-structured population model with state-dependent delay*, Internat. J. Differential Equations Appl. **6** (2002), 77–91.

[184] A.V. Rezounenko and J. Wu, *A non-local PDE model for population dynamics with state-selective delay: Local theory and global attractors*, J. Comput. Appl. Math. (2005), in press.

[185] M. Santillan, J. Mahaffy, J. Bélair and M.C. Mackey, *Regulation of platelet production: The normal response to perturbation and cyclical platelet disease*, J. Theoret. Biol. **206** (2000), 585–603.

[186] J.-G. Si and X.-P. Wang, *Analytic solutions of a second order functional differential equation with a state derivative dependent delay*, Colloq. Math. **79** (1999), 273–281.

[187] J.-G. Si and X.-P. Wang, *Analytic solutions of a second-order functional differential equation with a state dependent delay*, Result. Math. **39** (2001), 345–352.

[188] J.-G. Si, X. Wang and S.S. Cheng, *Analytic solutions of a functional differential equation with a state derivative dependent delay*, Aequationes Math. **57** (1999), 75–86.

[189] H.L. Smith, *Hopf bifurcation in a system of functional equations modeling the spread of infectious disease*, SIAM J. Appl. Math. **43** (1983), 370–385.

[190] H.L. Smith, *Some results on the existence of periodic solutions of state-dependent delay differential equations*, Ordinary and Delay Differential Equations, J. Wiener et al., eds., Pitman Res. Notes Math. Ser., Vol. 272, Longman Scientific & Technical, Harlow, Essex, UK (1993), 218–222.

[191] H.L. Smith, *Monotone Dynamical Systems. An Introduction to the Theory of Competitive and Cooperative Systems*, Amer. Math. Soc., Providence, RI (1995).

[192] L.R. Stanford, *Conduction velocity variations minimize conduction time differences among retinal ganglion cell axons*, Science **238** (1987), 358–360.

[193] B. Stevens, S. Tanner and R.D. Fields, *Control of myelination by specific patterns of neural impulses*, J. Neuroscience **18** (1998), 9303–9311.

[194] A. Stokes, *Stability of functional differential equations with perturbed lags*, J. Math. Anal. Appl. **47** (1974), 604–619.

[195] J. Sugie, *Oscillating solutions of scalar delay-differential equations with state dependence*, Appl. Anal. **27** (1998), 217–227.

[196] L. Tavernini, *The approximate solution of Volterra differential systems with state-dependent time lags*, SIAM J. Numer. Anal. **15** (1978), 1039–1052.

[197] S. Thompson and L.F. Shampine, *A friendly Fortran DDE solver*, Appl. Numer. Math., to appear.

[198] S.P. Travis, *A one-dimensional two-body problem of classical electrodynamics*, SIAM J. Appl. Math. **28** (1975), 611–632.

[199] K.P. Unnikrishnan, J.J. Hopfield and D.W. Tank, *Connected-digit speaker dependent speech recognition using a neural network with time-delayed connections*, IEEE Trans. Signal Processing **39** (1991), 698–713.

[200] E.I. Verriest, *Stability of systems with state-dependent and random delays*, IMA J. Math. Control Inform. **19** (2002), 103–114.

[201] P. Waltman, *Deterministic Threshold Models in the Theory of Epidemics*, Lecture Notes in Biomath., Vol. 1, Springer, New York (1974).

[202] H.O. Walther, *On instability, ω-limit sets, and periodic solutions of nonlinear autonomous differential delay equations*, Functional Differential Equation and Approximation of Fixed Points, H.O. Peitgen and H.O. Walther, eds., Lecture Notes in Math., Vol. 730, Springer, Heidelberg (1979), 489–503.

[203] H.O. Walther, *Delay equations: Instability and the trivial fixed point's index*, Abstract Cauchy Problems and Functional Differential Equations, F. Kappel and W. Schappacher, eds., Pitman Res. Notes Math. Ser., Vol. 48, Pitman, London (1981), 231–238.

[204] H.O. Walther, *The two-dimensional attractor of $x'(t) = -\mu x(t) + f(x(t-1))$*, Mem. Amer. Math. Soc. **113** (1995), no. 544.

[205] H.O. Walther, *Stable periodic motion of a system with state-dependent delay*, Differential Integral Equations **15** (2002), 923–944.

[206] H.O. Walther, *The solution manifold and C^1-smoothness of solution operators for differential equations with state dependent delay*, J. Differential Equations **195** (2003), 46–65.

[207] H.O. Walther, *Smoothness properties of semiflows for differential equations with state dependent delay*, Proceedings of the International Conference on Differential and Functional Differential Equations, Moscow, 2002, Vol. 1, Moscow State Aviation Institute (MAI), Moscow (2003), 40–55 (in Russian). English version: J. Math. Sci. **124** (2004), 5193–5207.

[208] H.O. Walther, *Stable periodic motion of a system using echo for position control*, J. Dynamics Differential Equations **15** (2003), 143–223.

[209] H.O. Walther, *Differentiable semiflows for differential equations with state-dependent delay*, Univ. Iagel. Acta Math. **XLI** (2003), 53–62.

[210] H.O. Walther and M. Yebdri, *Smoothness of the attractor of almost all solutions of a delay differential equation*, Dissertationes Mathematicae **CCCLXVIII** (1997).

[211] D.R. Willé and C.T.H. Baker, *Stepsize control and continuity consistency for state-dependent delay-differential equations*, J. Comput. Appl. Math. **53** (1994), 163–170.

[212] D.R. Willé and C.T.H. Baker, *A short note on the propagation of derivative discontinuities in Volterra-delay integro-differential equations*, NA Report 187, Department of Mathematics, University of Manchester (1990).

[213] D.R. Willé and C.T.H. Baker, *The tracking of derivative discontinuities in systems of delay differential equations*, Appl. Numer. Math. **9** (1992), 209–222.

[214] D.R. Willé and C.T.H. Baker, *Some issues in the detections and locations of derivative discontinuities in delay-differential equations*, NA Report 238, Department of Mathematics, University of Manchester (1994).

[215] S.R. Williams and G.J. Stuart, *Dependence of EPSP efficacy on synapse location in neocortical pyramidal neurons*, Science **295** (2002), 1907–1910.

[216] E. Winston, *Uniqueness of the zero solution for differential equations with state-dependence*, J. Differential Equations **7** (1970), 395–405.

[217] E. Winston, *Comparison theorems for scalar delay differential equations*, J. Math. Anal. Appl. **29** (1970), 455–463.

[218] E. Winston, *Uniqueness of solutions of state dependent delay differential equations*, J. Math. Anal. Appl. **47** (1974), 620–625.

[219] J. Wu, *Introduction to Neural Dynamics and Signal Transmission Delay*, Walter de Gruyter, Berlin (2001).

[220] W. Yan, A. Zhao and J. Yan, *Comparison theorem and oscillation results on state-dependent delay differential equations*, Appl. Anal. **69** (1998), 199–205.

[221] Z. Yang and J. Cao, *Existence of periodic solutions in neutral state-dependent delays equations and models*, J. Comput. Appl. Math. **174** (2005), 179–199.

[222] T. Yoneyama, *On the $\frac{3}{2}$ stability theorem for one-dimensional delay differential equations*, J. Math. Anal. Appl. **125** (1987), 161–173.

[223] T. Yoneyama, *Uniform stability for n-dimensional delay-differential equations*, Funkcial. Ekvac. **34** (1991), 495–504.

[224] A.A.S. Zaghrout and S.H. Attalah, *Analysis of a model of stage-structured population dynamics growth with time state-dependent time delay*, Appl. Math. Comput. **77** (1996), 185–194.

[225] Z. Zhang, X. Zheng and Z. Wang, *Periodic solutions of two-patches predator-prey dispersion-delay models with functional response*, Bull. Inst. Math. Acad. Sinica **31** (2003), 117–133.

This page intentionally left blank

CHAPTER 6

Global Solution Branches and Exact Multiplicity of Solutions for Two Point Boundary Value Problems

Philip Korman

Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221-0025, USA

Dedicated to the memory of my father Lev Korman

Contents

1. Introduction . . . 549
2. Bifurcation theory approach . . . 553
    2.1. Some general properties of solutions of autonomous problems . . . 553
    2.2. Convex nonlinearities . . . 559
    2.3. S-shaped solution curves . . . 561
    2.4. Cubic-like nonlinearities . . . 568
    2.5. Computing the location of bifurcation . . . 573
    2.6. Computing the direction of bifurcation . . . 575
    2.7. Pitchfork bifurcation and symmetry breaking . . . 575
    2.8. Sign-changing solutions . . . 577
    2.9. The Neumann problem . . . 579
    2.10. Similarity of the solution branches . . . 582
3. A class of symmetric nonlinearities . . . 585
4. General nonlinearities . . . 590
5. Time maps . . . 594
    5.1. There are several different formulas for the time map . . . 594
    5.2. Time map formula through global linearization . . . 595
    5.3. Variational formula for the time map . . . 601
    5.4. A nonlocal problem . . . 601
6. Numerical computation of solutions . . . 602
Acknowledgement . . . 604
References . . . 604

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 3
Edited by A. Cañada, P. Drábek and A. Fonda
© 2006 Elsevier B.V. All rights reserved

This page intentionally left blank

1. Introduction

We consider solutions of the two point boundary value problems

$$u''(x)+\lambda f\big(x,u(x)\big)=0 \quad \text{for } a<x<b, \qquad u(a)=u(b)=0, \tag{1.1}$$

depending on a parameter λ. We wish to know how many *exactly* solutions does problem (1.1) have, and how these solutions change with λ. What is the role of the parameter λ? Of course, it could be absorbed into the nonlinearity f. However, as is often the case, it is helpful to have something "extra" in the statement of the problem. Consider for example the problem

$$u''(x)+4\mathrm{e}^{\frac{5u(x)}{5+u(x)}}=0 \quad \text{for } 0<x<1, \qquad u(0)=u(1)=0. \tag{1.2}$$

Problems of this type come up in combustion theory, referred to as "perturbed Gelfand problem", see, e.g., J. Bebernes and D. Eberly [7]. It will follow from a result we present below that this problem has *exactly* three positive solutions. It appears next to impossible to establish this result directly. We introduce a parameter λ, and consider

$$u''(x)+\lambda\mathrm{e}^{\frac{5u(x)}{5+u(x)}}=0 \quad \text{for } 0<x<1, \qquad u(0)=u(1)=0. \tag{1.3}$$

We now study *curves of solutions*, $u=u(x,\lambda)$. The advantage of this approach is that some parts of the solution curve are easy to understand, and it also becomes clear what are the tougher parts of the solution curve that we need to study—the *turning points*. For example, it is easy to understand the "small" solutions of (1.3), by applying the implicit function theorem (in Banach spaces) in the neighborhood of the trivial solution $\lambda=0$, $u=0$. We then continue this curve of solutions for increasing $\lambda>0$ until a critical solution is reached, i.e. the implicit function theorem is no longer applicable. We show that at the critical solution the Crandall–Rabinowitz Theorem 1.2 (see below) applies. It implies that either the solution curve continues forward in λ through the critical solution, or it just bends back (no secondary bifurcations or other eccentric behaviour is possible). We then show that the global solution curve makes exactly two turns, and the value of $\lambda=4$ from (1.2) lies between the turns, thus establishing the existence of three solutions. The bifurcation approach, just described, has been developed in the recent years by Y. Li, T. Ouyang, J. Shi and the present author. It applies also to the semilinear elliptic problems for balls in $\mathbb{R}^n$, however in the present paper we restrict to the ODE case (1.1).

The most detailed results are obtained when one considers positive solutions of autonomous problems, i.e. when $f=f(u)$. Since in that case both the length and the position of the interval (a,b) are irrelevant, and since positive solutions are symmetric with respect to the midpoint of the interval, it is convenient to pose the problem on the interval $(-1,1)$, i.e. we consider

$$u''(x)+\lambda f\big(u(x)\big)=0 \quad \text{for } -1<x<1, \qquad u(-1)=u(1)=0. \tag{1.4}$$

It turns out that convexity properties of $f(u)$ are important for determining the direction of the turn for solution curves. Accordingly, in the simplest case $f''(u) > 0$ and $f(u) > 0$, for $u > 0$, we can give an exhaustive analysis of the problem. (In case $f''(u) < 0$ and $f(u) > 0$, for $u > 0$, it is easy to prove uniqueness of solutions.) The next case in order of complexity is when $f(u)$ changes concavity exactly once. The prominent case is when $f(u)$ is modelled on a cubic with simple roots:

$$u'' + \lambda(u-a)(u-b)(c-u) = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0. \tag{1.5}$$

We assume that $0 \leqslant a < b < c$, since the analysis is easier if some root(s) is negative. We wish to describe how many *exactly* positive solutions does problem (1.5) have for various λ.

This problem was studied in a 1981 paper by J. Smoller and A. Wasserman [57]. They succeeded in solving the problem for $a = 0$, while their proof for $a > 0$ case contained an error. This error was discovered by S.-H. Wang [59], who was able to solve the problem under some restriction on $a > 0$. Both papers used the phase-plane analysis. Then P. Korman, Y. Li and T. Ouyang [30] used bifurcation theory to attack the problem, but again some restrictions were necessary (all of the above mentioned papers covered more general $f(u)$, behaving like cubic). Very recently, P. Korman, Y. Li and T. Ouyang [33], building on their previous work, have given a computer assisted proof for general cubic. It turns out that the set of all positive solutions consists of two curves, with the lower curve monotone in λ, and the upper curve having exactly one turn. The computations in P. Korman, Y. Li and T. Ouyang [30] also showed that the approach in J. Smoller and A. Wasserman [57] could not possibly cover the general cubic. (That approach required a certain integral to be positive, in order to derive a differential inequality for a time map. However, that integral changes sign for some cubics.) In the next section we state the optimal result, and describe the approach taken in [30] and [33].

Another prominent class of problems where $f(u)$ changes concavity exactly once is

$$u'' + \lambda e^{\frac{au}{u+a}} = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0, \tag{1.6}$$

from combustion theory. Here a is a second parameter. In case $a = 5$, we have the problem (1.3), discussed above. If $a \leqslant 4$, the problem is easy. In that case the solution curve is monotone, and it continues for all $\lambda > 0$ without any turns. Following some earlier results of K.J. Brown, M.M.A. Ibrahim and R. Shivaji [9] and others (see [60] and [29] for the earlier references), S.-H. Wang [60] has proved existence of a constant a_0, so that for $a > a_0$ the solution curve of (2.21) is exactly S-shaped, i.e. it starts at $\lambda = 0$, $u = 0$, it makes exactly two turns, and then it continues for all $\lambda > 0$ without any more turns. S.-H. Wang [60] gave an approximation of the constant $a_0 \simeq 4.4967$. That paper, as well as all previous ones, used a time map approach. P. Korman and Y. Li [29] have applied the bifurcation approach to the problem. Since bifurcation approach is more general, this opened a way to do other problems. In fact, Y. Du and Y. Lou [14] have used a similar approach, with several additional tricks of their own, to prove that for a ball in two dimensions a similar result holds for sufficiently large a.

P. Korman and Y. Li [29] had also improved the value of the constant to $a_0 \simeq 4.35$, i.e. for $a > a_0$ the solution sets are S-shaped curves. But what about $4 < a < a_0$? S.-H. Wang [60] has conjectured existence of a critical number $\bar{a}$, so that for $a \leqslant \bar{a}$ the solution curve is monotone, while for $a > \bar{a}$ the solution curve is exactly S-shaped (the number a_0, mentioned above, is just an upper bound for $\bar{a}$). Recently, P. Korman, Y. Li and T. Ouyang [33] has given a computer assisted proof of the S.-H. Wang's conjecture. Numerical calculations show that $\bar{a} \simeq 4.07$.

Other topics we discuss using the bifurcation approach involve pitchfork bifurcation and symmetry breaking, sign changing solutions, and the Neumann problem. We also present a recent formula from P. Korman, Y. Li and T. Ouyang [33], which allows one to compute all possible values of $\alpha = u(0)$, at which solution of (1.4), with the maximal value equal to α, is singular.

The case when $f = f(x,u)$ is much harder than the autonomous case. In particular, the time map method does not apply. Bifurcation approach works, but it becomes much more complicated. For example, solutions of the corresponding linearized problem need not be of one sign (an implicit example of that is provided by the Theorem 1.10 in W.-M. Ni and R.D. Nussbaum [46]). In the papers P. Korman and T. Ouyang a class of $f(x,u)$ has been identified, for which the theory of positive solutions is very similar to that for the autonomous case, see, e.g., [34–36]. Further results in this direction have been given in P. Korman, Y. Li and T. Ouyang [30], and P. Korman and J. Shi [40]. Namely, assume that $f \in C^2$ satisfies

$$f(-x,u) = f(x,u) \quad \text{for all } -1 < x < 1 \text{ and } u > 0,$$

$$f_x(x,u) \leqslant 0 \quad \text{for all } 0 < x < 1 \text{ and } u > 0.$$

Under the above conditions any positive solution of (1.1) is an even function, with $u'(x) < 0$ for all $x \in (0,1]$, see B. Gidas, W.-M. Ni and L. Nirenberg [15]. We show that any solution of the corresponding linearized problem is of one sign, and then outline a number of exact multiplicity results.

Without symmetry assumption on $f(x,u)$ things are even more hard. In Section 4 we present extensions of the previous results in P. Korman and T. Ouyang [38]. The notion of Schwarzian derivative from Complex Analysis turns out to play a role here.

The bifurcation approach is effective for other problems, in addition to the two point problems that we discuss in the present paper. Most notably, similar results were developed for PDE's on a ball or annulus in $\mathbb{R}^n$, see, e.g., P. Korman, Y. Li and T. Ouyang [31] or T. Ouyang and J. Shi [50]. It was also used for systems of equations in P. Korman [22], for fourth order equations in P. Korman [26], and for periodic problems in P. Korman and T. Ouyang [37].

In Section 5 we give a brief review of time map method. Let $u = u(t)$ be solution of the initial value problem,

$$u'' + f(u) = 0, \qquad u(0) = 0, \quad u'(0) = p.$$

Using ballistic analogy, we can interpret this as "shooting" from the ground level, at an angle $p > 0$. Let $T/2$ denote the time it takes for the projectile to reach its maximum amplitude. By symmetry of positive solutions, $T = T(p)$ is then the time when the projectile falls back to the ground, the *time map*. The function $u(t)$ then satisfies the two point Dirichlet problem

$$u'' + f(u) = 0 \quad \text{for } 0 < t < T, \qquad u(0) = u(T) = 0,$$

which by rescaling is equivalent to (1.4). There are two completely different formulas for the same time map $T = T(p)$. The first one is obtained by direct integration, see, e.g., W.S. Loud [43] for an early reference, while the second one was derived by R. Schaaf [53] through a change of variables, which converts the problem into a harmonic oscillator. Both formulas for the time map are nontrivial to use. The first one involves improper integrals, while the second one is highly implicit. ("Name your poison", so to say.) However, both approaches are well developed by now, see the book by R. Schaaf [53], and the papers of S.-H. Wang and his coworkers, of I. Addou, and many other papers, including J. Cheng [10,11], and K.J. Brown et al. [9]. We give an exposition of the second approach, and connect it to the notion of *generalized averages* from P. Korman and Y. Li [28].

In the final Section 6 we discuss numerical computation of solutions of (1.4). Again, the autonomous case is much easier. We describe two efficient ways to compute the solutions, and explain why finite differences (or finite elements) are not appropriate for autonomous problems.

The basic tool for continuation of solutions is the implicit function theorem in Banach spaces. We present it here in the formulation of M.G. Crandall and P.H. Rabinowitz [12], see also L. Nirenberg [45], and A. Ambrosetti and G. Prodi [5].

THEOREM 1.1. *Let X, Λ and Z be Banach spaces, and $f(x, \lambda)$ a continuous mapping of an open set $U \subset X \times \Lambda \to Z$. Assume that f has a Frechet derivative with respect to x, $f_x(x, \lambda)$ which is continuous on U. Assume that*

$$f(x_0, \lambda_0) = 0 \quad \textit{for some } (x_0, \lambda_0) \in U.$$

If $f_x(x_0, \lambda_0)$ is an isomorphism (i.e. $1:1$ and onto) of X onto Z, then there is a ball $B_r(\lambda_0) = \{\lambda \colon \|\lambda - \lambda_0\| < r\}$ and a unique continuous map $x(\lambda) : B_r(\lambda_0) \to X$, such that

$$f\big(x(\lambda), \lambda\big) \equiv 0, \quad x(\lambda_0) = x_0.$$

If f is of class C^p, so is $x(\lambda)$, $p \geqslant 1$.

In the conditions of the above theorem, we refer to (x_0, λ_0) as a regular solution, otherwise we call a solution singular. What happens at a singular solution? (I.e. when $f_x(x_0, \lambda_0)$ is not an isomorphism.) In general, practically anything imaginable may happen, as one can see even for functions of two variables. However, in a lucky case solution will continue through a critical point, either by making a simple turn there, or maybe it even continues forward in λ (the critical point is then like a point of inflection). M.G. Crandall and P.H. Rabinowitz [13] have given conditions for that to occur. The following result is one of our principal tools.

THEOREM 1.2. [13] *Let X and Y be Banach spaces. Let $(\overline{\lambda}, \overline{x}) \in \mathbb{R} \times X$ and let F be a continuously differentiable mapping of an open neighborhood of $(\overline{\lambda}, \overline{x})$ into Y. Let the null-space $N(F_x(\overline{\lambda}, \overline{x})) = \operatorname{span} x_0$ be one-dimensional and* $\operatorname{codim} R(F_x(\overline{\lambda}, \overline{x})) = 1$. *Let $F_\lambda(\overline{\lambda}, \overline{x}) \notin R(F_x(\overline{\lambda}, \overline{x}))$. If Z is a complement of $\operatorname{span} x_0$ in X, then the solutions of $F(\lambda, x) = F(\overline{\lambda}, \overline{x})$ near $(\overline{\lambda}, \overline{x})$ form a curve $(\lambda(s), x(s)) = (\overline{\lambda} + \tau(s), \overline{x} + s x_0 + z(s))$, where $s \to (\tau(s), z(s)) \in \mathbb{R} \times Z$ is a continuously differentiable function near $s = 0$ and* $\tau(0) = \tau'(0) = 0$, $z(0) = z'(0) = 0$.

Except for a brief discussion on p-Laplace equations, we consider only the classical solutions throughout this paper. We shall denote the derivatives of $u(x)$ by either $u'(x)$ or u_x, and mix both notations sometimes to make our discussion more transparent.

Most of the results in the present paper are based on our joint papers with Y. Li, T. Ouyang and J. Shi. Working with these talented colleagues has been a wonderful experience for me, and I wish to thank them for this opportunity. I also wish to thank Professors A. Canada, P. Drabek and A. Fonda for inviting me to write this review paper.

2. Bifurcation theory approach

2.1. *Some general properties of solutions of autonomous problems*

We will consider positive, negative and sign-changing solutions of the Dirichlet problem (for $u = u(x)$)

$$u'' + \lambda f(u) = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0, \tag{2.1}$$

depending on a parameter λ. We assume throughout this section that $f(u) \in C^2(\overline{\mathbb{R}}_+)$. We choose to consider the problem on the interval $(-1, 1)$ for convenience (which is related to the symmetry of solutions). By shifting and scaling, we can replace the interval $(-1, 1)$ by any other interval (a, b).

LEMMA 2.1. *Let $\xi \in (-1, 1)$ be any critical point of $u(x)$, i.e. $u'(\xi) = 0$. Then $u(x)$ is symmetric with respect to ξ.*

PROOF. Let $v(x) \equiv u(2\xi - x)$. Then $v(x)$ satisfies the same equation (2.1), and, moreover, $v(\xi) = u(\xi)$ and $v'(\xi) = u'(\xi) = 0$. By uniqueness of initial value problems, $u(x) \equiv v(x)$, and the proof follows. □

LEMMA 2.2. *Solution of* (2.1) *cannot have points of positive minimum, and of negative maximum.*

PROOF. Let us rule out the case of positive minimums, with the other case being similar. Assume on the contrary that there are points of positive minimums, and let ξ be the largest such point. Since $u(\xi) > 0$ and $u(1) = 0$, we can find a point $\eta \in (\xi, 1)$, so that $u(\xi) = u(\eta)$. Observe that $u'(\eta) < 0$. Indeed, if we had $u'(\eta) = 0$, then by the preceding lemma,

η would have to be a point of minimum, contradicting the maximality of ξ. We know that the energy $E(x) = \frac{1}{2}u'(x)^2 + \lambda F(u(x))$ is constant, but by above, $E(\eta) > E(\xi)$, a contradiction. ☐

We now consider positive solutions of (2.1). It follows from the lemmas above, that any positive solution is an even function, with $u'(x) > 0$ on $(-1, 0)$, and $u'(x) < 0$ on $(0, 1)$. (Of course, by the classical theorem of B. Gidas, W.-M. Ni and L. Nirenberg [15] this result holds for balls in $\mathbb{R}^n$ for any $n \geqslant 1$.) Hence $\alpha \equiv u(0)$ is the maximal value of solution. We show next that it is impossible for two solutions of (2.1) to share the same α.

LEMMA 2.3. *The value of $u(0) = \alpha$ uniquely identifies the solution pair $(\lambda, u(x))$ (i.e. there is at most one λ, with at most one solution $u(x)$, so that $u(0) = \alpha$).*

PROOF. Assume on the contrary that we have two solution pairs $(\lambda, u(x))$ and $(\mu, v(x))$, with $u(0) = v(0) = \alpha$. Clearly, $\lambda \neq \mu$, since otherwise we have a contradiction with uniqueness of initial value problems. (Recall that $u'(0) = v'(0) = 0$.) Then $u(\frac{1}{\sqrt{\lambda}}x)$ and $v(\frac{1}{\sqrt{\mu}}x)$ are both solutions of the same initial value problem

$$u'' + f(u) = 0, \qquad u(0) = \alpha, \quad u'(0) = 0,$$

and hence $u(\frac{1}{\sqrt{\lambda}}x) \equiv v(\frac{1}{\sqrt{\mu}}x)$, but that is impossible, since the first function vanishes at $x = \sqrt{\lambda}$, while the second one at $x = \sqrt{\mu}$. ☐

Bifurcation theory approach revolves around the study of the linearized equation for (2.1)

$$w'' + \lambda f'\bigl(u(x)\bigr)w = 0 \quad \text{for } -1 < x < 1, \qquad w(-1) = w(1) = 0, \tag{2.2}$$

where $u(x)$ is a solution of (2.1). If this problem has a nontrivial solution, we call $u(x)$ a singular solution of (2.1). We say that the solution $u(x)$ is nonsingular, if $w(x) \equiv 0$ is the only solution of (2.2). The following lemma is easy to prove in the autonomous case.

LEMMA 2.4. *Let $u(x)$ be a positive solution of* (2.1), *with*

$$u'(1) < 0. \tag{2.3}$$

If the problem (2.2) *admits a nontrivial solution, then it does not change sign, i.e. we may assume that $w(x) > 0$ on* $(-1, 1)$.

PROOF. The function $u'(x)$ also satisfies the linear equation (2.2). By the condition (2.3), $u'(x)$ is not a multiple of $w(x)$. Hence its roots are interlaced with those of $w(x)$. If $w(x)$ had a root ξ inside say $(-1, 0)$, then $u'(x)$ would have to vanish on $(-1, \xi)$, which is impossible by the remarks following Lemma 2.2. ☐

Condition (2.3) will hold for any positive solution, provided that

$$f(0) \geqslant 0, \tag{2.4}$$

see, e.g., p. 107 in M. Renardi and R.C. Rogers [52]. If $f(0) < 0$ it is possible to have $u'(1) = 0$. We shall encounter such a situation later, in connection with symmetry-breaking bifurcation. What we see here is a manifestation of the "divide" between the problems when (2.4) holds, and the case of $f(0) < 0$.

LEMMA 2.5. *If problem* (2.2) *admits nontrivial solutions, then the solution set is one dimensional. If moreover $u(x)$ is a positive solution, satisfying* (2.3), *then $w(x)$ is an even function.*

PROOF. By uniqueness of initial value problem the value of $w'(1)$ uniquely determines $w(x)$, and hence the null space is one dimensional. Turning to the second claim, if $u(x)$ is positive, then it is even. Hence $w(-x)$ also solves (2.2). Since the null space is one dimensional, $w(-x) = cw(x)$ for some constant c. Evaluating this relation at $x = 0$, we conclude that $c = 1$ (since $w(0) > 0$ by the previous lemma), which is the desired symmetry. ☐

The following lemma gives a simple condition for positive solutions of (2.1) to be nonsingular.

LEMMA 2.6. *Assume that either*

$$f'(u) > \frac{f(u)}{u} \quad \textit{for all } u > 0,$$

or the opposite inequality holds. Then the linearized problem (2.2) *has only the trivial solution.*

PROOF. If we rewrite Eq. (2.1) in the form

$$u'' + \lambda \frac{f(u)}{u} u = 0,$$

and use the Sturm comparison theorem, we conclude that the positive solution $u(x)$ oscillates faster than $w(x)$, and hence it must vanish on $(-1, 1)$, which is impossible. ☐

Another very simple condition is the following.

LEMMA 2.7. *Assume that*

$$f'(u) < 0 \quad \textit{for all } u > 0.$$

Then the linearized problem (2.2) *has only the trivial solution.*

PROOF. Multiplying Eq. (2.2) by w, and integrating, we conclude that problem (2.2) can have only the trivial solution. □

We shall need the following lemma, which "connects" the solutions of (2.1) and (2.2). It will allow us to verify the crucial condition of the Crandall–Rabinowitz Theorem 1.1 for both positive and sign changing solutions.

LEMMA 2.8. *If problem* (2.2) *admits a nontrivial solution*, *then*

$$\int_{-1}^{1} f(u)w \, dx = \frac{1}{\lambda} u'(1)w'(1). \tag{2.5}$$

PROOF. The quantity $u''(x)w(x) - u'(x)w'(x)$ is a constant, and hence

$$u''(x)w(x) - u'(x)w'(x) = -u'(1)w'(1).$$

Integrating over $(-1, 1)$ (by parts), we conclude the lemma. □

Very often one is looking for positive solutions of (2.1). A possible reason for this emphasis, is that only positive solutions have a chance to be stable, a property significant for applications. Let us recall the notion of stability. For any solution $u(x)$ of (2.1) let $(\mu, w(x))$ denote the principal eigenpair of the corresponding linearized equation, i.e. $w(x) > 0$ satisfies

$$w'' + \lambda f'(u)w + \mu w = 0 \quad \text{for } -1 < x < 1, \qquad w(-1) = w(1) = 0. \tag{2.6}$$

The solution $u(x)$ of (2.1) is called unstable if $\mu < 0$, it is called stable if $\mu > 0$, and neutrally stable in case $M = 0$. (This is so-called linear stability. It means, roughly, that solutions of the corresponding heat equation, with the initial data near $u(x)$ will tend to $u(x)$, as $t \to \infty$, see the book by D. Henry [16].)

PROPOSITION 1. *Let* $u(x)$ *be a solution of* (2.1) *that changes sign on* $(-1, 1)$. *Then* $u(x)$ *is unstable.*

PROOF. Let $(\mu, w(x))$ denote the principal eigenpair of (2.6). Assume that on the contrary $\mu \geqslant 0$. Since $u(x)$ changes sign, we can find $-1 < x_1 < x_2 < 1$, such that $u'(x_1) = u'(x_2) = 0$ and say $u'(x) < 0$ on (x_1, x_2) (the other case is similar). Observe that $u''(x_1) < 0$ and $u''(x_2) > 0$ ($u'(x)$ satisfies a linear equation, it cannot vanish together with its derivative). Denoting $p(x) = u''(x)w(x) - u'(x)w'(x)$, we have

$$p'(x) = \mu u'(x)w(x) \leqslant 0 \quad \text{for } x \in (x_1, x_2).$$

We see that $p(x)$ is nonincreasing on (x_1, x_2). But, $p(x_1) = u''(x_1)w(x_1) < 0$ and $p(x_2) = u''(x_2)w(x_2) > 0$, a contradiction. □

This result was also proved by R. Schaaf [53]. A similar result for balls in $\mathbb{R}^n$ can be found in C.S. Lin and W.-M. Ni [42].

LEMMA 2.9. *Any two positive solutions of* (2.1) *do not intersect inside* $(-1,1)$ (*i.e. they are strictly ordered on* $(-1,1)$).

PROOF. Let $u(x)$ and $v(x)$ be two intersecting solutions. Since both of them are even functions, they intersect on the half-interval $(0,1)$ as well. Let $0<\xi<\eta<1$ be two consecutive intersection points. If $v(x)>u(x)$ on (ξ,η), then $|u'(\xi)|>|v'(\xi)|$, while $|u'(\eta)|<|v'(\eta)|$. The energy $E(x)=\frac{1}{2}u'(x)^2+\lambda F(u(x))$ is constant for any solution $u(x)$. But at ξ, $u(x)$ has higher energy than $v(x)$, and at η the order is reversed, a contradiction. □

The following result from P. Korman [21] gives a detailed description of the solution shape for large λ. (In [21] we proved this result for balls in $\mathbb{R}^n$.) If for some reason solutions cannot be of that shape, it follows that there are no positive solutions of (2.1) for large λ. Recall that root α of $f(u)$ is called *stable* if $f(\alpha)=0$ and $f'(\alpha)<0$.

THEOREM 2.1. *Let* $u(x,\lambda)$ *be a positive solution of* (2.1), *that exists for all* $\lambda>\bar{\lambda}$, *for some* $\bar{\lambda}>0$. *Assume that either* $\lim_{u\to\infty}\frac{f(u)}{u}=\infty$, *or there is* $u_0>0$, *so that* $f(u)\leqslant 0$ *for* $u\geqslant u_0$. *Then the interval* $(-1,1)$ *can be decomposed into a union of open intervals, whose total length* $=2$, *so that on each such subinterval* $u(x,\lambda)$ *tends to a stable root of* $f(u)$, *as* $\lambda\to\infty$.

EXAMPLE. Assume that $f(u)<0$ for $0<u<\bar{u}$, with some $\bar{u}>0$, $f(u)>0$ for $u>\bar{u}$, and $\lim_{u\to\infty}\frac{f(u)}{u}=\infty$, e.g., $f(u)=u^p-1$, with $p>1$. Then the problem

$$u''+\lambda\big(u^p-1\big)=0,\quad -1<x<1,\qquad u(-1)=u(1)=0$$

has no positive solution for λ large enough. Indeed, since $f(u)$ has no stable roots, solution cannot exhibit the behaviour described in the above theorem, and hence the solution cannot exist for all large λ.

The bifurcation approach applies also to the quasilinear problems of the type

$$\big(\varphi(u')\big)'+\lambda f(u)=0\quad\text{for } -1<x<1,\qquad u(-1)=u(1)=0. \tag{2.7}$$

The prominent example is that of p-Laplacian $\varphi(t)=t|t|^{p-2}$, with $p\geqslant 2$. Motivated by this example, we assume that $\varphi(t)\in C^2(\mathbb{R}\setminus\{0\})$ satisfies

$$\varphi'(t)>0\quad\text{for all } t\neq 0. \tag{2.8}$$

We consider weak solutions of (2.7), which are of class $C[-1,1]\cap C^1(-1,1)\cap C^2((-1,1)\setminus\{0\})$. (Since the problem is degenerate elliptic, the value of $u''(0)$ might be infinite.)

Assuming condition (2.8), we will show that the conclusions of Lemmas 2.1, 2.2, 2.4 and 2.5 hold verbatim for problem (2.7). In case $\varphi(t) = t|t|^{p-2}$, Lemmas 2.3 and 2.8 hold too. The proofs are basically the same, but there are some difficulties due to degeneracy. For example, in the proof of Lemma 2.1 it is not apparent that $u(x)$ is symmetric with respect to ξ in case $\xi = 0$. (If $\xi \neq 0$, the proof is as before.) We therefore combine the first two lemmas to assert that the solution has the same shape as before.

LEMMA 2.10. *Assuming condition* (2.8), *any positive solution of* (2.7) *is an even function*, *with* $u'(x) > 0$ *on* $(-1, 0)$, *and* $u'(x) < 0$ *on* $(0, 1)$.

PROOF. We need to adjust the definition of energy. Define $\Phi(z) = \int_0^z t\varphi'(t)\,\mathrm{d}t$. Then the energy $\Phi(u'(x)) + \lambda F(u(x))$ is constant (in case of p-Laplacian, $\frac{p}{p-1}|u'(x)|^p + \lambda F(u(x)) = constant$). Using the energy, we conclude as before that $u(x)$ cannot have any points of minimum inside the interval $(-1, 1)$. Also, since the energy is constant, it follows that

$$u'(-1) = -u'(1). \tag{2.9}$$

To prove the symmetry, we consider $v(x) \equiv u(-x)$. The function $v(x)$ satisfies the same problem (2.7). By (2.9) it has the same initial conditions at $x = 1$ as $u(x)$. Hence $u(x) \equiv v(x)$. And finally, observe that an even function with no interior minimums has the desired shape. □

It is easy to see that Lemma 2.3 holds in case of p-Laplacian. Since we need homogeneity for rescaling, we cannot assert it for the general problem (2.7). Next we consider the linearized problem for (2.7)

$$\begin{aligned} &\big(\varphi'(u')w'\big)' + \lambda f'(u)w = 0 \quad \text{for } -1 < x < 1, \\ &w(-1) = w(1) = 0. \end{aligned} \tag{2.10}$$

LEMMA 2.11. *Assume that condition* (2.3) *holds. If problem* (2.10) *admits a nontrivial solution*, *then it does not change sign*, *i.e. we may assume that* $w(x) > 0$ *on* $(-1, 1)$. *Moreover*, *in the case of p-Laplacian*, *the following generalization of the formula* (2.5) *holds*:

$$\int_{-1}^{1} f(u)w\,\mathrm{d}x = \frac{2}{p\lambda}\varphi'\big(u'(1)\big)u'(1)w'(1). \tag{2.11}$$

PROOF. The proof of the first statement is exactly the same as before. Turning to the other one, we differentiate Eq. (2.10)

$$\big(\varphi'(u')u_x'\big)' + \lambda f'(u)u_x = 0. \tag{2.12}$$

Combining problems (2.10) and (2.12), we have

$$\varphi'\big(u'(x)\big)\big(u''(x)w(x) - u'(x)w'(x)\big) = -\varphi'\big(u'(1)\big)u'(1)w'(1). \tag{2.13}$$

We now integrate over $(-1,1)$. In the case of p-Laplacian, $\varphi(t)=t|t|^{p-2}$, $\varphi'(t)=(p-1)|t|^{p-2}$, and $t\varphi'(t)=(p-1)\varphi(t)$. Hence the second term on the left is equal to $-(p-1)\int_{-1}^{1}\varphi(u'(x))w'(x)\,\mathrm{d}x$. Integrating this term by parts, we can combine it with the first term of the resulting equation. Finally, we observe that $\varphi'(u'(x))u''(x)=-\lambda f(u(x))$. □

2.2. *Convex nonlinearities*

For convex nonlinearities one can give an exhaustive description of the bifurcation diagrams for problem (2.1), since we are able to show that the solution curve cannot turn more than once. Namely, we assume that $f(u)\in C^2(\overline{\mathbb{R}}_+)$ satisfies

$$f(0)>0,\quad \text{and}\quad f(u)>0\quad \text{for } u>0, \tag{2.14}$$

$$f''(u)>0\quad \text{for } u>0, \tag{2.15}$$

$$f(u)\geqslant au-b\quad \text{for } u>0 \text{ and some constants } a>0 \text{ and } b\geqslant 0. \tag{2.16}$$

THEOREM 2.2. *Problem* (2.1), *under the above conditions*, *has at most two positive solutions for any* λ. *Moreover, all positive solutions lie on a unique curve in the* $(\lambda,u(0))$ *plane. This curve begins at the point* $(\lambda=0,u(0)=0)$, *and either it tends to infinity at some* $\lambda_0>0$, *or else it bends back at some* $\lambda_0>0$, *and then continues without any more turns*, *and tends to infinity at some* $\bar{\lambda}$, $0\leqslant\bar{\lambda}<\lambda_0$.

PROOF. When $\lambda=0$ we have a trivial solution $u=0$. It follows by the implicit function theorem that for small $\lambda>0$ there is a continuous in λ curve of solutions, passing through $(0,0)$. We claim that this solution curve cannot be continued indefinitely for all $\lambda>0$. Assume on the contrary that solutions can be continued as $\lambda\to\infty$. Write problem (2.1) in the corresponding integral form,

$$u(x)=\lambda\int_{-1}^{1}G(x,\xi)f\bigl(u(\xi)\bigr)\,\mathrm{d}\xi, \tag{2.17}$$

where $G(x,\xi)$ is the corresponding Green's function. It is well known that $G(x,\xi)>0$ for all $0<x,\xi<1$. Since by our assumptions $f(u)$ is bounded from below by a positive constant, it follows that $u(x)$ will become uniformly large, as $\lambda\to\infty$. Let $\phi_1(x)$ be the principal eigenvalue of $-u''$ on the interval $(-1,1)$ subject to zero boundary conditions, and λ_1 the corresponding principal eigenvalue (here $\phi_1(x)=\cos\frac{\pi}{2}x$, and $\lambda_1=\frac{\pi^2}{4}$). Multiplying Eq. (2.1) by $\phi_1(x)$ and integrating, we have

$$\begin{aligned}\lambda_1\int_{-1}^{1}u\phi_1\,\mathrm{d}x&=-\int_{-1}^{1}u''\phi_1\,\mathrm{d}x=\lambda\int_{-1}^{1}f(u)\phi_1\,\mathrm{d}x\\&\geqslant\lambda a\int_{-1}^{1}u\phi_1\,\mathrm{d}x-\lambda b\int_{-1}^{1}\phi_1\,\mathrm{d}x.\end{aligned}$$

But this leads to a contradiction, as $\lambda \to \infty$, since $\int_{-1}^{1} u\phi_1 \, \mathrm{d}x \to \infty$. ($u(x)$ is a convex function, tending to infinity.)

Let λ_0 denote the supremum of λ, for which the solution curve continues to the right. It is possible that solutions become unbounded as $\lambda \to \lambda_0$ (this is one of the possibilities discussed in the statement of the theorem). So assume that the solutions stay bounded, as $\lambda \to \lambda_0$. Passing to the limit in the integral form of the equation, see (2.17), we conclude the existence of a bounded solution $u_0(x)$, which our solution curve enters at $\lambda = \lambda_0$. Clearly the pair $(\lambda_0, u_0(x))$ is a singular solution of (2.1) (since it cannot be continued to the right in λ). We show next that the Crandall–Rabinowitz Theorem 1.2 applies at $(\lambda_0, u_0(x))$.

We begin by recasting the equation in operator form $F(\lambda, u) = 0$, where the map $F(\lambda, u) : \mathbb{R}_+ \times C^2(-1, 1) \to C(-1, 1)$ is defined by $F(\lambda, u) = u''(x) + \lambda f(u(x))$. Observe that $F_u(\lambda, u)w$ is given by the left-hand side of the linearized equation (2.2). Since the point (λ_0, u_0) is singular, it follows that the linearized equation (2.2) has a nontrivial solution $w(x)$, which is positive by Lemma 2.4. By Lemma 2.5 it follows that the null-space $N(F_u(\lambda_0, u_0)) = \mathit{span}\{w(x)\}$ is one-dimensional, and then $\mathit{codim} R(F_u(\lambda_0, u_0)) = 1$, since $F_u(\lambda_0, u_0)$ is a Fredholm operator of index zero. To apply the Crandall–Rabinowitz Theorem 1.2, it remains to show that $F_\lambda(\lambda_0, u_0) \notin R(F_u(\lambda_0, u_0))$. Assuming the contrary would imply existence of a nontrivial $v(x)$, a solution of

$$v'' + \lambda_0 f_u(\lambda_0, u_0)v = f(\lambda_0, u_0) \quad \text{for } x \in (-1, 1), \qquad v(-1) = v(1) = 0.$$

By the Fredholm alternative (or just multiplying this equation by w, Eq. (2.2) by v, subtracting and integrating)

$$\int_{-1}^{1} f(\lambda_0, u_0) w(x) \, \mathrm{d}x = 0,$$

which contradicts Lemma 2.8. (Since $f(0) > 0$, we have $u'(1) < 0$, and also $w'(1) < 0$ by uniqueness of initial value problems. Hence by Lemma 2.8, the above integral is positive.) Hence the Crandall–Rabinowitz Theorem 1.2 applies at $(\lambda_0, u_0(x))$.

Next we compute the direction of bifurcation at the point $(\lambda_0, u_0(x))$. According to the Crandall–Rabinowitz Theorem 1.2, the solution set near the point $(\lambda_0, u_0(x))$ is a curve $\lambda = \lambda(s)$, $u = u(s)$, with $\lambda(0) = \lambda_0$ and $u(0) = u_0(x)$. Observe that $\lambda'(0) = 0$, and $u_s(0) = w(x)$, according to the Crandall–Rabinowitz theorem. If we can show that $\lambda''(0) < 0$, it would follow that the solution curve turns to the left at $(\lambda_0, u_0(x))$, since the function $\lambda(s)$ has a maximum at $s = 0$. To express $\lambda''(s)$, we differentiate Eq. (2.1) twice in s, obtaining

$$u_{ss}'' + \lambda f_u u_{ss} + \lambda f_{uu} u_s^2 + 2\lambda' f_u u_s + \lambda'' f = 0, \qquad u_{ss}(-1) = u_{ss}(1) = 0.$$

Letting $s = 0$, we have by the above remarks

$$u_{ss}'' + \lambda_0 f_u u_{ss} + \lambda_0 f_{uu} w^2 + \lambda''(0) f = 0, \qquad u_{ss}(-1) = u_{ss}(1) = 0. \tag{2.18}$$

Multiplying this equation by w, Eq. (2.2) by u_{ss}, subtracting and integrating

$$\lambda''(0) = -\lambda_0 \frac{\int_{-1}^{1} f_{uu}(\lambda_0, u_0(x)) w^3(x)\, \mathrm{d}x}{\int_{-1}^{1} f(\lambda_0, u_0) w(x)\, \mathrm{d}x} < 0, \tag{2.19}$$

with the last inequality due to convexity of $f(u)$ and Lemma 2.8.

The above analysis is valid not only at the point $(\lambda_0, u_0(x))$, but also at any other critical point. Hence, locally near any critical point, the solution set consists of a parabola-like curve, facing to the left in the $(\lambda, u(0))$ plane. Hence, after bending back at the point $(\lambda_0, u_0(x))$, our solution curve continues for decreasing λ, without ever encountering critical points. (At any critical point, we could not possibly have a parabola-like curve, described above, since our curve has arrived from the right.) Hence, the solution curve continues globally, without any turns, and it has to go to infinity at some $\bar{\lambda} \geqslant 0$. We then have one of the solution curves, described in the theorem, and the maximum value of solutions on this curve, $u(0)$ varies from zero to infinity. Hence all possible maximum values are "taken", and so by Lemma 2.3 this curve exhausts the solution set. □

REMARKS.

1. If, moreover, we have

$$\lim_{u \to \infty} \frac{f(u)}{u} = \infty, \tag{2.20}$$

   then the curve cannot go to infinity at a finite λ. Hence, it will bend back at some λ_0, and go to infinity at $\lambda = 0$. Indeed, since $f(u) > 0$, the solutions $u(x, \lambda)$ are concave in x. So that if $u(x, \lambda)$ gets large near some $\lambda = \lambda_1$, it would have to get uniformly large on some interval, say on $(-1/2, 1/2)$. Writing our equation in the form $u'' + \lambda \frac{f(u)}{u} u = 0$, and using the Sturm comparison theorem, we conclude that the positive solution $u(x)$ has to vanish on $(-1/2, 1/2)$, which is impossible.
2. One can show that the solutions on the lower branch are increasing in λ, for all $x \in (-1, 1)$, see P. Korman and T. Ouyang, [34,35]. (On the upper branch this is no longer true, but the maximal value $u(0)$ is increasing as we trace the branch, i.e. it is decreasing in λ.)
3. Our assumptions did not require for $f(u)$ to be increasing.
4. All three possibilities, mentioned in the theorem, can actually occur, see Fig. 1 for the results of numerical computations. Observe that in the second and third cases $f(u)$ is asymptotically linear, and bifurcation from infinity happens.

2.3. *S-shaped solution curves*

We saw in the previous section that solution curves are relatively simple for convex $f(u)$. If $f(u)$ changes concavity, then the solution curve may admit more than one turn. One

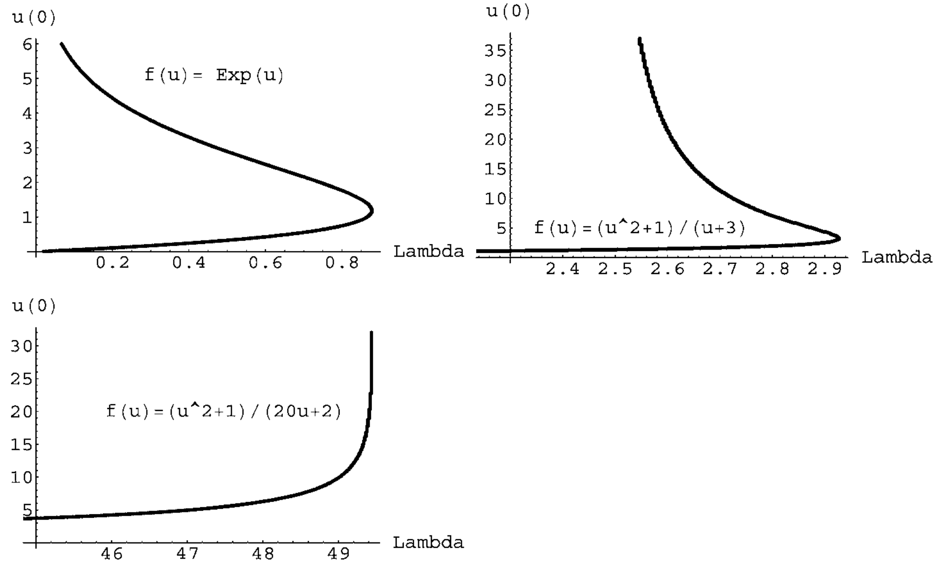


Fig. 1. Three types of solution curves for convex $f(u)$.

prominent nonlinearity, with change in concavity, is connected to combustion theory, see the nice book of J. Bebernes and D. Eberly [7]. Namely, we consider the problem

$$u'' + \lambda e^{\frac{au}{u+a}} = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0, \tag{2.21}$$

where a is a constant (this problem is referred to as "the perturbed Gelfand problem" in [7]). In case $a \leqslant 4$, the problem is easy. In that case $uf'(u) - f(u) < 0$ for all $u > 0$, and hence all positive solutions are nonsingular. This means that the solution curve is monotone, i.e. it continues for all $\lambda > 0$ without any turns. S.-H. Wang [59] has proved existence of a constant a_0, so that for $a > a_0$ the solution curve of (2.21) is exactly S-shaped, i.e. it starts at $\lambda = 0$, $u = 0$, it makes exactly two turns, and then it continues for all $\lambda > 0$ without any more turns. S.-H. Wang [59] gave an approximation of the constant $a_0 \simeq 4.4967$. That paper, as well as all previous ones, used a time map approach. P. Korman and Y. Li [28] have applied the bifurcation approach to the problem to show the exactness of the S-shaped curves, and they also improved the value of the constant to $a_0 \simeq 4.35$, i.e. for $a > a_0$ the solution sets are S-shaped curves. S.-H. Wang [59] has conjectured existence of critical number $\bar{a}$, so that for $a \leqslant \bar{a}$ the solution curve is monotone, while for $a > \bar{a}$ the solution curve is exactly S-shaped. Recently, P. Korman, Y. Li and T. Ouyang [33] has given a computer assisted proof of the S.-H. Wang's conjecture. Numerical calculations show that $\bar{a} \simeq 4.07$.

We are going to discuss the S-shaped curves, mostly following P. Korman and Y. Li [29]. However, in that paper time maps were still used at one point. Subsequently, in P. Korman and J. Shi [40] an argument not using time maps was given. Next we present this result, dealing with instability of solutions (it also turned out to be of independent interest), after we recall the notion of stability. For any solution $u(x)$ of (2.1) let $(\mu, w(x))$ denote the principal eigenpair of the corresponding linearized equation, i.e. $w(x) > 0$ satisfies

$$w'' + \lambda f'(u)w + \mu w = 0 \quad \text{for } -1 < x < 1, \qquad w(-1) = w(1) = 0. \tag{2.22}$$

The solution $u(x)$ of (2.1) is called unstable if $\mu < 0$, otherwise it is stable.

Let $F(u) = \int_0^u f(t)\,\mathrm{d}t$, $h(u) = 2F(u) - uf(u)$. The instability result from P. Korman and J. Shi [40] is

THEOREM 2.3. *Assume that* $f \in C^1[0, \infty)$, $f(0) > 0$, *and for some* $\alpha > \beta > 0$ *we have*:

$$h'(u) \geqslant 0 \quad \textit{for } 0 < u < \beta, \qquad h'(u) \leqslant 0 \quad \textit{for } \beta < u < \alpha, \tag{2.23}$$

$$h(\alpha) \leqslant 0. \tag{2.24}$$

Then the solution of (2.1) *with* $u(0) = \alpha$ *is unstable*, *if it exists*.

PROOF. We have $h(0) = 0$, $h'(u) = f(u) - uf'(u)$, $h'(0) = f(0) > 0$. It follows from our conditions that $h(u)$ is unimodular on $[0, \alpha]$, and it takes its positive maximum at $u = \beta$. Define $x_0 \in (0, 1)$ by $u(x_0) = \beta$. We then conclude

$$\begin{aligned} &f\big(u(x)\big) - u(x) f'\big(u(x)\big) \leqslant 0 \quad \text{on } (0, x_0), \\ &f\big(u(x)\big) - u(x) f'\big(u(x)\big) \geqslant 0 \quad \text{on } (x_0, 1). \end{aligned} \tag{2.25}$$

We also remark that by the condition (2.24),

$$\int_0^1 \big[f(u) - uf'(u)\big]u'(x)\,\mathrm{d}x = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}x} h\big(u(x)\big)\,\mathrm{d}x = -h(\alpha) \geqslant 0. \tag{2.26}$$

Assume now that $u(x)$ is stable, i.e. $\mu \geqslant 0$ in (2.22). Without loss of generality, we assume that $w > 0$ in $(-1, 1)$. By the maximum principle, $u'(1) < 0$, so near $x = 1$ we have $-u'(x) > w(x)$. Since $-u'(0) = 0$, while $w(0) > 0$, the functions $w(x)$ and $-u'(x)$ change their order at least once on $(0, 1)$. We claim that the functions $w(x)$ and $-u'(x)$ change their order exactly once on $(0, 1)$. (We ignore the points where these functions merely “touch”.) Observe that $-u'(x)$ satisfies

$$(-u')'' + \lambda f'(u)(-u') = 0 \quad \text{on } (0, 1), \tag{2.27}$$

while $w(x)$ (and any of its positive multiples) is a supersolution of the same equation. Let $x_3 \in (0, 1)$ be the largest point where $w(x)$ and $-u'(x)$ change the order. Assuming the claim to be false, let x_2, with $0 < x_2 < x_3$, be the next point where the order changes. We have $w > -u'$ on (x_2, x_3), and the opposite inequality to the left of x_2. Since $w(0) > -u'(0)$, there is another point $x_1 < x_2$, where the order is changed. We can now find a constant $\gamma > 1$, and a point $x_0 \in (x_1, x_2)$ so that a $\gamma w(x)$, a supersolution of (2.27), touches at x_0 from above a solution $-u'(x)$ of the same equation, a contradiction.

Since the point of changing of order is unique, by scaling of $w(x)$ we can achieve

$$\begin{aligned} -u'(x) &\leqslant w(x) \quad \text{on } (0, x_0), \\ -u'(x) &\geqslant w(x) \quad \text{on } (x_0, 1). \end{aligned} \tag{2.28}$$

Using (2.25), (2.28), and also (2.26), we have

$$\int_0^1 \bigl[f(u) - uf'(u)\bigr] w(x)\,\mathrm{d}x < \int_0^1 \bigl[f(u) - uf'(u)\bigr]\bigl(-u'(x)\bigr)\,\mathrm{d}x \leqslant 0, \tag{2.29}$$

since the integrand on the left is pointwise smaller than the one on the right. On the other hand, multiplying Eq. (2.22) by u, Eq. (2.1) by w, subtracting and integrating over (0, 1), we have

$$\int_0^1 \bigl[f(u) - uf'(u)\bigr] w(x)\,\mathrm{d}x = \frac{\mu}{\lambda}\int_0^1 uw\,\mathrm{d}x \geqslant 0,$$

which contradicts (2.29). So $\mu < 0$. □

We will consider a class of nonlinearities, including $f(u) = \mathrm{e}^{\frac{au}{u+a}}$, so let us list our assumptions. We assume that $f(u) \in C^2[0, \bar{u}]$ for some $0 < \bar{u} \leqslant \infty$, and that it satisfies

$$f(u) > 0 \quad \text{for all } 0 \leqslant u < \bar{u}. \tag{2.30}$$

We assume $f(u)$ to be convex-concave, i.e. there an $\alpha \in (0, \bar{u})$, such that

$$f''(u) > 0 \quad \text{for } u \in (0, \alpha), \qquad f''(u) < 0 \quad \text{for } u \in (\alpha, \bar{u}). \tag{2.31}$$

We define a function $I(u) = f^2(u) - 2F(u)f'(u)$, where as before $F(u) = \int_0^u f(t)\,\mathrm{d}t$. Assume there is a $\beta > \alpha$, such that

$$I(\beta) = f^2(\beta) - 2F(\beta)f'(\beta) \geqslant 0. \tag{2.32}$$

The following lemma has originated from P. Korman, Y. Li and T. Ouyang [30].

LEMMA 2.12. *Assume that $f(u)$ satisfies conditions* (2.30)–(2.32). *Let (λ, u) be any critical point of* (2.1), *such that $u(0) \geqslant \beta$, and let $w(x)$ be the corresponding solution of the linearized problem* (2.2). *Then*

$$\int_0^1 f''\bigl(u(x)\bigr)u_x(x)w^2(x)\,\mathrm{d}x > 0. \tag{2.33}$$

PROOF. We shall derive a convenient expression for the integral in (2.33). Differentiate (2.2)

$$w_x'' + \lambda f'(u)w_x + \lambda f''(u)u_x w = 0. \tag{2.34}$$

Multiplying Eq. (2.34) by w, Eq. (2.2) by w_x, integrating and subtracting, we express

$$\lambda \int_0^1 f''(u) u_x w^2 \, \mathrm{d}x = w'^2(1) - \lambda w^2(0) f'\bigl(u(0)\bigr). \tag{2.35}$$

By differentiation, we verify that $u''(x)w(x) - u'(x)w'(x)$ is constant for all x, and hence

$$u''(x)w(x) - u'(x)w'(x) = -\lambda w(0) f\bigl(u(0)\bigr) \quad \text{for all } x \in [-1, 1]. \tag{2.36}$$

Evaluating this expression at $x = 1$, we obtain

$$w'(1) = \frac{\lambda w(0) f(u(0))}{u'(1)}. \tag{2.37}$$

Multiplying (2.1) by u', and integrating over $(0, 1)$, we have

$$u'^2(1) = 2\lambda F\bigl(u(0)\bigr). \tag{2.38}$$

Using (2.38) and (2.37) in (2.35), we finally express

$$\lambda \int_0^1 f''(u) u_x w^2 \, \mathrm{d}x = \frac{w^2(0)}{2F(\rho)} I(\rho), \tag{2.39}$$

where we denote $\rho = u(0)$. By our assumption, $I(\beta) \geqslant 0$. Since

$$I'(\rho) = -2F(\rho) f''(\rho) > 0 \quad \text{for } \rho \geqslant \beta,$$

we conclude that $I(\rho) > I(\beta) \geqslant 0$, and the lemma follows. ☐

The following lemma contains the crucial trick, which has originated from P. Korman, Y. Li and T. Ouyang [30]. It says that for convex-concave problems only turns to the right are possible in the (λ, α) plane, once the maximum value of the solution, $u(0)$, has reached a certain level.

LEMMA 2.13. *In the conditions of the preceding Lemma* 2.12, *assume again that* $u(0) \geqslant \beta$, *and* $w(x)$ *the corresponding solution of the linearized problem* (2.2). *Then*

$$\int_0^1 f''\bigl(u(x)\bigr) w^3(x) \, \mathrm{d}x < 0. \tag{2.40}$$

PROOF. Let $(\lambda, u(x))$ be a critical point of (2.1). Since $u(0) \geqslant \beta > \alpha$, it follows that the function $f''(u(x))$ changes sign exactly once on $(0, 1)$, say at x_0. Then we have

$$f''\bigl(u(x)\bigr) < 0 \quad \text{for } x \in (0, x_0), \qquad f''\bigl(u(x)\bigr) > 0 \quad \text{for } x \in (x_0, 1). \tag{2.41}$$

We have proved in Theorem 2.3 that the functions $w(x)$ and $-u'(x)$ intersect exactly once on $(0,1)$. By scaling $w(x)$ we may assume that they intersect at x_0. ($w(x)$ is a solution of a linear problem, and hence it is defined up to a constant multiple.) In view of Lemma 2.12, we then have

$$\int_0^1 f''\big(u(x)\big)w^3(x)\,\mathrm{d}x < \int_0^1 f''\big(u(x)\big)w^2(-u_x)\,\mathrm{d}x < 0,$$

since by (2.41) the integrand on the right is pointwise greater than the one the left. □

The same approach can be used to prove the following more general theorem, which was implicit in P. Korman, Y. Li and T. Ouyang [30], see also T. Ouyang and J. Shi [50].

THEOREM 2.4 [30].
(i) *Assume that* $f(0) \geqslant 0$, $f''(u) < 0$ *for* $0 < u < u_0$, $f''(u) > 0$ *for* $u > u_0$. *Then only turns to the left are possible on the solution curve.*

(ii) *Assume that* $f(0) \leqslant 0$, $f''(u) > 0$ *for* $0 < u < u_0$, $f''(u) < 0$ *for* $u > u_0$. *Then only turns to the right are possible on the solution curve.*

(Of course, in both cases we conclude existence of at most two positive solutions, with the maximum values lying in the first positive hump of $f(u)$.)

We are now ready for the main result of this section, see P. Korman and Y. Li [29].

THEOREM 2.5. *Assume that* $f(u)$ *satisfies conditions* (2.30) *and* (2.31), *and moreover,*

$$\lim_{u\to\infty} \frac{f(u)}{u} = 0. \tag{2.42}$$

With $h(u) \equiv 2F(u) - uf(u)$, *assume that*

$$h(\alpha) < 0. \tag{2.43}$$

Then the solution set of (2.1) *consists of one curve, which is exactly S-shaped, i.e. it starts at* $\lambda = 0$, $u = 0$, *it makes exactly two turns, and then it continues for all* $\lambda < \infty$, *without any more turns.*

PROOF. By the implicit function theorem there is a curve of positive solutions of (2.1), starting at $\lambda = 0$, $u = 0$. As in Theorem 2.2, this curve continues for increasing λ, until a possible singular solution (λ_0, u_0) is reached, at which point the Crandall–Rabinowitz Theorem 1.2 applies. By (2.19) it follows that only turns to the left are possible if $u(0) < \alpha$, since $f(u)$ is convex for $u \in (0,\alpha)$. Until the first critical point (λ_0, u_0) is reached, the solutions are stable. Indeed, the solution curve starts at $(\lambda = 0, u = 0)$, which is a stable solution (the principal eigenvalue of the corresponding linearized problem $= \frac{\pi^2}{2}$), while any change of stability requires a passage through a singular point. By Theorem 2.3 when

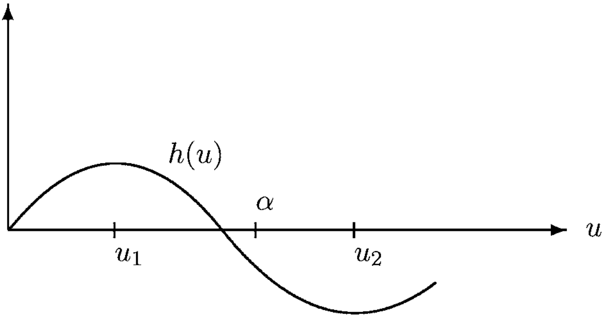


Fig. 2. The function $h(u)$.

$u(0) = \alpha$ the solution is unstable. Hence a singular solution was reached before that, and since only turns to the left are possible when $u(x) < \alpha$, it follows that *exactly one* turn has occurred, and at $u(0) = \alpha$ the solution curve travels to the left.

We now show that the solution curve keeps traveling to the left, until $u(0)$ increases to the level when only turns to the right are possible. For that we take a close look at the function $h(u) = 2F(u) - uf(u)$. Since

$$h'(u) = f(u) - uf'(u), \qquad h''(u) = -uf''(u),$$

it follows that the function $h'(u)$ is decreasing on $(0, \alpha)$ and increasing on (α, ∞). We have $h'(0) = f(0) > 0$, and so $h'(u)$ can have at most two roots. We claim that it has exactly two roots, u_1 and u_2 with $h'(u)$ being positive on $(0, u_1) \cup (u_2, \infty)$, and negative on (u_1, u_2). Indeed, existence of the first root is clear, since $h(0) = 0$ and $h(\alpha) < 0$. As for the second root u_2, if it did not exist, we would have

$$uf'(u) > f(u) \quad \text{for all } u > \alpha. \tag{2.44}$$

Integrating (2.44),

$$f(u) > \frac{f(\alpha)}{\alpha} u \quad \text{for all } u > \alpha,$$

contradicting the assumption (2.42). So that the function $h(u)$ starts with $h(0) = 0$, it is increasing on $(0, u_1)$, decreasing on (u_1, u_2), with absolute minimum at u_2, and then it increases on the interval (u_2, ∞) (see Fig. 2). By Theorem 2.3 the solution curve keeps traveling to the left, while $u(0) \in (\alpha, u_2)$.

We claim that for $u(0) > u_2$ Lemma 2.13 applies. For that we need to check that for $\beta = u_2$ the condition (2.32) holds. Indeed, since $h(u_2) < 0$, we have $f(u_2)u_2 > 2F(u_2)$. Hence

$$I(u_2) = f^2(u_2) - 2F(u_2)f'(u_2) > f^2(u_2) - u_2 f(u_2) f'(u_2) = 0,$$

and the claim follows (observe that $f'(u_2) = \frac{f(u_2)}{u_2} > 0$). By Lemma 2.13, only turns to the right are possible when $u(0) > u_2$.

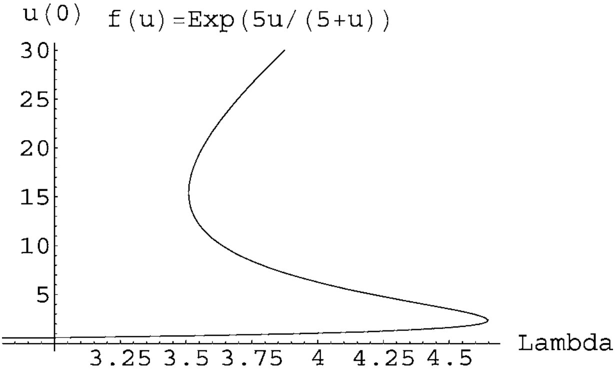


Fig. 3. An S-shaped solution curve.

Let us now put it all together. We have a curve of solutions, which starts at $(\lambda = 0, u = 0)$. As we travel on this curve, $u(0)$ is always increasing. By the time we reach $u(0) = \alpha$ level, the solution curve has made exactly one turn to the left. When $\alpha < u(0) < u_2$ the solution curve travels to the left. When $u(0) > u_2$, the solution curve cannot travel to the left indefinitely, since it is easy to see that solutions are bounded for bounded λ. Hence, the curve must turn to the right. Since for $u(0) > u_2$ only turns to the right are possible, exactly one such turn occurs. It follows that the solution curve is exactly S-shaped. $\square$

In Fig. 3 we give an example of an S-shaped solution curve. Notice that *Mathematica* has drawn the vertical axis around $\lambda = 3$. Also observe that an actual S-shaped solution curve is way different from what most people would draw by hand.

2.4. *Cubic-like nonlinearities*

We again consider the problem

$$u'' + \lambda f(u) = 0, \quad x \in (-1, 1), \qquad u(-1) = u(1) = 0, \tag{2.45}$$

where $f(u)$ behaves like a cubic with three distinct roots, with a model example $f(u) = (u - a)(u - b)(c - u)$. Namely, we assume that the function $f(u) \in C^2(\overline{\mathbb{R}}_+)$ has three nonnegative roots at $0 \leqslant a < b < c$, and

$$\begin{gathered} f(u) > 0 \quad \text{on } [0, a) \cup (b, c), \qquad f(u) < 0 \quad \text{on } (a, b) \cup (c, \infty), \\ \int_a^c f(u)\,\mathrm{d}u > 0. \end{gathered} \tag{2.46}$$

Moreover, we assume there is an $\alpha > b$, so that

$$f''(u) > 0 \quad \text{for } 0 \leqslant u < \alpha, \qquad f''(u) < 0 \quad \text{for } u > \alpha. \tag{2.47}$$

This problem was originally studied using time-maps, see J. Smoller and A. Wasserman [57] and S.-H. Wang [59,60]. In P. Korman, Y. Li and T. Ouyang [30] the bifurcation approach was applied. The case of $a = 0$ turned out to be easier for both time-maps and bifurcation approaches, while in case $a > 0$ some restriction on a (a bound from above) was necessary for both approaches.

We shall do the case $a = 0$ first, after two simple lemmas. (I.e. $f(u)$ is modeled on $f(u) = u(u-b)(c-u)$.) Let $\beta \in (b, c)$ be the unique point satisfying

$$f'(\beta) = \frac{f(\beta)}{\beta}. \tag{2.48}$$

(I.e. the point where the straight line through the origin is tangent to the graph of $y = f(u)$.) Clearly, $\beta > \alpha$. The following lemma shows that no turns of the solution curve are possible until the maximum value of the solution reaches a certain level.

LEMMA 2.14. *Assume that $f(u) \in C^2$ satisfies conditions* (2.46) *and* (2.47). *If $u(x)$ is a critical solution of* (2.45) *then*

$$u(0) > \beta. \tag{2.49}$$

PROOF. We claim that

$$f'(u) > \frac{f(u)}{u} \quad \text{for } 0 < u < \beta. \tag{2.50}$$

Indeed, denote $p(u) = uf'(u) - f(u)$. Then $p(0) = p(\beta) = 0$, and $p'(u) = uf''(u)$, which implies that $p(u)$ is increasing on $(0, \alpha)$ and decreasing on (α, β). Then (2.50) follows, and hence by Lemma 2.6 the linearized equation has only the trivial solution, in case $u(0) \leqslant \beta$. □

LEMMA 2.15. *Assume that $f(u) \in C^2$ satisfies conditions* (2.46) *and* (2.47), *with $a = 0$. Assume $u(x)$ is a critical solution of* (2.45), *and let $w(x)$ be solution of the corresponding linearized problem* (2.2). *Then*

$$\int_0^1 f''(u) w^3 \, \mathrm{d}x < 0. \tag{2.51}$$

PROOF. We begin by showing that

$$\int_0^1 f''(u) u_x^2 w \, \mathrm{d}x = 0. \tag{2.52}$$

We have (using Eqs. (2.45) and (2.2))

$$(u''w' - u'w'')' = \lambda f''(u) u_x^2 w.$$

Integrating over $(0,1)$, and using that $w''(1) = -\lambda f'(u(1))w(1) = 0$ and $u''(1) = -\lambda f(u(1)) = -\lambda f(0) = 0$, we conclude (2.52) (it is here that we use that $f(0) = 0$, i.e. $a = 0$).

We now proceed similarly to Lemma 2.13. Similarly to that lemma, we show that the functions $w(x)$ and $-u'(x)$ intersect exactly once on $(0,1)$. Observe that by Lemma 2.14, we have $u(0) > \beta > \alpha$ at any critical solution $u(x)$. Let $\xi \in (0,1)$ be the point where $u(\xi) = \alpha$. By scaling $w(x)$ we may assume that $w(x)$ and $-u'(x)$ intersect at ξ. Then on the interval $(0,\xi)$, where $f''(u(x))$ is negative, we have $u_x^2 < w^2$, while on the interval $(\xi,1)$, where $f''(u(x))$ is positive, we have $u_x^2 > w^2$. We then have, in view of (2.52),

$$\int_0^1 f''(u)w^3\,\mathrm{d}x < \int_0^1 f''(u)u_x^2 w\,\mathrm{d}x = 0,$$

since the integral on the left is pointwise smaller than the one on the right. □

The following theorem is from P. Korman, Y. Li and T. Ouyang [30].

THEOREM 2.6. *Assume that* $f(u) \in C^2$ *satisfies conditions* (2.46) *and* (2.47), *with* $a = 0$. *Then there is a critical* λ_0 *such that for* $\lambda < \lambda_0$ *problem* (2.45) *has no positive solutions, it has exactly one positive solution at* $\lambda = \lambda_0$, *and exactly two positive solutions for* $\lambda > \lambda_0$. *Moreover, all solutions lie on a single solution curve, which for* $\lambda > \lambda_0$ *has two branches* $0 < u^-(x,\lambda) < u^+(x,\lambda)$, *with* $u^+(x,\lambda)$ *strictly monotone increasing in* λ, *and* $\lim_{\lambda\to\infty} u^+(x,\lambda) = c$. *On the lower branch,* $u^-(0,\lambda)$ *is monotone decreasing,* $\lim_{\lambda\to\infty} u^-(x,\lambda) = 0$ *for all* $x \neq 0$, *while* $u^-(0,\lambda) > b$ *for all* λ. *We also have* $\lim_{\lambda\to\infty} u^-(0,\lambda) = \theta$, *where* θ *is defined by the relation* $\int_0^\theta f(u)\,\mathrm{d}u = 0$.

PROOF. If $\lambda c > \frac{\pi^2}{4}$, then existence of positive solutions follows by monotone iterations. Indeed, $\phi = c$ is a supersolution of (2.45), while $\psi = \epsilon \cos\frac{\pi}{2}x$ is a subsolution of the same problem, if ϵ is chosen sufficiently small ($\lambda = \frac{\pi^2}{4}$ and $\phi_1 = \cos\frac{\pi}{2}x$ give, of course, the principal eigenpair of the Laplacian on $(-1,1)$). We now continue the positive solution (any one) for decreasing λ. At regular points we use the implicit function theorem for continuation, while the singular point(s) will be discussed below. We cannot continue this curve indefinitely for decreasing λ, since it has no place to go. Indeed, solutions are bounded by c, and so the right hand side of Eq. (2.45) goes to zero, and hence $u(x) \to 0$ as $\lambda \to 0$. But that is impossible, since $f(u)$ is negative near $u = 0$, while at the point of maximum $u''(0) \leqslant 0$. Hence at some critical $\lambda = \lambda_0$ and $u = u_0$ the solution curve cannot be continued further for decreasing λ.

As before, we show that the Crandall–Rabinowitz Theorem 1.2 applies at (λ_0, u_0). According to Lemma 2.15 a turn to the right must occur at this, and any other critical point. Hence, exactly one turn happens, and the solution curve has exactly two branches.

The properties of the solution branches are easy to prove, see [30]. □

In Fig. 4 we give an example for the above theorem. Notice that *Mathematica* has chosen the point $(6,2)$ as the point where the axes intersect.

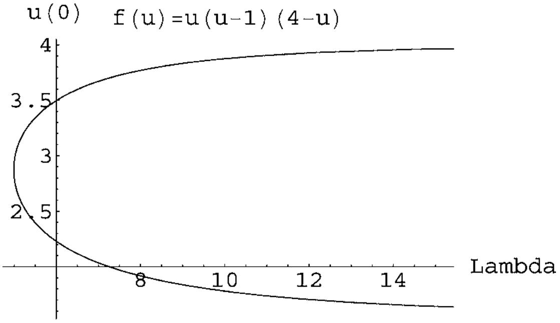


Fig. 4. A parabola-like solution curve, when $a = 0$.

Next we turn to the case when $a > 0$. I.e. we assume that $f(u)$ satisfies conditions (2.46) and (2.47), with the cubic $f(u) = (u - a)(u - b)(c - u)$ being our model example. Similarly to the above, we denote by $\beta \in (b, c)$ the unique point satisfying

$$f'(\beta) = \frac{f(\beta)}{\beta - a}. \tag{2.53}$$

(I.e. the point where the straight line through the point $(a, 0)$ is tangent to the graph of $y = f(u)$.) Clearly, $\beta > \alpha$. The proof of the following lemma is similar to that of Lemma 2.14, and so we omit it (see [33]).

LEMMA 2.16. *Assume that* $f(u) \in C^2$ *satisfies conditions* (2.46) *and* (2.47). *If* $u(x)$ *is a critical solution of* (2.45) *then*

$$u(0) > \beta. \tag{2.54}$$

We define a constant $\tau \in (b, c)$ by $f'(\tau) = 0$, i.e. τ is the second root of $f'(u)$. We also recall the function $I(u) = f^2(u) - 2F(u)f'(u)$, defined previously.

LEMMA 2.17. *Assume that* $f(u) \in C^2$ *satisfies conditions* (2.46) *and* (2.47). *Assume that either*

$$\int_a^\tau f(u)\,\mathrm{d}u \leqslant 0, \tag{2.55}$$

or

$$I(\beta) \geqslant 0. \tag{2.56}$$

If $u(x))$ *is a critical solution of* (2.45), *and* $w(x)$ *is a solution of the corresponding linearized problem*, *then*

$$\int_0^1 f''\bigl(u(x)\bigr)u'(x)w^2(x)\,\mathrm{d}x > 0. \tag{2.57}$$

PROOF. For any solution of (2.45) we have $\int_a^{u(0)} f(u)\,\mathrm{d}u > 0$ (just multiply the equation by u' and integrate between $x = 0$ and the point $x = \xi$, such that $u(\xi) = a$). So that if (2.55) holds, then $u(0) > \tau$, i.e. $f'(u(0)) < 0$. Then (2.57) follows from formula (2.35) for the integral (2.57). In case condition (2.56) holds, the proof proceeds the same way as in Lemma 2.12. □

REMARK. We can replace condition (2.55) by requiring that $u(0) > \tau$.

The following result was essentially proved in P. Korman, Y. Li and T. Ouyang [33].

THEOREM 2.7. *Assume that* $f(u) \in C^2$ *satisfies conditions* (2.46) *and* (2.47). *Assume either the condition* (2.56) *is satisfied, or else assume that any solution of problem* (2.45), *with* $u(0) \in (\beta, \tau)$ *is noncritical. Then there exists a critical* λ_0, *such that problem* (2.45) *has exactly one positive solution for* $0 < \lambda < \lambda_0$, *exactly two positive solutions at* $\lambda = \lambda_0$, *and exactly three positive solutions for* $\lambda_0 < \lambda < \infty$. *Moreover, all solutions lie on two smooth solution curves. One of the curves, referred to as the lower curve, starts at* ($\lambda = 0, u = 0$), *it is increasing in* λ, *and* $\lim_{\lambda\to\infty} u(x,\lambda) = a$ *for* $x \in (-1, 1)$. *The upper curve is a parabola-like curve with exactly one turn to the right.*

PROOF. The properties of the lower curve are easy to prove. According to the implicit function theorem there is a curve of positive solutions, starting at $\lambda = 0$ and $u = 0$. Since $f'(u) < 0$ when $u < a$, it follows by Lemma 2.7 that solutions are nondegenerate, and hence they can be continued for all $\lambda > 0$. It is easy to see that the solutions on this curve are increasing in λ, and $\lim_{\lambda\to\infty} u(x,\lambda) = a$ for all $x \in (-1, 1)$, see [30].

Turning to the upper curve, recall that critical solutions are possible only if $u(0) > \beta$. If condition (2.56) holds then we have (2.57). The same way as in Lemma 2.12 we show that at any critical point

$$\int_0^1 f''\bigl(u(x)\bigr) w^3(x)\,\mathrm{d}x < 0, \tag{2.58}$$

which means that only turns to the right are possible on the upper curve. Similarly to Theorem 2.6 for the $a = 0$ case, we show existence of solutions on the upper curve, and that the upper curve has to turn. Hence, exactly one turn occurs on the upper curve, and its other properties are proved similarly to the $a = 0$ case. In the other case, when (2.55) holds, we know that no critical points are possible, until $u(0) > \tau$. But then again (2.57) holds, which implies (2.58), and we proceed the same way as in the first case. □

The above result shows that either one gets "lucky" at the level $u(0) = \beta$, i.e. condition (2.56) holds, and the above Theorem 2.7 applies, or else the interval (β, τ) is "dangerous", i.e. we need to rule out the possibility of any turns when $\beta < u(0) < \tau$ (since we cannot tell their direction). For that computer assisted proofs are feasible. In fact, in a recent paper P. Korman, Y. Li and T. Ouyang [33] have given three independent computer assisted proofs in case of a cubic. We describe their result next. Let $f(u) = (u - a)(u - b)(c - u)$,

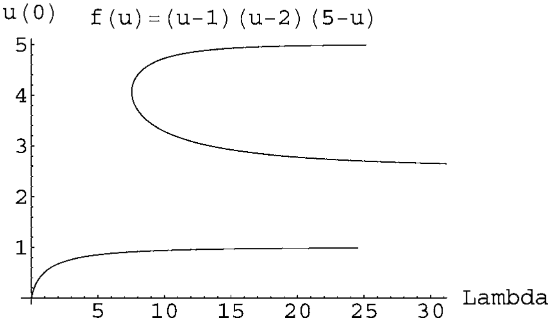


Fig. 5. A two-piece solution curve for a cubic, with $a > 0$.

with $0 < a < b < c$. For problem (2.45) to have a positive solution it is necessary that $\int_a^c f(u)\,\mathrm{d}u > 0$, i.e.

$$b < \frac{a+c}{2}. \tag{2.59}$$

It was shown in [33] that under the necessary condition (2.59) the above Theorem 2.7 applies, providing an exact multiplicity result for the general cubic. In the next section we present a new tool, used in [33] to give one of the computer assisted proofs.

2.5. *Computing the location of bifurcation*

Assume that for the problem

$$u''(x) + f\bigl(u(x)\bigr) = 0, \quad x \in (-1,1), \qquad u(-1) = u(1) = 0 \tag{2.60}$$

bifurcation occurs at $u(0) = \alpha$, i.e. the corresponding linearized problem

$$w''(x) + f'\bigl(u(x)\bigr)w(x) = 0, \quad x \in (-1,1), \qquad w(-1) = w(1) = 0 \tag{2.61}$$

admits a nontrivial solution. The following result of P. Korman, Y. Li and T. Ouyang [33] provides a way to determine all possible α's at which bifurcation may occur, i.e. the corresponding solution of (2.60) is singular.

THEOREM 2.8. *A positive solution of problem* (2.60) *with the maximal value* $\alpha = u(0)$ *is singular if and only if*

$$G(\alpha) \equiv F(\alpha)^{1/2} \int_0^\alpha \frac{f(\alpha) - f(\tau)}{[F(\alpha) - F(\tau)]^{3/2}}\,\mathrm{d}\tau - 2 = 0. \tag{2.62}$$

PROOF. We need to show that the problem (2.61) has a nontrivial solution. By direct verification the function $w(x) = -u'(x)\int_x^1 \frac{1}{u'^2(t)}\,\mathrm{d}t$ satisfies Eq. (2.61). Also $w(1) = 0$. If we also have

$$w'(0) = 0, \tag{2.63}$$

then since $u(x)$ is an even function, the function $w(x)$ is also even (by uniqueness for initial value problems), and hence $w(-1) = 0$, which gives us a nontrivial solution of (2.61). Conversely, every nontrivial solution of (2.61) is an even function, and hence (2.63) is satisfied.

Using Eq. (2.60), we compute

$$w'(x) = f\bigl(u(x)\bigr)\int_x^1 \frac{1}{u'^2(t)}\,\mathrm{d}t + \frac{1}{u'(x)}.$$

Since the energy $\frac{u'^2}{2}(x) + F(u(x))$ is constant,

$$\frac{u'^2}{2}(x) + F\bigl(u(x)\bigr) = F\bigl(u(0)\bigr) = F(\alpha).$$

On the interval (0, 1) we express

$$u'(x) = -\sqrt{2}\sqrt{F(\alpha) - F(u(x))}. \tag{2.64}$$

We use this formula in the integral $\int_x^1 \frac{1}{u'(t)^2}\,\mathrm{d}t$, and then we make a change of variables $t \to s$, by letting $s = u(t)$. We have

$$\begin{aligned} 2^{3/2}\int_x^1 \frac{1}{u'^2(t)}\,\mathrm{d}t &= -\int_x^1 \frac{u'(t)\,\mathrm{d}t}{[F(\alpha) - F(u(t))]^{3/2}} \\ &= -\int_{u(x)}^0 \frac{1}{[F(\alpha) - F(s)]^{3/2}}\,\mathrm{d}s. \end{aligned} \tag{2.65}$$

Using formulas (2.64) and (2.65), we express

$$2^{3/2}w'(x) = \int_0^{u(x)} \frac{f(u(x))}{[F(\alpha) - F(\tau)]^{3/2}}\,\mathrm{d}\tau - \frac{2}{[F(\alpha) - F(u(x))]^{1/2}}. \tag{2.66}$$

If we try to set here $x = 0$, then both terms on the right are infinite. Instead, we observe that

$$\begin{aligned} -\frac{2}{[F(\alpha) - F(u)]^{1/2}} &= -\int_0^u \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{2}{[F(\alpha) - F(\tau)]^{1/2}}\,\mathrm{d}\tau - \frac{2}{F(\alpha)^{1/2}} \\ &= -\int_0^u \frac{f(\tau)}{[F(\alpha) - F(\tau)]^{3/2}}\,\mathrm{d}\tau - \frac{2}{F(\alpha)^{1/2}}. \end{aligned} \tag{2.67}$$

Using (2.67) in (2.66), we obtain

$$2^{3/2}w'(x) = \int_0^{u(x)} \frac{f(u(x)) - f(\tau)}{[F(\alpha) - F(\tau)]^{3/2}}\,\mathrm{d}\tau - \frac{2}{F(\alpha)^{1/2}}. \tag{2.68}$$

The integral on the right is now nonsingular, as we let $x \to 0$. At $x = 0$ we see that (2.63) is equivalent to (2.62). □

In case of a cubic $f(u) = (u - a)(u - b)(c - u)$, P. Korman, Y. Li and T. Ouyang [33] have used formula (2.62) to give a computer assisted proof that there are no turning points in the dangerous region, $u(0) \in (\beta, \tau)$, thus establishing the exact multiplicity result (Theorem 2.7) from the previous section.

2.6. *Computing the direction of bifurcation*

We have seen that at a critical solution $u(x)$ of (2.60) the integral $I = \int_0^1 f''(u)w^3\,\mathrm{d}x$ governs the direction of bifurcation. Also, it is known that in case $I \neq 0$ a critical solution $u(x)$ is *nondegenerate*, i.e. it persists when the equation is perturbed slightly (i.e. the turning points persist under perturbations), see, e.g., [32]. The following result from P. Korman, Y. Li and T. Ouyang [33] allows one to compute the integral I as a function of $\alpha = u(0)$.

THEOREM 2.9. *At any critical solution $u(x)$ of* (2.60), *with* $u(0) = \alpha$,

$$I = c\int_0^\alpha f''(u)\left(\int_u^\alpha f(s)\,\mathrm{d}s\right)\left(\int_0^u \frac{\mathrm{d}s}{(\int_s^\alpha f(t)\,\mathrm{d}t)^{3/2}}\right)^3 \mathrm{d}u, \tag{2.69}$$

where $c = \frac{1}{4\sqrt{2}}u'^3(1)w'^3(1) > 0$.

This formula is rather involved, but using *Mathematica* it can be evaluated numerically. In a future paper, with Y. Li and T. Ouyang, we use this result to handle equations modeled on polynomials of arbitrary power.

2.7. *Pitchfork bifurcation and symmetry breaking*

So far for the problem

$$u'' + \lambda f(u) = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0 \tag{2.70}$$

we have considered the cases when $f(0) \geqslant 0$. As we have observed earlier, this condition implies that $|u_x(\pm 1, \lambda)| \neq 0$ for any positive solution $u(x, \lambda)$. Since we also know that

$u_x(x, \lambda) < 0$ for all $x \in (0, 1)$, there is no way for a positive solution to become sign-changing, as we vary λ (no interior roots, or zero slope at the boundary are possible). The situation changes drastically in case

$$f(0) < 0. \tag{2.71}$$

A solution may develop a zero slope at the boundary, and become sign-changing. In fact, a pitchfork bifurcation usually happens. In addition to the sign-changing symmetric solution, two symmetry-breaking solution emerge, see P. Korman [20].

Let us consider the problem

$$u''(x) + u^{2k}(x) - \lambda = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0, \tag{2.72}$$

which we will relate to the problems of type (2.70) shortly. Here $k \geqslant 1$ is an integer. For $k = 1$ this problem was exhaustively analyzed in J.C. Scovel's Ph.D. thesis [54], and in H.P. McKean and J.C. Scovel [44]. They used explicit integration via elliptic functions, which means that their method does not work for $k > 1$. It turned out that the solution set of (2.72) for $k = 1$ consists of infinitely many identically looking curves. Each curve is a parabola like curve, with pitchfork bifurcation on one of the branches. (I.e. there is exactly one turn, and exactly one point of pitchfork bifurcation on each curve, see Fig. 6.) V. Anuradha and R. Shivaji [6] have studied a related problem. Using the quadrature technique, they showed existence of infinitely many points of bifurcation. In [20] P. Korman had used bifurcation theory to approach problem (2.72), and in particular the case of $f(u) = u^{2k}$, with $k > 1$. We were able to generalize some, but not all, of the results of H.P. McKean and J.C. Scovel [44].

It is well known that at $\lambda = 0$ there exists a unique positive solution of (2.72). This solution is known to be nondegenerate, so that we can continue it for small $\lambda > 0$. Setting $u(x) = \mu v(x)$, with μ determined by the relation $\mu^{2k} = \lambda$, we convert problem (2.72) into (a particular case of the problem (2.70))

$$\begin{aligned} &v''(x) + \lambda\big(v^{2k}(x) - 1\big) = 0, \quad \text{for } -1 < x < 1, \\ &v(-1) = v(1) = 0, \end{aligned} \tag{2.73}$$

where λ is a new parameter (equal to μ^{2k-1}). With the parameter now in front of the nonlinearity, Lemma 2.8 applies, and hence we can always continue both positive and sign-changing solutions of (2.73) (and also of (2.72). Observe that the curve of positive solutions does not turn for $\lambda > 0$ (for $g(v) \equiv v^{2k} - 1$, we have $vg'(v) > g(v)$ for all $v > 0$). By Theorem 2.1 this curve of positive solutions cannot be continued for all $\lambda > 0$ (the function $g(v) = v^{2k} - 1$ has no "stable" roots, i.e. roots where derivative is negative). By the Sturm's comparison theorem, it is easy to see that positive solutions cannot become unbounded at a finite λ. Hence, solutions on this curve must eventually stop being positive, and the only way this can happen is that $u'(\pm 1) = 0$ at some λ_0 (in view of the symmetry of positive solutions).

We now outline the pitchfork bifurcation analysis for the general problem (2.70), and more details can be found in P. Korman [20]. So suppose problem (2.70) has a curve of positive solutions $u(x,\lambda)$, so that for $\lambda < \lambda_0$ we have $u_x(1,\lambda) < 0$, while at $\lambda = \lambda_0$ we have $u_x(1,\lambda_0) = 0$. The function $u_x(x,\lambda_0)$ is then a solution of the corresponding linearized problem ($u_0 = u(x,\lambda_0)$)

$$w'' + \lambda_0 f'(u_0)w = 0 \quad \text{for } -1 < x < 1, \qquad w(-1) = w(1) = 0. \tag{2.74}$$

The null-space of the linearized problem is one-dimensional (by Lemma 2.5), and it is spanned by the odd function $u_x(x,\lambda_0)$. If we now restrict to the space of even functions, the null-space will be empty, and hence by the implicit function theorem the solution curve $u(x,\lambda)$ continues for $\lambda > \lambda_0$, as sign-changing symmetric (even) solutions. We can compute the tangential direction for this curve at $\lambda = \lambda_0$:

$$u_\lambda(x,\lambda_0) = xu_x(x,\lambda_0). \tag{2.75}$$

Indeed, the function $u_\lambda(x,\lambda_0) - xu_x(x,\lambda_0)$ is an even function, solving the linearized problem (2.74). Hence it must be zero, justifying (2.75). If we let $v = u - u(x,\lambda)$, where $u(x,\lambda)$ is the curve of sign-changing symmetric solutions, then for $\lambda > \lambda_0$ we have a trivial solution $v = 0$. We showed in [20] that the conditions of the Crandall–Rabinowitz theorem on bifurcation from the trivial solutions are satisfied at $\lambda = \lambda_0$, giving rise to a parabola-like curve of symmetry breaking solutions (see also M. Ramaswamy [51], which we used in [20]). Their tangential direction is given by $u_x(x,\lambda_0)$.

One of the reasons we were not able to fully recover the beautiful results of McKean and Scovel [44], is that we could not tell the direction of the pitchfork bifurcation: which way the symmetry breaking solutions bifurcate, toward $\lambda > \lambda_0$ or $\lambda < \lambda_0$? Recently X. Hou, P. Korman and Y. Li [17] has given a computer assisted way (again computer assisted!) to settle this question. Here is their result, which says that the pitchfork opens forward.

THEOREM 2.10. *Consider problem* (2.72), *with* $1 \leqslant k \leqslant 720$. *Let* λ_0 *be the point of pitchfork bifurcation.* (*The value of* λ_0 *was explicitly computed in* [17].) *Then there is a negative* $\bar{\lambda} = \bar{\lambda}(k) < 0$, *so that problem* (2.72) *has exactly two positive solutions for* $\bar{\lambda} < \lambda < 0$, *it has exactly one positive and one negative solution on* $(0,\lambda_0)$. *Moreover, there is a* $\lambda_1(k) > \lambda_0$, *so that problem* (2.72) *has four solutions on* (λ_0,λ_1), *one negative* (*and symmetric*), *one sign-changing and symmetric* (*with* $u(0) > 0$), *and two asymmetric solutions.*

In Fig. 6 we present a picture of pitchfork bifurcation from [17] (produced by X. Hou). We draw $u'(-1)$ as a function of λ. In that figure solid lines denote positive and negative solutions, the dashed line denotes sign-changing symmetric solutions, and the doted lines stand for the symmetry breaking solutions.

2.8. *Sign-changing solutions*

We consider sign-changing solutions of the two point problem

$$u'' + f(u) = 0 \quad \text{for } x \in (0,1), \qquad u(0) = u(1) = 0. \tag{2.76}$$

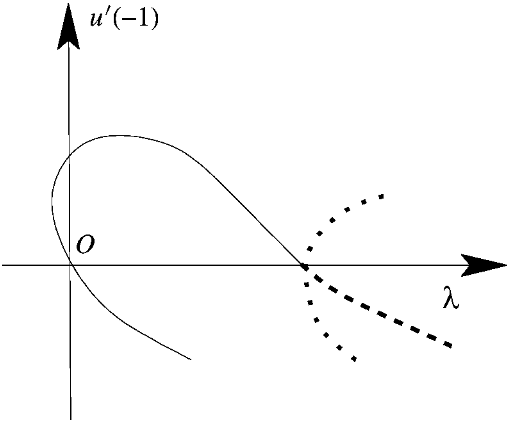


Fig. 6. Pitchfork bifurcation.

Notice that we pose the problem over the interval (0, 1), since sign-changing solutions need not be symmetric. Also we do not have λ in front of $f(u)$ (one can think that λ is absorbed into $f(u)$). Corresponding linearized problem is

$$w'' + f'(u)w = 0 \quad \text{for } x \in (0,1), \qquad w(0) = w(1) = 0. \tag{2.77}$$

The following result is from P. Korman and T. Ouyang [39].

THEOREM 2.11. *Let $f \in C^2(\mathbb{R})$, and assume that either one of the following two inequalities holds*:

$$\frac{f(u)}{u} - f'(u) > 0\ (< 0) \quad \textit{for almost every } u \in \mathbb{R}. \tag{2.78}$$

Then any solution of problem (2.76), *satisfying $u'(0) \neq 0$, is nonsingular* (*i.e.* (2.77) *admits only the trivial solution*).

PROOF. Assume on the contrary that problem (2.77) admits a nontrivial solution $w(x)$.

Step 1. We show that the number of roots of u and w inside (0, 1) differs by one. Assume for definiteness that $f'(u) > \frac{f(u)}{u}$ for almost every $u \in \mathbb{R}$. If we regard (2.76) as a linear equation $u'' + \frac{f(u)}{u}u = 0$, then by the Sturm comparison theorem the function $w(x)$ has a root between any two roots of $u(x)$. Since both functions vanish at the endpoints, $x = 0$ and $x = 1$, it follows that w has one more interior root than u.

Step 2. We will show that u and w have the same number of interior roots. This will result in a contradiction, proving the theorem. We denote by n_u the number of interior roots of u, and use the same notation for other functions. The functions w and u' satisfy the same linear equation, and hence their roots are interlaced. Since w vanishes at the endpoints and u' does not, it follows that $n_{u'} = n_w + 1$. Since $n_{u'} = n_u + 1$, it follows that $n_u = n_w$. □

Corresponding to the solution $u(x)$ of (2.76) we may consider an eigenvalue problem

$$\varphi'' + f'(u)\varphi + \mu\varphi = 0 \quad \text{for } x \in (0,1), \qquad \varphi(0) = \varphi(1) = 0. \tag{2.79}$$

The eigenvalues of (2.79) form a sequence $\mu_1 < \mu_2 < \cdots < \mu_n < \cdots$, tending to infinity. The number of negative eigenvalues is called the *Morse index* of $u(x)$. The following theorem is from P. Korman and T. Ouyang [39].

THEOREM 2.12. *Let $u(x)$ any solution of problem* (2.76), *with k interior roots and satisfying $u'(0) \neq 0$. Then the Morse index of $u(x)$ is either k or $k+1$. Moreover, the Morse index equals k if the first inequality in* (2.78) *holds, and it equals $k+1$ if the second inequality in* (2.78) *holds.*

Similar results for balls in $\mathbb{R}^n$ have been given in J. Shi and J. Wang [56].

2.9. *The Neumann problem*

Consider the Neumann problem

$$u'' + \lambda f(u) = 0 \quad \text{for } 0 < x < 1, \qquad u'(0) = u'(1) = 0. \tag{2.80}$$

We are interested in the solution branches bifurcating off constant solutions. By translation we may assume the constant solution to be zero, i.e. we assume that

$$f(0) = 0, \tag{2.81}$$

and that $f(u) \in C^2(a^-, a^+)$ for some $-\infty \leqslant a^- < 0 < a^+ \leqslant \infty$. We assume that

$$uf(u) > 0 \quad \text{on } \bigl(a^-, a^+\bigr). \tag{2.82}$$

We consider solutions of (2.80) such that $u(x) \in (a^-, a^+)$. It suffices to consider only the *increasing* solutions of (2.80), i.e. $u'(x) > 0$ on $(0,1)$, since other solutions can be produced from them by reflection, pasting and scaling. Clearly we have $u(0) < 0 < u(1)$, since $x = 0$ and $x = 1$ are points of minimum and maximum, respectively.

The corresponding linearized problem is

$$w'' + \lambda f'(u)w = 0 \quad \text{for } 0 < x < 1, \qquad w'(0) = w'(1) = 0. \tag{2.83}$$

If this problem has only the trivial solution, then the solution branches bifurcating from zero do not turn. A simple condition for this to happen goes back to Z. Opial [47], see also R. Schaaf [53]. Namely, we assume that either one of the following two inequalities holds:

$$\frac{f(u)}{u} - f'(u) > 0\ (< 0) \quad \text{for every } u \in \bigl(a^-, a^+\bigr) \setminus \{0\}. \tag{2.84}$$

It is an elementary exercise to show that (2.84) will follow if either one of the following two inequalities holds:

$$uf''(u) > 0\ (< 0) \quad \text{for every } u \in \left(a^-, a^+\right) \setminus \{0\}. \tag{2.85}$$

The following result is of course known, see [47] and [53], although previously it was stated in different terms (involving monotonicity of time maps), and proved by different methods.

THEOREM 2.13. *Assume that conditions* (2.81), (2.82) *and* (2.84) (*or* (2.85)) *hold, and* $u(x) \in (a^-, a^+)$ *for all* $x \in (0, 1)$. *Then the linearized problem* (2.83) *admits only the trivial solution.*

PROOF. Assume on the contrary that $w(x)$ is a nontrivial solution of (2.83). Observe that $u'(x)$ satisfies the same equation (2.83), $u'(x) > 0$ for $x \in (0, 1)$ and $u'(0) = u'(1) = 0$. It follows by the Sturm comparison theorem that $w(x)$ has exactly one root on (0, 1); we call it η, i.e. $w(\eta) = 0$. We may assume (by scaling) that $w(0) < 0$ and $w(1) > 0$, and hence $w'(\eta) > 0$. Let ξ denote the unique root of the increasing solution $u(x)$, i.e. $u(\xi) = 0$ and $u'(\xi) > 0$. Writing Eq. (2.80) in the form $u'' + \lambda\frac{f(u)}{u}u = 0$, and combining it with (2.83), we have

$$(u'w - uw')' + \lambda\left[\frac{f(u)}{u} - f'(u)\right]uw = 0. \tag{2.86}$$

We now consider two cases.

Case 1. $\xi \leqslant \eta$. Assume that the first inequality holds in (2.84), i.e. the quantity in the square bracket in (2.86) is positive. We integrate (2.86) over the interval $(\eta, 1)$, where both $u(x)$ and $w(x)$ are positive

$$u(\eta)w'(\eta) + \lambda\int_\eta^1\left[\frac{f(u)}{u} - f'(u)\right]uw\,\mathrm{d}x = 0.$$

We have a contradiction, since both terms on the left are positive.

If the second inequality holds in (2.84), i.e. the quantity in the square bracket in (2.86) is negative, we integrate (2.86) over the interval $(0, \xi)$, where both $u(x)$ and $w(x)$ are negative

$$u(\xi)w'(\xi) + \lambda\int_0^\xi\left[\frac{f(u)}{u} - f'(u)\right]uw\,\mathrm{d}x = 0.$$

Again, we have a contradiction, since both terms on the left are negative.

Case 2. $\xi \geqslant \eta$. If the first inequality holds in (2.84), we integrate over $(0, \eta)$, where both $u(x)$ and $w(x)$ are negative

$$-u(\eta)w'(\eta) + \lambda \int_0^\eta \left[\frac{f(u)}{u} - f'(u)\right] uw \, \mathrm{d}x = 0.$$

Both terms on the left are positive, a contradiction. If the second inequality holds in (2.84), we integrate over $(\xi, 1)$, where both $u(x)$ and $w(x)$ are positive

$$-u'(\xi)w(\xi) + \lambda \int_\xi^1 \left[\frac{f(u)}{u} - f'(u)\right] uw \, \mathrm{d}x = 0.$$

Both terms on the left are negative, again we have a contradiction. □

Beyond this simple theorem, we know of only two results on the Neumann problem. The first one is due to R. Schaaf [53]. It dealt with monotonicity of time maps, here we rephrase it in terms of nondegeneracy of solutions.

THEOREM 2.14 [53]. *Assume that the function $f(u)$ is either an $A - B$ or C function on the interval (a^-, a^+). Then the linearized problem* (2.83) *admits only the trivial solution.*

The other one is from P. Korman [23].

THEOREM 2.15. *Assume that $f(u)$ satisfies $f'(u) > 0$ and $f'''(u) < 0$ on the interval $(0, a^+)$, and $f''(u) > 0$ on (a^-, a^+). Then the linearized problem* (2.83) *admits only the trivial solution.*

The last result is not very satisfactory. Its only advantage is that no third order assumptions on $f(u)$ are made on $(a^-, 0)$, while on $(0, a^+)$ such functions are of class $A - B$. R. Schaaf's result is better.

According to the condition (2.85), we can handle the cases when $f(u)$ changes concavity at its root $u = 0$. But what if it keeps the same concavity? We wish to pose the following problem.

PROBLEM. Assume that $f(0) = 0$, and

$$f''(u) > 0 \quad \text{for } u \in \left(a^-, a^+\right).$$

Is it true that any increasing solution of the Neumann problem (2.80), with values in (a^-, a^+), is nondegenerate (i.e. (2.83) has only the trivial solution)?

Using bifurcation approach, we can also treat some nonautonomous problems. For example,

$$u'' + \lambda b(x) f(u) = 0 \quad \text{for } 0 < x < 1, \qquad u'(0) = u'(1) = 0, \tag{2.87}$$

with the given function $b(x)$ being positive and continuous. The corresponding linearized problem is now

$$w'' + \lambda b(x) f'(u) w = 0 \quad \text{for } 0 < x < 1, \qquad w'(0) = w'(1) = 0. \tag{2.88}$$

Formula (2.86) still holds here (with an extra factor of $b(x)$ in front of the square bracket), and hence the arguments of the above theorem can be used unchanged. In particular, we conclude as above that any nontrivial solution of (2.88) cannot vanish exactly once. Unlike the autonomous problem, we cannot yet conclude that $w(x)$ is zero, since we cannot automatically exclude the possibilities that $w(x)$ has no roots, or at least two roots. We need to introduce another condition:

$$f'(u) \geqslant 0 \quad \text{for every } u \in \left(a^-, a^+\right). \tag{2.89}$$

THEOREM 2.16. *Assume that* $b(x)$ *is positive and continuous on* $[0, 1]$, *and* $f(u) \in C^1[a^-, a^+]$ *satisfies conditions* (2.81), (2.82), (2.89), *and the first inequality holds in* (2.84). *Let* $u(x)$ *be an increasing solution of the Neumann problem* (2.80), *satisfying* $u(x) \in (a^-, a^+)$ *for all* $x \in (0, 1)$. *Then the linearized problem* (2.88) *admits only the trivial solution.*

PROOF. As we mentioned above, the arguments used in proof of Theorem 2.13 apply here as well. In case the first inequality holds in (2.84), we have proved in Theorem 2.13 that $w(x)$ cannot vanish on either side of ξ, the root of $u(x)$. Hence $w(x)$ keeps the same sign over $(0, 1)$. But then integrating the linearized equation (2.88),

$$\int_0^1 b(x) f'(u) w \, \mathrm{d}x = 0,$$

which is a contradiction, since the integrand is of one sign. □

As an example, the function $f(u) = u - u^3$ satisfies the conditions of this theorem on the interval $(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$.

2.10. *Similarity of the solution branches*

We saw in the previous two sections that under the same condition (2.84) we could prove nondegeneracy for both sign-changing solutions, and for the Neumann problem. It turns out that a curve of solutions with an odd number of sign changes is always similar to curves of solutions of Neumann problems. (I.e. both curves have the same number of critical points, with the same direction of turns.) Let us fix the notation, before we state the result. We consider the problem

$$u'' + \lambda f(u) = 0 \quad \text{on } (0, 1), \tag{2.90}$$

subject to either Dirichlet

$$u(0) = u(1) = 0, \tag{2.91}$$

or Neumann

$$u'(0) = u'(1) = 0 \tag{2.92}$$

boundary conditions. We shall consider the solutions as the positive parameter λ varies, and refer to the solution curves as either Dirichlet or Neumann branches, depending on the boundary conditions used. Recall that by Lemma 2.1 any solution of Eq. (2.90) is symmetric with respect to any of its critical points. This implies, in particular, that either minimum or maximum occurs at any critical point. It follows that any solution of Neumann problem is determined by its values on any subinterval $I \subset (0, 1)$, whose end-points are two consecutive critical points of $u(x)$. We can then obtain the solution on the entire interval $(0, 1)$ through reflections and translations. We refer to I, and any other interval uniquely determining the solution through reflections and translations, as a *determining interval*. The interval I, joining two consecutive critical points of $u(x)$, is also a determining interval for the Dirichlet problem. Another determining interval for Dirichlet problem is (ξ, η), where $0 \leqslant \xi < \theta < \eta \leqslant 1$ are three consecutive roots of $u(x)$. This interval contains both positive and negative humps (and all positive (negative) humps are translations of one another).

As we vary λ the number of roots on Dirichlet branches, as well as the number of monotonicity changes on Neumann branches, remain constant. Indeed, by Lemma 2.2 solutions of (2.90) cannot have points of positive maximum and negative minimum, and there is no other mechanism by which extra roots (or monotonicity changes) may be created.

The natural way to distinguish the Dirichlet branches is by the number of interior roots, and the Neumann branches can be identified by the number of changes of monotonicity (both properties remain constant on the solution curves). Any solution of the Dirichlet problem with at least one interior root contains a solution of the Neumann problem on a subinterval of $(0, 1)$. Indeed, just consider the solution between two consecutive critical points. In order for solutions of the Neumann problem to contain in turn a solution of the Dirichlet problem, we need to impose some conditions on $f(u)$. Namely, we assume that

$$f(0) = 0, \tag{2.93}$$

and there exist two constants $-\infty \leqslant m < 0 < M \leqslant \infty$ so that

(f1) $f(u) > 0$ for $u \in (0, M)$,
(f2) $f(u) < 0$ for $u \in (m, 0)$.

LEMMA 2.18. *Under conditions* (2.93), (f1) *and* (f2) *any solution of the Neumann problem for* (2.90), *satisfying*

$$m < u(x) < M \quad \textit{for all } x \tag{2.94}$$

has a root between any two critical points.

PROOF. Follows immediately, by multiplying the equation (2.90) by u', and integrating between any two consecutive critical points. □

DEFINITION. We call two solution branches of (2.90) to be *similar* if for any solution on the either branch there is a determining interval so that by stretching of x, or by reflection $x \to 2a - x$, for some $a \in (0, 1)$, we obtain a solution from the other branch on a (different) determining interval. Clearly, if solution branches are similar then the corresponding solution curves in (λ, u) "plane" have the same shape.

The following result was proved in this form by P. Korman [24], although it can also be found in R. Schaaf [51].

THEOREM 2.17. *All Neumann branches of* (2.90) *are similar, and if* $f(u)$ *satisfies conditions* (2.93), (f1) *and* (f2), *while all solutions satisfy* (2.94), *then the Neumann branches are similar to the Dirichlet ones with an odd number of interior roots* (*and these Dirichlet branches are also all similar*).

PROOF. We begin with Neumann branches. If a Neumann solution changes monotonicity twice, then its increasing part is a reflection of its decreasing part with respect to $x = \frac{1}{2}$. If a Neumann solution changes monotonicity n times, then all critical points occur at i/n, $i = 1, \dots, n-1$, and the graphs of solution on all intervals where it is increasing (decreasing) are translations of one another. Since an interval connecting any two critical points is a determining interval, the equivalence of the Neumann branches follows (via rescaling). □

If a Dirichlet solution has $2k-1$ interior roots, it has k identical positive humps and k identical negative humps. Assume for definiteness that solution starts with a negative hump, followed by a positive one, and so on. If ξ is the first point of (negative) minimum of $u(x)$, then the first interior root occurs at 2ξ. If $2\xi + \eta$ is the point of the first (positive) maximum, then the second interior root occurs at $2\xi + 2\eta$. The last critical point, a positive maximum, occurs at $1 - \eta$. Observe that $k(2\xi + 2\eta) = 1$, i.e. $\xi + \eta = \frac{1}{2k}$. So while both ξ and η vary with λ, $u(x)$ solves the Neumann problem on the interval $(\xi, 1 - \eta)$, and this interval has a *fixed* length of

$$1 - \eta - \xi = \frac{2k-1}{2k}.$$

So that any Dirichlet solution curve "carries" inside it a solution of a Neumann problem on a fixed interval (which can be made to be (0, 1) by rescaling), and hence the Dirichlet branch cannot have any more complexity (like extra turns) than any Neumann branch.

Conversely, consider the Neumann problem with $2k+1$ changes of monotonicity. Assume for definiteness that $u(0) < 0$. Then $u(1) > 0$. Assume that $\xi = \xi(\lambda)$ is the smallest interior root, and $1 - \eta$ is the largest one, $\eta = \eta(\lambda)$. On the interval $(0, 1)$ we then have $2k+1$ negative half-humps, each of width ξ, and $2k+1$ positive ones, each of width η. So that $\xi + \eta = \frac{1}{2k+1}$. On the interval $(\xi, 1 - \eta)$ we have a solution of the Dirichlet problem with $2k-1$ interior roots, and the length of this interval is

$$1 - \eta - \xi = \frac{2k}{2k+1},$$

which does not vary with λ. So that any Neumann branch "carries" inside it a solution of a Dirichlet problem on a fixed interval, and hence the Neumann branch cannot have any more complexity than the corresponding Dirichlet branch with an odd number of interior zeroes.

Finally, the Dirichlet branches with odd number of interior zeroes are all similar, since any two such branches are similar to a pair of Neumann branches, but Neumann branches are all similar.

The Dirichlet branches with even number of interior zeroes may behave differently, as the following example due to R. Schaaf [53] shows.

EXAMPLE [53]. For the problem

$$u'' + \lambda\left(e^u - 1\right) = 0 \quad \text{on } (0,1), \qquad u(0) = u(1) = 0$$

the branch bifurcating from the principal eigenvalue does not turn, while all other branches have exactly one turn.

3. A class of symmetric nonlinearities

For the autonomous equation (2.1) both phase-plane analysis and bifurcation theory apply. If we allow explicit dependence of the nonlinearity on x, i.e. consider

$$u'' + \lambda f(x,u) = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0, \tag{3.1}$$

then the problem becomes much more complicated. For example, solutions of the corresponding linearized problem need not be of one sign. In the papers P. Korman and T. Ouyang a class of $f(x,u)$ has been identified, for which the theory of positive solutions is very similar to that for the autonomous case, see, e.g., [34–36]. Further results in this direction have been given in P. Korman, Y. Li and T. Ouyang [30], and P. Korman and J. Shi [40]. Namely, we assume that $f \in C^2$ satisfies

$$f(-x,u) = f(x,u) \quad \text{for all } -1 < x < 1 \text{ and } u > 0, \tag{3.2}$$

$$f_x(x,u) \leqslant 0 \quad \text{for all } 0 < x < 1 \text{ and } u > 0. \tag{3.3}$$

Under the above conditions any positive solution of (3.1) is an even function, with $u'(x) < 0$ for all $x \in (0,1]$, see B. Gidas, W.-M. Ni and L. Nirenberg [15]. (For the one-dimensional problem (3.1) a different proof of the symmetry of solutions is given in P. Korman [18]. It is a little simpler than the moving plane method of [15], and it allows to relax somewhat the condition (3.3).) As before the linearized problem

$$w'' + \lambda f_u(x,u)w = 0 \quad \text{for } -1 < x < 1, \qquad w(-1) = w(1) = 0 \tag{3.4}$$

will be important for the multiplicity results.

LEMMA 3.1 [34]. *Under conditions* (3.2) *and* (3.3) *any nontrivial solution of* (3.4) *is of one sign. Moreover,* $w(x)$ *is an even function, and it spans the null set of* (3.4).

PROOF. Assume that $w(x)$ has a root on [0, 1) (the case when $w(x)$ vanishes on $(-1,0]$ is similar). We may assume (taking $-w$ if necessary) that there is a subinterval (x_1,x_2), $0 \leqslant x_1 < x_2 \leqslant 1$, so that $w(x)>0$ on (x_1,x_2), and $w(x_1)=w(x_2)=0$. Integrating the relation $[u'w'-u''w]'=\lambda f_x w$ over (x_1,x_2),

$$u'(x_2)w'(x_2)-u'(x_1)w'(x_1)=\lambda\int_{x_1}^{x_2} f_x w\,\mathrm{d}x.$$

We have a contradiction, since the quantity on the left is positive, while the one on the right is nonpositive.

The null set of (3.4) is one dimensional, since it can be parameterized by $w'(1)$. To prove that $w(x)$ is even, observe that $w(-x)$ is also a solution of (3.1), and hence $w(-x)=cw(x)$ for some constant c (since the null set of (3.4) is one dimensional). Evaluating this at $x=0$, we conclude that $c=1$ (since $w(0)>0$), and the claim follows. □

The next lemma shows that the Crandall–Rabinowitz Theorem 1.2 applies at any critical solution.

LEMMA 3.2. *Under conditions* (3.2) *and* (3.3) *let* $u(x)$ *be a critical solution of* (3.1), *and* $w(x)$ *a solution of the corresponding linearized problem. Then we have*

$$\int_0^1 f(x,u)w\,\mathrm{d}x > \frac{1}{2\lambda}u'(1)w'(1)>0. \tag{3.5}$$

PROOF. By the preceding lemma we may assume that $w(x)>0$. We then have

$$(u''w-u'w')'=-\lambda f_x w>0 \quad \text{for } x>0.$$

So that the function $u''w-u'w'$ is increasing on (0, 1), and then

$$u''w-u'w'<-u'(1)w'(1) \quad \text{for } x>0.$$

Integrating this over (0, 1), and expressing u'' from Eq. (3.1), we conclude (3.5). □

The following result from P. Korman and J. Shi [40] is an extension of Lemma 2.3. Unlike the autonomous case, several conditions are now needed.

THEOREM 3.1 See [40]. *In addition to* (3.2) *and* (3.3) *assume that*

$$f(x,u)>0 \quad \textit{for all } -1<x<1 \textit{ and } u>0. \tag{3.6}$$

Then the set of positive solutions of (3.1) *can be parameterized by their maximum values* $u(0)$. (*I.e.* $u(0)$ *uniquely determines the pair* $(\lambda,u(x))$.)

PROOF. Assume on the contrary $v(x)$ is another solution of (3.1), corresponding to some parameter $\mu \geqslant \lambda$, but $u(0) = v(0)$. The case of $\mu = \lambda$ is not possible in view of uniqueness of initial value problems, so assume that $\mu > \lambda$. Then $v(x)$ is a supersolution of (3.1), i.e.

$$v'' + \lambda f(x, v) < 0 \quad \text{for } -1 < x < 1, \qquad v(-1) = v(1) = 0. \tag{3.7}$$

Since $v''(0) < u''(0)$, it follows that $v(x) < u(x)$ for $x > 0$ small. Let $0 < \xi \leqslant 1$ be the first point where the graphs of $u(x)$ and $v(x)$ intersect (i.e. $v(x) < u(x)$ on $(0, \xi)$). We now multiply Eq. (3.1) by u', and integrate over $(0, \xi)$. Denoting by $x_2(u)$ the inverse function of $u(x)$ on $(0, \xi)$, we have

$$\frac{1}{2}u'^2(\xi) + \int_{u(0)}^{u(\xi)} f\big(x_2(u), u\big)\,\mathrm{d}u = 0. \tag{3.8}$$

Similarly denoting by $x_1(u)$ the inverse function of $v(x)$ on $(0, \xi)$, we have from (3.7)

$$\frac{1}{2}v'^2(\xi) + \int_{u(0)}^{u(\xi)} f\big(x_1(u), u\big)\,\mathrm{d}u > 0. \tag{3.9}$$

Subtracting (3.9) from (3.8), noticing that $x_2(u) > x_1(u)$ for all $u \in (u(\xi), u(0))$, and using condition (3.3), we have

$$\frac{1}{2}\big[u'^2(\xi) - v'^2(\xi)\big] + \int_{u(\xi)}^{u(0)} \big[f\big(x_1(u), u\big) - f\big(x_2(u), u\big)\big]\,\mathrm{d}u < 0. \tag{3.10}$$

Since both terms on the left are positive, we obtain a contradiction. □

Next we consider positive solutions of the boundary value problem

$$u'' + \lambda b(x) f(u) = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0. \tag{3.11}$$

We assume that $b(x) \in C^1[-1, 1]$ satisfies $b(x) > 0$ for $x \in [-1, 1]$, and $b(x) = b(-x)$, $b'(x) < 0$ for $x \in (0, 1)$. We also assume that $f(u) > 0$, so that this problem belongs to the class discussed above. For any solutions $u(x)$ let $(\mu, w(x))$ denote the principal eigenpair of the corresponding linearized equation, i.e. we assume that $w(x) > 0$ satisfies

$$\begin{aligned} &w'' + \lambda b(x) f'(u) w + \mu w = 0 \quad \text{for } -1 < x < 1, \\ &w(-1) = w(1) = 0. \end{aligned} \tag{3.12}$$

The following theorem is taken from P. Korman and J. Shi [40].

THEOREM 3.2. *Assume $f \in C^2[0, \infty)$, $f(u) > 0$, $f'(u) > 0$ and $f''(u) > 0$ for all $u > 0$, and for some $\alpha > 0$ condition* (2.43) *is satisfied. Then the solution of* (3.11) *with $u(0) = \alpha$ is unstable if it exists.*

PROOF. In the proof of Theorem 2.3, (2.25) and (2.26) are still true. Assume now that $u(x)$ is stable, i.e. $\mu \geqslant 0$ in (3.12). Then $w(x)$ is a positive solution of the problem

$$w'' + g(x, w) = 0 \quad \text{for } -1 < x < 1, \qquad w(-1) = w(1) = 0, \tag{3.13}$$

with $g(x, w) = \lambda b(x) f'(u(x))w + \mu w$. Since $g(x, w)$ is even in x, and

$$g_x = \lambda b'(x) f'(u)w + \lambda b(x) f''(u)u'w < 0 \quad \text{on } (0, 1),$$

the theorem of B. Gidas, W.-M. Ni and L. Nirenberg [15] applies to (3.13). It follows that $w(x)$ is an even function with $w'(x) < 0$ on $(0, 1)$. Recall that $w(x)$ is determined up to a constant multiple. Since $w(x)$ is decreasing, while $-u'(x)$ is increasing on $(0, 1)$, by scaling $w(x)$ we can achieve (2.28). Using (2.25), (2.28), and also (2.26), we have (2.29).

Since $b(x) > 0$, $b'(x) < 0$ in $(0, 1)$ using (2.25) and (2.29), we have

$$\begin{aligned}
&\int_0^1 b(x)\bigl[f(u) - uf'(u)\bigr]w(x)\,\mathrm{d}x \\
&\quad = \int_0^{x_0} b(x)\bigl[f(u) - uf'(u)\bigr]w(x)\,\mathrm{d}x \\
&\qquad + \int_{x_0}^1 b(x)\bigl[f(u) - uf'(u)\bigr]w(x)\,\mathrm{d}x \\
&\quad < \int_0^{x_0} b(x_0)\bigl[f(u) - uf'(u)\bigr]w(x)\,\mathrm{d}x \\
&\qquad + \int_{x_0}^1 b(x_0)\bigl[f(u) - uf'(u)\bigr]w(x)\,\mathrm{d}x \\
&\quad = b(x_0)\int_0^1 \bigl[f(u) - uf'(u)\bigr]w(x)\,\mathrm{d}x \leqslant 0.
\end{aligned} \tag{3.14}$$

On the other hand, multiplying Eq. (3.12) by u, Eq. (3.1) by w, subtracting and integrating over $(0, 1)$, we have

$$\int_0^1 b(x)\bigl[f(u) - uf'(u)\bigr]w(x)\,\mathrm{d}x = \frac{\mu}{\lambda}\int_0^1 uw\,\mathrm{d}x \geqslant 0. \tag{3.15}$$

We reach a contradiction by combining (3.14) and (3.15). □

As an application we have the following exact multiplicity result from P. Korman and J. Shi [40]. It extends the corresponding result in [34] by not restricting the behavior of $f(u)$ at infinity. Theorem 3.1 above allows us to conclude the uniqueness of the solution curve.

THEOREM 3.3. *We assume that* $b(x) \in C^1[-1, 1]$ *satisfies* $b(x) > 0$ *for* $x \in [-1, 1]$, *and* $b(x) = b(-x)$, $b'(x) < 0$ *for* $x \in (0, 1)$. *Assume* $f \in C^2[0, \infty)$, $f(u) > 0$, $f'(u) > 0$ *and*

$f''(u) > 0$ *for all* $u > 0$, *while* $h(\alpha) \leqslant 0$ *for some* $\alpha > 0$. *Then there exist two constants* $0 \leqslant \bar{\lambda} < \lambda_0$, *so that problem* (3.11) *has no solution for* $\lambda > \lambda_0$, *exactly two solutions for* $\bar{\lambda} < \lambda < \lambda_0$, *and in case* $\bar{\lambda} > 0$ *it has exactly one solution for* $0 < \lambda < \bar{\lambda}$. *Moreover, all solutions lie on a unique smooth solution curve. If we moreover assume that* $\lim_{u\to\infty} \frac{f(u)}{u} = \infty$, *then* $\bar{\lambda} = 0$.

EXAMPLE. Theorem 3.3 applies (with $\bar{\lambda} = 0$) to an example from combustion theory

$$u'' + \lambda b(x)\mathrm{e}^u = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0,$$

where $b(x)$ satisfies the above conditions.

P. Korman and T. Ouyang [34] have considered a class of *indefinite* problems

$$\begin{aligned} &u''(x) + \lambda u(x) + h(x)u^p(x) = 0, \quad -1 < x < 1, \\ &u(-1) = u(1) = 0. \end{aligned} \tag{3.16}$$

Here $p > 1$, and λ a real parameter. The given function $h(x)$ is assumed to be even, and it is allowed to change sign on $(-1, 1)$. By using bifurcation analysis, as above, as well as earlier work of T. Ouyang [48,49], it was possible to give an exhaustive description of the set of positive solutions of (3.16).

We denote by $\phi_1 = \cos\frac{\pi}{2}x$, the principal eigenfunction of $-u''$ on $(-1, 1)$, corresponding to the principal eigenvalue $\lambda_1 = \frac{\pi^2}{4}$. We assume that $h(x) \in C^1(-1, 1) \cap C^0[-1, 1]$ is an even function, and moreover,

$$h(0) > 0, \quad \text{and} \quad h'(x) < 0 \quad \text{for } x \in (0, 1), \tag{3.17}$$

$$\int_{-1}^{1} h(x)\phi_1^{p+1}(x)\,\mathrm{d}x < 0. \tag{3.18}$$

(Notice that the last assumption implies that $h(x)$ changes sign.)

THEOREM 3.4 [34]. *Assume that conditions* (3.17) *and* (3.18) *hold for problem* (3.16). *Then there is a critical* λ_0, $\lambda_0 > \lambda_1$, *so that for* $-\infty < \lambda \leqslant \lambda_1$ *problem* (3.16) *has exactly one positive solution, it has exactly two positive solutions for* $\lambda_1 < \lambda < \lambda_0$, *exactly one at* $\lambda = \lambda_0$, *and no positive solutions for* $\lambda > \lambda_0$. *Moreover, all positive solutions lie on a unique continuous in* λ *curve, which bifurcates from zero at* $\lambda = \lambda_1$ (*to the right*), *it continues without any turns to* $\lambda = \lambda_0$, *at which it turns to the left, and then continues without any more turns for all* $-\infty < \lambda \leqslant \lambda_0$. *We also have* $\max_x u(x) \to \infty$ *as* $\lambda \to -\infty$.

As far as we know, this is still the only known exact multiplicity result for indefinite problems.

Let us mention next the cubic problems

$$\begin{aligned} &u'' + \lambda\big(u - a(x)\big)\big(u - b(x)\big)\big(c(x) - u\big) = 0 \quad \text{for } -1 < x < 1, \\ &u(-1) = u(1) = 0, \end{aligned} \tag{3.19}$$

with given even functions $0 \leqslant a(x) \leqslant b(x) \leqslant c(x)$. As we discussed above, for constant a, b and c, the exact multiplicity question has been settled only recently by P. Korman, Y. Li and T. Ouyang [33], via a computer assisted proof. One can expect that under some conditions the same global picture holds for variable coefficients. This was established for several special cases. P. Korman, Y. Li and T. Ouyang [30] have given an exact multiplicity result in case $a = b = 0$. P. Korman and T. Ouyang [36] had done the same in case $a = 0$, and P. Korman and T. Ouyang [38] had given an exact multiplicity result in case when $a > 0$ is a constant.

Next we indicate an extension. Consider a problem with a variable diffusion coefficient

$$\big(a(x)u'\big)' + \lambda f(u) = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0. \tag{3.20}$$

We assume that a given function $a(x) \in C^1[-1, 1]$ is even, and it satisfies

$$a(x) > 0 \quad \text{and} \quad xa'(x) \leqslant 0, \quad \text{for } x \in [-1, 1]. \tag{3.21}$$

We perform a change of variables $x \to s$, given by

$$s = \int_0^x \frac{\mathrm{d}t}{a(t)}.$$

If we denote by $s_0 = \int_0^1 \frac{\mathrm{d}t}{a(t)}$, then this transformation gives a one-to-one map of the interval $(-1, 1)$ onto $(-s_0, s_0)$. Moreover, $s > 0$ (< 0) iff $x > 0$ (< 0). The problems (3.20) transforms into

$$\begin{aligned} &u_{ss} + \lambda a\big(x(s)\big) f(u) = 0 \quad \text{for } -s_0 < x < s_0, \\ &u(-s_0) = u(s_0) = 0. \end{aligned} \tag{3.22}$$

Observe that the function $s = s(x)$ is odd, and hence its inverse $x = x(s)$ is also odd, and then $a(x(s))$ is even. In view of (3.21)

$$\frac{\mathrm{d}}{\mathrm{d}s} a\big(x(s)\big) = a'\big(x(s)\big) a\big(x(s)\big) \leqslant 0 \ (\geqslant 0) \quad \text{if } s > 0 \ (< 0).$$

If we now assume that $f(u) > 0$ for $u > 0$, then the problem (3.22) satisfies conditions (3.2) and (3.3). Hence, we can translate our results, in particular Theorem 3.3, to problem (3.20).

4. General nonlinearities

Without the symmetry assumptions on $f(x, u)$ the problem is much harder. We restrict to a subclass of such problems, i.e. we now consider positive solutions of the boundary value problem

$$u'' + \lambda \alpha(x) f(u) = 0 \quad \text{for } a < x < b, \qquad u(a) = u(b) = 0, \tag{4.1}$$

on an arbitrary interval (a, b). We assume that $f(u)$ and $\alpha(x)$ are positive functions of class C^2, i.e.

$$f(u) > 0 \quad \text{for } u > 0, \qquad \alpha(x) > 0 \quad \text{for } x \in [a, b]. \tag{4.2}$$

As before, it will be crucial for bifurcation analysis to prove positivity for the corresponding linearized problem

$$w'' + \lambda\alpha(x) f'(u)w = 0 \quad \text{for } a < x < b, \qquad w(a) = w(b) = 0. \tag{4.3}$$

The following result was proved in P. Korman and T. Ouyang [38], although our exposition here is a little different.

LEMMA 4.1. *In addition to conditions* (4.2), *assume that*

$$\frac{3}{2}\frac{\alpha'^2}{\alpha} - \alpha'' < 0 \quad \textit{for all } x \in (a, b). \tag{4.4}$$

If the linearized problem (4.3) *admits a nontrivial solution*, *then we may assume that* $w(x) > 0$ *on* (a, b).

PROOF. Let $z(x) = g(x)u'(x)$, with $g(x)$ to be chosen shortly. Then $z(x)$ satisfies the equation

$$z'' + \lambda\alpha(x) f'(u)z = g''(x)u'(x) - \lambda\big(2g'(x)\alpha(x) + \alpha'(x)g(x)\big) f.$$

We now chose $g(x) = \alpha(x)^{-1/2}$. Then $2g'(x)\alpha(x) + \alpha'(x)g(x) = 0$, while

$$g''(x) < 0 \quad \text{for } a < x < b, \tag{4.5}$$

in view of the condition (4.4).

Notice that any positive solution of (4.1) is a concave function, and hence it has only one critical point, the point of global maximum. Let x_0 be the point of maximum of $u(x)$. We have

$$z'' + \lambda\alpha(x) f'(u)z = g''(x)u'(x), \tag{4.6}$$

with the right-hand side negative on (a, x_0) and positive on (x_0, b). This will make it impossible for $w(x)$ to vanish inside (a, b). Indeed, if we assume that $w(x)$ vanishes on say (x_0, b), we could find two consecutive roots of $w(x)$, $x_0 \leqslant x_1 < x_2 \leqslant b$ so that $w(x_1) = w(x_2) = 0$, while $w(x) > 0$ on (x_1, x_2). We now multiply the equation (4.6) by $w(x)$, Eq. (4.3) by $z(x)$, subtract and integrate over (x_1, x_2), obtaining

$$-g(x_2)u'(x_2)w'(x_2) + g(x_1)u'(x_1)w'(x_1) = \int_{x_1}^{x_2} g''(x)u'(x)w(x)\,\mathrm{d}x.$$

We have a contradiction, since the quantity on the left is negative, and the integral on the right is positive. □

REMARKS.

1. Recall the Schwarzian derivative from Complex Analysis and Dynamical Systems

$$(Sf)(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2.$$

   If one denotes $A(x) = \int \alpha(x)\,\mathrm{d}x$, then our condition (4.4) says that the Schwarzian derivative of $A(x)$ is positive.
2. Semilinear equations on an annulus in $\mathbb{R}^n$, $n > 2$, can be reduced by a standard change of variables to the problem (4.1), with $\alpha(x) = x^{-2k}$ and $k = 1 + \frac{1}{n-2}$, see, e.g., [19]. One sees that our condition (4.4) just misses this kind of functions. In [19] positivity of $w(x)$ was proved under an extra assumption that the annulus is "thin".

We shall present a new result on positivity of $w(x)$, after we prove a simple lemma.

LEMMA 4.2. *Assuming conditions* (4.2), *let* x_0 *be the unique point of maximum of the positive solution of* (4.1). *Assume that*

$$\alpha'(x) < 0 \quad \text{on } (x_0, b). \tag{4.7}$$

If the corresponding linearized problem (4.3) *admits a nontrivial solution* $w(x)$, *then this solution cannot vanish inside* (x_0, b).

PROOF. Assuming the contrary, let γ be the largest root of $w(x)$ on (x_0, b), and assume that $w(x) > 0$ on (γ, b). (The number of roots of $w(x)$ inside (a, b) is at most finite, as follows by the Sturm's comparison theorem, since both functions $f'(u(x))$ and $\alpha(x)$ are bounded on $[a, b]$, and λ is fixed. Hence, there is a largest root γ.) Differentiate Eq. (4.1)

$$u_x'' + \lambda\alpha(x) f'(u)u_u + \lambda\alpha'(x) f(u) = 0. \tag{4.8}$$

Multiplying Eq. (4.8) by $w(x)$, Eq. (4.3) by $u'(x)$, subtracting and integrating, we have

$$-u'(b)w'(b) + u'(\gamma)w'(\gamma) + \lambda\int_\gamma^b \alpha'(x) f\big(u(x)\big)w(x)\,\mathrm{d}x = 0.$$

This results in a contradiction, since all terms on the left are negative. □

REMARK. If $\alpha'(x) > 0$ on (a, x_0), then a similar proof shows that $w(x)$ cannot vanish inside (a, x_0).

LEMMA 4.3. *Assume conditions* (4.2) *hold*, *and in addition assume that*

$$\alpha'(x) < 0 \quad \text{on } (a, b), \tag{4.9}$$

and

$$2\alpha(x) + x\alpha'(x) > 0 \quad \text{on } (a,b). \tag{4.10}$$

If the linearized problem (4.3) *admits a nontrivial solution, then we may assume that* $w(x) > 0$ *on* (a,b).

PROOF. Let x_0 be the unique point of maximum of the solution $u(x)$. By the previous Lemma 4.2 it follows that $w(x)$ cannot vanish on (x_0, b). Assuming that $w(x)$ vanishes on $(a, x_0]$, let $\gamma \in (a, x_0]$ be the first root of $w(x)$, and we may assume that $w(x) > 0$ on (a, γ). We consider the function $\zeta(x) = x[u'(x)w'(x) + \lambda\alpha(x) f(u(x))w(x)] - u'(x)w(x)$, introduced by M. Tang [58]. One computes

$$\zeta'(x) = \lambda\big[2\alpha(x) + x\alpha'(x)\big] f(u)w. \tag{4.11}$$

Integrating over (a, γ),

$$\gamma u'(\gamma)w'(\gamma) - au'(a)w'(a) = \lambda \int_a^\gamma \big[2\alpha(x) + x\alpha'(x)\big] f(u)w \,\mathrm{d}x.$$

We have a contradiction, since the quantity on the left is negative, while the integral on the right is positive. □

REMARK. Our condition (4.10) again just misses the case of an annulus.

Positivity of $w(x)$ can be used to prove uniqueness and exact multiplicity results. For example, we can prove the following theorem.

THEOREM 4.1. *For problem* (4.1) *assume that conditions* (4.2) *hold, and that either condition* (4.4) *holds, or conditions* (4.9) *and* (4.10) *hold. In addition assume that* $f''(u) > 0$ *for all* $u > 0$, *and* $\lim_{u\to\infty} \frac{f(u)}{u} = \infty$. *Then there is a critical* $\lambda_0 > 0$, *so that problem* (4.1) *has exactly two positive solutions for* $0 < \lambda < \lambda_0$, *exactly one positive solution at* $\lambda = \lambda_0$, *and no positive solutions for* $\lambda > \lambda_0$. *Moreover, all solutions lie on a unique smooth solution curve, which starts at* $(\lambda = 0, u = 0)$, *bends back at* $\lambda = \lambda_0$, *and tends to infinity as* $\lambda \to 0$.

PROOF. The proof is similar to that of Theorem 2.2, except for proving the uniqueness of the solution curve (since the maximum value of the solution no longer identifies that solution). However, if another solution curve existed, one of its ends would have to go through the point $(\lambda = 0, u = 0)$, contradicting the uniqueness of solutions near regular points, which follows by the implicit function theorem. (Since $w > 0$ at the turning point, one of the branches is increasing in λ, i.e. it is decreasing for decreasing λ. By the arguments of [34], or [35], the monotonicity is preserved along the branch, and hence this branch must go into the origin, as $\lambda \to 0$.) □

EXAMPLE. The theorem applies to the problem

$$u'' + \lambda\alpha(x)e^u = 0, \quad a < x < b, \qquad u(a) = u(b) = 0,$$

if $\alpha(x) > 0$ satisfies either condition (4.4), or conditions (4.9) and (4.10).

5. Time maps

5.1. *There are several different formulas for the time map*

Let $u = u(t)$ be solution of the initial value problem,

$$u'' + f(u) = 0, \qquad u(0) = 0, \quad u'(0) = p.$$

Using ballistic analogy, we can interpret this as "shooting" from the ground level, at an angle $p > 0$. Let $T/2$ denote the time it takes for the projectile to reach its maximum amplitude α, $\alpha = \alpha(p)$. By symmetry of positive solutions, $T = T(p)$ is then the time when the projectile falls back to the ground, the *time map*. Since the energy is constant (as before, $F(u) = \int_0^u f(t)\,dt$)

$$\frac{1}{2}\left(\frac{du}{dt}\right)^2 + F\big(u(t)\big) = F(\alpha) = \frac{1}{2}p^2.$$

Solving this for $\frac{dt}{du}$, and integrating

$$T/2 = \frac{1}{\sqrt{2}}\int_0^{\alpha} \frac{du}{\sqrt{F(\alpha) - F(u)}}, \tag{5.1}$$

which lets us compute $T = T(\alpha)$ (or $T = T(p)$, since $\alpha = \alpha(p)$). This formula has been used extensively for a long time, see, e.g., W.S. Loud [43], T. Laetsch [41], K.J. Brown et al. [9], J. Smoller and A. Wasserman [57], S.-H. Wang [59–61], I. Addou [1,2], I. Addou and S.-H. Wang [3], S.-H. Wang and T.S. Yeh [63], and J. Cheng [10,11]. It is not easy to use this formula. The integral is improper at $u = \alpha$, so that one needs a regularizing substitution before differentiating in α. One regularizing substitution is $u = \alpha \sin\theta$, which gives

$$T/2 = \frac{\alpha}{\sqrt{2}}\int_0^{\pi/2} \frac{\cos\theta\,d\theta}{\sqrt{F(\alpha) - F(\alpha\sin\theta)}}. \tag{5.2}$$

The integrand is now bounded. Formula (5.2) can be used for numerical computations, as well as for proving theorems. For more information we refer the reader to the above mentioned papers, particularly to the recent papers of S.-H. Wang and his coworkers.

5.2. *Time map formula through global linearization*

We are interested in positive solutions of the two point problem for $u = u(t)$

$$u'' + f(u) = 0, \quad 0 < t < T, \qquad u(0) = u(T) = 0. \tag{5.3}$$

We do not consider the end point T to be fixed, but rather depending on $p = u'(0)$ (or on the maximum value of the solution α, $\alpha = u(T/2)$). To obtain the formula for $T = T(p)$, we begin by transforming (5.3) into the system form

$$\begin{aligned} u' &= y, \\ y' &= -f(u), \end{aligned} \tag{5.4}$$

together with the initial conditions

$$u(0) = 0, \quad y(0) = p. \tag{5.5}$$

Let $F(u) = \int_0^u f(t)\,\mathrm{d}t$. In the linear case when $f(u) = u$, we have $F(u) = \frac{1}{2}u^2$. We now define the function $g(x)$, for $x \geqslant 0$, by

$$F\bigl(g(x)\bigr) = \frac{1}{2}x^2. \tag{5.6}$$

In other words, $g(x) = F^{-1}(\frac{1}{2}x^2)$, and the inverse function F^{-1} is defined, provided we assume throughout this section that $f(u) \in C^2(0,a) \cap C[0,a]$ for some $0 < a \leqslant \infty$, and

$$f(u) > 0 \quad \text{for } u \in (0,a). \tag{5.7}$$

We assume also that

$$\text{either } f(0) > 0, \quad \text{or} \quad f(0) = 0 \text{ and } f'(0) > 0. \tag{5.8}$$

Differentiate (5.6)

$$f\bigl(g(x)\bigr)g'(x) = x. \tag{5.9}$$

In (5.4) we let $u = g(x)$, then multiply the second equation by $g'(x)$, and use (5.9)

$$\begin{aligned} g'(x)x' &= y, \\ g'(x)y' &= -f\bigl(g(x)\bigr)g'(x) = -x. \end{aligned} \tag{5.10}$$

We now change the independent variable in (5.10), $t \to \theta$, by solving

$$\frac{\mathrm{d}t}{\mathrm{d}\theta} = g'\bigl(x(t)\bigr), \qquad t(0) = 0. \tag{5.11}$$

Then the system (5.4) is linearized, and problem (5.4), (5.5) transforms into

$$\begin{aligned}
&\frac{\mathrm{d}x}{\mathrm{d}\theta} = y,\\
&\frac{\mathrm{d}y}{\mathrm{d}\theta} = -x,\\
&x(0) = 0, \quad y(0) = p.
\end{aligned} \tag{5.12}$$

Solution of (5.12) is

$$x = p\sin\theta, \quad y = p\cos\theta.$$

Using this in (5.11) and integrating, we have the formula for the time map

$$T = \int_0^{\pi} g'(p\sin\theta)\,\mathrm{d}\theta. \tag{5.13}$$

This formula was derived by R. Schaaf [53], and was used by her to obtain a number of uniqueness and multiplicity results.

Separating variables in (5.11) and integrating

$$\int_0^T \frac{f(g(x(t)))}{x(t)}\,\mathrm{d}t = \pi, \tag{5.14}$$

where we have used (5.9) to express g'. From the definition of $g(x)$ we have

$$x = \sqrt{2F(g(x))} = \sqrt{2F(u)},$$

and hence we can rewrite (5.14) as

$$\int_0^T \frac{f(u(t))}{\sqrt{F(u(t))}}\,\mathrm{d}t = \sqrt{2}\pi. \tag{5.15}$$

This formula was derived in a different way by P. Korman and Y. Li [28], where the quantity on the left was referred to as "generalized average" of the solution of (5.3). The reason why this term was chosen is that in case $f(u) = u^3$, this formula gives the average value of the solution: $\int_0^T u(t)\,\mathrm{d}t = \frac{\pi}{\sqrt{2}}$.

REMARKS.

1. We needed the positivity of $f(u)$ so that the inverse function F^{-1} is defined, however there is no need to distinguish between $f(0) = 0$ and $f(0) > 0$ cases for both formulas (5.13) and (5.15). In case $f(0) > 0$ the integral in (5.15) (and in (5.14)) is improper at both end points, however since $u'(0) \neq 0$ and $u'(T) \neq 0$, the integral converges. (For small t, $u(t) \sim u'(0)t$, $F(u(t)) \sim f(0)u(t) \sim f(0)u'(0)t$.)

2. Similarly, in the derivation of the time map formula (5.13), we run into an improper integral in case $f(0) > 0$. Indeed, when solving for $\theta = \theta(t)$ in (5.11), we have $\theta = \int_0^t \frac{\mathrm{d}s}{g'(x(s))}$, which is an improper integral at $s = 0$. However, as we have just seen, it is a convergent integral. Hence the time map formula in (5.13) is valid in both cases $f(0) = 0$, and $f(0) > 0$.
3. Let us collect the properties of the function $g(x)$. We have $g(0) = 0$, $g'(x) = \frac{\sqrt{2F(g(x))}}{f(g(x))} > 0$ for $x > 0$. We also have $g'(0) = 0$ in case $f(0) > 0$, and, by L'Hopital's rule as was observed in R. Schaaf [53], $g'(0) = \frac{1}{\sqrt{f'(0)}}$ in case $f(0) = 0$. Observe that $g'(0)$ is defined, thanks to condition (5.8).

It is sometimes more convenient to express $T = T(\alpha)$, where α is the maximum value of the solution, $\alpha = u(T/2)$. Since the energy $\frac{1}{2}u'^2(x) + F(u(x))$ is constant, it follows that $p = \sqrt{2F(\alpha)}$, and hence

$$T/2 = \int_{\pi/2}^{\pi} g'(p\sin\theta)\,\mathrm{d}\theta = \int_0^{\pi/2} g'\Bigl(\sqrt{2F(\alpha)}\cos\theta\Bigr)\,\mathrm{d}\theta. \tag{5.16}$$

$T/2$ is, of course, the time it takes the solution to travel from its maximum to zero.

EXAMPLE. Consider the problem ($u = u(x)$)

$$u'' + \lambda \mathrm{e}^u = 0, \quad x \in (0,1), \qquad u(0) = u(1) = 0. \tag{5.17}$$

By a change of independent variable we convert it to

$$u'' + \mathrm{e}^u = 0, \quad x \in (0,1), \qquad u(0) = u(T) = 0, \tag{5.18}$$

where $T = \sqrt{\lambda}$. Here $f(u) = \mathrm{e}^u$, $F(u) = \mathrm{e}^u - 1$, and $g(x) = \ln(\frac{1}{2}x^2 + 1)$. The integral in (5.16) is then relatively simple, and in fact *Mathematica* gives

$$\sqrt{\lambda}/2 = \sqrt{\frac{2}{\mathrm{e}^\alpha}}\,\mathrm{ArcTanh}\left[\sqrt{\frac{\mathrm{e}^\alpha - 1}{\mathrm{e}^\alpha}}\right]. \tag{5.19}$$

Plotting this formula (with λ along the horizontal axis and α along the vertical one), we obtain the same bifurcation diagram as obtained by standard integration. Computation this way is considerably faster than by integration, and we also observe that here $f(0) > 0$. Formula (5.19) can also be obtained by explicit integration of problem (5.17).

We can proceed similarly for the general case

$$u'' + \lambda f(u) = 0, \quad x \in (0,1), \qquad u(0) = u(1) = 0. \tag{5.20}$$

We do not have a simple formula for $g(x)$ anymore, however from (5.16) (see also (5.9)) we obtain (as before, $T = \sqrt{\lambda}$)

$$\sqrt{\lambda}/2 = \int_0^{\pi/2} \frac{\sqrt{2F(\alpha)}\cos\theta}{f(F^{-1}(F(\alpha)\cos^2\theta))}\,\mathrm{d}\theta. \tag{5.21}$$

This formula will provide probably one of the most efficient ways to compute the bifurcation diagrams, once the evaluation of the inverse function F^{-1} is numerically implemented.

Assume now there is an $a > 0$, so that $f(a) = 0$, and $f(u) > 0$ for $u > a$, while no assumptions on the sign of $f(u)$ are made when $u \in (0, a)$. If we denote by $T_1/2$ the time it takes the solution to travel from its maximum to $u(x) = a$, then

$$T_1/2 = \int_{\theta_0}^{\pi/2} g'\Big(\sqrt{2F(\alpha)}\cos\theta\Big)\,\mathrm{d}\theta, \quad \text{where } \theta_0 = \sin^{-1}\sqrt{\frac{F(a)}{F(\alpha)}}. \tag{5.22}$$

The following theorem we proved in [27].

THEOREM 5.1. *Assume that for some* $0 < a < b \leqslant \infty$ *we have*

$$f(u) > 0 \quad \textit{for } a < u < b, \tag{5.23}$$

$$f'(u)\int_a^u f(t)\,\mathrm{d}t - \frac{1}{2}f^2(u) > 0 \quad \textit{for } a < u < b. \tag{5.24}$$

(*Observe that we implicitly assume that* $f(a) = 0$.) *Then the problem*

$$u'' + f(u) = 0, \quad x \in (0,1), \qquad u(0) = u(1) = 0$$

has at most one positive solution, *with* $a < \alpha = u(1/2) < b$.

What is remarkable here is than no assumptions whatsoever are made on $f(u)$ when $u \in (0, a)$. We used generalized averages to prove this result, but it should be possible to obtain it from formula (5.22) too. In fact, there is a similar result in R. Schaaf's book [53] (a little less general than the above theorem). Observe that (5.24) will follow if $f(a) = 0$ and $f''(u) > 0$ for $a < u < b$. A more general result for p-Laplacian case has been given recently by J. Cheng [10].

We now show how the time map formula gives rise to uniqueness and multiplicity results for the Dirichlet problem. Compute

$$\frac{\mathrm{d}T}{\mathrm{d}p} = \int_0^{\pi} g''(p\sin\theta)\sin\theta\,\mathrm{d}\theta, \tag{5.25}$$

where $g''(x)$ is computed from (5.9) (written in the form $g'(x) = \frac{\sqrt{2F(u)}}{f(u)}$, $u = g(x)$)

$$g''(x) = \frac{f^2 - 2Ff'}{f^3}(u), \quad \text{with } u = g(x). \tag{5.26}$$

If the time map $T(p)$ is monotone, then clearly the positive solution of the Dirichlet problem (5.3) for any *fixed* T is unique. Hence, we have uniqueness of solutions if either

$$I(u) \equiv f'(u)F(u) - \frac{1}{2}f^2(u) > 0 \quad \text{for almost all } u > 0, \tag{5.27}$$

or the opposite inequality holds. This condition was derived by R. Schaaf [53], and it also follows from the generalized averages in [28].

We observe next that this condition does not add anything to the standard uniqueness condition

$$uf'(u) - f(u) \quad \text{does not change sign for } u > 0. \tag{5.28}$$

I.e. (5.28) holds whenever (5.27) does, and so condition (5.28) is both simpler and more general.

Indeed, we begin by observing

$$\frac{\mathrm{d}^2}{\mathrm{d}u^2}\Bigl(\sqrt{F(u)}\Bigr) = \frac{I(u)}{2F^{3/2}(u)} \equiv J(u),$$

where $J(u)$ has the same sign as $I(u)$. Integrating between some $a > 0$ and $u > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}u}\Bigl(\sqrt{F(u)}\Bigr) = \int_a^u J(\xi)\,\mathrm{d}\xi + c > 0, \tag{5.29}$$

where $c = \frac{f(a)}{2\sqrt{F(a)}} > 0$. Integrating (5.29),

$$\sqrt{F(u)} = cu + c_1 + \int_a^u (u-\xi)J(\xi)\,\mathrm{d}\xi, \tag{5.30}$$

where $c_1 = -\int_0^a \xi J(\xi)\,\mathrm{d}\xi$. From (5.30) we find $F(u)$, and then $f(u)$ and $f'(u)$ by differentiation. We then have

$$uf'(u) - f(u) = 2\sqrt{F(u)}J(u)u + 2\int_0^u \xi J(\xi)\,\mathrm{d}\xi\left(\int_a^u J(\xi)\,\mathrm{d}\xi + c\right).$$

In view of (5.29), the quantity in the bracket is positive, and it follows that if $J(u)$ is positive (negative), so is $uf'(u) - f(u)$.

We now consider the problem

$$u'' + \lambda f(u) = 0, \quad x \in (0,1), \qquad u(0) = u(1) = 0, \tag{5.31}$$

depending on a positive parameter λ. As before, we can convert it to problem (5.3), with $T = \sqrt{\lambda}$. If we can show that $T''(p) > 0$ (or $T''(p) < 0$) for all $p > 0$, it will follow that for any λ there is at most two p's with $T(p) = \sqrt{\lambda}$, i.e. at most two solutions of (5.31). Since

$$T''(p) = \int_0^\pi g'''(p\sin\theta)\sin^2\theta\,\mathrm{d}\theta,$$

it suffices to show that the function $g'''(u)$ keeps the same sign. By formula (1-1-15) in R. Schaaf [53]

$$g'''(u) = -g'(x)\frac{3f'(u)(f^2(u) - 2F(u)f'(u)) + 2F(u)f(u)f''(u)}{f^4(u)},$$
$$\text{with } u = g(x),$$

which led her to the following condition: if

$$3f'(u)\bigl(f^2(u) - 2F(u)f'(u)\bigr) + 2F(u)f(u)f''(u) > 0 \text{ (or } < 0),$$
$$\text{for all } u > 0 \tag{5.32}$$

then problem (5.31) has at most two positive solutions. Since condition (5.32) is not easy to verify, R. Schaaf [53] went on to develop her $A - B$ and C conditions, which are sufficient for (5.32) to hold.

Condition (5.32) says that $g'(u)$ is either convex or concave. Working with the generalized averages, P. Korman and Y. Li [28] have shown that the same result is true if $\frac{1}{g'(u)}$ is convex (also in the case $\frac{1}{g'(u)}$ concave, but this possibility is included in the case when $g'(u)$ is concave). This led them to the following condition: if

$$\frac{1}{2}f''(u)F^2(u) + \frac{3}{8}f^3(u) - \frac{3}{4}f(u)f'(u)F(u) > 0 \quad \text{for all } u > 0, \tag{5.33}$$

then problem (5.31) has at most two positive solutions. Observe that this condition is different from (5.32). Conditions (5.32) and (5.33) work in both cases $f(0) = 0$ and $f(0) > 0$. Also, computer algebra can help in verifying these conditions.

EXAMPLE. The function $f(u) = 2 + e^{-u}\sin u$ satisfies (5.33). This function changes concavity infinitely many times. A straightforward computation, using *Mathematica*, shows that for this function the left-hand side of (5.33) is positive, tending to 10.125 as $u \to \infty$. Hence, problem (5.31) with this $f(u)$ has at most two positive solutions for any $\lambda > 0$.

REMARKS.

1. The time map formula can be also developed for the p-Laplacian case. Actually, even more general case is developed in Section 2.5 of R. Schaaf's book [53].
2. Finally, we mention why we constantly stress that all results about the time map hold in both cases $f(0) = 0$ and $f(0) > 0$. The important book by R. Schaaf [53] treats the $f(0) = 0$ case in Chapter 1, while the case $f(0) > 0$ (and also the case $f(0) < 0$) is postponed to Chapter 3. Some readers might form an incorrect impression that the book covers only the $f(0) = 0$ case (as in the MathSciences Review of that book).

5.3. *Variational formula for the time map*

In addition to the two formulas for the time map, discussed above, a curious variational formula has been discovered by R. Benguria and M.C. Depassier, see [8], which has also references to their earlier papers. If $u(t)$ is a solution of

$$u'' + \lambda f(u) = 0, \quad 0 < t < 1, \qquad u(0) = \alpha, \quad u'(0) = u(1) = 0, \tag{5.34}$$

it is shown by R. Benguria and M.C. Depassier that

$$\lambda = \max_{g \in D} \frac{1}{2} \frac{(\int_0^\alpha g'(y)^{1/3}\,\mathrm{d}y)^3}{\int_0^\alpha f(y)g(y)\,\mathrm{d}y},$$

where $D = \{g | g \in C^1(0, \alpha), g' > 0, g(0) = 0\}$. By rescaling, this formula is of course equivalent to a time map formula. It was used in [8] to obtain lower and upper bounds for time maps.

5.4. *A nonlocal problem*

Using the generalized inverses, we now give a complete description of the solution set of a nonlocal problem. We begin with a simple observation. It is well known that for any $L > 0$ the problem (here $u = u(x)$)

$$u'' + u^3 = 0, \quad 0 < x < L, \qquad u(0) = u(L) = 0$$

has a unique positive solution, and a unique negative solution. If we now take a positive solution on the interval $(0, L/k)$, followed by the negative solution on $(L/k, 2L/k)$, and so on, then we obtain a solution with $k - 1$ sign changes, for any positive integer k.

We now consider a nonlocal problem, where instead of a second boundary condition we prescribe the average value of the solution on some fixed interval $(0, L)$

$$\begin{aligned} &u'' + u^3 = 0, \quad 0 < x < L, \\ &u(0) = 0, \\ &\int_0^L u(s)\,\mathrm{d}s = \alpha, \end{aligned} \tag{5.35}$$

where α is a prescribed constant. We are interested in both positive, negative and sign-changing solutions, i.e. we shall talk of solutions with k sign changes, where $k \geqslant 0$. Without loss of generality we may assume $\alpha \geqslant 0$ (otherwise, consider $v = -u$). If $\alpha = 0$, it is clear that there exists exactly two solution of (5.35) with k sign changes, for any odd $k \geqslant 1$. Indeed, a solution of Eq. (5.35) with $u(0) = u(L) = 0$ having an odd number of roots inside $(0, L)$, and its negative, provide the desired solutions of (5.35). So that we may assume $\alpha > 0$.

THEOREM 5.2 [27]. *For any* $0 < \alpha < \frac{\pi}{\sqrt{2}}$ *there exists exactly one solution of* (5.35) *with k sign changes, for any* $k \geqslant 0$. *For* $\alpha = \frac{\pi}{\sqrt{2}}$ *there exists exactly one solution with k sign changes, for any even* $k \geqslant 0$, *and no solutions if k is odd. For any* $\alpha > \frac{\pi}{\sqrt{2}}$ *problem* (5.35) *has no solutions.*

PROOF. The problem "scales right". Setting $x = bt$, and $u = \frac{1}{b}v$, we see that $v = v(t)$ satisfies

$$
\begin{aligned}
&v'' + v^3 = 0, \quad 0 < t < \frac{L}{b}, \\
&v(0) = 0, \\
&\int_0^{L/b} v(s)\,\mathrm{d}s = \alpha.
\end{aligned}
\tag{5.36}
$$

Comparing with (5.35), we see that only the length of the interval has changed. Hence we have a one-to-one map between the solution sets on any two intervals. So consider a solution $U(x)$ of the equation $u'' + u^3 = 0$, with $u(0) = 0$, which has k sign changes, whose roots are $x = 1, 2, \ldots$, and such that $U(x) > 0$ on $(0, 1)$, $U(x) < 0$ on $(1, 2)$, and so on. According to formula (5.15), the integral of $U(x)$ over any of its positive humps is equal to $\frac{\pi}{\sqrt{2}}$, while the integral of $U(x)$ over any of its negative humps is $-\frac{\pi}{\sqrt{2}}$. Imagine cutting this solution with a sliding vertical line $x = \xi$. By continuity, for any $\alpha \in (0, \frac{\pi}{\sqrt{2}}]$ we can find a unique $\xi \in (0, 1]$ so that $U(x)$ is positive solution of (5.35) on the interval $(0, \xi)$. We then map this solution to the original interval $(0, L)$ by the above transformation. Similarly, for any $\alpha \in (0, \frac{\pi}{\sqrt{2}})$ we can find a unique $\xi \in (1, 2)$ so that we have a solution of (5.35) on the interval $(0, \xi)$, with exactly one sign change. We then map $U(x)$ to the original interval, as before. Similarly we construct solutions with arbitrarily many sign changes.

By (5.15), no solution is possible in case $\alpha > \frac{\pi}{\sqrt{2}}$. □

6. Numerical computation of solutions

Good analytical understanding of a problem goes hand in hand with efficient numerical calculation of its solution. We know that for positive solutions the maximum value $u(0) = \alpha$ uniquely determines the solution pair $(\lambda, u(x))$ of the problem

$$
u'' + \lambda f(u) = 0 \quad \text{for } -1 < x < 1, \qquad u(-1) = u(1) = 0, \tag{6.1}
$$

see Lemma 2.3 above. We also know that the parameter λ in (6.1) can be "scaled out", i.e. $v(x) \equiv u(\frac{1}{\sqrt{\lambda}}x)$ solves the equation $v'' + f(v) = 0$, while $v(0) = u(0) = \alpha$, and $v'(0) = u'(0) = 0$. The root of $v(x)$ is $r = \sqrt{\lambda}$. We therefore solve the initial value problem

$$
v'' + f(v) = 0, \qquad v(0) = \alpha, \quad v'(0) = 0, \tag{6.2}
$$

```
Clear ["Global`*"]
f[u_] = Exp [5u/(5+u)];
t = {};
Δα = 0.2; α0 = 0; nsteps = 150;
Do[up = α0 + j * Δα;
  If[(NIntegrate[f[t], {t, 0, up}] > 0 && f[up] > 0),
    α[j] = α0 + j * Δα;
    Int[θ_] := NIntegrate[f[u], {u, α[j] Sin[θ], α[j]}];
    λ[j] = 2 α[j]^2 (NIntegrate[Cos[θ]/√Int[θ], {θ, 0, Pi/2}])^2;
    t = Append[t, {λ[j], α[j]}]
  ]
  , {j, 1, nsteps}
]
ListPlot[t , PlotJoined → True,
  AxesLabel → {"Lambda", "u(0)"}, PlotLabel → "f(u)=Exp(5u/(5+u))"]
```

and find its first positive root r. Then $\lambda = r^2$ by the above remarks. This way for each α we can find the corresponding λ. After we choose sufficiently many α_n and compute the corresponding λ_n, we can plot the pairs (λ_n, α_n), obtaining a bifurcation diagram in (λ, α) plane. We stress that the resulting two-dimensional bifurcation curve gives a faithful representation of the solution set of (6.1), since the value $u(0) = \alpha$ uniquely determines the solution pair $(\lambda, u(x))$. The program for solving (6.1) is essentially one short loop, involving the **NDSolve** command in *Mathematica*. It can be found at the author's webpage: http://math.uc.edu/~kormanp/.

An equally good way to do numerical computations is by direct integration. For problem (6.2) we have $r = T/2$, where as before r is the first positive root, and T is the time map. I.e. $\lambda = T^2/4$. Using formula (5.2) for the time map, we have

$$\lambda = \frac{1}{2}\alpha^2 \left(\int_0^{\pi/2} \frac{\cos\theta}{\sqrt{\int_{\alpha\sin\theta}^{\alpha} f(u)\,\mathrm{d}u}}\,\mathrm{d}\theta \right)^2. \tag{6.3}$$

The *Mathematica* program based on (6.3) is so short and simple, that we include its listing here. It solves problem (6.1) for $f(u) = \mathrm{e}^{5u/(5+u)}$, and produces an S-shaped bifurcation curve, in agreement with our results. (Our program is solving the Dirichlet problem on the interval (0, 1), rather than (−1, 1), which accounts for the extra factor of 4.)

We see absolutely no need to ever use finite differences (or finite elements) for problem (6.1). If we divide the interval (0, 1) into n pieces, with step $h = 1/n$ and subdivision points $x_i = ih$, and denote by u_i the numerical approximation of $u(x_i)$, the finite difference approximation of (6.1) is

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + \lambda f(u_i) = 0, \quad 1 \leqslant i \leqslant n-1, \qquad u_0 = u_n = 0. \tag{6.4}$$

This is a system of *nonlinear* algebraic equations, more complicated in every way than the original problem (6.1). In particular, this system often has more solutions than the corresponding differential equation (6.1). The existence of the extra solutions (not corresponding to the solutions of (6.1)) has been recognized for a while, and a term *spurious* *solutions* has been used. For example in case $f(u) = \mathrm{e}^u$ the solution curve of (6.1) has exactly one turn (as we proved before), while the solution curve of (6.4) has three turns, see P. Korman [25]. Increasing the number of subdivision points n does not remove the two spurious turns, it just moves them closer to $\lambda = 0$. Actually the spurious turns are avoided in the opposite direction, when $n \leqslant 6$. We found this hard to prove, even when $n = 2$. When studying problem (6.4), we can no longer rely on the familiar tools from differential equations. Even in the case $f(u) = u^k$ the analysis of problem (6.4) is very involved, see E.L. Allgower [4].

For the general problem (1.1) (with $f = f(x, u)$) we suggest using the predictor-corrector method. If solution $u(x, \lambda)$ is known, one approximates

$$u(x, \lambda + \Delta\lambda) \simeq u(x, \lambda) + u_\lambda(x, \lambda)\Delta\lambda, \tag{6.5}$$

and then a very accurate approximation of $u(x, \lambda + \Delta\lambda)$ can be usually obtained in around 4 steps of Newton's iteration, with the initial guess given by (6.5). This way we can continue the solution in λ. To find $u_\lambda(x, \lambda)$ one solves a linear problem

$$\begin{aligned} &u_\lambda'' + f_u(x, u)u_\lambda + f(x, u) = 0 \quad \text{for } -1 < x < 1, \\ &u_\lambda(-1) = u_\lambda(1) = 0. \end{aligned} \tag{6.6}$$

To solve (6.6) one uses finite differences. (There are no spurious solutions for linear problems!) The resulting tri-diagonal system is easily solved by Gaussian elimination.

Acknowledgement

It is a pleasure to thank Idris Addou for many useful comments.

References

[1] I. Addou, *Multiplicity of solutions for quasilinear elliptic boundary-value problems*, Electron. J. Differential Equations **21** (1999).

[2] I. Addou, *Multiplicity results for classes of one-dimensional p-Laplacian boundary-value problems with cubic-like nonlinearities*, Electron. J. Differential Equations **52** (2000).

[3] I. Addou and S.-H. Wang, *Exact multiplicity results for a p-Laplacian problem with concave-convex-concave nonlinearities*, Nonlinear Anal. **53** (2003), 111–137.

[4] E.L. Allgower, *On a discretization of* $y'' + \lambda y^k = 0$, Topics in Numerical Analysis, Vol. II, Proc. Roy. Irish Acad. Conf., Univ. College, Dublin, 1974, Academic Press, London (1975), 1–15.

[5] A. Ambrosetti and G. Prodi, *A Primer of Nonlinear Analysis*, Cambridge University Press (1993).

[6] V. Anuradha and R. Shivaji, *Existence of infinitely many non-trivial bifurcation points*, Results Math. **22** (1992), 641–650.

[7] J. Bebernes and D. Eberly, *Mathematical Problems from Combustion Theory*, Springer (1989).

[8] R. Benguria and M.C. Depassier, *Variational calculation of the period of nonlinear oscillators*, J. Statist. Phys. **116** (2004), 923–931.

[9] K.J. Brown, M.M.A. Ibrahim and R. Shivaji, *S-shaped bifurcation curves*, Nonlinear Anal. **5** (1981), 475–486.

[10] J. Cheng, *Uniqueness results for the one-dimensional p-Laplacian*, J. Math. Anal. Appl. **311** (2005), 381–388.

[11] J. Cheng, *Exact number of positive solutions for semipositone problems*, J. Math. Anal. Appl. **313** (2006), 322–341.

[12] M.G. Crandall and P.H.Rabinowitz, *Bifurcation from simple eigenvalues*, J. Funct. Anal. **8** (1971), 321–340.

[13] M.G. Crandall and P.H.Rabinowitz, *Bifurcation, perturbation of simple eigenvalues and linearized stability*, Arch. Rational Mech. Anal. **52** (1973), 161–180.

[14] Y. Du and Y. Lou, *Proof of a conjecture for the perturbed Gelfand equation from combustion theory*, J. Differential Equations **173** (2001), 213–230.

[15] B. Gidas, W.-M. Ni and L. Nirenberg, *Symmetry and related properties via the maximum principle*, Comm. Math. Phys. **68** (1979), 209–243.

[16] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Math., Vol. 840, Springer, New York (1981).

[17] X. Hou, P. Korman and Y. Li, *On the direction of the pitchfork bifurcation*, Preprint.

[18] P. Korman, *Symmetry of positive solutions for elliptic problems in one dimension*, Appl. Anal. **58** (1995), 351–365.

[19] P. Korman, *Uniqueness and exact multiplicity results for two classes of semilinear problems*, Nonlinear Anal. **31** (1998), 849–865.

[20] P. Korman, *The global solution set for a class of semilinear problems*, J. Math. Anal. Appl. **226** (1998), 101–120.

[21] P. Korman, *Curves of sign-changing solutions for semilinear equations*, Nonlinear Anal. TMA **51** (2002), 801–820.

[22] P. Korman, *Global solution curves for semilinear systems*, Math. Methods Appl. Sci. **25** (2002), 3–20.

[23] P. Korman, *Exact multiplicity of solutions for a class of semilinear Neumann problems*, Comm. Appl. Nonlinear Anal. **9** (2002), 1–17.

[24] P. Korman, *Similarity of solution branches for two-point semilinear problems*, Proceedings of the Fifth Mississippi State Conference on Differential Equations and Computational Simulations, Mississippi State, MS, 2001, Electron. J. Differential Equations **10**, Southwest Texas State Univ., San Marcos, TX (2003), 187–191 (electronic).

[25] P. Korman, *An accurate computation of the global solution curve for the Gelfand problem through a two point approximation*, Appl. Math. Comput. **139** (2003), 363–369.

[26] P. Korman, *Uniqueness and exact multiplicity of solutions for a class of fourth-order semilinear problems*, Proc. Roy. Soc. Edinburgh Sect. A **134** (2004), 179–190.

[27] P. Korman, *Application of generalized averages to uniqueness of solutions, and to a non-local problem*, Comm. Appl. Nonlinear Anal. **12** (2005), 69–78.

[28] P. Korman and Y. Li, *Generalized averages of solutions of two-point Dirichlet problems*, J. Math. Anal. Appl. **239** (1999), 478–484.

[29] P. Korman and Y. Li, *On the exactness of an S-shaped bifurcation curve*, Proc. Amer. Math. Soc. **127** (1999), 1011–1020.

[30] P. Korman, Y. Li and T. Ouyang, *Exact multiplicity results for boundary-value problems with nonlinearities generalising cubic*, Proc. Roy. Soc. Edinburgh, Sect. A **126** (1996), 599–616.

[31] P. Korman, Y. Li and T. Ouyang, *An exact multiplicity result for a class of semilinear equations*, Comm. Partial Differential Equations **22** (1997), 661–684.

[32] P. Korman, Y. Li and T. Ouyang, *Perturbation of global solution curves for semilinear problems*, Adv. Nonlinear Stud. **3** (2003), 289–299.

[33] P. Korman, Y. Li and T. Ouyang, *Computing the location and the direction of bifurcation*, Math. Res. Lett. **12** (2005), 933–944.

[34] P. Korman and T. Ouyang, *Exact multiplicity results for two classes of boundary value problems*, Differential Integral Equations **6** (1993), 1507–1517.

[35] P. Korman and T. Ouyang, *Multiplicity results for two classes of boundary-value problems*, SIAM J. Math. Anal. **26** (1995), 180–189.

[36] P. Korman and T. Ouyang, *Exact multiplicity results for a class of boundary-value problems with cubic nonlinearities*, J. Math. Anal. Appl. **194** (1995), 328–341.

[37] P. Korman and T. Ouyang, *Exact multiplicity results for two classes of periodic equations*, J. Math. Anal. Appl. **194** (1995), 763–779.

[38] P. Korman and T. Ouyang, *Solution curves for two classes of boundary-value problems*, Nonlinear Anal. **27** (1996), 1031–1047.

[39] P. Korman and T. Ouyang, *Multiplicity and Morse indices of sign-changing solutions for semilinear equations*, Canad. Appl. Math. Quart. **7** (1999), 239–250.

[40] P. Korman and J. Shi, *Instability and exact multiplicity of solutions of semilinear equations*, Proceedings of the Conference on Nonlinear Differential Equations, Coral Gables, FL, 1999, Electron. J. Differential Equations **5**, Southwest Texas State Univ., San Marcos, TX (2000), 311–322 (electronic).

[41] T. Laetsch, *The number of solutions of a nonlinear two point boundary value problem*, Indiana Univ. Math. J. **20** (1970/1971).

[42] C.S. Lin and W.-M. Ni, *A counterexample to the nodal domain conjecture and a related semilinear equation*, Proc. Amer. Math. Soc. **102** (1988), 271–277.

[43] W.S. Loud, *Periodic solutions of* $x'' + cx' + g(x) = \varepsilon f(t)$, Mem. Amer. Math. Soc. **31** (1959), 58 pp.

[44] H.P. McKean and J.C. Scovel, *Geometry of some simple nonlinear differential operators*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **13** (1986), 299–346.

[45] L. Nirenberg, *Topics in Nonlinear Functional Analysis*, Courant Institute Lecture Notes, Amer. Math. Soc. (1974).

[46] W.-M. Ni and R.D. Nussbaum, *Uniqueness and nonuniqueness of positive radial solutions of* $\Delta u + f(u,r)$, Comm. Pure Appl. Math. **38** (1985), 69–108.

[47] Z. Opial, Sur les periodes des solutions de l'equation differentielle $x'' + g(x) = 0$, Ann. Polon. Math. **10** (1961), 49–72 (in French).

[48] T. Ouyang, *On the positive solutions of semilinear equations* $\Delta u + \lambda u - hu^p = 0$ *on the compact manifolds*, Trans. Amer. Math. Soc. **331** (1992), 503–527.

[49] T. Ouyang, *On the positive solutions of semilinear equations* $\Delta u + \lambda u + hu^p = 0$ *on compact manifolds*, II, Indiana Univ. Math. J. **40** (1991), 1083–1141.

[50] T. Ouyang and J. Shi, *Exact multiplicity of positive solutions for a class of semilinear problems*, II, J. Differential Equations **158** (1999), 94–151.

[51] M. Ramaswamy, *On the global set of solutions of a nonlinear ODE: Theoretical and numerical description*, J. Differential Equations **65** (1986), 1–48.

[52] M. Renardi and R.C. Rogers, *An Introduction to Partial Differential Equations*, Springer (1992).

[53] R. Schaaf, *Global Solution Branches of Two Point Boundary Value Problems*, Lecture Notes in Math., Vol. 1458, Springer (1990).

[54] J.C. Scovel, *Geometry of some nonlinear differential operators*, Ph.D. thesis, Courant Institute, NYU (1983).

[55] J. Shi, *Semilinear Neumann boundary value problems on a rectangle*, Trans. Amer. Math. Soc. **354** (2002), 3117–3154.

[56] J. Shi and J. Wang, *Morse indices and exact multiplicity of solutions to semilinear elliptic problems*, Proc. Amer. Math. Soc. **127** (1999), 3685–3695.

[57] J. Smoller and A. Wasserman, *Global bifurcation of steady-state solutions*, J. Differential Equations **39** (1981), 269–290.

[58] M. Tang, *Uniqueness of positive radial solutions for* $\Delta u - u + u^p = 0$ *on an annulus*, J. Differential Equations **189** (2003), 148–160.

[59] S.-H. Wang, *A correction for a paper by J. Smoller and A. Wasserman*, J. Differential Equations **77** (1989), 199–202.

[60] S.-H. Wang, *On S-shaped bifurcation curves*, Nonlinear Anal. TMA **22** (1994), 1475–1485.

[61] S.-H. Wang, *On the time map of a nonlinear two-point boundary value problem*, Differential Integral Equations **7** (1994), 49–55.

[62] S.-H. Wang and N.D. Kazarinoff, *Bifurcation of steady-state solutions for a scalar reaction–diffusion equation in one space variable*, J. Austral. Math. Soc. Ser. A **52** (1992), 343–355.

[63] S.-H. Wang and T.S. Yeh, *A complete classification of bifurcation diagrams of a Dirichlet problem with concave–convex nonlinearities*, J. Math. Anal. Appl. **291** (2004), 128–153.

CHAPTER 7

Singularities and Laplacians in Boundary Value Problems for Nonlinear Ordinary Differential Equations

Irena Rachůnková* and Svatoslav Staněk*

Department of Mathematics, Palacký University, 779 00 Olomouc, Tomkova 40, Czech Republic

Milan Tvrdý†

Mathematical Institute, Academy of Sciences of the Czech Republic, 115 67 Praha 1, Žitná 25, Czech Republic

Contents

0. Notation 609
1. Principles of solvability of singular higher order BVPs 610
   1.1. Regular and singular BVPs 610
   1.2. Singularities in time variable 612
   1.3. Singularities in space variables 612
   1.4. Existence principles for BVPs with time singularities 613
   1.5. Existence principles for BVPs with space singularities 617
2. Existence results for singular two-point higher order BVPs 619
   2.1. Focal conditions 620
   2.2. Conjugate conditions 624
   2.3. (n, p) boundary conditions 627
   2.4. Sturm–Liouville conditions 628
   2.5. Lidstone conditions 631
   2.6. Historical and bibliographical notes 633
3. Principles of solvability of singular second order BVPs with ϕ-Laplacian 634
   3.1. Regularization of singular problems with ϕ-Laplacian 635
   3.2. General existence principle for singular BVPs with ϕ-Laplacian 638
   3.3. Nonlocal singular BVPs 641
   3.4. Historical and bibliographical notes 647

---

[*] Supported by the Council of Czech Government MSM 6198959214.

[†] Supported by the grant No. 201/04/1077 of the Grant Agency of the Czech Republic and by the Academy of Sciences of the Czech Republic, Institutional Research Plan No. AV0Z10190503.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 3
Edited by A. Cañada, P. Drábek and A. Fonda
© 2006 Elsevier B.V. All rights reserved

4. Singular Dirichlet BVPs with ϕ-Laplacian . . . 648
  4.1. Method of lower and upper functions . . . 649
  4.2. Positive nonlinearities . . . 655
  4.3. Sign-changing nonlinearities . . . 659
  4.4. Sign-changing solutions and w-solutions . . . 666
  4.5. Historical and bibliographical notes . . . 667
5. Singular periodic BVPs with ϕ-Laplacian . . . 668
  5.1. Method of lower and upper functions for regular problems . . . 670
  5.2. Method of lower and upper functions for singular problems . . . 677
  5.3. Attractive singular forces . . . 680
  5.4. Repulsive singular forces . . . 683
  5.5. Historical and bibliographical notes . . . 705
6. Other types of two-point second order BVPs . . . 708
References . . . 717

Abstract

In this text we investigate solvability of various nonlinear singular boundary value problems for ordinary differential equations on the compact interval $[0, T]$. The nonlinearities in differential equations may be singular both in the time and space variables. Location of all singular points in $[0, T]$ need not be known.

The work is divided into 6 sections. Sections 1 and 2 are devoted to singular higher order boundary value problems. The remaining ones deal with the second order case. Motivated by various applications in physics we admit here the left-hand sides of the equations under consideration containing the ϕ-Laplacian or p-Laplacian operator. The special attention is paid to Dirichlet and periodic problems.

Usually, the main ideas of the proofs of the results mentioned are described. More detailed proofs are included in the cases where no proofs are available in literature or where the details are needed later.

0. Notation

Let $J \subset \mathbb{R}$, $k \in \mathbb{N}$, $p \in (1, \infty)$. Then we will write:

- $L_\infty(J)$ for the set of functions essentially bounded and (Lebesgue) measurable on J; the corresponding norm is $\|u\|_\infty = \sup\operatorname{ess}\{|u(t)|\colon\ t \in J\}$.
- $L_1(J)$ for the set of functions (Lebesgue) integrable on J; the corresponding norm is $\|u\|_1 = \int_J |u(t)|\,\mathrm{d}t$.
- $L_p(J)$ for the set of functions whose pth powers of modulus are integrable on J; the corresponding norm is $\|u\|_p = (\int_J |u(t)|^p\,\mathrm{d}t)^{1/p}$.
- $C(J)$ and $C^k(J)$ for the sets of functions continuous on J and having continuous kth derivatives on J, respectively.
- $AC(J)$ and $AC^k(J)$ for the sets of functions absolutely continuous on J and having absolutely continuous kth derivatives on J, respectively.
- $AC_{loc}(J)$ and $AC^k_{loc}(J)$ for the sets of functions absolutely continuous on each compact interval $I \subset J$ and having absolutely continuous kth derivatives on each compact interval $I \subset J$, respectively.
- If $J = [a, b]$, we will simply write $C[a, b]$ instead of $C([a, b])$ and similarly for other types of intervals and other functional sets defined above.

Further, we use the following notation:

- If $u \in L_\infty[a, b]$ is continuous on $[a, b]$, then $\max\{|u(t)|\colon\ t \in [a, b]\} = \sup\operatorname{ess}\{|u(t)|\colon\ t \in [a, b]\}$. Therefore the norm in $C[a, b]$ will be denoted by $\|u\|_\infty = \max\{|u(t)|\colon\ t \in [a, b]\}$ and the norm in $C^k[a, b]$ by $\|u\|_{C^k} = \sum_{i=0}^k \|u^{(i)}\|_\infty$.
- Let $n \in \mathbb{N}$ and $\mathcal{M} \subset \mathbb{R}^n$. Then $\overline{\mathcal{M}}$ will denote the closure of $\mathcal{M}$, $\partial\mathcal{M}$ the boundary of $\mathcal{M}$ and meas $(\mathcal{M})$ the Lebesgue measure of $\mathcal{M}$.
- $\deg(\mathcal{I} - \mathcal{F}, \Omega)$ stands for the Leray–Schauder degree of $\mathcal{I} - \mathcal{F}$ with respect to Ω, where $\mathcal{I}$ denotes the identity operator.

We say that a function f satisfies *the Carathéodory conditions* on the set $[a, b] \times \mathcal{M}$ if:

$$f(\cdot, x_0, \dots, x_{n-1})\colon [a, b] \to \mathbb{R} \text{ is measurable}$$
$$\text{for all } (x_0, x_1, \dots, x_{n-1}) \in \mathcal{M}; \tag{0.1}$$

$$f(t, \cdot, \dots, \cdot)\colon \mathcal{M} \to \mathbb{R} \text{ is continuous for a.e. } t \in [a, b]; \tag{0.2}$$

$$\left\{\begin{array}{l} \text{for each compact set } \mathcal{K} \subset \mathcal{M} \text{ there is a function } m_{\mathcal{K}} \in L_1[a, b] \text{ such that} \\ \quad \left|f(t, x_0, \dots, x_{n-1})\right| \leqslant m_{\mathcal{K}}(t) \quad \text{for a.e. } t \in [a, b] \\ \text{and all } (x_0, x_1, \dots, x_{n-1}) \in \mathcal{K}. \end{array}\right. \tag{0.3}$$

In this case we will write $f \in Car([a, b] \times \mathcal{M})$. If $J \subset [a, b]$ and $J \neq \overline{J}$, then $f \in Car(J \times \mathcal{M})$ will mean that $f \in Car(I \times \mathcal{M})$ for each compact interval $I \subset J$.

1. Principles of solvability of singular higher order BVPs

1.1. *Regular and singular BVPs*

For $n \in \mathbb{N}$, $[0,T] \subset \mathbb{R}$, $i \in \{0,1,\dots,n-1\}$ and a closed set $\mathcal{B} \subset C^i[0,T]$ consider the boundary value problem

$$u^{(n)} = f\big(t, u, \dots, u^{(n-1)}\big), \tag{1.1}$$

$$u \in \mathcal{B}. \tag{1.2}$$

In what follows, we will investigate the solvability of problem (1.1), (1.2) on the set $[0,T] \times \mathcal{A}$, where $\mathcal{A}$ is a closed subset of $\mathbb{R}^n$ or $\mathcal{A} = \mathbb{R}^n$. The classical existence results are based on the assumption $f \in Car([0,T] \times \mathcal{A})$. In this case we will say that problem (1.1), (1.2) is *regular on* $[0,T] \times \mathcal{A}$. If $f \notin Car([0,T] \times \mathcal{A})$ we will say that problem (1.1), (1.2) is *singular on* $[0,T] \times \mathcal{A}$.

Motivated by the following applications we will mainly address singular problems.

EXAMPLE 1. In certain problems in fluid dynamics and boundary layer theory (see, e.g., Callegari, Friedman and Nachman [43–45]) the second order differential equation

$$u'' + \frac{\psi(t)}{u^{\lambda}} = 0 \tag{1.3}$$

arose. Here $\lambda \in (0,\infty)$ and $\psi \in C(0,1)$, $\psi \notin L_1[0,1]$. Equation (1.3) is known as the generalized Emden–Fowler equation. Its solvability with the Dirichlet boundary conditions

$$u(0) = u(1) = 0 \tag{1.4}$$

was investigated by Taliaferro [141] in 1979 and then by many other authors. Problem (1.3), (1.4) has been studied on the set $[0,1] \times [0,\infty)$ because positive solutions have been sought. We can see that $f(t,x) = \psi(t)x^{-\lambda}$ does not fulfill conditions (0.2) and (0.3) with $[a,b] = [0,1]$ and $\mathcal{M} = [0,\infty)$. Hence problem (1.3), (1.4) is singular on $[0,1] \times [0,\infty)$.

EXAMPLE 2. Consider the fourth order degenerate parabolic equation

$$U_t + \big(|U|^{\mu} U_{yyy}\big)_y = 0,$$

which arises in droplets and thin viscous flows models (see, e.g., [32,33]). The source-type solutions of this equation have the form

$$U(y,t) = t^{-b} u\big(yt^{-b}\big), \quad b = \frac{1}{\mu + 4},$$

which leads to the study of the third order ordinary differential equation on $[-1,1]$,

$$u''' = btu^{1-\mu}.$$

We see that $f(t,x) = btx^{1-\mu}$ is singular on $[-1,1] \times [0,\infty)$ if $\mu > 1$.

EXAMPLE 3. Similarly to Example 2, the sixth order degenerate equation

$$U_t - \left(|U|^{\mu} U_{yyyyy}\right)_y = 0$$

which arises in semiconductor models (Bernis [30,31]) leads to the fifth order ordinary differential equation

$$-u^{(5)} = \frac{t}{u^{\lambda}}$$

which is singular for $\lambda > 0$.

A solvability decision for singular boundary value problems requires an exact definition of a solution to such problems. Here, we will work with the same definition of a solution both for regular problems and for singular ones.

DEFINITION 1.1. A function $u \in AC^{n-1}[0,T] \cap \mathcal{B}$ is said to be a *solution of problem* (1.1), (1.2), if it satisfies the equality $u^{(n)}(t) = f(t, u(t), \dots, u^{(n-1)}(t))$ a.e. on $[0,T]$. If we investigate problem (1.1), (1.2) on $\mathcal{A} \neq \mathbb{R}^n$, we moreover require $(u(t), \dots, u^{(n-1)}(t)) \in \mathcal{A}$ for $t \in [0,T]$.

In literature, *an alternative approach to solvability* of singular problems can be found. In this approach, solutions are defined as functions whose $(n-1)$st derivatives can have discontinuities at some points in $[0,T]$. Here we will call them the w-solutions. According to Kiguradze [92] or Agarwal and O'Regan [3] we define them as follows. In contrast to our starting setting, to define w-solutions we assume that $i \in \{0, 1, \dots, n-2\}$ and $\mathcal{B}$ is a closed subset in $C^i[0,T]$.

DEFINITION 1.2. We say that u is *a w-solution of problem* (1.1), (1.2) if there exists a finite number of points $t_\nu \in [0,T]$, $\nu = 1, 2, \dots, r$, such that if we denote $J = [0,T] \setminus \{t_\nu\}_{\nu=1}^{r}$, then $u \in C^{n-2}[0,T] \cap AC_{loc}^{n-1}(J) \cap \mathcal{B}$ satisfies $u^{(n)}(t) = f(t, u(t), \dots, u^{(n-1)}(t))$ a.e. on $[0,T]$. If $\mathcal{A} \neq \mathbb{R}^n$ we require $(u(t), \dots, u^{(n-1)}(t)) \in \mathcal{A}$ for $t \in J$.

Clearly each solution is a w-solution and each w-solution which belongs to $AC^{n-1}[0,T]$ is a solution.

In the study of singular problem (1.1), (1.2) we will focus our attention on two types of singularities of the function f:

Let $J \subset [0,T]$. We say that $f : J \times \mathcal{A} \to \mathbb{R}$ has singularities in its *time variable* t, if $J \neq \overline{J} = [0,T]$ and

$$f \in Car(J \times \mathcal{A}) \quad \text{and} \quad f \notin Car\big([0,T] \times \mathcal{A}\big). \tag{1.5}$$

Let $\mathcal{D} \subset \mathcal{A}$. We say that $f : [0,T] \times \mathcal{D} \to \mathbb{R}$ has singularities in its *space variables* $x_0, x_1, \dots, x_{n-1}$, if $\mathcal{D} \neq \overline{\mathcal{D}} = \mathcal{A}$ and

$$f \in Car\big([0,T] \times \mathcal{D}\big) \quad \text{and} \quad f \notin Car\big([0,T] \times \mathcal{A}\big). \tag{1.6}$$

We will study particular cases of (1.5) and (1.6) which will be described in Section 1.2 and Section 1.3, respectively.

1.2. *Singularities in time variable*

According to (0.3) and (1.5) a function f has a singularity in its *time variable* t, if f is not integrable on $[0,T]$. Let us define it more precisely. Let $k \in \mathbb{N}$, $t_i \in [0,T]$, $i = 1, \dots, k$, $J = [0,T] \setminus \{t_1, t_2, \dots, t_k\}$ and let $f \in Car(J \times \mathcal{A})$. Assume that for each $i \in \{1, \dots, k\}$ there exists $(x_0, \dots, x_{n-1}) \in \mathcal{A}$ such that

$$\int_{t_i}^{t_i+\varepsilon} \big|f(t, x_0, \dots, x_{n-1})\big| \, \mathrm{d}t = \infty \quad \text{or} \quad \int_{t_i-\varepsilon}^{t_i} \big|f(t, x_0, \dots, x_{n-1})\big| \, \mathrm{d}t = \infty \tag{1.7}$$

for any sufficiently small $\varepsilon > 0$. Then f does not fulfill (0.3) with $\mathcal{M} = \mathcal{A}$ and, according to (1.5), function f has singularities in its time variable t, namely at the values $t_1, \dots, t_k$. We will call these values the *singular points* of f.

EXAMPLE. Let $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \dots, k$, be continuous. Then the function

$$f(t, x_0, \dots, x_{n-1}) = \sum_{i=1}^{k} \frac{1}{t - t_i} f_i(x_0, \dots, x_{n-1}),$$

has singular points $t_1, t_2, \dots, t_k$.

1.3. *Singularities in space variables*

By virtue of (0.2) and (1.6) we see that if f has a singularity in some of its *space variables* then f is not continuous in this variable on a region where f is studied. Motivated by Eq. (1.3) we will consider the following case. Let $\mathcal{A}_i \subset \mathbb{R}$ be a closed interval and let $c_i \in \mathcal{A}_i$, $\mathcal{D}_i = \mathcal{A}_i \setminus \{c_i\}$, $i = 0, 1, \dots, n-1$. Assume that there exists $j \in \{0, 1, \dots, n-1\}$ such that

$$\begin{cases} \limsup\limits_{x_j \to c_j, x_j \in \mathcal{D}_j} \big|f(t, x_0, \dots, x_j, \dots, x_{n-1})\big| = \infty \\ \text{for a.e. } t \in [0,T] \text{ and for some } x_i \in \mathcal{D}_i, \ i = 0, 1, \dots, n-1, \ i \neq j. \end{cases} \tag{1.8}$$

If we put $\mathcal{A} = \mathcal{A}_0 \times \cdots \times \mathcal{A}_{n-1}$, we see that f does not fulfill (0.2) with $\mathcal{M} = \mathcal{A}$ and, according to (1.6), the function f has a singularity in its space variable x_j, namely at the value c_j.

Let u be a solution of (1.1), (1.2) and let a point $t_u \in [0, T]$ be such that $u^{(j)}(t_u) = c_j$ for some $j \in \{0, \dots, n-1\}$. Then t_u is called *a singular point corresponding to the solution u*.

Now, let u be a w-solution of (1.1), (1.2). Assume that a point $t_u \in [0, T]$ is such that $u^{(n-1)}(t_u)$ does not exist or there is a $j \in \{0, \dots, n-1\}$ such that $u^{(j)}(t_u) = c_j$. Then t_u is called *a singular point corresponding to the w-solution u*.

EXAMPLE. Let $h_1, h_2, h_3 \in L_1[0, T]$, $h_2 \neq 0$, $h_3 \neq 0$ a.e. on $[0, T]$. Consider the Dirichlet problem

$$u'' + h_1(t) + \frac{h_2(t)}{u} + \frac{h_3(t)}{u'} = 0, \quad u(0) = u(T) = 0. \tag{1.9}$$

Let u be a solution of (1.9). Then 0 and T are singular points corresponding to u. Moreover, there exists at least one point $t_u \in (0, T)$ satisfying $u'(t_u) = 0$, which means that t_u is also a singular point corresponding to u. Note that (in contrast to the points 0 and T) we do not know the location of t_u in $(0, T)$.

In accordance with this example, we will distinguish two types of singular points corresponding to solutions or to w-solutions: *singular points of type I*, where we know their location in $[0, T]$ and *singular points of type II* whose location is not known.

1.4. *Existence principles for BVPs with time singularities*

Singular problems are usually investigated by means of auxiliary regular problems. To establish the existence of a solution of a singular problem we introduce a sequence of approximating regular problems which are solvable. Then we pass to the limit for the sequence of approximate solutions to get a solution of the original singular problem. Here we provide existence principles which contain the main rules for a construction of such sequences to get either w-solutions or solutions.

Consider problem (1.1), (1.2) on $[0, T] \times \mathcal{A}$. For the sake of simplicity assume that f has only one time singularity at $t = t_0$, $t_0 \in [0, T]$. It means that

$$\left\{\begin{array}{l} J = [0, T] \setminus \{t_0\},\ f \in Car(J \times \mathcal{A}) \text{ satisfies at least one of the conditions} \\ \text{(i)} \quad \displaystyle\int_{t_0-\varepsilon}^{t_0} \left| f(t, x_0, \dots, x_{n-1}) \right| \mathrm{d}t = \infty, \quad t_0 \in (0, T], \\ \text{(ii)} \quad \displaystyle\int_{t_0}^{t_0+\varepsilon} \left| f(t, x_0, \dots, x_{n-1}) \right| \mathrm{d}t = \infty, \quad t_0 \in [0, T), \\ \text{for some } (x_0, x_1, \dots, x_{n-1}) \in \mathcal{A} \text{ and each sufficiently small } \varepsilon > 0. \end{array}\right. \tag{1.10}$$

Let us have a sequence of regular problems

$$u^{(n)} = f_k\big(t, u, \ldots, u^{(n-1)}\big), \quad u \in \mathcal{B}, \tag{1.11}$$

where $f_k \in Car([0,T] \times \mathbb{R}^n)$, $k \in \mathbb{N}$.

THEOREM 1.3 (First existence principle for w-solutions of (1.1), (1.2)). *Let* (1.10) *hold and let $\mathcal{B}$ be a closed subset in $C^{n-2}[0,T]$. Assume that the conditions*

$$\begin{cases} \textit{for each } k \in \mathbb{N} \textit{ and each } (x_0, \ldots, x_{n-1}) \in \mathcal{A}, \\ f_k(t, x_0, \ldots, x_{n-1}) = f(t, x_0, \ldots, x_{n-1}) \quad \textit{a.e. on } [0,T] \setminus \Delta_k, \\ \textit{where } \Delta_k = \big(t_0 - \frac{1}{k}, t_0 + \frac{1}{k}\big) \cap [0,T], \end{cases} \tag{1.12}$$

and

$$\begin{cases} \textit{there exists a bounded set } \Omega \subset C^{n-1}[0,T] \textit{ such that} \\ \textit{for each } k \in \mathbb{N} \textit{ the regular problem } (1.11) \textit{ has a solution } u_k \in \Omega \\ \textit{and } \big(u_k(t), \ldots, u_k^{(n-1)}(t)\big) \in \mathcal{A} \textit{ for } t \in [0,T] \end{cases} \tag{1.13}$$

are fulfilled. Then

$$\begin{cases} \textit{there exist a function } u \in C^{n-2}[0,T] \textit{ and a subsequence} \\ \{u_{k_\ell}\} \subset \{u_k\} \textit{ such that } \lim_{\ell \to \infty} \|u_{k_\ell} - u\|_{C^{n-2}} = 0; \end{cases} \tag{1.14}$$

$$\begin{cases} \lim_{\ell \to \infty} u_{k_\ell}^{(n-1)}(t) = u^{(n-1)}(t) \textit{ locally uniformly on } J \\ \textit{and } \big(u(t), \ldots, u^{(n-1)}(t)\big) \in \mathcal{A} \textit{ for } t \in J; \end{cases} \tag{1.15}$$

$$\textit{the function } u \in AC_{loc}^{n-1}(J) \textit{ is a w-solution of problem } (1.1), (1.2). \tag{1.16}$$

SKETCH OF THE PROOF. *Step* 1. *Convergence of the sequence of approximate solutions.*

Condition (1.13) implies that the sequences $\{u_k^{(i)}\}$, $0 \leqslant i \leqslant n-2$, are bounded and equicontinuous on $[0,T]$. By the Arzelà–Ascoli theorem the assertion (1.14) is true and $u \in \mathcal{B} \subset C^{n-2}[0,T]$. Since $\{u_k^{(n-1)}\}$ is bounded on $[0,T]$, we get, due to (1.11) and (1.12), that for each $t \in [0, t_0)$ the sequence $\{u_k^{(n-1)}\}$ is equicontinuous on $[0,t]$ and so the same holds on $[t,T]$ if $t \in (t_0, T]$. The Arzelà–Ascoli theorem and the diagonalization principle yield (1.15).

Step 2. *Properties of the limit u.*

By virtue of (1.12), (1.14) and (1.15) we have

$$\lim_{k_\ell \to \infty} f_{k_\ell}\big(t, u_{k_\ell}(t), \ldots, u_{k_\ell}^{(n-1)}(t)\big) = f\big(t, u(t), \ldots, u^{(n-1)}(t)\big)$$

$$\text{a.e. on } [0,T]. \tag{1.17}$$

Hence, using the Lebesgue convergence theorem, we can deduce that if $t_0 \neq 0$ the limit u solves the equation

$$u^{(n-1)}(t) = u^{(n-1)}(0) + \int_0^t f\bigl(s, u(s), \dots, u^{(n-1)}(s)\bigr)\,\mathrm{d}s \quad \text{for } t \in [0, t_0) \tag{1.18}$$

and if $t_0 \neq T$ the limit u solves the equation

$$u^{(n-1)}(t) = u^{(n-1)}(T) - \int_t^T f\bigl(s, u(s), \dots, u^{(n-1)}(s)\bigr)\,\mathrm{d}s \quad \text{for } t \in (t_0, T], \tag{1.19}$$

which immediately yields (1.16). □

For the existence of a solution $u \in AC^{n-1}[0, T]$ of problem (1.1), (1.2) we will impose additional conditions on f on some neighbourhood of t_0.

THEOREM 1.4 (First existence principle for solutions of (1.1), (1.2)). *Let all assumptions of Theorem* 1.3 *be fulfilled. Further, assume that*

$$\begin{cases} \textit{there exist } \psi \in L_1[0,T],\ \eta > 0 \textit{ and } \lambda_1, \lambda_2 \in \{-1, 1\} \textit{ such that} \\ \quad \lambda_1 f_{k_\ell}\bigl(t, u_{k_\ell}(t), \dots, u_{k_\ell}^{(n-1)}(t)\bigr) \geqslant \psi(t) \\ \textit{for all } \ell \in \mathbb{N} \textit{ and for a.e. } t \in [t_0 - \eta, t_0) \cap [0, T] \textit{ provided } (1.10)(\mathrm{i}) \textit{ holds}, \\ \quad \lambda_2 f_{k_\ell}\bigl(t, u_{k_\ell}(t), \dots, u_{k_\ell}^{(n-1)}(t)\bigr) \geqslant \psi(t) \\ \textit{for all } \ell \in \mathbb{N} \textit{ and for a.e. } t \in (t_0, t_0 + \eta] \cap [0, T] \textit{ provided } (1.10)(\mathrm{ii}) \textit{ is true}. \end{cases} \tag{1.20}$$

Then the assertions (1.14) *and* (1.15) *are valid and* $u \in AC^{n-1}[0, T]$ *is a solution of problem* (1.1), (1.2).

SKETCH OF THE PROOF. *Step* 1. As in the proof of Theorem 1.3 we get that (1.14)–(1.16) hold.

Step 2. Since u is a w-solution of problem (1.1), (1.2), it remains to prove that $u \in AC^{n-1}[0, T]$. Assume that condition (1.10)(i) holds. Since

$$u_{k_\ell}^{(n-1)}(t) - u_{k_\ell}^{(n-1)}(t_0 - \eta) = \int_{t_0-\eta}^t f_{k_\ell}\bigl(s, u_{k_\ell}(s), \dots, u_{k_\ell}^{(n-1)}(s)\bigr)\,\mathrm{d}s \tag{1.21}$$

for $t \in (0, t_0)$, we get due to (1.13) that there is a $c \in (0, \infty)$ such that

$$\lambda_1 \int_{t_0-\eta}^{t_0} f_{k_\ell}\big(s, u_{k_\ell}(s), \ldots, u_{k_\ell}^{(n-1)}(s)\big)\,\mathrm{d}s \leqslant c \tag{1.22}$$

for each $\ell \in \mathbb{N}$. By the Fatou lemma, having in mind conditions (1.17), (1.20) and (1.22), we deduce that $f(t, u(t), \ldots, u^{(n-1)}(t)) \in L_1[0, t_0]$. Similarly, if condition (1.10)(ii) holds, we deduce that $f(t, u(t), \ldots, u^{(n-1)}(t)) \in L_1[t_0, T]$. Therefore $f(t, u(t), \ldots, u^{(n-1)}(t)) \in L_1[0, T]$ and due to (1.18) and (1.19) we have that $u \in AC^{n-1}[0, T]$ is a solution of problem (1.1), (1.2). □

In the sequel we will need the following definition:

DEFINITION 1.5. Let $[a, b] \subset \mathbb{R}$ and $\{g_k\} \subset L_1[a, b]$. We say that the sequence $\{g_k\}$ is *uniformly integrable on* $[a, b]$ if

$$\begin{cases} \text{for each } \varepsilon > 0 \text{ there exists } \delta > 0 \text{ such that} \\ \displaystyle\sum_{j=1}^{\infty}(b_j - a_j) < \delta \quad \Longrightarrow \quad \sum_{j=1}^{\infty}\int_{a_j}^{b_j} \big|g_k(t)\big|\,\mathrm{d}t < \varepsilon \\ \text{for each } k \in \mathbb{N} \text{ and each sequence of intervals } \{(a_j, b_j)\} \text{ in } [a, b]. \end{cases} \tag{1.23}$$

Note that condition (1.23) is satisfied for example if there exists $\psi \in L_1[a, b]$ such that $|g_k(t)| \leqslant \psi(t)$ for a.e. $t \in [a, b]$ and all $k \in \mathbb{N}$.

THEOREM 1.6 (Second existence principle for solutions of (1.1), (1.2)). *Let all assumptions of Theorem* 1.3 *be fulfilled and assume in addition that $\mathcal{B}$ is a closed subset in $C^{n-1}[0, T]$ and that*

$$\begin{cases} \textit{there exists } \eta > 0 \textit{ such that the sequence } \big\{f_k\big(t, u_k(t), \ldots, u_k^{(n-1)}(t)\big)\big\} \\ \textit{is uniformly integrable on } [t_0 - \eta, t_0 + \eta] \cap [0, T]. \end{cases} \tag{1.24}$$

Then

$$\begin{cases} \textit{there exist a function } u \in \overline{\Omega} \textit{ and a subsequence } \{u_{k_\ell}\} \subset \{u_k\} \textit{ such that} \\ \quad \lim\limits_{\ell\to\infty} \|u_{k_\ell} - u\|_{C^{n-1}} = 0 \textit{ and } \big(u(t), \ldots, u^{(n-1)}(t)\big) \in \mathcal{A} \quad \textit{for } t \in [0, T] \end{cases} \tag{1.25}$$

and $u \in AC^{n-1}[0, T]$ is a solution of problem (1.1), (1.2).

SKETCH OF THE PROOF. *Step* 1. By (1.13) we get that the sequences $\{u_k^{(i)}\}$, $0 \leqslant i \leqslant n-2$, are bounded in $C[0,T]$ and equicontinuous on $[0,T]$ and $\{u_k^{(n-1)}\}$ is bounded in $C[0,T]$. Using (1.24) one can show that $\{u_k^{(n-1)}\}$ is also equicontinuous on $[0,T]$. The Arzelà–Ascoli theorem yields (1.25) and $u \in \mathcal{B} \subset C^{n-1}[0,T]$.

Step 2. As in Step 2 of the proof of Theorem 1.3 we get that u is a w-solution of problem (1.1), (1.2).

Step 3. It remains to prove that $u \in AC^{n-1}[0,T]$. Since $u \in AC_{loc}^{n-1}(J)$, it is sufficient to prove

$$u^{(n-1)} \in AC\big([t_0-\eta, t_0+\eta] \cap [0,T]\big). \tag{1.26}$$

Assume that (1.10)(i) holds and $[t_0-\eta, t_0] \subset [0,T]$. By virtue of (1.17) and (1.24), applying Vitali's convergence theorem we obtain $f(t, u(t), \dots, u^{(n-1)}(t)) \in L_1[t_0-\eta, t_0]$. If (1.10)(ii) holds, we can assume $[t_0, t_0+\eta] \subset [0,T]$ and deduce similarly that

$$f\big(t, u(t), \dots, u^{(n-1)}(t)\big) \in L_1[t_0, t_0+\eta].$$

Hence, we get (1.26). □

1.5. *Existence principles for BVPs with space singularities*

Similarly to Section 1.4 we will establish sufficient properties for an approximate sequence of regular problems and of their solutions to pass to a limit and yield a solution of the original singular problem (1.1), (1.2). Let $\mathcal{A}_i \subset \mathbb{R}$, $i = 0, \dots, n-1$, be closed intervals and let $\mathcal{A} = \mathcal{A}_0 \times \cdots \times \mathcal{A}_{n-1}$. Consider problem (1.1), (1.2) on $[0,T] \times \mathcal{A}$ and assume that f has only one singularity at each x_i, namely at the values $c_i \in \mathcal{A}_i$, $i = 0, \dots, n-1$. Denoting $\mathcal{D} = \mathcal{D}_0 \times \cdots \times \mathcal{D}_{n-1}$, $\mathcal{D}_i = \mathcal{A}_i \setminus \{c_i\}$, $i = 0, \dots, n-1$, we will assume that

$$f \in Car\big([0,T] \times \mathcal{D}\big) \text{ satisfies (1.8)} \quad \text{for } j = 0, \dots, n-1. \tag{1.27}$$

Consider a sequence of regular problems (1.11) where $f_k \in Car([0,T] \times \mathbb{R}^n)$, $k \in \mathbb{N}$. We will use the approach used by Rachůnková and Staněk in [121] and [122].

THEOREM 1.7 (Second existence principle for w-solutions of (1.1), (1.2)). *Let* (1.13), (1.27) *hold and let $\mathcal{B}$ be a closed subset in $C^{n-2}[0,T]$. Assume that*

$$\begin{cases} \text{for each } k \in \mathbb{N}, \text{ for a.e. } t \in [0,T] \text{ and each } (x_0, \dots, x_{n-1}) \in \mathcal{D} \\ \text{we have} \\ \quad f_k(t, x_0, \dots, x_{n-1}) = f(t, x_0, \dots, x_{n-1}) \\ \quad \text{if } |x_i - c_i| \geqslant \frac{1}{k},\ 0 \leqslant i \leqslant n-1. \end{cases} \tag{1.28}$$

Then the assertion (1.14) *is valid.*

If, moreover, the set

$$\Sigma = \big\{ s \in [0,T] \colon u^{(i)}(s) = c_i \text{ for some } i \in \{0, \dots, n-1\}$$
$$\text{or } u^{(n-1)}(s) \text{ does not exist} \big\}$$

is finite, then the assertion (1.15) *is valid for* $J = [0,T] \setminus \Sigma$. *If, in addition,*

$$\begin{cases} \text{the sequence } \left\{ f_{k_\ell}\big(t, u_{k_\ell}(t), \dots, u_{k_\ell}^{(n-1)}(t)\big)\right\} \\ \text{is uniformly integrable on each interval } [a,b] \subset J \end{cases} \tag{1.29}$$

then $u \in AC_{loc}^{n-1}(J)$ *is a w-solution of problem* (1.1), (1.2).

SKETCH OF THE PROOF. *Step* 1. *Convergence of the sequence of approximate solutions.*

As in Step 1 of the proof of Theorem 1.3 we get (1.14) and $u \in \mathcal{B} \subset C^{n-2}[0,T]$. Assume that Σ is finite and choose an arbitrary $[a,b] \subset J$. According to (1.27) and (1.28) we can prove that the sequence $\{u_{k_\ell}^{(n-1)}\}$ is equicontinuous on $[a,b]$. Using the Arzelà–Ascoli theorem and the diagonalization principle we deduce that the subsequence $\{u_{k_\ell}\}$ can be chosen so that it fulfills (1.15).

Step 2. *Convergence of the sequence of regular right-hand sides.*

Consider sets

$$\mathcal{V}_1 = \big\{t \in [0,T]\colon\ f(t,\cdot,\dots,\cdot)\colon \mathcal{D} \to \mathbb{R} \text{ is not continuous}\big\},$$
$$\mathcal{V}_2 = \big\{t \in [0,T]\colon\ \text{the equality in (1.28) is not valid}\big\}.$$

We can see that $\operatorname{meas}(\mathcal{V}_1) = \operatorname{meas}(\mathcal{V}_2) = 0$. Denote $\mathcal{U} = \Sigma \cup \mathcal{V}_1 \cup \mathcal{V}_2$ and choose an arbitrary $t \in [0,T] \setminus \mathcal{U}$. By (1.14) and (1.15) there exists $\ell_0 \in \mathbb{N}$ such that for each $\ell \in \mathbb{N}$, $\ell \geqslant \ell_0$,

$$\big|u^{(i)}(t) - c_i\big| > \frac{1}{k_\ell}, \quad \big|u_{k_\ell}^{(i)}(t) - c_i\big| \geqslant \frac{1}{k_\ell} \quad \text{for } i \in \{0,\dots,n-1\}.$$

According to (1.28) we have

$$f_{k_\ell}\big(t, u_{k_\ell}(t), \dots, u_{k_\ell}^{(n-1)}(t)\big) = f\big(t, u_{k_\ell}(t), \dots, u_{k_\ell}^{(n-1)}(t)\big)$$

and, by (1.14), (1.15),

$$\lim_{\ell\to\infty} f_{k_\ell}\big(t, u_{k_\ell}(t), \dots, u_{k_\ell}^{(n-1)}(t)\big) = f\big(t, u(t), \dots, u^{(n-1)}(t)\big). \tag{1.30}$$

Since $\operatorname{meas}(\mathcal{U}) = 0$, (1.30) holds for a.e. $t \in [0,T]$.

Step 3. *Existence of a w-solution.*

Choose an arbitrary interval $[a,b] \subset J$. By (1.29) and (1.30) we can use Vitali's convergence theorem [23] to show that $f(t, u(t), \dots, u^{(n-1)}(t)) \in L_1[a,b]$ and if we pass to the limit in the sequence

$$u_{k_\ell}^{(n-1)}(t) = u_{k_\ell}^{(n-1)}(a) + \int_a^t f_{k_\ell}\big(s, u_{k_\ell}(s), \dots, u_{k_\ell}^{(n-1)}(s)\big)\,\mathrm{d}s, \quad t \in [a,b],$$

we get

$$u^{(n-1)}(t) = u^{(n-1)}(a) + \int_a^t f\big(s, u(s), \ldots, u^{(n-1)}(s)\big)\,\mathrm{d}s, \quad t \in [a, b].$$

Since $[a, b] \subset J$ is an arbitrary interval, we conclude that $u \in AC_{loc}^{n-1}(J)$ satisfies Eq. (1.1) for a.e. $t \in [0, T]$. □

THEOREM 1.8 (Third existence principle for solutions of (1.1), (1.2)). *Let* (1.13), (1.27), (1.28) *hold and let* $\mathcal{B}$ *be a closed subset of* $C^{n-1}[0, T]$. *Further, assume that the sequence*

$$\big\{f_k\big(t, u_k(t), \ldots, u_k^{(n-1)}(t)\big)\big\} \quad \textit{is uniformly integrable on } [0, T]. \tag{1.31}$$

Then the assertion (1.25) *is valid. If, moreover, the functions* $u^{(i)} - c_i$, $0 \leqslant i \leqslant n-1$, *have at most a finite number of zeros in* $[0, T]$, *then* $u \in AC^{n-1}[0, T]$ *is a solution of* (1.1), (1.2).

SKETCH OF THE PROOF. *Step* 1. As in Step 1 in the proof of Theorem 1.6 we get that (1.25) is valid and $u \in \mathcal{B} \subset C^{n-1}[0, T]$.

Step 2. As in Step 2 in the proof of Theorem 1.7 we get that (1.30) is valid.

Step 3. We can argue as in Step 3 in the proof of Theorem 1.7 if we take $[0, T]$ instead of $[a, b]$ and (1.31) instead of (1.29). □

2. Existence results for singular two-point higher order BVPs

In this section we are interested in problems for higher order differential equations having singularities in their space variables only (see Section 1.3). We consider the focal, conjugate, (n, p), Sturm–Liouville and Lidstone boundary conditions which appear most frequently in literature. Boundary conditions considered are two-point, linear and homogeneous.

Existence results for the above singular problems are proved by regularization and sequential techniques which consist in the construction of a proper sequence of auxiliary regular problems and in limit processes (see Section 1.5). To prove solvability of the auxiliary regular problems we use the Nonlinear Fredholm Alternative (see, e.g., [95, Theorem 4] or [146, p. 25]) which we formulate in the form convenient for our problems. In particular, we consider the differential equation

$$u^{(n)} + \sum_{i=0}^{n-1} a_i(t)u^{(i)} = g\big(t, u, \ldots, u^{(n-1)}\big) \tag{2.1}$$

and the corresponding linear homogeneous differential equation

$$u^{(n)} + \sum_{i=0}^{n-1} a_i(t)u^{(i)} = 0 \tag{2.2}$$

where $a_i \in L_1[0,T]$, $0 \leqslant i \leqslant n-1$, $g \in Car([0,T] \times \mathbb{R}^n)$. Further, we introduce boundary conditions

$$\mathcal{L}_j(u) = r_j, \quad 1 \leqslant j \leqslant n, \tag{2.3}$$

and the corresponding homogeneous boundary conditions

$$\mathcal{L}_j(u) = 0, \quad 1 \leqslant j \leqslant n, \tag{2.4}$$

where $\mathcal{L}_j : C^{n-1}[0,T] \to \mathbb{R}$ are linear and continuous functionals and $r_j \in \mathbb{R}$, $1 \leqslant j \leqslant n$.

THEOREM 2.1 (Nonlinear Fredholm Alternative). *Let the linear homogeneous problem* (2.2), (2.4) *have only the trivial solution and let there exist a function* $\psi \in L_1[0,T]$ *such that*

$$\left|g(t, x_0, \ldots, x_{n-1})\right| \leqslant \psi(t) \quad \textit{for a.e. } t \in [0,T] \textit{ and all } x_0, \ldots, x_{n-1} \in \mathbb{R}.$$

Then the nonlinear problem (2.1), (2.3) *has a solution* $u \in AC^{n-1}[0,T]$.

2.1. *Focal conditions*

We discuss the singular $(p, n-p)$ right focal problem

$$(-1)^{n-p} u^{(n)} = f\big(t, u, \ldots, u^{(n-1)}\big), \tag{2.5}$$

$$u^{(i)}(0) = 0, \quad 0 \leqslant i \leqslant p-1, \qquad u^{(j)}(T) = 0, \ p \leqslant j \leqslant n-1, \tag{2.6}$$

where $n \geqslant 2$, $p \in \mathbb{N}$ is fixed, $1 \leqslant p \leqslant n-1$, $f \in Car([0,T] \times \mathcal{D})$ with

$$\mathcal{D} = \begin{cases} \underbrace{\mathbb{R}_+^{p+1} \times \mathbb{R}_- \times \mathbb{R}_+ \times \mathbb{R}_- \times \cdots \times \mathbb{R}_+}_{n} & \text{if } n-p \text{ is odd,} \\ \underbrace{\mathbb{R}_+^{p+1} \times \mathbb{R}_- \times \mathbb{R}_+ \times \mathbb{R}_- \times \cdots \times \mathbb{R}_-}_{n} & \text{if } n-p \text{ is even} \end{cases}$$

and f may be singular at the value 0 of all its space variables. Here $\mathbb{R}_- = (-\infty, 0)$ and $\mathbb{R}_+ = (0, \infty)$. Notice that if $f > 0$ then the singular points corresponding to the solutions of problem (2.5), (2.6) are only of type I. The Green function of problem $u^{(n)} = 0$, (2.6) is presented in [18] and [19].

The existence result for the singular problem (2.5), (2.6) is given in the following theorem.

THEOREM 2.2 [120, Theorem 4.3]. *Let* $f \in Car([0,T] \times \mathcal{D})$ *and let there exist positive constants* ε *and* r *such that*

$$\varepsilon (T-t)^r \leqslant f(t, x_0, \ldots, x_{n-1})$$

for a.e. $t \in [0, T]$ *and each* $(x_0, \ldots, x_{n-1}) \in \mathcal{D}$.

Also assume that for a.e. $t \in [0, T]$ *and each* $(x_0, \ldots, x_{n-1}) \in \mathcal{D}$ *we have*

$$f(t, x_0, \ldots, x_{n-1}) \leqslant \varphi(t) + \sum_{i=0}^{n-1} \omega_i\bigl(|x_i|\bigr) + \sum_{i=0}^{n-1} h_i(t)|x_i|^{\alpha_i},$$

where $\alpha_i \in (0, 1)$, $\varphi, h_i \in L_1[0, T]$ *are nonnegative*, $\omega_i : \mathbb{R}_+ \to \mathbb{R}_+$ *are nonincreasing*, $0 \leqslant i \leqslant n-1$, *and*

$$\int_0^T \omega_i\bigl(t^{r+n-i}\bigr)\,\mathrm{d}t < \infty \quad \textit{for } 0 \leqslant i \leqslant n-1.$$

Then there exists a solution $u \in AC^{n-1}[0, T]$ *of problem* (2.5), (2.6) *with*

$$\begin{cases} u^{(i)} > 0 \quad \textit{on } (0, T] & \textit{for } 0 \leqslant i \leqslant p-1, \\ (-1)^{j-p} u^{(j)} > 0 \quad \textit{on } [0, T) & \textit{for } p \leqslant j \leqslant n-1. \end{cases} \tag{2.7}$$

SKETCH OF PROOF. *Step* 1. *Construction of a sequence of regular differential equations related to Eq.* (2.5).

Put

$$\varphi^*(t) = \varphi(t) + \sum_{i=0}^{n-1} \omega_i(1) + \sum_{i=0}^{n-1} h_i(t) \quad \text{for a.e. } t \in [0, T].$$

Then $\varphi^* \in L_1[0, T]$ and there exists $r^* > 0$ such that the estimate $\|u^{(n-1)}\|_\infty < r^*$ is valid for any function $u \in AC^{n-1}[0, T]$ satisfying (2.6), $(-1)^{n-p} u^{(n)}(t) \geqslant \varepsilon (T-t)^r$ and

$$(-1)^{n-p} u^{(n)}(t) \leqslant \varphi^*(t) + \sum_{i=0}^{n-1} \omega_i\bigl(\bigl|u^{(i)}(t)\bigr|\bigr) + \sum_{i=0}^{n-1} h_i(t)\bigl|u^{(i)}(t)\bigr|^{\alpha_i}$$

for a.e. $t \in [0, T]$. Now for $m \in \mathbb{N}$, $0 \leqslant i \leqslant n-1$ and $x \in \mathbb{R}$, put $\varrho_i = 1 + r^* T^{n-i-1}$ and

$$\sigma_i\bigl(\tfrac{1}{m}, x\bigr) = \begin{cases} \frac{1}{m} \operatorname{sign} x & \text{for } |x| < \frac{1}{m}, \\ x & \text{for } \frac{1}{m} \leqslant |x| \leqslant \varrho_i, \\ \varrho_i \operatorname{sign} x & \text{for } \varrho_i < |x|. \end{cases}$$

Extend f onto $[0, T] \times (\mathbb{R} \setminus \{0\})^n$ as an even function in each of its space variables x_i, $0 \leqslant i \leqslant n-1$, and for a.e. $t \in [0, T]$ and each $(x_0, \ldots, x_{n-1}) \in \mathbb{R}^n$, define auxiliary functions

$$f_m(t, x_0, \ldots, x_{n-1}) = f\left(t, \sigma_0\left(\frac{1}{m}, x_0\right), \ldots, \sigma_{n-1}\left(\frac{1}{m}, x_{n-1}\right)\right), \quad m \in \mathbb{N}.$$

In this way we get the family of regular differential equations

$$(-1)^{n-p}u^{(n)} = f_m\bigl(t, u, \ldots, u^{(n-1)}\bigr) \tag{2.8}$$

depending on $m \in \mathbb{N}$.

Step 2. *Properties of solutions to problems* (2.8), (2.6).

By Theorem 2.1 we show that for any $m \in \mathbb{N}$, problem (2.8), (2.6) has a solution $u_m \in AC^{n-1}[0,T]$ satisfying (for $t \in [0,T]$)

$$\begin{cases} u_m^{(i)}(t) \geqslant ct^{r+n-i} & \text{if } 0 \leqslant i \leqslant p-1, \\ (-1)^{i-p}u_m^{(i)}(t) \geqslant c(T-t)^{r+n-i} & \text{if } p \leqslant i \leqslant n-1, \end{cases} \tag{2.9}$$

where c is a positive constant and $\|u_m^{(n-1)}\|_\infty < r^*$. Moreover, the sequence $\{u_m^{(n-1)}\}$ is equicontinuous on $[0,T]$. By virtue of the Arzelà–Ascoli theorem, a convergent subsequence $\{u_{k_m}\}$ exists and let $\lim_{m\to\infty} u_{k_m} = u$. Then $u \in C^{n-1}[0,T]$, u satisfies (2.6) and, because of (2.9),

$$\begin{aligned} &u^{(i)}(t) \geqslant ct^{r+n-i} && \text{for } t \in [0,T] \text{ and } 1 \leqslant i \leqslant p-1, \\ &(-1)^{i-p}u^{(i)}(t) \geqslant c(T-t)^{r+n-i} && \text{for } t \in [0,T] \text{ and } p \leqslant i \leqslant n-1. \end{aligned}$$

Also,

$$\bigl|f_{k_m}\bigl(t, u_{k_m}(t), \ldots, u_{k_m}^{(n-1)}(t)\bigr)\bigr| \leqslant \varrho(t) \quad \text{for a.e. } t \in [0,T] \text{ and all } m \in \mathbb{N},$$

where $\varrho \in L_1[0,T]$ and

$$\lim_{m\to\infty} f_{k_m}\bigl(t, u_{k_m}(t), \ldots, u_{k_m}^{(n-1)}(t)\bigr) = f\bigl(t, u(t), \ldots, u^{(n-1)}(t)\bigr) \quad \text{for a.e. } t \in [0,T].$$

Now, letting $m \to \infty$ in

$$u_{k_m}^{(n-1)}(t) = u_{k_m}^{(n-1)}(0) + (-1)^{n-p}\int_0^t f_{k_m}\bigl(s, u_{k_m}(s), \ldots, u_{k_m}^{(n-1)}(s)\bigr)\,\mathrm{d}s$$

we conclude

$$u^{(n-1)}(t) = u^{(n-1)}(0) + (-1)^{n-p}\int_0^t f\bigl(s, u(s), \ldots, u^{(n-1)}(s)\bigr)\,\mathrm{d}s, \quad t \in [0,T].$$

Hence $u \in AC^{n-1}[0,T]$ and u is a solution of problem (2.5), (2.6). □

EXAMPLE. Let $\gamma \in (0,1)$, $\alpha_i \in [0,1)$, $c_i \in (0,\infty)$, $\beta_i \in (0, \frac{1}{n+\gamma-i})$ and let $h_i \in L_1[0,T]$ be nonnegative for $0 \leqslant i \leqslant n-1$. Then, by Theorem 2.2, there exists a solution $u \in AC^{n-1}[0,T]$ of the differential equation

$$(-1)^{n-p}u^{(n)} = \left(\frac{T-t}{t}\right)^{\gamma} + \sum_{i=0}^{n-1} \frac{c_i}{|u^{(i)}|^{\beta_i}} + \sum_{i=0}^{n-1} h_i(t)\bigl|u^{(i)}\bigr|^{\alpha_i}$$

satisfying the $(p, n-p)$ right focal boundary conditions (2.6) and (2.7).

REMARK 2.3. Substituting $t = T - s$ in (2.5), (2.6) and using Theorem 2.2 we can also give results for the existence of solutions to singular differential equations satisfying $(n-p, p)$ left focal boundary conditions

$$u^{(i)}(0) = 0, \quad p \leqslant i \leqslant n-1, \qquad u^{(j)}(T) = 0, \quad 0 \leqslant j \leqslant p-1$$

(see [120, Theorem 4.4]).

The singular problem (2.5), (2.6) was also considered on the interval [0, 1] by Agarwal, O'Regan and Lakshmikantham in [12] where f is assumed to be continuous and independent of space variables $x_p, \dots, x_{n-1}$ and may be singular at $x_i = 0$, $0 \leqslant i \leqslant p-1$. They examined the problem

$$\begin{cases} (-1)^{n-p}u^{(n)} = \varphi(t)h\bigl(t, u, \dots, u^{(p-1)}\bigr), \\ u^{(i)}(0) = 0, \quad 0 \leqslant i \leqslant p-1, \qquad u^{(j)}(1) = 0, \quad p \leqslant j \leqslant n-1, \end{cases} \tag{2.10}$$

under the assumptions

$$\varphi \in C^0(0,1) \quad \text{with } \varphi > 0 \text{ on } (0,1) \text{ and } \varphi \in L_1[0,1], \tag{2.11}$$

$$h : [0,1] \times \mathbb{R}_+^p \to \mathbb{R}_+ \text{ is continuous,} \tag{2.12}$$

$$\begin{cases} h(t, x_0, \dots, x_{p-1}) \leqslant \sum_{i=0}^{p-1} g_i(x_i) + r\bigl(\max\{x_0, \dots, x_{p-1}\}\bigr) \quad \text{on } [0,1] \times \mathbb{R}_+^p \\ \text{with } g_i > 0 \text{ continuous and nonincreasing on } \mathbb{R}_+ \text{ for each } i = 0, \dots, p-1 \\ \text{and } r \geqslant 0 \text{ continuous and nondecreasing on } [0,\infty), \end{cases} \tag{2.13}$$

$$\begin{cases} h(t, x_0, \dots, x_{p-1}) \geqslant \sum_{i=0}^{p-1} h_i(x_i) \quad \text{on } [0,1] \times \mathbb{R}_+^p \\ \text{with } h_i > 0 \text{ continuous and nonincreasing on } \mathbb{R}_+ \text{ for each } i = 0, \dots, p-1, \end{cases} \tag{2.14}$$

$$\begin{cases} \int_0^1 \varphi(t) g_i\big(k_i t^{p-i}\big)\,\mathrm{d}t < \infty \quad \text{for each } i = 0, \ldots, p-1, \\ \text{where } k_i > 0 \ (i = 0, \ldots, p-1) \text{ are constants} \end{cases} \tag{2.15}$$

and

$$\begin{cases} \text{if } z > 0 \text{ satisfies } z \leqslant a_0 + b_0 r(z) \text{ for constants } a_0 \geqslant 0 \text{ and } b_0 \geqslant 0, \\ \text{then there exists a constant } K \text{ (which may depend only on } a_0 \text{ and } b_0) \\ \text{such that } z \leqslant K. \end{cases} \tag{2.16}$$

The next result was proved by sequential technique and a nonlinear alternative of Leray–Schauder type [77, Theorem 2.3].

THEOREM 2.4 [12, Theorem 2.1]. *Suppose* (2.11)–(2.16) *hold. Then problem* (2.10) *has a solution* $u \in C^{n-1}[0,1] \cap C^n(0,1)$ *with* $u^{(i)} > 0$ *on* $(0,1]$ *for* $0 \leqslant i \leqslant p-1$.

EXAMPLE. Consider the problem

$$\begin{cases} (-1)^{n-p} u^{(n)} = \displaystyle\sum_{i=0}^{p-1} \left(\frac{1}{(u^{(i)})^{\beta_i}} + \mu_i \big(u^{(i)}\big)^{\alpha_i} + \tau_i \right), \\ u^{(i)}(0) = 0, \quad 0 \leqslant i \leqslant p-1, \qquad u^{(j)}(1) = 0, \quad p \leqslant j \leqslant n-1 \end{cases} \tag{2.17}$$

with $\beta_i \in (0,\infty)$, $\mu_i, \tau_i \in [0,\infty)$, $\alpha_i \in [0,1)$ for $0 \leqslant i \leqslant p-1$. In addition, assume $\beta_i(p-i) < 1$ for $i = 0, \ldots, p-1$. Theorem 2.4 guarantees that problem (2.17) has a solution $u \in C^{n-1}[0,1] \cap C^n(0,1)$ with $u^{(i)} > 0$ on $(0,1]$ for $0 \leqslant i \leqslant p-1$.

2.2. *Conjugate conditions*

Let $1 \leqslant p \leqslant n-1$ be a fixed natural number. Consider the $(p, n-p)$ conjugate problem

$$(-1)^p u^{(n)} = f\big(t, u, \ldots, u^{(n-1)}\big), \tag{2.18}$$

$$u^{(i)}(0) = 0, \quad 0 \leqslant i \leqslant n-p-1, \qquad u^{(j)}(T) = 0, \quad 0 \leqslant j \leqslant p-1 \tag{2.19}$$

where $n \geqslant 2$, $f \in Car([0,T] \times \mathcal{D})$, $\mathcal{D} = (0,\infty) \times (\mathbb{R} \setminus \{0\})^{n-1}$ and f may be singular at the value 0 of all its space variables.

Replacing t by $T - t$, if necessary, we may assume that $p \in \{1, \ldots, \frac{n}{2}\}$ for n even and $p \in \{1, \ldots, \frac{n+1}{2}\}$ for n odd. We observe that the larger p is chosen, the more complicated structure of the set of all singular points of a solution to (2.18), (2.19) and its derivatives is obtained. We note that solutions of problem (2.18), (2.19) have singular points of type I at $t = 0$ and/or $t = T$ and also singular points of type II. Since the singular problem (2.18), (2.19) for $n = 2$ is the Dirichlet problem discussed in Section 4, we assume that $n > 2$.

THEOREM 2.5 ([121, Theorems 2.1 and 2.7] and [123]). *Let* $n > 2$ *and* $1 \leqslant p \leqslant n-1$ *be fixed natural numbers. Suppose that the following conditions are satisfied*:

$$\begin{cases} f \in Car\big([0,T] \times \mathcal{D}\big) \textit{ and there exists } c > 0 \textit{ such that} \\ \quad c \leqslant f(t, x_0, \dots, x_{n-1}) \\ \textit{for a.e. } t \in [0,T] \textit{ and all } (x_0, \dots, x_{n-1}) \in \mathcal{D}, \end{cases} \tag{2.20}$$

$$\begin{cases} h \in Car\big([0,T] \times [0,\infty)\big) \textit{ is nondecreasing in its second variable and} \\ \quad \limsup_{z\to\infty} \dfrac{1}{z} \displaystyle\int_0^T h(t,z)\,\mathrm{d}t < \left(1 + \sum_{i=0}^{n-2} \frac{T^{n-i-1}}{(n-i-2)!}\right)^{-1}, \end{cases} \tag{2.21}$$

$$\begin{cases} \omega_i : (0,\infty) \to (0,\infty) \textit{ are nonincreasing and} \\ \quad \displaystyle\int_0^T \omega_i\big(t^{n-i}\big)\,\mathrm{d}t < \infty \quad \textit{for } 0 \leqslant i \leqslant n-1, \end{cases} \tag{2.22}$$

$$\begin{cases} f(t, x_0, \dots, x_{n-1}) \leqslant h\left(t, \displaystyle\sum_{i=0}^{n-1} |x_i|\right) + \displaystyle\sum_{i=0}^{n-1} \omega_i\big(|x_i|\big) \\ \textit{for a.e. } t \in [0,T] \textit{ and all } (x_0, \dots, x_{n-1}) \in \mathcal{D}. \end{cases} \tag{2.23}$$

Then the $(p, n-p)$ *conjugate problem* (2.18), (2.19) *has a solution* $u \in AC^{n-1}[0,T]$ *and* $u > 0$ *on* $(0,T)$.

SKETCH OF PROOF. *Step* 1. *Uniform integrability.*

Put

$$\mathcal{B} = \big\{u \in AC^{n-1}[0,T]\colon u \text{ satisfies (2.19) and } (-1)^p u^{(n)}(t) \geqslant c \\ \text{for a.e. } t \in [0,T]\big\},$$

where $c > 0$ is taken from (2.20). Conditions (2.20) and (2.22) guarantee that there exists a positive constant A such that

$$\int_0^T \omega_i\big(\big|u^{(i)}(t)\big|\big)\,\mathrm{d}t \leqslant A \quad \text{for each } u \in \mathcal{B} \text{ and } 0 \leqslant i \leqslant n-1$$

and that the set of functions $\{\omega_i(|u^{(i)}(t)|)\colon u \in \mathcal{B},\ 0 \leqslant i \leqslant n-1\}$ is uniformly integrable on $[0,T]$. Also, $u > 0$ on $(0,T)$ for each $u \in \mathcal{B}$.

Step 2. *Estimates of functions belonging to* $\mathcal{B}$.

By virtue of (2.21), there exists $r^* > 1$ such that the estimate $\|u\|_{C^{n-1}} < r^*$ holds for each function $u \in \mathcal{B}$ satisfying

$$u^{(n)}(t) \leqslant h\left(t, n + \sum_{i=0}^{n-1} \big|u^{(i)}(t)\big|\right) + \sum_{i=0}^{n-1} \big[\omega_i\big(\big|u^{(i)}(t)\big|\big) + \omega_i(1)\big] \\ \text{for a.e. } t \in [0,T].$$

Step 3. *Construction of regular problems to* (2.18), (2.19) *and properties of their solutions.*

For $m \in \mathbb{N}$, let $h_m \in Car([0,T] \times ([0,\infty) \times \mathbb{R}^{n-1}))$ be such that $h_m(t, x_0, \dots, x_{n-1}) = f(t, x_0, \dots, x_{n-1})$ for a.e. $t \in [0,T]$ and any $x_0 \geqslant \frac{1}{m}$, $|x_j| \geqslant \frac{1}{m}$, $1 \leqslant j \leqslant n-1$. Put

$$f_m(t, x_0, x_1, \dots, x_{n-1}) = h_m\bigl(t, \sigma_0(x_0), \sigma(x_1), \dots, \sigma(x_{n-1})\bigr)$$

for a.e. $t \in [0,T]$ and each $(x_0, x_1, \dots, x_{n-1}) \in \mathbb{R}^n$, where

$$\sigma_0(x) = \begin{cases} |x| & \text{if } |x| \leqslant r^*, \\ r^* & \text{if } |x| > r^*, \end{cases} \qquad \sigma(x) = \begin{cases} x & \text{if } |x| \leqslant r^*, \\ r^* \operatorname{sign} x & \text{if } |x| > r^*. \end{cases}$$

Now, the sequence of regular differential equations

$$(-1)^p u^{(n)} = f_m\bigl(t, u, \dots, u^{(n-1)}\bigr) \tag{2.24}$$

is considered. It follows from Theorem 2.1 that for each $m \in \mathbb{N}$ there exists a solution u_m of problem (2.24), (2.19) and $\|u_m\|_{C^{n-1}} < r^*$. Moreover, the sequence of functions $\{f_m(t, u_m(t), \dots, u_m^{(n-1)}(t))\}$ is uniformly integrable on $[0,T]$. By the Arzelà–Ascoli theorem there exists a subsequence $\{u_{k_m}\}$ converging in $C^{n-1}[0,T]$, $\lim_{m\to\infty} u_{k_m} = u$. Then $u \in C^{n-1}[0,T]$ satisfies (2.19) and for each $i \in \{1, \dots, n-1\}$, the function $u^{(n-i)}$ has a finite number of zeros $0 \leqslant a_{i1} < \dots < a_{i,p_i} \leqslant T$ and satisfies

$$\bigl|u^{(n-i)}(t)\bigr| \geqslant \frac{c}{i!}(t - a_{ik})^i \quad \text{for } t \in [a_{ik}, a_{i,k+1}]$$

or

$$\bigl|u^{(n-i)}(t)\bigr| \geqslant \frac{c}{i!}(a_{i,k+1} - t)^i \quad \text{for } t \in [a_{ik}, a_{i,k+1}]$$

(see [123]). Therefore $u \in AC^{n-1}[0,T]$ and u is a solution of problem (2.18), (2.19) due to Theorem 1.8. From Step 1 and (2.20) it follows that $u > 0$ on $(0,T)$. □

EXAMPLE. Let $p \in \{1, \dots, n-1\}$. Let $\alpha_i \in (0,1)$, $\beta_i \in (0, \frac{1}{n-i})$ and $b_i \in L_1[0,T]$, $c_i \in L_\infty[0,T]$ be nonnegative for $0 \leqslant i \leqslant n-1$. Also, let $\varphi \in L_1[0,T]$ and $\varphi(t) \geqslant c$ for a.e. $t \in [0,T]$ with $c > 0$. Then the differential equation

$$(-1)^p u^{(n)} = \varphi(t) + \sum_{i=0}^{n-1} \left(b_i(t) \bigl|u^{(i)}\bigr|^{\alpha_i} + \frac{c_i(t)}{|u^{(i)}|^{\beta_i}} \right)$$

has a solution $u \in AC^{n-1}[0,T]$ satisfying (2.19) and $u > 0$ on $(0,T)$.

2.3. (n, p) *boundary conditions*

Here we are concerned with the singular (n, p) problem

$$-u^{(n)} = f\big(t, u, \dots, u^{(n-1)}\big), \tag{2.25}$$

$$\begin{cases} u^{(i)}(0) = 0, & 0 \leqslant i \leqslant n-2, \\ u^{(p)}(T) = 0, & p \text{ fixed}, 0 \leqslant p \leqslant n-1, \end{cases} \tag{2.26}$$

where $n \geqslant 2$, $f \in Car([0, T] \times \mathcal{D})$, $\mathcal{D} = (0, \infty) \times (\mathbb{R} \setminus \{0\})^{n-2} \times \mathbb{R}$ and f may be singular at the value 0 of its space variables $x_0, \dots, x_{n-2}$. Notice that the $(n, 0)$ problem is simultaneously also the $(1, n-1)$ conjugate problem. For $f > 0$, solutions of problem (2.25), (2.26) have singular points of type I at $t = 0$, $t = T$ and also singular points of type II.

THEOREM 2.6 [14, Theorem 4.2]. *Suppose*

$$\begin{cases} f \in Car([0, T] \times \mathcal{D}) \textit{ and there exist positive } \psi \in L_1[0, T] \textit{ and} \\ K \in (0, \infty) \textit{ such that } \psi(t) \leqslant f(t, x_0, \dots, x_{n-1}) \quad \textit{for a.e. } t \in [0, T] \\ \textit{and each } (x_0, \dots, x_{n-1}) \in (0, K] \times (\mathbb{R} \setminus \{0\})^{n-2} \times \mathbb{R}, \end{cases} \tag{2.27}$$

$$\begin{cases} h_j \in L_1[0, T] \textit{ is nonnegative}, \omega_i : (0, \infty) \to (0, \infty) \textit{ is nonincreasing}, \\ \displaystyle\sum_{k=0}^{n-1} \frac{1}{(n-k-1)!} \int_0^T h_k(t) t^{n-k-1} \, \mathrm{d}t < 1, \quad \int_0^T \omega_i\big(t^{n-i-1}\big) \, \mathrm{d}t < \infty \\ \textit{for } 0 \leqslant j \leqslant n-1 \textit{ and } 0 \leqslant i \leqslant n-2, \end{cases} \tag{2.28}$$

$$\begin{cases} \textit{for a.e. } t \in [0, T] \textit{ and each } (x_0, \dots, x_{n-1}) \in \mathcal{D} \textit{ we have} \\ \qquad f(t, x_0, \dots, x_{n-1}) \leqslant \varphi(t) + \displaystyle\sum_{i=0}^{n-2} \omega_i\big(|x_i|\big) + \sum_{j=0}^{n-1} h_j(t)|x_j| \\ \textit{where } \varphi \in L_1[0, T] \textit{ is nonnegative.} \end{cases} \tag{2.29}$$

Then there exists a solution $u \in AC^{n-1}[0, T]$ *of problem* (2.25), (2.26) *with* $u^{(i)} > 0$ *on* $(0, T]$ *for* $0 \leqslant i \leqslant p-1$ (*if* $p \geqslant 1$) *and* $u^{(p)} > 0$ *on* $(0, T)$.

SKETCH OF PROOF. *Step* 1. *A priori bounds for solutions of problem* (2.25), (2.26).

Upper and lower bounds for solutions of problem (2.25), (2.26) and their derivatives are given by means of the Green function of the problem $-u^{(n)} = 0$, (2.26) (see, e.g., [1]).

Step 2. *Construction of auxiliary regular problems and properties of their solutions.*

Using Step 1, a sequence of regular differential equations

$$-u^{(n)} = f_m\big(t, u, \dots, u^{(n-1)}\big), \quad m \in \mathbb{N}, \ m \geqslant m_0 \geqslant \frac{1}{K}, \tag{2.30}$$

with $f_m \in Car([0, T] \times \mathbb{R}^n)$ is constructed. By the Leray–Schauder degree theory, we prove that for any $m \geqslant m_0$ problem (2.30), (2.26) has a solution u_m. The sequence

$\{u_m\}_{m=m_0}^{\infty}$ is bounded in $C^{n-1}[0,T]$ and $\{u_m^{(n-1)}\}_{m=m_0}^{\infty}$ is equicontinuous on $[0,T]$. The Arzelà–Ascoli theorem guarantees the existence of a subsequence $\{u_k\}_{k=1}^{\infty}$ of $\{u_m\}_{m=m_0}^{\infty}$ converging in $C^{n-1}[0,T]$ to a function u. Then $u \in C^{n-1}[0,T]$ satisfies (2.26) and $u^{(i)} > 0$ on $(0,T]$ for $0 \leqslant i \leqslant p-1$ (if $p \geqslant 1$) and $u^{(p)} > 0$ on $(0,T)$. Moreover, for each $i \in \{p+1,\dots,n-2\}$, the function $u^{(i)}$ has a unique zero ξ_i in $(0,T)$ $(0 < \xi_{n-2} < \xi_{n-1} < \cdots < \xi_{p+1} < T)$ and satisfies

$$u^{(i)}(t) \geqslant \begin{cases} ct^{n-i-1} & \text{for } t \in [0,\xi_{i+1}], \\ c(\xi_i - t) & \text{for } t \in [\xi_{i+1},\xi_i], \end{cases} \qquad u^{(i)}(t) \leqslant c(\xi_i - t) \quad \text{for } t \in [\xi_i, T]$$

where c is a positive constant. Since $\{f_k(t,u_k(t),\dots,u_k^{(n-1)}(t))\}$ is uniformly integrable on $[0,T]$, we can use Theorem 1.8 concluding that $u \in AC^{n-1}[0,T]$ and u is a solution of problem (2.25), (2.26). □

A related existence result for the singular (n,p) problem

$$\begin{cases} -u^{(n)} = \varphi(t)h\big(t,u,\dots,u^{(p-1)}\big), \\ u^{(i)}(0) = 0, \quad 0 \leqslant i \leqslant n-2, \\ u^{(p)}(1) = 0, \quad p \text{ fixed}, \ 1 \leqslant p \leqslant n-1 \end{cases} \tag{2.31}$$

was presented in [12] with h continuous and positive on $[0,1] \times (0,\infty)^p$ and $\varphi \in C^0(0,1) \cap L_1[0,1]$ positive on $(0,1)$. In this setting solutions of problem (2.31) cannot have singular points of type II. The result is the following.

THEOREM 2.7 [12, Theorem 3.1]. *Suppose that* (2.11)–(2.14) *and* (2.16) *hold and*

$$\begin{cases} \displaystyle\int_0^1 \varphi(t) g_i\big(k_i t^{n-1-i}\big)\,\mathrm{d}t < \infty \quad \textit{for each } i = 0,\dots,p-1, \\ \textit{where } k_i > 0, \ i = 0,\dots,p-1, \ \textit{are constants.} \end{cases}$$

Then problem (2.31) *has a solution* $u \in C^{n-1}[0,1] \cap C^n(0,1)$ *with* $u^{(j)} > 0$ *on* $(0,1]$ *for* $0 \leqslant j \leqslant p-1$.

2.4. *Sturm–Liouville conditions*

We are now concerned with the Sturm–Liouville problem for the nth-order differential equation (2.25), $n \geqslant 3$, and the boundary conditions

$$\begin{cases} u^{(i)}(0) = 0, \quad 0 \leqslant i \leqslant n-3, \\ \alpha u^{(n-2)}(0) - \beta u^{(n-1)}(0) = 0, \\ \gamma u^{(n-2)}(T) + \delta u^{(n-1)}(T) = 0, \end{cases} \tag{2.32}$$

where $\alpha, \gamma > 0$ and $\beta, \delta \geqslant 0$. Notice that the function f in Eq. (2.25) may be singular at the value 0 of its space variables $x_0, \dots, x_{n-1}$. If $f > 0$, solutions of problem (2.25), (2.32) have singular points of type I at the end points of the interval $[0, T]$ and also singular points of type II.

We will impose the following conditions on the function f in (2.25):

$$\begin{cases} f \in Car\big([0,T] \times \mathcal{D}\big) \text{ where } \mathcal{D} = (0,\infty)^{n-1} \times \big(\mathbb{R} \setminus \{0\}\big) \\ \text{and there exist positive constants } \varepsilon \text{ and } r \text{ such that} \\ \quad \varepsilon t^r \leqslant f(t, x_0, \dots, x_{n-1}) \\ \text{for a.e. } t \in [0,T] \text{ and each } (x_0, \dots, x_{n-1}) \in \mathcal{D}; \end{cases} \tag{2.33}$$

$$\begin{cases} \text{for a.e. } t \in [0,T] \text{ and each } (x_0, \dots, x_{n-1}) \in \mathcal{D}, \\ \quad f(t, x_0, \dots, x_{n-1}) \leqslant \varphi(t) + \displaystyle\sum_{i=0}^{n-1} \omega_i\big(|x_i|\big) + \sum_{i=0}^{n-1} h_i(t)|x_i|^{\alpha_i} \\ \text{with } \alpha_i \in (0,1),\ \varphi, h_i \in L_1[0,T] \text{ nonnegative}, \\ \omega_i : (0,\infty) \to (0,\infty) \text{ nonincreasing, } 0 \leqslant i \leqslant n-1, \text{ and} \\ \quad \displaystyle\int_0^T \omega_{n-1}\big(t^{r+1}\big)\,\mathrm{d}t < \infty, \quad \int_0^T \omega_i\big(t^{n-i-1}\big)\,\mathrm{d}t < \infty \quad \text{for } 0 \leqslant i \leqslant n-2; \end{cases} \tag{2.34}$$

$$\begin{cases} \text{for a.e. } t \in [0,T] \text{ and each } (x_0, \dots, x_{n-1}) \in \mathcal{D}, \\ \quad f(t, x_0, \dots, x_{n-1}) \leqslant \varphi(t) + \displaystyle\sum_{i=0, i \neq n-2}^{n-1} \omega_i\big(|x_i|\big) + q(t)\omega_{n-2}\big(|x_{n-2}|\big) \\ \qquad + \displaystyle\sum_{i=0}^{n-1} h_i(t)|x_i|^{\alpha_i} \\ \text{with } \alpha_i \in (0,1),\ \varphi, q, h_i \in L_1[0,T] \text{ nonnegative}, \\ \omega_i : (0,\infty) \to (0,\infty) \text{ nonincreasing, } 0 \leqslant i \leqslant n-1, \text{ and} \\ \quad \displaystyle\int_0^T \omega_{n-1}\big(t^{r+1}\big)\,\mathrm{d}t < \infty, \quad \int_0^T \omega_j\big(t^{n-j-2}\big)\,\mathrm{d}t < \infty \quad \text{for } 0 \leqslant j \leqslant n-3. \end{cases} \tag{2.35}$$

The next two theorems show that our sufficient conditions for the solvability of problem (2.25), (2.32) with $\min\{\beta, \delta\} > 0$ are weaker then those for this problem with $\min\{\beta, \delta\} = 0$.

THEOREM 2.8 [120, Theorem 4.1]. *Let conditions* (2.33) *and* (2.34) *be satisfied and let* $\min\{\beta, \delta\} = 0$. *Then problem* (2.25), (2.32) *has a solution* $u \in AC^{n-1}[0,T]$, $u^{(n-2)} > 0$ *on* $(0,T)$ *and* $u^{(j)} > 0$ *on* $(0,T]$ *for* $0 \leqslant j \leqslant n-3$.

SKETCH OF PROOF. *Step* 1. *A priori bounds for the solution of* (2.25), (2.32).

By (2.33) and by the properties of the Green function to problem

$$-u''=0,\quad \alpha u(0)-\beta u'(0)=0,\quad \gamma u(T)+\delta u'(T)=0,$$

the existence of a positive constant A is proved such that for any function $u\in AC^{n-1}[0,T]$ satisfying (2.32) and $-u^{(n)}(t)\geqslant \varepsilon t^r$ for a.e. $t\in[0,T]$ we have

$$u^{(n-2)}(t)\geqslant\begin{cases}At & \text{for } t\in[0,\frac{T}{2}],\\ A(T-t) & \text{for } t\in(\frac{T}{2},T],\end{cases}\tag{2.36}$$

$$u^{(j)}(t)\geqslant\frac{A}{4(n-j-1)!}t^{n-j-1}\quad\text{for } t\in[0,T] \text{ and } 0\leqslant j\leqslant n-3\tag{2.37}$$

and

$$u^{(n-1)}(t)\begin{cases}\geqslant\dfrac{\varepsilon}{r+1}(\xi-t)^{r+1} & \text{for } t\in[0,\xi],\\ <-\dfrac{\varepsilon}{r+1}(t-\xi)^{r+1} & \text{for } t\in(\xi,T],\end{cases}\tag{2.38}$$

where $\xi\in(0,T)$ (depending on the solution u) is the unique zero of $u^{(n-1)}$. Condition (2.34) guarantees the existence of a positive constant S such that $\|u\|_{C^{n-1}}\leqslant S$ for any solution u to (2.25), (2.32).

Step 2. *Construction of regular differential equations.*

Using Step 1, a sequence of regular differential equations (2.30) is constructed where $f_m\in Car([0,T]\times\mathbb{R}^n)$,

$$f_m(t,x_0,\dots,x_{n-1})=f(t,x_0,\dots,x_{n-1})$$

for a.e. $t\in[0,T]$ and all $(x_0,\dots,x_{n-1})\in\mathbb{R}^n$ such that

$$\frac{1}{m}\leqslant x_j\leqslant S+1\quad\text{if } 0\leqslant j\leqslant n-2,\qquad \frac{1}{m}\leqslant|x_{n-1}|\leqslant S+1$$

and

$$\sup\bigl\{f_m(t,x_0,\dots,x_{n-1})\colon (x_0,\dots,x_{n-1})\in\mathbb{R}^n\bigr\}\in L_1[0,T]\quad\text{for all } m\in\mathbb{N}.$$

Then Theorem 2.1 guarantees that the regular problem (2.30), (2.32) has a solution u_m which satisfies (2.36)–(2.38) (with u_m instead of u).

Step 3. *Properties of solutions to regular problems* (2.30), (2.32).

The sequence $\{u_m\}$ is considered. It is proved that $\{u_m\}$ is bounded in $C^{n-1}[0,T]$ and, by (2.34), the sequence of functions $\{f_m(t,u_m(t),\dots,u_m^{(n-1)}(t))\}$ is uniformly integrable on $[0,T]$, which implies that $\{u_m^{(n-1)}\}$ is equicontinuous on $[0,T]$. Hence a subsequence

$\{u_{k_m}\}$ converging in $C^{n-1}[0,T]$ exists and let $\lim_{m\to\infty} u_{k_m} = u$. Since u satisfies (2.36)–(2.38), the functions $u^{(i)}$, $0 \leqslant i \leqslant n-1$, have a finite number of zeros in $[0,T]$. Therefore, by Theorem 1.8, $u \in AC^{n-1}[0,T]$ and u is a solution of problem (2.25), (2.32) such that $u^{(j)} > 0$ on $(0,T]$ for $0 \leqslant j \leqslant n-3$ and $u^{(n-2)} > 0$ on $(0,T)$. ☐

THEOREM 2.9 [120, Theorem 4.2]. *Let conditions* (2.33) *and* (2.35) *be satisfied and let* $\min\{\beta,\delta\} > 0$. *Then there exists a solution* $u \in AC^{n-1}[0,T]$ *of problem* (2.25), (2.32) *such that* $u^{(n-2)} > 0$ *on* $[0,T]$ *and* $u^{(j)} > 0$ *on* $(0,T]$ *for* $0 \leqslant j \leqslant n-3$.

SKETCH OF PROOF. Since $\min\{\beta,\delta\} > 0$, there is a positive constant B such that $u^{(n-2)} \geqslant B$ on $[0,T]$ for any solution u of problem (2.25), (2.32). Further, the inequalities (2.37) with B instead of A and (2.38) hold. Next we argue as in the sketch of proof to Theorem 2.8. ☐

2.5. *Lidstone conditions*

Let $\mathbb{R}_- = (-\infty,0)$, $\mathbb{R}_+ = (0,\infty)$ and $\mathbb{R}_0 = \mathbb{R}\setminus\{0\}$. Here we will consider the singular problem

$$(-1)^n u^{(2n)} = f\big(t,u,\dots,u^{(2n-2)}\big), \tag{2.39}$$

$$u^{(2j)}(0) = u^{(2j)}(T) = 0, \quad 0 \leqslant j \leqslant n-1, \tag{2.40}$$

where $n \geqslant 1$ and $f \in Car([0,T]\times\mathcal{D})$ with

$$\mathcal{D} = \begin{cases} \underbrace{\mathbb{R}_+ \times \mathbb{R}_0 \times \mathbb{R}_- \times \mathbb{R}_0 \times \cdots \times \mathbb{R}_+}_{4k-3} & \text{if } n = 2k-1,\\ \underbrace{\mathbb{R}_+ \times \mathbb{R}_0 \times \mathbb{R}_- \times \mathbb{R}_0 \times \cdots \times \mathbb{R}_-}_{4k-1} & \text{if } n = 2k \end{cases}$$

(for $n = 1, 2$ and 3 we have $\mathcal{D} = \mathbb{R}_+$, $\mathcal{D} = \mathbb{R}_+ \times \mathbb{R}_0 \times \mathbb{R}_-$ and $\mathcal{D} = \mathbb{R}_+ \times \mathbb{R}_0 \times \mathbb{R}_- \times \mathbb{R}_0 \times \mathbb{R}_+$, respectively). The function f may be singular at the value 0 of its space variables $x_0,\dots,x_{2n-2}$. If f is positive on $[0,T]\times\mathcal{D}$, solutions of problem (2.39), (2.40) have singular points of type I at $t=0$ and $t=T$ as well as singular points of type II.

THEOREM 2.10 [14, Theorem 4.1]. *Let the following conditions be satisfied*:

$$\begin{cases} f \in Car([0,T]\times\mathcal{D}) \textit{ and there exists } \varphi \in L_1[0,T] \textit{ such that}\\ \quad 0 < \varphi(t) \leqslant f(t,x_0,\dots,x_{2n-2})\\ \textit{for a.e. } t \in [0,T] \textit{ and each } (x_0,\dots,x_{2n-2}) \in \mathcal{D}; \end{cases} \tag{2.41}$$

$$\begin{cases} \textit{for } \ 0 \leqslant j \leqslant 2n-2, \ h_j \in L_1[0,T] \textit{ are nonnegative and} \\ \displaystyle\sum_{j=0}^{n-1} \frac{T^{2(n-j)-3}}{6^{n-j-1}} \int_0^T t(T-t)h_{2j}(t)\,\mathrm{d}t \\ \qquad \displaystyle + \sum_{j=0}^{n-2} \frac{T^{2(n-j-2)}}{6^{n-j-2}} \int_0^T t(T-t)h_{2j+1}(t)\,\mathrm{d}t < 1 \\ (\textit{here } \sum_{j=0}^{n-2} = 0 \textit{ if } n = 1); \end{cases} \tag{2.42}$$

$$\begin{cases} \omega_j : \mathbb{R}_+ \to \mathbb{R}_+ \textit{ are nonincreasing, } \Lambda > 0 \textit{ and} \\ \displaystyle\int_0^T \omega_j(s)\,\mathrm{d}s < \infty, \qquad \omega_j(uv) \leqslant \Lambda\omega_j(u)\omega_j(v) \\ \textit{for } 0 \leqslant j \leqslant 2n-2 \textit{ and } u, v \in \mathbb{R}_+; \end{cases} \tag{2.43}$$

$$\begin{cases} \displaystyle f(t, x_0, \ldots, x_{2n-2}) \leqslant \psi(t) + \sum_{j=0}^{2n-2} \omega_j\big(|x_j|\big) + \sum_{j=0}^{2n-2} h_j(t)|x_j| \\ \textit{for a.e. } t \in [0,T] \textit{ and each } (x_0, \ldots, x_{2n-2}) \in \mathcal{D}, \\ \textit{where } \psi \in L_1[0,T] \textit{ is nonnegative.} \end{cases} \tag{2.44}$$

Then problem (2.39), (2.40) *has a solution* $u \in AC^{2n-1}[0,T]$ *and*

$$(-1)^j u^{(2j)} > 0 \quad \textit{on } (0,T) \quad \textit{for } 0 \leqslant j \leqslant n-1.$$

SKETCH OF PROOF. *Step* 1. *A priori bounds for solutions of problem* (2.39), (2.40).

Using (2.41) and the properties of the Green functions to problems $u^{(2j)} = 0$, $u^{(2i)}(0) = u^{(2i)}(T) = 0$, $0 \leqslant i \leqslant j-1$, it is proved that for any solution u of problem (2.39), (2.40) and for each j, $0 \leqslant j \leqslant n-1$, the inequality $(-1)^j u^{(2j)} > 0$ holds on $(0,T)$ and the function $(-1)^j u^{(2j+1)}$ is decreasing on $[0,T]$ and vanishes at a unique $\xi_j \in (0,T)$ (depending on u). Moreover,

$$\big|u^{(2j)}(t)\big| \geqslant A\frac{T^{2(n-j)-5}}{30^{n-j-1}} t(T-t), \quad t \in [0,T],\ 0 \leqslant j \leqslant n-1,$$

and (if $n > 1$)

$$\big|u^{(2j+1)}(t)\big| \geqslant A\frac{T^{2(n-j)-7}}{30^{n-j-2}} \left| \int_{\xi_j}^t s(T-s)\,\mathrm{d}s \right|, \quad t \in [0,T],\ 0 \leqslant j \leqslant n-2,$$

where $A = \int_0^T t(T-t)\varphi(t)\,\mathrm{d}t$.

Step 2. *Construction of a sequence of regular problems.*

For each $m \in \mathbb{N}$, define $f_m \in Car([0,T] \times \mathbb{R}^{2n-1})$ satisfying

$$f_m(t, x_0, \ldots, x_{2n-2}) = f(t, x_0, \ldots, x_{2n-2})$$

for a.e. $t \in [0,T]$ and each $(x_0,\dots,x_{2n-2}) \in \mathcal{D}$, $|x_j| \geqslant \frac{1}{m}$, $0 \leqslant j \leqslant 2n-2$. By virtue of (2.44) and a fixed point theorem of Leray–Schauder type (see, e.g., [50, Corollary 8.1]), for each $m \in \mathbb{N}$ there exists a solution u_m of the regular differential equation $(-1)^n u^{(2n)} = f_m(t,u,\dots,u^{(2n-2)})$ satisfying (2.40). Further, $\|u_m\|_{C^{2n-1}} \leqslant B$ for each $m \in \mathbb{N}$ where B is a positive constant and the sequence $\{f_m(t,u_m(t),\dots,u_m^{(2n-2)}(t))\}$ is uniformly integrable on $[0,T]$ due to (2.43) and (2.44).

Step 3. *Limit processes.*

From Step 2 it follows that $\{u_m\}$ is bounded in $C^{2n-1}[0,T]$. Hence, by the Arzelà–Ascoli theorem and a compactness principle, there exists its subsequence $\{u_{k_m}\}$ which converges in $C^{2n-2}[0,T]$ and $\{u_{k_m}^{(2n-1)}(0)\}$ converges in $\mathbb{R}$. Let $\lim_{m\to\infty} u_{k_m} = u$, $\lim_{m\to\infty} u_{k_m}^{(2n-1)}(0) = C$. Then $u \in C^{2n-2}[0,T]$ satisfies (2.40) and, by Step 1, the functions $u^{(i)}$, $0 \leqslant i \leqslant 2n-1$, have a finite number of zeros on $[0,T]$. Therefore, by Theorem 1.8, $u \in AC^{2n-1}[0,T]$ and u is a solution of (2.39). Moreover, $(-1)^j u^{(2j)} > 0$ on $(0,T)$ for $0 \leqslant j \leqslant n-1$. □

2.6. *Historical and bibliographical notes*

Higher order boundary value problems with space singularities have been mostly studied by Agarwal, Eloe, Henderson, Lakshmikantham, O'Regan, Rachůnková and Staněk.

Positive solutions in the set $C^{n-1}[0,1] \cap C^n(0,1)$ were obtained in [7] for the singular $(p,n-p)$ right focal problem $(-1)^{n-p}u^{(n)} = \varphi(t)f(t,u)$, (2.6) on the interval $[0,1]$ where $f \in C^0([0,1]\times(0,\infty))$ is positive and may be singular at $u=0$. In [7] the authors also discussed applications in fluid theory and boundary layer theory.

Singular $(p,n-p)$ conjugate problems were studied in [5], [58] (with $p=1$) and [59] for the differential equation $(-1)^{n-p}u^{(n)} = f(t,u)$ where $f \in C^0((0,1)\times(0,\infty))$ and may be singular at $u=0$. Here positive solutions on $(0,1)$ belong to the class $C^{n-1}[0,1] \cap C^n(0,1)$. Existence results in [58] and [59] are proved by a fixed point theorem for operators that are decreasing with respect to a cone and those in [5] by a nonlinear alternative of Leray–Schauder.

The existence of positive solutions on $(0,1)$ to singular Sturm–Liouville problems for the differential equation $-u^{(n)} = f(t,u,\dots,u^{(n-2)})$ can be found in [20]. There $f \in C((0,1)\times(0,\infty)^{n-1})$ is positive and may be singular at the value 0 of all its space variables. The results are proved by a fixed point theorem for mappings that are decreasing with respect to a cone in a Banach space.

Existence results for positive solutions to singular $(p,n-p)$ focal, conjugate and (n,p) problems are given in [8,9] for differential equations with the right-hand side $\varphi(t)f(t,u)$ where $f \in C([0,1]\times(0,\infty))$ and may be singular at $u=0$. The paper [8] is the first to establish conditions for the existence of two solutions to singular $(p,n-p)$ focal and (n,p) problems. Further multiple solutions for singular $(p,n-p)$ focal, conjugate and (n,p) problems are established in [9]. The technique presented in [8,9] to guarantee the existence of twin solutions to the singular problems combines (i) a nonlinear alternative of Leray–Schauder, (ii) the Krasnoselskii fixed point theorem in a cone, and (iii) lower type inequalities.

Notice that in all cited papers singular points corresponding to solutions of the singular problems under discussion are only of type I.

3. Principles of solvability of singular second order BVPs with ϕ-Laplacian

In the theory of partial differential equations, the p-Laplace equation

$$\operatorname{div}\bigl(|\nabla v|^{p-2}\nabla v\bigr) = h\bigl(|x|, v\bigr) \tag{3.1}$$

is considered. Here ∇ is the gradient, $p > 1$ and $|x|$ is the Euclidean norm in $\mathbb{R}^n$ of $x = (x_1, \dots, x_n)$, $n > 1$. Radially symmetric solutions of (3.1) (i.e., solutions that depend only on the variable $r = |x|$) satisfy the ordinary differential equation

$$r^{1-n}\bigl(r^{n-1}|v'|^{p-2}v'\bigr)' = h(r, v), \quad ' = \frac{\mathrm{d}}{\mathrm{d}r}. \tag{3.2}$$

If $p = n$, the change of variables $t = \ln r$ transforms (3.2) into the equation

$$\bigl(|u'|^{p-2}u'\bigr)' = \mathrm{e}^{nt} h(\mathrm{e}^t, u), \quad ' = \frac{\mathrm{d}}{\mathrm{d}t}$$

and for $p \neq n$, the change of variables $t = r^{(p-n)/(p-1)}$ transforms (3.2) into the equation

$$\bigl(|u'|^{p-2}u'\bigr)' = \left|\frac{p-1}{p-n}\right|^p t^{\frac{p-n}{p(1-n)}} h\bigl(t^{\frac{p-1}{p-n}}, u\bigr), \quad ' = \frac{\mathrm{d}}{\mathrm{d}t}.$$

The operator $u \to (|u'|^{p-2}u')'$ is called the (one-dimensional) *p-Laplacian*. Its natural generalization is the *ϕ-Laplacian*

$$u \to \bigl(\phi(u')\bigr)',$$
$$\text{where } \phi : \mathbb{R} \to \mathbb{R} \text{ is an increasing homeomorphism and } \phi(\mathbb{R}) = \mathbb{R}. \tag{3.3}$$

Therefore Eq. (3.1) was a motivation for discussing the solutions to the differential equations

$$\bigl(|u'|^{p-2}u'\bigr)' = f(t, u, u')$$

and

$$\bigl(\phi(u')\bigr)' = f(t, u, u')$$

with the p-Laplacian and the ϕ-Laplacian, respectively.

In the next part of this section, we treat problems for second order differential equations with the ϕ-Laplacian on the left-hand side and with nonlinearities on the right-hand sides which can have singularities in their space variables. Boundary conditions under discussion are generally nonlinear and nonlocal. Using regularization and sequential techniques we present general existence principles for solvability of regular and singular problems.

3.1. *Regularization of singular problems with ϕ-Laplacian*

We discuss singular differential equations of the form

$$\big(\phi(u')\big)' = f(t, u, u') \tag{3.4}$$

with the ϕ-Laplacian. Here $f \in Car([0,T] \times \mathcal{D})$, the set $\mathcal{D} = \mathcal{D}_1 \times \mathcal{D}_2 \subset \mathbb{R}^2$ is not necessarily closed, $\mathcal{D}_1$, $\mathcal{D}_2$ are intervals and f may have singularities in its space variables on the boundary $\partial\mathcal{D}_j$ of $\mathcal{D}_j$ $(j = 1, 2)$. We note that f has a singularity on $\partial\mathcal{D}_j$ in its space variable x_j if there is an $a_j \in \partial\mathcal{D}_j$ such that for a.e. $t \in [0,T]$ and some $x_{3-j} \in \mathcal{D}_{3-j}$,

$$\limsup_{x_j \to a_j,\ x_j \in \mathcal{D}_j} \big|f(t, x_1, x_2)\big| = \infty.$$

Let $\mathcal{A}$ denote the set of functionals $\alpha : C^1[0,T] \to \mathbb{R}$ which are
(a) continuous and
(b) bounded, that is, $\alpha(\Omega)$ is bounded (in $\mathbb{R}$) for any bounded $\Omega \subset C^1[0,T]$.
For $\alpha, \beta \in \mathcal{A}$, consider the (generally nonlinear and nonlocal) boundary conditions

$$\alpha(u) = 0, \quad \beta(u) = 0. \tag{3.5}$$

DEFINITION 3.1. A function $u : [0,T] \to \mathbb{R}$ is said to be a *solution of problem* (3.4), (3.5) if $\phi(u') \in AC[0,T]$, u satisfies the boundary conditions (3.5) and $(\phi(u'(t)))' = f(t, u(t), u'(t))$ holds for a.e. $t \in [0,T]$.

Special cases of the boundary conditions (3.5) are the Dirichlet (Neumann; mixed; periodic and Sturm–Liouville type) boundary conditions which we get setting $\alpha(x) = x(0)$, $\beta(x) = x(T)$ ($\alpha(x) = x'(0)$, $\beta(x) = x'(T)$; $\alpha(x) = x(0)$, $\beta(x) = x'(T)$; $\alpha(x) = x(0) - x(T)$, $\beta(x) = x'(0) - x'(T)$ and $\alpha(x) = a_0 x(0) + a_1 x'(0)$, $\beta(x) = b_0 x(T) + b_1 x'(T)$).

In order to obtain an existence result for problem (3.4), (3.5), we use regularization and sequential techniques. For this purpose consider a sequence of regular differential equations

$$\big(\phi(u')\big)' = f_n(t, u, u') \tag{3.6}$$

where $f_n \in Car([0,T] \times \mathbb{R}^2)$. The function f_n is constructed in such a way that

$$f_n(t, x, y) = f(t, x, y) \quad \text{for a.e. } t \in [0,T] \text{ and all } (x, y) \in \mathcal{Q}_n$$

where $\mathcal{Q}_n \subset \mathcal{D}$ and roughly speaking $\mathcal{Q}_n$ converges to $\mathcal{D}$ as $n \to \infty$.

Let $h \in Car([0,T] \times \mathbb{R}^2)$ and consider the regular differential equation

$$\big(\phi(u')\big)' = h(t,u,u'). \tag{3.7}$$

A function $u : [0,T] \to \mathbb{R}$ is called a *solution of the regular problem* (3.7), (3.5) if $\phi(u') \in AC[0,T]$, u satisfies (3.5) and $(\phi(u'(t)))' = h(t,u(t),u'(t))$ for a.e. $t \in [0,T]$.

The next general existence principle can be used for solving the regular problem (3.7), (3.5).

THEOREM 3.2 (General existence principle for regular problems). *Assume* (3.3), $h \in Car([0,T] \times \mathbb{R}^2)$ *and* $\alpha, \beta \in \mathcal{A}$. *Suppose there exist positive constants* S_0 *and* S_1 *such that*

$$\|u\|_\infty < S_0, \quad \|u'\|_\infty < S_1$$

for all solutions u to the problem

$$\big(\phi(u')\big)' = \lambda h(t,u,u'), \quad \alpha(u) = 0, \quad \beta(u) = 0 \tag{3.8}$$

and each $\lambda \in [0,1]$. *Also assume there exist positive constants* Λ_0 *and* Λ_1 *such that*

$$|A| < \Lambda_0, \quad |B| < \Lambda_1 \tag{3.9}$$

for all solutions $(A,B) \in \mathbb{R}^2$ *of the system*

$$\begin{cases} \alpha(A+Bt) - \mu\alpha(-A-Bt) = 0, \\ \beta(A+Bt) - \mu\beta(-A-Bt) = 0 \end{cases} \tag{3.10}$$

and each $\mu \in [0,1]$. *Then problem* (3.7), (3.5) *has a solution.*

PROOF. Set

$$\Omega = \big\{x \in C^1[0,T]:\ \|x\|_\infty < \max\{S_0, \Lambda_0 + \Lambda_1 T\},\ \|x'\|_\infty < \max\{S_1, \Lambda_1\}\big\}.$$

Then Ω is an open, bounded and symmetric with respect to $0 \in C^1[0,T]$ subset of the Banach space $C^1[0,T]$. Define an operator $\mathcal{P} : [0,1] \times \overline{\Omega} \to C^1[0,T]$ by the formula

$$\begin{aligned} \mathcal{P}(\lambda,x)(t) &= x(0) + \alpha(x) \\ &\quad + \int_0^t \phi^{-1}\left(\phi\big(x'(0) + \beta(x)\big) + \lambda \int_0^s h\big(v, x(v), x'(v)\big)\,\mathrm{d}v\right)\mathrm{d}s. \end{aligned} \tag{3.11}$$

A standard argument shows that $\mathcal{P}$ is a continuous operator. We claim that $\mathcal{P}([0,1] \times \overline{\Omega})$ is compact in $C^1[0,T]$. Indeed, since $\overline{\Omega}$ is bounded in $C^1[0,T]$, we have

$$\big|\alpha(x)\big| \leqslant r, \quad |\beta(x)| \leqslant r, \quad \big|h\big(t, x(t), x'(t)\big)\big| \leqslant \varrho(t)$$

for a.e. $t \in [0,T]$ and $x \in \overline{\Omega}$, where r is a positive constant and $\varrho \in L_1[0,T]$. Set $K = \max\{S_1, \Lambda_1\} + r$ and $V = \max\{|\phi(-K)|, |\phi(K)|\}$. Then

$$\begin{aligned} \left|\mathcal{P}(\lambda,x)(t)\right| &\leqslant \max\{S_0, \Lambda_0 + \Lambda_1 T\} + r \\ &\quad + T\max\left\{\left|\phi^{-1}\left(-V - \|\varrho\|_1\right)\right|, \left|\phi^{-1}\left(V + \|\varrho\|_1\right)\right|\right\}, \\ \left|\mathcal{P}(\lambda,x)'(t)\right| &\leqslant \max\left\{\left|\phi^{-1}\left(-V - \|\varrho\|_1\right)\right|, \left|\phi^{-1}(V + \|\varrho\|_1)\right|\right\} \end{aligned}$$

and

$$\left|\phi\left(\mathcal{P}(\lambda,x)'(t_2)\right) - \phi\left(\mathcal{P}(\lambda,x)'(t_1)\right)\right| \leqslant \left|\int_{t_1}^{t_2} \varrho(t)\,\mathrm{d}t\right|$$

for $t, t_1, t_2 \in [0,T]$ and $(\lambda, x) \in [0,1] \times \overline{\Omega}$. Hence $\mathcal{P}([0,1] \times \overline{\Omega})$ is bounded in $C^1[0,T]$ and $\{\phi[\mathcal{P}(\lambda,x)']\}$ is equicontinuous on $[0,T]$. The mapping ϕ^{-1} being an increasing homeomorphism from $\mathbb{R}$ onto $\mathbb{R}$, we deduce from

$$\left|\mathcal{P}(\lambda,x)'(t_2) - \mathcal{P}(\lambda,x)'(t_1)\right| = \left|\phi^{-1}\left(\phi\left(\mathcal{P}(\lambda,x)'(t_2)\right)\right) - \phi^{-1}\left(\phi\left(\mathcal{P}(\lambda,x)'(t_1)\right)\right)\right|$$

that $\{\mathcal{P}(\lambda,x)'\}$ is also equicontinuous on $[0,T]$. Now the Arzelà–Ascoli theorem shows that $\mathcal{P}([0,1] \times \overline{\Omega})$ is compact in $C^1[0,T]$. Thus $\mathcal{P}$ is a compact operator.

Suppose that x_0 is a fixed point of the operator $\mathcal{P}(1,\cdot)$. Then

$$\begin{aligned} x_0(t) &= x_0(0) + \alpha(x_0) \\ &\quad + \int_0^t \phi^{-1}\left(\phi\left(x_0'(0) + \beta(x_0)\right) + \int_0^s h\left(v, x_0(v), x_0'(v)\right)\mathrm{d}v\right)\mathrm{d}s. \end{aligned}$$

Hence $\alpha(x_0) = 0$, $\beta(x_0) = 0$ and x_0 is a solution of the differential equation (3.7). Therefore x_0 is a solution of problem (3.7), (3.5) and to prove our theorem, it suffices to show that

$$\deg\left(\mathcal{I} - \mathcal{P}(1,\cdot), \Omega\right) \neq 0 \tag{3.12}$$

where "deg" stands for the Leray–Schauder degree and $\mathcal{I}$ is the identity operator on $C^1[0,T]$. To see this let a compact operator $\mathcal{K}\colon [0,1] \times \overline{\Omega} \to C^1[0,T]$ be given by

$$\mathcal{K}(\mu,x)(t) = x(0) + \alpha(x) - \mu\alpha(-x) + \left[x'(0) + \beta(x) - \mu\beta(-x)\right]t.$$

Then $\mathcal{K}(1,\cdot)$ is odd (i.e. $\mathcal{K}(1,-x) = -\mathcal{K}(1,x)$ for $x \in \overline{\Omega}$) and

$$\mathcal{K}(0,\cdot) = \mathcal{P}(0,\cdot). \tag{3.13}$$

If $\mathcal{K}(\mu_0, x_0) = x_0$ for some $\mu_0 \in [0,1]$ and $x_0 \in \overline{\Omega}$, then

$$\begin{aligned} x_0(t) &= x_0(0) + \alpha(x_0) - \mu_0\alpha(-x_0) + \left[x_0'(0) + \beta(x_0) - \mu_0\beta(-x_0)\right]t, \\ &\quad t \in [0,T]. \end{aligned}$$

Thus $x_0(t) = A_0 + B_0 t$ where $A_0 = x_0(0) + \alpha(x_0) - \mu_0\alpha(-x_0)$ and $B_0 = x_0'(0) + \beta(x_0) - \mu_0\beta(-x_0)$, so $\alpha(x_0) - \mu_0\alpha(-x_0) = 0$, $\beta(x_0) - \mu_0\beta(-x_0) = 0$. Hence

$$\alpha(A_0 + B_0 t) - \mu_0\alpha(-A_0 - B_0 t) = 0,$$
$$\beta(A_0 + B_0 t) - \mu_0\beta(-A_0 - B_0 t) = 0.$$

Therefore $|A_0| < \Lambda_0$, $|B_0| < \Lambda_1$ and $\|x_0\|_\infty < \Lambda_0 + \Lambda_1 T$, $\|x_0'\|_\infty < \Lambda_1$, which gives $x_0 \notin \partial\Omega$. Now, by the Borsuk antipodal theorem and a homotopy property,

$$\deg\bigl(\mathcal{I} - \mathcal{K}(0,\cdot), \Omega\bigr) = \deg\bigl(\mathcal{I} - \mathcal{K}(1,\cdot), \Omega\bigr) \neq 0. \tag{3.14}$$

Finally, assume that $\mathcal{P}(\lambda_*, x_*) = x_*$ for some $\lambda_* \in [0, 1]$ and $x_* \in \overline{\Omega}$. Then x_* is a solution of problem (3.8) with $\lambda = \lambda_*$ and, by our assumptions, $\|x_*\|_\infty < S_0$ and $\|x_*'\|_\infty < S_1$. Hence $x_* \notin \partial\Omega$ and the homotopy property yields

$$\deg\bigl(\mathcal{I} - \mathcal{P}(0,\cdot), \Omega\bigr) = \deg\bigl(\mathcal{I} - \mathcal{P}(1,\cdot), \Omega\bigr).$$

This, with (3.13) and (3.14), implies (3.12). Therefore, problem (3.7), (3.5) has a solution. ☐

REMARK 3.3. If functionals $\alpha, \beta \in \mathcal{A}$ are linear, then system (3.10) has the form

$$A\alpha(1) + B\alpha(t) = 0,$$
$$A\beta(1) + B\beta(t) = 0.$$

All of its solutions (A, B) are bounded if and only if $\alpha(1)\beta(t) - \alpha(t)\beta(1) \neq 0$ (and then $(A, B) = (0, 0)$). This is satisfied for example for the Dirichlet conditions but not for the periodic ones.

3.2. *General existence principle for singular BVPs with ϕ-Laplacian*

Let us consider the singular problem (3.4), (3.5). By regularization and sequential techniques, we construct an approximating sequence of regular problems (3.6), (3.5) for whose solvability Theorem 3.2 can be used. Existence results for the singular problem (3.4), (3.5) can be proved by the following two general existence principles. The first principle uses the Vitali convergence theorem, the other is based on a combination of the Lebesgue dominated convergence theorem and the Fatou theorem.

THEOREM 3.4 (General existence principle for singular problems I). *Assume* (3.3). *Let there exist a bounded set* $\Omega \subset C^1[0, T]$ *such that*

(i) *for each* $n \in \mathbb{N}$, *the regular problem* (3.6), (3.5) *has a solution* $u_n \in \Omega$,

(ii) *the sequence* $\{f_n(t, u_n(t), u_n'(t))\}$ *is uniformly integrable on* $[0, T]$.

Then

(a) *there exist* $u \in \overline{\Omega}$ *and a subsequence* $\{u_{k_n}\}$ *of* $\{u_n\}$ *such that* $\lim_{n\to\infty} u_{k_n} = u$ *in* $C^1[0,T]$,

(b) *u is a solution of problem* (3.4), (3.5) *if*

$$\lim_{n\to\infty} f_{k_n}\big(t, u_{k_n}(t), u'_{k_n}(t)\big) = f\big(t, u(t), u'(t)\big)$$

for a.e. $t \in [0,T]$.

PROOF. Since Ω is bounded in $C^1[0,T]$ and $\{u_n\} \subset \Omega$, we have

$$\|u_n\|_\infty \leqslant r, \quad \|u'_n\|_\infty \leqslant r, \quad n \in \mathbb{N}, \tag{3.15}$$

where r is a positive constant. Now (ii) guarantees that for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\big|\phi\big(u'_n(t_2)\big) - \phi\big(u'_n(t_1)\big)\big| \leqslant \left|\int_{t_1}^{t_2} \big|f_n\big(t, u_n(t), u'_n(t)\big)\big|\,\mathrm{d}t\right| < \varepsilon$$

for each t_1, $t_2 \in [0,T]$, $|t_1 - t_2| < \delta$ and $n \in \mathbb{N}$. Therefore $\{\phi(u'_n)\}$ is equicontinuous on $[0,T]$, and by virtue of (3.15) and the fact that ϕ is continuous and increasing on $\mathbb{R}$, $\{u'_n\}$ is equicontinuous on $[0,T]$ as well. The Arzelà–Ascoli theorem guarantees the existence of a subsequence $\{u_{k_n}\}$ of $\{u_n\}$ converging in $C^1[0,T]$ to some $u \in \overline{\Omega}$.

Suppose that $\lim_{n\to\infty} f_{k_n}(t, u_{k_n}(t), u'_{k_n}(t)) = f(t, u(t), u'(t))$ for a.e. $t \in [0,T]$. By (ii), $\{f_{k_n}(t, u_{k_n}(t), u'_{k_n}(t))\}$ is uniformly integrable on $[0,T]$. Therefore, by Vitali's convergence theorem, $f(t, u(t), u'(t)) \in L_1[0,T]$ and letting $n \to \infty$ in

$$\phi\big(u'_{k_n}(t)\big) = \phi\big(u'_{k_n}(0)\big) + \int_0^t f_{k_n}\big(s, u_{k_n}(s), u'_{k_n}(s)\big)\,\mathrm{d}s, \quad t \in [0,T],\ n \in \mathbb{N},$$

we arrive at

$$\phi\big(u'(t)\big) = \phi\big(u'(0)\big) + \int_0^t f\big(s, u(s), u'(s)\big)\,\mathrm{d}s, \quad t \in [0,T].$$

Consequently, $\phi(u') \in AC[0,T]$ and u is a solution of (3.4). In addition, since

$$\lim_{n\to\infty} u_{k_n} = u \quad \text{in } C^1[0,T]$$

and α and β are continuous in $C^1[0,T]$, it follows that $\alpha(u) = 0$, $\beta(u) = 0$. Hence u is a solution of problem (3.4), (3.5). □

REMARK 3.5. Let f in (3.4) have singularities only at the value 0 of its space variables and let f_n in (3.6) satisfy $f_n(t,x,y) = f(t,x,y)$ for a.e. $t \in [0,T]$ and all $(x,y) \in \mathcal{D}$, $n \in \mathbb{N}$, $|x| \geqslant \frac{1}{n}$ and $|y| \geqslant \frac{1}{n}$. Then the condition

$$\lim_{n\to\infty} f_{k_n}\big(t, u_{k_n}(t), u'_{k_n}(t)\big) = f\big(t, u(t), u'(t)\big)$$

for a.e. $t \in [0, T]$ is satisfied if u and u' have a finite number of zeros.

THEOREM 3.6 (General existence principle for singular problems II). *Assume* (3.3). *Let f have singularities only at the value* 0 *of its space variables. Let f_n in equation* (3.6) *satisfy*

$$\begin{cases} \text{for a.e. } t \in [0, T] \text{ and each } x, y \in \mathbb{R} \setminus \{0\}, \\ 0 \leqslant f_n(t, x, y) \leqslant p(|x|, |y|) \quad \text{where } p \in C((0, \infty) \times (0, \infty)). \end{cases} \tag{3.16}$$

Suppose that for each $n \in \mathbb{N}$, the regular problem (3.6), (3.5) *has a solution u_n and there exists a subsequence $\{u_{k_n}\}$ of $\{u_n\}$ converging in $C^1[0, T]$ to some u. Then u is a solution of the singular problem* (3.4), (3.5) *if u and u' have a finite number of zeros and*

$$\lim_{n \to \infty} f_{k_n}\big(t, u_{k_n}(t), u'_{k_n}(t)\big) = f\big(t, u(t), u'(t)\big) \quad \text{for a.e. } t \in [0, T]. \tag{3.17}$$

PROOF. Assume that (3.17) is true and $0 \leqslant \xi_1 < \xi_2 < \cdots < \xi_m \leqslant T$ are all the zeros of u and u'. Since $\|u_{k_n}\|_\infty \leqslant L$ and $\|u'_{k_n}\|_\infty \leqslant L$ for each $n \in \mathbb{N}$ where L is a positive constant, it follows from (3.16), (3.17),

$$\phi\big(u'_{k_n}(T)\big) - \phi\big(u'_{k_n}(0)\big) = \int_0^T f_{k_n}\big(t, u_{k_n}(t), u'_{k_n}(t)\big)\,\mathrm{d}t, \quad n \in \mathbb{N},$$

and the Fatou theorem that

$$\int_0^T f\big(t, u(t), u'(t)\big)\,\mathrm{d}t \leqslant \phi(L) - \phi(-L).$$

Hence $f(t, u(t), u'(t)) \in L_1[0, T]$. Set $\xi_0 = 0$ and $\xi_{m+1} = T$. We claim that for all $j \in \{0, 1, \dots, m\}$, $\xi_j < \xi_{j+1}$, the equality

$$\phi\big(u'(t)\big) = \phi\left(u'\left(\frac{\xi_j + \xi_{j+1}}{2}\right)\right) + \int_{\frac{\xi_j+\xi_{j+1}}{2}}^t f\big(s, u(s), u'(s)\big)\,\mathrm{d}s \tag{3.18}$$

is satisfied for $t \in [\xi_j, \xi_{j+1}]$. Indeed, let $j \in \{0, 1, \dots, m\}$ and $\xi_j < \xi_{j+1}$. Let us look at the interval $[\xi_j + \delta, \xi_{j+1} - \delta]$ where $\delta \in (0, \frac{\xi_j+\xi_{j+1}}{2})$. We know that $|u| > 0$ and $|u'| > 0$ on (ξ_j, ξ_{j+1}) and therefore $|u(t)| \geqslant \varepsilon$, $|u'(t)| \geqslant \varepsilon$ for $t \in [\xi_j + \delta, \xi_{j+1} - \delta]$ with a positive constant ε. Hence there exists $n_0 \in \mathbb{N}$ such that $|u_{k_n}(t)| \geqslant \frac{\varepsilon}{2}$, $|u'_{k_n}(t)| \geqslant \frac{\varepsilon}{2}$ for $t \in [\xi_j + \delta, \xi_{j+1} - \delta]$ and $n \geqslant n_0$. This yields (see (3.16))

$$0 \leqslant f_{k_n}\big(t, u_{k_n}(t), u'_{k_n}(t)\big) \leqslant \max\left\{p(u, v)\colon u, v \in \left[\frac{\varepsilon}{2}, L\right]\right\}$$

for a.e. $t \in [\xi_j + \delta, \xi_{j+1} - \delta]$ and all $n \geqslant n_0$. Letting $n \to \infty$ in

$$\phi\big(u'_{k_n}(t)\big) = \phi\left(u'_{k_n}\left(\frac{\xi_j + \xi_{j+1}}{2}\right)\right) + \int_{\frac{\xi_j+\xi_{j+1}}{2}}^t f_{k_n}\big(s, u_{k_n}(s), u'_{k_n}(s)\big)\,\mathrm{d}s$$

gives (3.18) for $t \in [\xi_j + \delta, \xi_{j+1} - \delta]$ by the Lebesgue dominated convergence theorem. Since $\delta \in (0, \frac{\xi_j + \xi_{j+1}}{2})$ is arbitrary, (3.18) is true on the interval (ξ_j, ξ_{j+1}) and using the fact that $f(t, u(t), u'(t)) \in L_1[0, T]$, we conclude that (3.18) holds also at $t = \xi_j$ and ξ_{j+1}. From (3.18) it follows that $\phi(u') \in AC[0, T]$ and Eq. (3.4) is satisfied for a.e. $t \in [0, T]$. Finally, $\alpha(u_{k_n}) = 0$ and $\beta(u_{k_n}) = 0$ and the continuity of α and β yields $\alpha(u) = 0$ and $\beta(u) = 0$. Hence u is a solution of problem (3.4), (3.5). □

3.3. *Nonlocal singular BVPs*

We consider differential equations of the type

$$\big(\phi(u')\big)' = f(t, u, u') \tag{3.19}$$

where ϕ is an increasing and odd homeomorphism, $\phi(\mathbb{R}) = \mathbb{R}$, f satisfies the Carathéodory conditions on a subset of $[0, T] \times \mathbb{R}^2$ and f may be singular in its space variables.

We also discuss nonlinear nonlocal boundary conditions

$$u(0) = u(T), \quad \max\big\{u(t)\colon 0 \leqslant t \leqslant T\big\} = c, \ c \in \mathbb{R}, \tag{3.20}$$

$$u(0) = u(T) = -\gamma \min\big\{u(t)\colon 0 \leqslant t \leqslant T\big\}, \quad \gamma \in (0, \infty), \tag{3.21}$$

$$\min\big\{u(t)\colon 0 \leqslant t \leqslant T\big\} = 0, \quad \delta(u') = 0, \quad \delta \in \mathcal{B}, \tag{3.22}$$

where $\mathcal{B}$ denotes the set of functionals $\delta : C[0, T] \to \mathbb{R}$ which are

(a) continuous, $\delta(0) = 0$, and

(b) increasing, that is $x, y \in C[0, T]$, $x < y$ on $(0, T) \Rightarrow \delta(x) < \delta(y)$.

EXAMPLE. Let $n \in \mathbb{N}$ and $0 \leqslant a < b \leqslant T$. Then the functionals $\delta_1(x) = x(\frac{T}{2}) + \max\{x(t)\colon 0 \leqslant t \leqslant T\}$, $\delta_2(x) = \int_a^b x^{2n+1}(t)\,\mathrm{d}t$, $\delta_3(x) = \int_0^T \mathrm{e}^{x(t)}\,\mathrm{d}t - T$ belong to the set $\mathcal{B}$. The functionals $\delta_4(x) = x(0)$ and $\delta_5(x) = x(T)$ satisfy condition (a) of $\mathcal{B}$ but do not satisfy condition (b). Hence $\delta_4, \delta_5 \notin \mathcal{B}$.

The boundary conditions (3.20)–(3.22) are special cases of (3.5) where

$$\alpha(x) = x(0) - x(T), \qquad \beta(x) = \max\big\{x(t)\colon 0 \leqslant t \leqslant T\big\} - c \qquad \text{for (3.20)},$$

$$\alpha(x) = x(0) - x(T), \qquad \beta(x) = x(0) + \gamma \min\big\{x(t)\colon 0 \leqslant t \leqslant T\big\} \quad \text{for (3.21)},$$

and

$$\alpha(x) = \min\big\{x(t)\colon 0 \leqslant t \leqslant T\big\}, \qquad \beta(x) = \delta(x') \quad \text{for (3.22)}.$$

The next theorems give sufficient conditions for solvability of the three nonlocal singular problems given above. Their proofs are based on applying general existence principles presented in Theorems 3.2, 3.4 and 3.6. Notice that if $f < 0$ in Eq. (3.19) then the singular points corresponding to the solutions of problem (3.19), (3.20) are of type II and, if $f > 0$, the solutions of problems (3.19), (3.21) and (3.19), (3.22) have singular points of type II.

THEOREM 3.7 [137, Theorem 2.1]. *Suppose* $f \in Car([0,T] \times \mathbb{R} \times (\mathbb{R} \setminus \{0\})$,

$$\begin{cases} -q(x)\big(\omega_1(|y|) + \omega_2(|y|)\big) \leqslant f(t,x,y) \leqslant -a \\ \textit{for a.e. } t \in [0,T] \textit{ and each } (x,y) \in \mathbb{R} \times (\mathbb{R} \setminus \{0\}), \\ \textit{where } a > 0,\ q \in C(\mathbb{R}) \textit{ is positive}, \omega_1 \in C[0,\infty) \textit{ is nonnegative}, \\ \omega_2 \in C(0,\infty) \textit{ is positive and nonincreasing and} \\ \int_0^1 \omega_2\big(\phi^{-1}(s)\big)\,\mathrm{d}s < \infty \end{cases} \tag{3.23}$$

and

$$\int_{d-1}^{d} \frac{\mathrm{d}s}{H^{-1}(\int_s^d q(v)\,\mathrm{d}v)} < \infty \quad \textit{for any } d \in \mathbb{R},$$

where

$$H(x) = \int_0^{\phi(x)} \frac{\phi^{-1}(s)\,\mathrm{d}s}{\omega_1(1+\phi^{-1}(s)) + \omega_2(\phi^{-1}(s))} \quad \textit{for } x \in [0,\infty).$$

Let

$$\mathcal{S} = \left\{c \in \mathbb{R}\colon \lim_{x \to -\infty} \int_x^c \frac{\mathrm{d}s}{H^{-1}(\int_s^c q(v)\mathrm{d}v)} > \frac{T}{2}\right\}.$$

Then problem (3.19), (3.20) *has a solution for each* $c \in \mathcal{S}$.

SKETCH OF PROOF. *Step* 1. *Regularization.*

A sequence of auxiliary regular differential equations $(\phi(u'))' = f_n(t,u,u')$ is constructed where $f_n \in Car([0,T] \times \mathbb{R}^2)$ and

$$f_n(t,x,y) = f(t,x,y) \quad \text{for a.e. } t \in [0,T] \text{ and each } x \in \mathbb{R},\ |y| \geqslant \frac{1}{n},\ n \in \mathbb{N}.$$

Step 2. *Existence of solutions of regular problems* (3.6), (3.20).

Let $c \in \mathcal{S}$. By (3.23), the existence of positive constants S_0 and S_1 (independent of n and λ) is proved such that $\|u\|_\infty < S_0$ and $\|u'\|_\infty < S_1$ for any $\lambda \in [0,1]$, $n \in \mathbb{N}$ and each solution u of the differential equation

$$\big(\phi(u')\big)' = \lambda f_n(t,u,u') \tag{3.24}$$

satisfying the conditions (3.20). Put $\alpha(x) = x(0) - x(T)$ and $\beta(x) = \max\{x(t)\colon 0 \leqslant t \leqslant T\} - c$ for $x \in C^1[0,T]$. Then system (3.10) has a unique solution $(A,B) = (c\frac{1-\mu}{1+\mu}, 0)$

for each $\mu \in [0, 1]$ and therefore all solutions of this system are bounded in $\mathbb{R}^2$. Hence Theorem 3.2 guarantees that for each $n \in \mathbb{N}$, problem (3.6), (3.20) has a solution u_n and $\|u_n\|_\infty < S_0$, $\|u_n'\|_\infty < S_1$.

Step 3. *Properties of solutions of regular problems* (3.6), (3.20).

The sequence $\{u_n\}$ is considered. It is proved that

$$\begin{aligned} &u_n'(t) \geqslant \phi^{-1}\big(a(\xi_n - t)\big) && \text{for } t \in [0, \xi_n], \\ &\big|u_n'(t)\big| \geqslant \phi^{-1}\big(a(t - \xi_n)\big) && \text{for } t \in [\xi_n, T], \end{aligned}$$

where ξ_n is the unique zero of u_n' and $a > 0$ appears in (3.23). Next, it is shown that the sequence $\{f_n(t, u_n(t), u_n'(t))\}$ is uniformly integrable on $[0, T]$. Hence $\{u_n\}$ is bounded in $C^1[0, T]$ and $\{u_n'\}$ is equicontinuous on $[0, T]$ and, by the Arzelà–Ascoli theorem and the compactness principle, we can assume without loss of generality that $\{u_n\}$ converges in $C^1[0, T]$ and $\{\xi_n\}$ converges in $\mathbb{R}$. Let

$$\lim_{n\to\infty} u_n = u \quad \text{and} \quad \lim_{n\to\infty} \xi_n = \xi.$$

Then $u \in C^1[0, T]$ satisfies (3.20),

$$\begin{aligned} &u'(t) \geqslant \phi^{-1}\big(a(\xi - t)\big) && \text{for } t \in [0, \xi], \\ &\big|u'(t)\big| \geqslant \phi^{-1}\big(a(t - \xi)\big) && \text{for } t \in [\xi, T] \end{aligned}$$

and

$$\lim_{n\to\infty} f_n\big(t, u_n(t), u_n'(t)\big) = f\big(t, u(t), u'(t)\big) \quad \text{for a.e. } t \in [0, T].$$

Theorem 3.4 now guarantees that u is a solution of problem (3.19), (3.20). □

EXAMPLE. Let $p > 2$, $\alpha \in [0, p - 2)$ and $\beta \in (0, p - 1)$. Then for any $c \in \mathbb{R}$ there exists a solution of the differential equation

$$\big(|u'|^{p-2}u'\big)' + \big(2 + \sin(tu) + |u|\big)\left(|u'|^\alpha + \frac{1}{|u'|^\beta}\right) = 0$$

satisfying boundary conditions (3.20).

THEOREM 3.8 [138, Theorem 4.1]. *Let* $f \in Car([0,T] \times (\mathbb{R} \setminus \{0\})^2)$. *Let*

$$\begin{cases} a \leqslant f(t,x,y) \leqslant \big(h_1(|x|) + h_2(|x|)\big)\big(\omega_1(\phi(|y|)) + \omega_2(\phi(|y|))\big) \\ \text{for a.e. } t \in [0,T] \text{ and all } (x,y) \in (\mathbb{R} \setminus \{0\})^2, \\ \text{where } a > 0,\ h_1, \omega_1 \in C[0,\infty) \text{ are nonnegative and nondecreasing}, \\ h_2, \omega_2 \in C(0,\infty) \text{ are positive and nonincreasing}, \\ \displaystyle\int_0^1 h_2(s)\,\mathrm{d}s < \infty, \quad \int_0^1 \omega_2(s)\,\mathrm{d}s < \infty, \quad \int_0^\infty \frac{\mathrm{d}s}{\omega_2(s)} = \infty, \end{cases} \tag{3.25}$$

and let

$$\liminf_{x\to\infty} \int_0^x \frac{\mathrm{d}s}{K^{-1}(\frac{T}{2}(h_1(x) + h_2(s)))} > \frac{T}{2}, \tag{3.26}$$

where

$$K(x) = \int_0^{\phi(x)} \frac{\mathrm{d}s}{\omega_1(\phi(1) + s) + \omega_2(s)}, \quad x \in [0,\infty). \tag{3.27}$$

Then there exists a solution of problem (3.19), (3.21) *for each* $\gamma > 0$.

SKETCH OF PROOF. *Step* 1. *Regularization.*

A sequence of approximating differential equations

$$\big(\phi(u')\big)' = f_n(t,u,u')$$

is introduced where $f_n \in Car([0,T] \times \mathbb{R}^2)$ and

$$f_n(t,x,y) = f(t,x,y) \quad \text{for a.e. } t \in [0,T] \text{ and each } |x| \geqslant \frac{1}{n},\ |y| \geqslant \frac{1}{n},\ n \in \mathbb{N}.$$

Step 2. *Existence of solutions of regular problems* (3.6), (3.21).

Let $\gamma > 0$ in (3.21). Using (3.25) and (3.26), the existence of a positive constant P (depending on γ) is proved such that $\|u\|_\infty < PT$ and $\|u'\|_\infty < P$ for each solution u of problem (3.24), (3.21) with $\lambda \in [0,1]$ and $n \in \mathbb{N}$. Put $\alpha(x) = x(0) - x(T)$ and

$$\beta(x) = x(0) + \gamma \min\big\{x(t)\colon 0 \leqslant t \leqslant T\big\} \quad \text{for } x \in C^1[0,T].$$

The system (3.10) has a unique solution $(A,B) = (0,0)$ for each $\mu \in [0,1]$. Hence, by Theorem 3.2 for each $n \in \mathbb{N}$, there exists a solution u_n of problem (3.6), (3.21) and $\|u_n\|_\infty < PT$, $\|u_n'\|_\infty < P$.

Step 3. *Properties of solutions of regular problems* (3.6), (3.21).

The sequence $\{u_n\}$ is considered. From (3.25) it follows that u_n' is increasing on $[0, T]$ and has a unique zero $\xi_n \in (0, T)$ and u_n vanishes exactly at two points t_{1n}, t_{2n}, $0 < t_{1n} < \xi_n < t_{2n} < T$, $u_n > 0$ on $[0, t_{1n}) \cup (t_{2n}, T]$ and $u_n < 0$ on (t_{1n}, t_{2n}). Further, u_n satisfies the inequality

$$|u_n(t)| \geqslant \begin{cases} \dfrac{S|t - t_{1n}|}{\xi_n - t_{1n}} & \text{for } t \in [0, \xi_n], \\[2ex] \dfrac{S|t - t_{2n}|}{t_{2n} - \xi_n} & \text{for } t \in [\xi_n, T], \end{cases}$$

where S is a positive constant and the sequence $\{f_n(t, u_n(t), u_n'(t))\}$ is uniformly integrable on $[0, T]$, which implies that $\{u_n'\}$ is equicontinuous on $[0, T]$. Moreover, there exists a positive constant Δ such that

$$t_{1n} \geqslant \gamma\Delta, \quad \xi_n - t_{1n} > \Delta, \quad t_{2n} - \xi_n > \Delta, \quad T - t_{2n} > \gamma\Delta \quad \text{for } n \in \mathbb{N}.$$

Hence, by the Arzelà–Ascoli theorem, there exists a subsequence $\{u_{k_n}\}$ which converges in $C^1[0, T]$ and let $u = \lim_{n\to\infty} u_{k_n}$. Then u vanishes exactly at two points in $[0, T]$, u' has a unique zero and $\lim_{n\to\infty} f_{k_n}(t, u_{k_n}(t), u_{k_n}'(t)) = f(t, u(t), u'(t))$ for a.e. $t \in [0, T]$. Now Theorem 3.4 guarantees that u is a solution of problem (3.19), (3.21). $\square$

REMARK 3.9. If $\lim_{x\to\infty} h_2(x) < A$ for some $A > 0$ and

$$\liminf_{x\to\infty} \frac{x}{K^{-1}(\frac{T}{2}(h_1(x) + A))} > \frac{T}{2},$$

then condition (3.26) is satisfied.

EXAMPLE. Let $q_j \in L_\infty[0, T]$ be nonnegative $(1 \leqslant j \leqslant 6)$, $q_1(t) \geqslant a > 0$ for a.e. $t \in [0, T]$, $p > 1$, $\beta_1, \beta_2, \beta_3 \in (0, p - 1)$, $\alpha_1 \in (0, p - 1 + \beta_2)$, $\alpha_2, \alpha_3 \in (0, 1)$. Then for each $\gamma > 0$, there exists a solution of the differential equation

$$\left(|u'|^{p-2}u'\right)' = q_1(t) + q_2(t)|u|^{\alpha_1} + \frac{q_3(t)}{|u|^{\alpha_2}} + \frac{q_4(t)}{|u|^{\alpha_3}|u'|^{\beta_1}} + q_5(t)|u'|^{\beta_2} + \frac{q_6(t)}{|u'|^{\beta_3}}$$

satisfying boundary conditions (3.21).

THEOREM 3.10. *Suppose* $f \in Car([0,T] \times (0,\infty) \times (\mathbb{R} \setminus \{0\}))$ *and the following conditions are satisfied*:

$$\begin{cases} \varphi(t) \leqslant f(t,x,y) \leqslant \big(h_1(x) + h_2(x)\big)\big[\omega_1\big(\phi(|y|)\big) + \omega_2\big(\phi(|y|)\big)\big] \\ \text{for a.e. } t \in [0,T] \text{ and each } (x,y) \in (0,\infty) \times (\mathbb{R} \setminus \{0\}), \\ \text{where } \varphi \in L_\infty[0,T] \text{ is positive}, \\ h_1, \omega_1 \in C[0,\infty) \text{ are positive and nondecreasing}, \\ h_2, \omega_2 \in C(0,\infty) \text{ are positive and nonincreasing}, \\ \displaystyle\int_0^1 h_2(s)\,\mathrm{d}s < \infty \end{cases} \tag{3.28}$$

and

$$\liminf_{x\to\infty} \frac{V(x)}{H(Tx)} > 1 \tag{3.29}$$

where

$$V(x) = \int_0^{\phi(x)} \frac{\phi^{-1}(s)\,\mathrm{d}s}{\omega_1(s+1) + \omega_2(s)},$$

$$H(x) = \int_0^x \big(h_1(s+1) + h_2(s)\big)\,\mathrm{d}s \quad \text{for } x \in [0,\infty).$$

Then for each $\delta \in \mathcal{B}$, *problem* (3.19), (3.22) *has a solution.*

SKETCH OF PROOF. *Step* 1. *Regularization.*

A sequence of auxiliary regular differential equations $(\phi(u'))' = f_n(t,u,u')$ is constructed with $f_n \in Car([0,T] \times \mathbb{R}^2)$ satisfying

$$f_n(t,x,y) = f(t,x,y) \quad \text{for a.e. } t \in [0,T] \text{ and each } x \geqslant \frac{1}{n},\ |y| \geqslant \frac{1}{n},\ n \in \mathbb{N}.$$

Step 2. *Existence of solutions of regular problems* (3.6), (3.22).

Fix $\delta \in \mathcal{B}$. From (3.28), (3.29) and from the properties of δ we obtain the existence of positive constants M_0 and M_1 such that $\|u\|_\infty < M_0$, $\|u'\|_\infty < M_1$ for each solution u of problem (3.24), (3.22) with $\lambda \in [0,1]$ and $n \in \mathbb{N}$. Set $\alpha(x) = \min\{x(t)\colon\ 0 \leqslant t \leqslant T\}$ and $\beta(x) = \delta(x')$ for $x \in C^1[0,T]$. Then system (3.10) has a unique solution $(A,B) = (0,0)$ for each $\mu \in [0,1]$. Therefore, by Theorem 3.2, for each $n \in \mathbb{N}$ there exists a solution u_n of problem (3.6), (3.22) and $\|u_n\|_\infty < M_0$, $\|u'\|_\infty < M_1$.

Step 3. *Properties of solutions of regular problems* (3.6), (3.22).

By Step 2, $\{u_n\}$ is bounded in $C^1[0,T]$ and from (3.28) it follows that $\{u_n'\}$ is equicontinuous on $[0,T]$. The assumption (3.28) and the properties of δ show that u_n has a unique zero ξ_n, $\xi_n \in (0,T)$, u_n' is increasing on $[0,T]$, $u_n'(\xi_n) = 0$ and

$$\left|u_n'(t)\right| \geqslant \left|\int_{\xi_n}^t \varphi(s)\,\mathrm{d}s\right|, \qquad u_n(t) \geqslant \int_{\xi_n}^t (t-s)\varphi(s)\,\mathrm{d}s \tag{3.30}$$

for $t \in [0,T]$. According to the Arzelà–Ascoli theorem, there exists a subsequence $\{u_{k_n}\}$ converging in $C^1[0,T]$ to some u and from (3.30) we see that u and u' vanish at a unique point. Since

$$\lim_{n\to\infty} f_{k_n}\bigl(t, u_{k_n}(t), u_{k_n}'(t)\bigr) = f\bigl(t, u(t), u'(t)\bigr) \quad \text{for a.e. } t \in [0,T],$$

Theorem 3.6 gives that u is a solution of problem (3.19), (3.22). □

REMARK 3.11. Problem (3.19), (3.22) was investigated in [140]. The conditions for the solvability of this problem are stronger there than those in Theorem 3.10. This is due to the fact that [140] uses the Vitali convergence theorem in limit processes whereas Theorem 3.10 is proved by Theorem 3.6.

EXAMPLE. Let $\varphi \in L_\infty[0,T]$ be positive, $p > 1$, $c_j > 0$ $(1 \leqslant j \leqslant 4)$, $\beta \in (0,1)$, $\alpha, \gamma, \delta, \lambda \in (0,\infty)$ and $\alpha + \gamma < p - 1$. Then for each $\delta \in \mathcal{B}$, the differential equation

$$\bigl(|u'|^{p-2}u'\bigr)' = \varphi(t)\left(1 + c_1 u^\alpha + \frac{c_2}{u^\beta}\right)\left(1 + c_3|u'|^\gamma + \frac{c_4}{|u'|^\lambda}\right)$$

has a solution u satisfying boundary conditions (3.22).

3.4. *Historical and bibliographical notes*

The general existence principles presented in Theorems 3.2 and 3.4 are special cases of the principles stated by Agarwal, O'Regan and Staněk in [16] for a class of second-order functional differential equations. Some general existence principles for second-order regular differential equations with the ϕ-Laplacian and Dirichlet or mixed boundary data have been established using the nonlinear alternative of Leray–Schauder type by O'Regan [109].

Second-order differential equations with the p-Laplacian and the ϕ-Laplacian occur in the study of the p-Laplace equations [91], general diffusion theory [22,40], non-Newtonian fluid theory [81] and the turbulent flow of a polytropic gas in a porous medium [60,36].

In recent years problems for $p(t)$-Laplacian equations have been studied (e.g., [63,64]). The $p(t)$-Laplacian is defined by $u \to (|u'|^{p(t)-2}u')'$ where $p \in C[0,T]$ and $p > 1$ on $[0,T]$. The $p(t)$-Laplacian is a generalization of the p-Laplacian.

4. Singular Dirichlet BVPs with ϕ-Laplacian

Motivated by various significant applications to non-Newtonian fluid theory, diffusion of flows in porous media, nonlinear elasticity and theory of capillary surfaces (see [22,60, 112], and Section 3.4), several authors have proposed the study of equations $(\phi_p(u'))' + f(t,u,u') = 0$ with the p-Laplacian $(\phi_p(u'))'$, where $p \in (1,\infty)$ and $\phi_p(y) = |y|^{p-2}y$ for $y \in \mathbb{R}$. Usually the p-Laplacian is replaced by its abstract and more general version, which leads to clearer exposition and better understanding of the methods that are employed to derive existence results. Therefore, similarly to Section 3, we will work with a ϕ-Laplacian which satisfies (3.3), i.e. ϕ is an increasing homeomorphism with $\phi(\mathbb{R}) = \mathbb{R}$.

We will consider a singular Dirichlet problem of the form

$$\big(\phi(u')\big)' + f(t,u,u') = 0, \quad u(0) = u(T) = 0 \tag{4.1}$$

and its special cases, in particular, a problem of the form

$$u'' + f(t,u,u') = 0, \quad u(0) = u(T) = 0, \tag{4.2}$$

where $\phi(y) = y$ on $\mathbb{R}$.

We will investigate problems (4.1) and (4.2) on the set $[0,T] \times \mathcal{A}$. In general, the function f depends on a time variable $t \in [0,T]$ and on two space variables x and y, where $(x,y) \in \mathcal{A}$ and $\mathcal{A}$ is a closed subset of $\mathbb{R}^2$ or $\mathcal{A} = \mathbb{R}^2$.

We assume that problems (4.1) and (4.2) are singular, which means, by Section 1, that f does not satisfy the Carathéodory conditions on $[0,T] \times \mathcal{A}$. In what follows, the types of singularities of f will be exactly specified for each problem under consideration.

In accordance with Section 1 we define:

DEFINITION 4.1. A function $u : [0,T] \to \mathbb{R}$ with $\phi(u') \in AC[0,T]$ is *a solution of problem* (4.1) if u satisfies $(\phi(u'(t)))' + f(t,u(t),u'(t)) = 0$ a.e. on $[0,T]$ and fulfills the boundary conditions $u(0) = u(T) = 0$.

A function $u \in C[0,T]$ is *a w-solution of problem* (4.1) if there exists a finite number of singular points $t_\nu \in [0,T]$, $\nu = 1,\dots,r$, such that if we denote $J = [0,T] \setminus \{t_\nu\}_{\nu=1}^r$, then $\phi(u') \in AC_{loc}(J)$, u satisfies $(\phi(u'(t)))' + f(t,u(t),u'(t)) = 0$ a.e. on $[0,T]$ and fulfills the boundary conditions $u(0) = u(T) = 0$.

Note that the condition $\phi(u') \in AC[0,T]$ implies $u \in C^1[0,T]$ and the condition $\phi(u') \in AC_{loc}(J)$ implies $u \in C^1(J)$. We will mention some papers where f is supposed to be continuous on $(0,T) \times \mathbb{R}^2$ and can have only time singularities at $t = 0$ and $t = T$. Then any solution (any w-solution) u of (4.1) moreover satisfies $\phi(u') \in C^1(0,T)$. If we investigate the solvability of problem (4.1) or (4.2) on the set $[0,T] \times \mathcal{A}$ and $\mathcal{A} \neq \mathbb{R}^2$, we impose on its solution u in addition the condition

$$\big(u(t), u'(t)\big) \in \mathcal{A} \quad \text{for } t \in [0,T]. \tag{4.3}$$

If u is a w-solution, then one requires it to satisfy (4.3) for $t \in J$ only.

In some cases (see, e.g., (4.9)) f does not depend on y. Then we work with a set $\mathcal{A}$ which is a closed subset of $\mathbb{R}$ or $\mathcal{A} = \mathbb{R}$ and condition (4.3) has the form $u(t) \in \mathcal{A}$ for $t \in [0, T]$.

REMARK 4.2. We will carry out the investigation of the singular problem (4.1) in the spirit of the existence principles presented in Sections 1 and 3:

- the singular problem is approximated by a sequence of solvable regular problems;
- a sequence $\{u_n\}$ of approximate solutions is generated;
- a convergence of a suitable subsequence $\{u_{k_n}\}$ is investigated;
- the type of this convergence determines the properties of its limit u and, among others, determines whether u is a w-solution or a solution of the original singular problem.

There are more possibilities how to construct an approximating sequence of regular problems. Their choice depends on the type of singularities of the nonlinearity f in (4.1) (time, space), on the type of singular points corresponding to a solution (w-solution) of (4.1) (type I, type II), on the type of results desired (existence of a solution, a positive solution, a w-solution, uniqueness), and so on. A common idea is that approximate functions f_n have no singularities, $f_n \neq f$ on neighbourhoods U_n of singular points of f, $f_n = f$ elsewhere, and $\lim_{n\to\infty} \operatorname{meas}(U_n) = 0$.

Having such a sequence of $\{f_n\}$ we study problems

$$\big(\phi(u')\big)' + f_n(t, u, u') = 0, \quad u(0) = u(T) = \varepsilon_n,\ n \in \mathbb{N}, \tag{4.4}$$

where $\varepsilon_n \in \mathbb{R}$, $\lim_{n\to\infty} \varepsilon_n = 0$. In some proofs, one simply puts $\varepsilon_n = 0$ for $n \in \mathbb{N}$.

Solvability of (4.4) can be investigated by means of various methods which have been developed for regular Dirichlet problems (fixed point theorems, topological degree arguments, the topological transversality method, variational methods, lower and upper functions, the Fredholm nonlinear alternative, etc.). See also Section 3. Using one of the above methods we generate a sequence of approximate solutions $\{u_n\}$ of (4.4). The crucial information which enables us to realize the limit process concerns a priori estimates of the approximate solutions u_n.

4.1. *Method of lower and upper functions*

It is well known that for regular second order boundary value problems *the lower and upper functions method* is a profitable instrument for proofs of their solvability and for a priori estimates of their solutions. See, e.g., [47–49,93,94,114,124,146]. Hence, it seems to be a good idea to extend this method to the singular problem (4.1). In literature there are several definitions of lower and upper functions for regular boundary value problems. (Note that in some papers they are called lower and upper solutions.) Here we will use the following definition which is the same both for regular problems with $f \in Car([0, T] \times \mathbb{R}^2)$ and for singular ones with $f \in Car((0, T) \times \mathbb{R}^2)$ having time singularities at $t = 0$ and $t = T$.

DEFINITION 4.3. A function $\sigma : [0, T] \to \mathbb{R}$ with $\phi(\sigma') \in AC[0, T]$ is called *a lower function of* (4.1) if σ satisfies

$$\big(\phi\big(\sigma'(t)\big)\big)' + f\big(t, \sigma(t), \sigma'(t)\big) \geqslant 0 \quad \text{for a.e. } t \in [0, T] \tag{4.5}$$

and

$$\sigma(0) \leqslant 0, \quad \sigma(T) \leqslant 0. \tag{4.6}$$

If the inequalities in (4.5) and (4.6) are reversed, then σ is called *an upper function of* (4.1).

For the special case (4.2) we admit a more general definition.

DEFINITION 4.4. A function $\sigma \in C[0, T]$ is called *a lower function of* (4.2) if there exists a finite set $\Sigma \subset (0, T)$ such that $\sigma \in AC^1_{loc}([0, T] \setminus \Sigma)$, $\sigma'(\tau+)$, $\sigma'(\tau-) \in \mathbb{R}$ for each $\tau \in \Sigma$,

$$\sigma''(t) + f\big(t, \sigma(t), \sigma'(t)\big) \geqslant 0 \quad \text{for a.e. } t \in [0, T], \tag{4.7}$$

$$\sigma(0) \leqslant 0, \quad \sigma(T) \leqslant 0, \quad \sigma'(\tau-) < \sigma'(\tau+) \quad \text{for each } \tau \in \Sigma. \tag{4.8}$$

If the inequalities in (4.7) and (4.8) are reversed, then σ is called *an upper function of* (4.2).

REMARK 4.5. (i) If, moreover, f is continuous on $(0, T) \times \mathbb{R}^2$, then a lower (upper) function σ of (4.1) is supposed to satisfy $\phi(\sigma') \in C^1(0, T)$ and a lower (upper) function of (4.2) belongs also to $C^2(0, T)$.

(ii) If the boundary conditions in (4.1) or in (4.2) are replaced by inhomogeneous ones, i.e. they have the form

$$u(0) = a, \quad u(T) = b$$

for some $a, b \in \mathbb{R}$, then the corresponding boundary inequalities in (4.6) or in (4.8) are modified to

$$\sigma(0) \leqslant a, \quad \sigma(T) \leqslant b.$$

We present straightforward extensions of the classical lower and upper functions method to a singular problem with the p-Laplacian

$$\big(\phi_p(u')\big)' + f(t, u) = 0, \quad u(0) = a, \quad u(T) = b, \tag{4.9}$$

where

$$\begin{cases} \phi_p(y) = |y|^{p-2}y, \ p > 1, \ a, b \in \mathbb{R}, \ f \in Car\big((0, T) \times \mathbb{R}\big), \\ f \text{ can have time singularities at } t = 0 \text{ and } t = T. \end{cases} \tag{4.10}$$

Recall that f has time singularities at $t = 0$ and $t = T$ if there exist $x, y \in \mathbb{R}$ such that

$$\int_0^{\varepsilon} \big|f(t,x,y)\big|\,\mathrm{d}t = \infty, \qquad \int_{T-\varepsilon}^{T} \big|f(t,x,y)\big|\,\mathrm{d}t = \infty$$

for each sufficiently small $\varepsilon > 0$.

Making use of ideas of the papers [100] by Lomtatidze and Torres and [80] by Habets and Zanolin, one can prove the following result for the special case of (4.9) with $p = 2$.

THEOREM 4.6. *Let $p = 2$ and* (4.10) *hold. Let σ_1 and σ_2 be a lower and an upper function for problem* (4.9) *and $\sigma_1 \leqslant \sigma_2$ on* $[0, T]$. *Assume also that there is a function $h \in L_1(I)$ on each compact interval $I \subset (0, T)$ such that*

$$\big|f(t,x)\big| \leqslant h(t) \quad \textit{for a.e. } t \in (0,T) \textit{ and each } x \in \big[\sigma_1(t), \sigma_2(t)\big],$$

and

$$\int_0^T t(T-t)h(t)\,\mathrm{d}t < \infty.$$

Then problem (4.9) *has a w-solution $u \in C[0,T] \cap AC^1_{loc}(0,T)$ such that*

$$\sigma_1(t) \leqslant u(t) \leqslant \sigma_2(t) \quad \textit{for } t \in [0,T]. \tag{4.11}$$

If for a.e. $t \in (0,T)$ the function $f(t,x)$ is nonincreasing in x, then the w-solution is unique. If $h \in L_1[0,T]$, then u belongs to $AC^1[0,T]$, i.e. u is a solution of (4.9).

Theorem 4.6 can be proved by means of the Schauder fixed point theorem which is applied to the operator $\mathcal{T} : C[0,T] \to C[0,T]$, where

$$(\mathcal{T}u)(t) = a + \frac{t}{T}(b-a) + \int_0^T G(t,s) f^*\big(s, u(s)\big)\,\mathrm{d}s.$$

Here G is the Green function of the problem $-u'' = 0$, $u(0) = u(T) = 0$ and f^* is given by

$$f^*(t,x) = \begin{cases} f\big(t, \sigma_1(t)\big) & \text{if } x < \sigma_1(t), \\ f(t,x) & \text{if } \sigma_1(t) \leqslant x \leqslant \sigma_2(t), \\ f\big(t, \sigma_2(t)\big) & \text{if } x > \sigma_2(t) \end{cases}$$

for a.e. $t \in [0,T]$ and each $x \in \mathbb{R}$.

REMARK 4.7. By virtue of (4.11) we can investigate problem (4.9) on $[0,T] \times \mathcal{A}_t$, where $\mathcal{A}_t = [\sigma_1(t), \sigma_2(t)]$ for $t \in [0,T]$. Therefore, in Theorem 4.6, instead of $f \in Car((0,T) \times \mathbb{R})$ it is sufficient to assume $f \in Car((0,T) \times A_t)$.

Jiang in [86] dealt with problem (4.9) under the assumption (4.10) and, in addition,

$$f \in C\big((0,T)\times\mathbb{R}\big). \tag{4.12}$$

He modified Theorem 4.6 for $p \neq 2$.

THEOREM 4.8. *Let* (4.10) *and* (4.12) *hold. Let* σ_1 *and* σ_2 *be a lower and an upper function for problem* (4.9) *and* $\sigma_1 \leqslant \sigma_2$ *on* $[0,T]$. *Assume also that there is a function* $h \in C(0,T)$ *such that*

$$\big|f(t,x)\big| \leqslant h(t) \quad \textit{for } t \in (0,T),\ x \in \big[\sigma_1(t),\sigma_2(t)\big],$$

and that there exist $\mu, \nu \in [0, p-1)$ *such that*

$$\int_0^T t^{\mu}(T-t)^{\nu}h(t)\,\mathrm{d}t < \infty. \tag{4.13}$$

Then problem (4.9) *has a w-solution* $u \in C[0,T]$ *satisfying* $\phi_p(u') \in C^1(0,T)$ *and* (4.11).

In contrast to $p = 2$, there is no Green function for $p \neq 2$, which makes the proof of Theorem 4.8 more difficult and complicated than that for $p = 2$.

REMARK 4.9. Motivated by physical and technical problems, there is a lot of papers studying problems *with both time and space singularities*. If such a problem has a singularity at $x = 0$, one often searches for solutions (w-solutions) which are positive on $(0,T)$. Although they vanish at 0 and T, they are still called *positive solutions* (*positive w-solutions*) in literature. In this case, problems (4.1) and (4.2) are investigated on the set $[0,T]\times\mathcal{A}$, where $\mathcal{A} = [0,\infty)\times\mathbb{R}$, and f in (4.1) or (4.2) is supposed to satisfy $f \in Car((0,T)\times\mathcal{D})$, where $\mathcal{D} = (0,\infty)\times\mathbb{R}$ (or more specifically f is supposed to be continuous on $(0,T)\times\mathcal{D}$). In this case lower and upper functions have to be positive on $(0,T)$ and consequently a lower function σ_1 has to satisfy $\sigma_1(0) = \sigma_1(T) = 0$.

Having in mind Remarks 4.7 and 4.9 we will search for positive w-solutions of a singular problem

$$u'' + f(t,u) = 0, \quad u(0) = u(T) = 0, \tag{4.14}$$

where

$$\begin{cases} f \in Car\big((0,T)\times(0,\infty)\big) \text{ can have} \\ \text{time singularities at } t = 0,\ t = T \text{ and a space singularity at } x = 0. \end{cases} \tag{4.15}$$

Recall that f has a space singularity at $x = 0$ if

$$\limsup_{x\to 0+}\big|f(t,x)\big| = \infty \quad \text{for a.e. } t \in [0,T].$$

Let us present a simple application of Theorem 4.6 in the spirit of Habets and Zanolin [80].

THEOREM 4.10. *Let a function f be positive and satisfy* (4.15). *Assume that for a.e. $t \in (0,T)$ the function $f(t,x)$ is nonincreasing in x. Suppose that there exists a lower function σ_1 of problem* (4.14) *such that $\sigma_1 > 0$ on $(0,T)$ and*

$$\int_0^T f\big(s,\sigma_1(s)\big)\,\mathrm{d}s < \infty.$$

Then problem (4.14) *has a unique positive solution $u \in AC^1[0,T]$ such that $\sigma_1 \leqslant u$ on $[0,T]$.*

Theorem 4.10 follows from Theorem 4.6 and Remark 4.7 if we put $h(t) = f(t,\sigma_1(t))$ and

$$\sigma_2(t) = \int_0^T G(t,s) f(s,k)\,\mathrm{d}s + k,$$

where $k = \max\{\sigma_1(t)\colon t \in [0,T]\}$ and G is the Green function of the problem

$$-u'' = 0, \quad u(0) = u(T) = 0.$$

The next result can be viewed as a corollary of [100, Theorem 1.1], where Lomtatidze and Torres studied an equation including an additional term $g(t,u)u'$.

THEOREM 4.11. *Let* (4.15) *hold. Let σ_1 and σ_2 be a lower and an upper function of problem* (4.14) *and*

$$\sigma_2(0) > 0, \quad \sigma_2(T) > 0, \quad 0 < \sigma_1 \leqslant \sigma_2 \quad \text{on } (0,T).$$

Let, moreover, for every $0 < \eta < \min\{\sigma_2(t)\colon t \in [0,T]\}$ there exist $h_\eta \in C(0,T)$ such that

$$\big|f(t,x)\big| \leqslant h_\eta(t) \quad \text{for } t \in (0,T) \text{ and all } x \in \big[\sigma_{1\eta}(t), \sigma_2(t)\big],$$

where $\sigma_{1\eta}(t) = \max\{\eta, \sigma_1(t)\}$ and

$$\int_0^T t(T-t) h_\eta(t)\,\mathrm{d}t < \infty.$$

Then problem (4.14) *has a positive w-solution $u \in C[0,T] \cap AC^1_{loc}(0,T)$ such that*

$$\sigma_1(t) \leqslant u(t) \leqslant \sigma_2(t) \quad \text{for } t \in [0,T].$$

SKETCH OF THE PROOF. *Step* 1. *Construction of auxiliary intervals.*

A decreasing sequence $\{a_n\} \subset (0, T)$ and an increasing sequence $\{b_n\} \subset (0, T)$ are constructed such that, among other, $\lim_{n\to\infty} a_n = 0$, $\lim_{n\to\infty} b_n = T$.

Step 2. *Construction of auxiliary regular problems.*

For $t \in [0, T]$, $x \in \mathbb{R}$, $n \in \mathbb{N}$, functions χ_n, α_n, β, f_n are given by

$$\chi_n(x) = \begin{cases} \sigma_1(a_n) & \text{if } x < \sigma_1(a_n), \\ x & \text{if } x \geqslant \sigma_1(a_n), \end{cases}$$

$$\alpha_n(t) = \begin{cases} \sigma_1(a_n) & \text{if } 0 \leqslant t \leqslant a_n, \\ \sigma_1(t) & \text{if } a_n \leqslant t \leqslant b_n, \\ \sigma_1(b_n) & \text{if } b_n \leqslant t \leqslant T, \end{cases} \qquad \beta(t) = \begin{cases} v_1(t) & \text{if } 0 \leqslant t \leqslant a_1, \\ \sigma_2(t) & \text{if } a_1 \leqslant t \leqslant b_1, \\ v_2(t) & \text{if } b_1 \leqslant t \leqslant T, \end{cases}$$

$$f_n(t, x) = \begin{cases} \frac{1}{2}\big[f\big(t, \chi_n(x)\big) + \big|f\big(t, \chi_n(x)\big)\big|\big] & \text{if } t \in (0, a_n] \cup [b_n, T), \\ f\big(t, \chi_n(x)\big) & \text{if } t \in (a_n, b_n), \end{cases}$$

where v_1 and v_2 are solutions of some auxiliary linear Dirichlet problems.

Step 3. *Convergence of the sequence of approximating solutions.*

Solvability of a sequence of regular problems

$$u'' + f_n(t, u) = 0, \quad u(0) = \sigma_1(a_n), \quad u(T) = \sigma_1(b_n), \quad n \in \mathbb{N}, \tag{4.16}$$

is investigated. The functions α_1 and β are a lower and an upper function of (4.16) with $n = 1$, and hence, by Theorem 4.6, there is a w-solution u_1 of (4.16) with $n = 1$ such that $\alpha_1 \leqslant u_1 \leqslant \beta$ on $[0, T]$. Further, α_2 and u_1 are a lower and an upper function of (4.16) with $n = 2$, and so Theorem 4.6 guarantees the existence of a w-solution u_2 of (4.16) with $n = 2$ such that $\alpha_2 \leqslant u_2 \leqslant u_1$ on $[0, T]$. In this way a sequence of w-solutions is obtained and then a limit process is applied. □

At the end of this subsection we will show another existence assertion in terms of the lower and upper functions for a problem with the p-Laplacian of the form

$$\big(\phi_p(u')\big)' + \psi(t)g(t, u) = 0, \quad u(0) = u(1) = 0, \tag{4.17}$$

where $\phi_p(y) = |y|^{p-2}y$, $p > 1$. Here we assume that

$$\begin{cases} \psi : (0, 1) \to (0, \infty) \text{ is continuous} \\ \text{and can have time singularities at } t = 0 \text{ and } t = 1, \end{cases} \tag{4.18}$$

$$\begin{cases} g : [0, 1] \times (0, \infty) \to \mathbb{R} \text{ is continuous} \\ \text{and can have a space singularity at } x = 0. \end{cases} \tag{4.19}$$

In this setting, using the paper by Agarwal, Lü and O'Regan [2], we offer the following result about the existence of positive w-solutions of (4.17).

THEOREM 4.12. *Let* (4.18) *and* (4.19) *hold. Assume that the following conditions are satisfied*:

$$\begin{cases} \{\rho_n\} \textit{ is a nonincreasing sequence of real numbers}, \\ \lim\limits_{n\to\infty} \rho_n = 0 \textit{ and } n_0 \in \mathbb{N},\ n_0 \geqslant 3 \textit{ being fixed}, \end{cases} \tag{4.20}$$

$$\max\left\{\int_0^{1/2} \phi_p^{-1}\left(\int_s^{1/2} \psi(t)\,\mathrm{d}t\right)\mathrm{d}s, \int_{1/2}^1 \phi_p^{-1}\left(\int_{1/2}^s \psi(t)\,\mathrm{d}t\right)\mathrm{d}s\right\} \\ = b_0 < \infty \tag{4.21}$$

and

$$\psi(t)g(t,\rho_n) \geqslant 0 \quad \textit{for } t \in \left[\frac{1}{2^{n+1}}, 1\right), \quad n \geqslant n_0.$$

Further assume that σ_1 *and* σ_2 *are a lower and an upper function of problem* (4.17) *with* $\sigma_1 > 0$ *on* $(0,1)$, $\max\{\rho_{n_0}, \sigma_1(t)\} \leqslant \sigma_2(t)$ *for* $t \in [0,1]$ *and*

$$\bigl(\phi_p\bigl(\sigma_2'(t)\bigr)\bigr)' + \psi(t)g\left(\frac{1}{2^{n_0+1}}, \sigma_2(t)\right) \leqslant 0 \quad \textit{for } t \in \left(0, \frac{1}{2^{n_0+1}}\right).$$

Then problem (4.17) *has a positive w-solution* $u \in C[0,1]$ *with* $\phi_p(u') \in C^1(0,1)$ *and*

$$\sigma_1(t) \leqslant u(t) \leqslant \sigma_2(t) \quad \textit{for } t \in [0,1].$$

In Section 4.3 we will show how lower and upper functions for regular problems can be applied to get not only a w-solution but also a solution of a given singular problem (see Theorem 4.18).

4.2. *Positive nonlinearities*

Many papers studying problem (4.1) or (4.2) with a space singularity at $x = 0$ concern the case that the nonlinearity f is positive. Such problems are referred to as *positone* ones in literature, see [10,11,135]. The positivity of f implies that each solution is concave and hence positive on $(0,T)$, and if, moreover, f has a space singularity at $x = 0$ but not at y, then each solution has only two corresponding singular points 0, T which are of type I. This makes the study of such problems easier than of those having sign-changing f or space singularities at y.

First we will discuss mixed singularities at t and x. In Section 1.1 we have presented problem (1.3), (1.4) the solvability of which was investigated by Taliaferro [141]. This problem has mixed singularities: the time ones at $t = 0$ and $t = 1$ as well as the space one at $x = 0$. Among many papers generalizing Taliaferro's existence results we choose the paper by Tineo [142] devoted to the existence of positive solutions or w-solutions to a singular problem

$$u'' + f(t, u, u') = 0, \quad u(0) = u(1) = 0, \tag{4.22}$$

where

$$\begin{cases} f : (0, 1) \times (0, \infty) \times \mathbb{R} \to (0, \infty) \text{ is continuous and can have} \\ \text{time singularities at } t = 0,\ t = 1 \text{ and a space singularity at } x = 0. \end{cases} \tag{4.23}$$

THEOREM 4.13 [142, Theorem 0.1]. *Let* (4.23) *hold. Suppose that there are continuous functions* $\varphi : (0, 1) \to (0, \infty)$, $\psi : (0, 1) \to [0, \infty)$ *and* $g : (0, \infty) \to (0, \infty)$ *such that* g *is decreasing and*

$$f(t, x, y) \leqslant \varphi(t)g(x) + \psi(t)|y| \quad \textit{for } t \in (0, 1),\ x \in (0, \infty),\ y \in \mathbb{R},$$

$$\int_0^1 t(1-t)\varphi(t)\,\mathrm{d}t < \infty, \qquad \int_0^1 \psi(t)\,\mathrm{d}t < \infty.$$

Assume further that for each constant $M > 0$ *there exists a continuous function* $\varepsilon_M : (0, 1) \to (0, \infty)$ *such that*

$$\varepsilon_M(t) \leqslant f(t, x, y) \quad \textit{for } t \in (0, 1),\ x \in (0, M],\ y \in \mathbb{R}.$$

Then problem (4.22) *has a positive w-solution* $u \in C[0, 1] \cap C^2(0, 1)$. *If, moreover,*

$$\int_0^1 g\big(kt(1-t)\big)\varphi(t)\,\mathrm{d}t < \infty \quad \textit{for all } k > 0,$$

then u *belongs to* $AC^1[0, 1]$, *which means that* u *is a solution of* (4.22).

The proof of Theorem 4.13 proceeds according to Remark 4.2. Solvability of auxiliary regular problems is obtained by the Leray–Schauder degree argument and the limit process is guaranteed by means of a priori estimates of the approximate solutions.

EXAMPLE. Let $\alpha, \beta \in (0, 2)$, $k, \lambda \in (0, \infty)$, $\varepsilon \in C(0, 1)$, $\varepsilon > 0$ on $(0, 1)$. By Theorem 4.13 the problem

$$u'' + \frac{1}{t^\alpha(1-t)^\beta u^\lambda} + t^k|u'| + \varepsilon(t) = 0, \quad u(0) = u(1) = 0$$

has a positive w-solution $u \in C[0, 1] \cap C^2(0, 1)$. If $\alpha + \lambda, \beta + \lambda \in (0, 1)$ then, moreover, $u \in C^1[0, 1]$. Hence u is a solution.

Let us turn back to problem (4.14) with f satisfying (4.15). Wang in [148] considered f which can have at most linear growth in x. He illustrated his result by functions

$$f(t,x)=\frac{1}{t^{\alpha}x^{\beta}}\quad \text{with } \alpha\in(1,2),\ \beta>0$$

or

$$f(t,x)=\delta x\exp\left(\frac{1}{x}\right)\quad \text{with a sufficiently small } \delta>0.$$

For f which is moreover continuous on $(0,T)\times(0,\infty)$, Agarwal and O'Regan [3,4,11] proved the existence of a positive w-solution of (4.14) with f increasing in x for large x. An example of such f is

$$f(t,x)=\delta\left(\frac{1}{x^{\alpha}}+x^{\beta}+1\right),\quad \alpha,\beta,\delta\in(0,\infty).$$

If f has sublinear growth in x (i.e. $\beta\in(0,1)$), then (4.14) has a positive w-solution for each $\delta>0$. If f has linear or superlinear growth in x (i.e. $\beta=1$ or $\beta\in(1,\infty)$), then (4.14) has a positive w-solution for any sufficiently small $\delta>0$. A formula for an upper bound of δ is also given.

Now let us consider space singularities at x and y. We will present conditions ensuring solvability of problems with singularities in space variables x and y and with singular points both of type I and of type II. The main difficulty in the study of singular points of type II is the fact that their location in $[0,T]$ is not known. This is the reason why in mathematical literature there are only few papers concerning solvability of such problems and no results about w-solutions are known.

The first existence result in this direction was reached by Staněk [135] in 2001. The following theorem can be viewed as a corollary of [135, Theorem 1].

For a fixed $A>0$ we consider a singular problem

$$u''+\mu f(t,u,u')=0,\quad u(0)=u(T)=0,\tag{4.24}$$

with a positive real parameter μ, where

$$\begin{cases} f \text{ is continuous on } \mathcal{D}_A=[0,T]\times(0,A)\times\left[-\frac{2A}{T},0\right)\cup\left(0,\frac{2A}{T}\right] \\ \text{and can have space singularities at } x=0, x=A, y=0. \end{cases}\tag{4.25}$$

Sufficient conditions on μ and f for the solvability of (4.24) in the set $[0,T]\times\mathcal{A}$, where $\mathcal{A}=[0,A]\times[-\frac{2A}{T},\frac{2A}{T}]$, are given in the next theorem.

THEOREM 4.14. *Let* (4.25) *hold. Suppose that there exists* $\delta>0$ *such that* f *satisfies*

$$\delta\leqslant f(t,x,y)\leqslant g(x)\omega\big(|y|\big)\quad \text{on } \mathcal{D}_A,$$

where $\omega \geqslant \delta$ *is continuous on* $(0, \frac{2A}{T}]$ *and* $g \in C(0, A) \cap L_1[0, A]$. *Let*

$$\mu_T = \left(\int_0^{2A/T} \frac{y}{\omega(y)}\,\mathrm{d}y\right)\left(\int_0^A g(x)\,\mathrm{d}x\right)^{-1}.$$

Then for any $\mu \in (0, \mu_T]$ *problem* (4.24) *has a solution* $u_\mu \in AC^1[0, T]$ *satisfying*

$$0 < u_\mu(t) < A \quad \textit{for } t \in (0, T). \tag{4.26}$$

Take notice of the fact that for any solution u of problem (4.24) there is a point $t_u \in (0, T)$ with $u'(t_u) = 0$. Since f has a singularity at $y = 0$ and we do not know the position of $t_u \in (0, T)$, t_u is a singular point of type II.

To prove Theorem 4.14 a two-parameter family of regular problems is constructed and their solvability is established by the topological transversality method. Then, a priori bounds for approximate solutions of regular problems are derived. Using these bounds and the Arzelà–Ascoli theorem, a solution of (4.24) is obtained by a limiting process.

EXAMPLE. Let $A > 0, a, b, c, d, \gamma \in [0, \infty)$, $a + b + d > 0$, $\alpha, \beta \in (0, 1)$. Then for a sufficiently small $\mu > 0$ the problem

$$\begin{aligned} &u'' + \mu\left(1 + \frac{a}{u^\alpha (A-u)^\beta} + bu^\gamma\right)\left(1 + cu'^2\right)\left(1 + \frac{d}{u'^2}\right) = 0, \\ &u(0) = u(T) = 0 \end{aligned}$$

has a solution u_μ satisfying (4.26). The upper bound μ_T is explicitly expressed in [135].

The next existence result is in the spirit of Staněk [139], where a more general state-dependent functional differential equation was studied. Here we consider problem (4.1) with the ϕ-Laplacian and a function f satisfying

$$\begin{cases} f \in Car\big([0, T] \times \mathcal{D}\big),\ \mathcal{D} = (0, \infty) \times \big(\mathbb{R} \setminus \{0\}\big), \\ \text{and } f \text{ can have space singularities at } x = 0,\ y = 0. \end{cases} \tag{4.27}$$

THEOREM 4.15. *Let* (4.27) *hold and let* ϕ *be odd. Suppose that there exists* $\delta \in (0, \infty)$ *such that* f *satisfies*

$$\delta \leqslant f(t, x, y) \leqslant \big(h_1(x) + h_2(x)\big)\big(\omega_1\big(\phi\big(|y|\big)\big) + \omega_2\big(\phi\big(|y|\big)\big)\big)$$

for a.e. $t \in [0, T]$ *and all* $(x, y) \in \mathcal{D}$, *where* $h_1, \omega_1 \in C[0, \infty)$ *are positive and nondecreasing*, $h_2, \omega_2 \in C(0, \infty) \cap L_1[0, 1]$ *are positive and nonincreasing and*

$$\int_0^\infty \frac{\mathrm{d}s}{\omega_1(s)} = \infty.$$

Let

$$\frac{T}{2} < \liminf_{u\to\infty} \frac{u}{K^{-1}(T(h_1(u)+h_2(u)))},$$

where K^{-1} denotes the inverse function to $K:[0,\infty)\to[0,\infty)$,

$$K(u) = \int_0^{\phi(u)} \frac{\mathrm{d}s}{\omega_1(s)+\omega_2(s)}.$$

Then problem (4.1) *has a positive solution $u \in AC^1[0,T]$.*

In the proof of this result the solvability of a sequence of regular problems is obtained by the Leray–Schauder degree theory. Limit processes are guaranteed by Vitali's convergence theorem.

EXAMPLE. Let $c \in L_\infty[0,T]$, $p \in (1,\infty)$, $\beta \in (0,1)$, $\gamma,\eta \in (0,p)$, $\delta \in (0,\infty)$ and $\alpha \in (0, p-\gamma)$. Further, let $c(t) \geqslant \delta$ a.e. on $[0,T]$. Then the problem

$$\begin{aligned}&\big(|u'|^{p-2}u'\big)' + c(t)\left(1+u^\alpha+\frac{1}{u^\beta}\right)\left(1+|u'|^\gamma+\frac{1}{|u'|^\eta}\right)=0,\\ &u(0)=u(T)=0\end{aligned}$$

has a positive solution.

4.3. *Sign-changing nonlinearities*

Results about the solvability of singular Dirichlet problems with sign-changing nonlinearities mostly concern w-solutions. Making use of the arguments of Section 1 we can show a new existence principle giving positive solutions to singular Dirichlet problems of the form

$$u'' + f(t,u,u') = 0, \quad u(0)=u(T)=0, \tag{4.28}$$

where

$$\begin{cases} f \in Car\big((0,T)\times\mathcal{D}\big) \text{ can change its sign, } \mathcal{D}=(0,\infty)\times\mathbb{R}, \\ \text{and } f \text{ can have mixed singularities at } t=0,\ t=T,\ x=0. \end{cases} \tag{4.29}$$

For $k\in\mathbb{N}$, $k\geqslant \frac{3}{T}$, $t\in[0,T]$, $x\in\mathbb{R}$, put $\Delta_k = [0,\frac{1}{k})\cup(T-\frac{1}{k},T]$,

$$\gamma_k(t)=\begin{cases}\frac{1}{k} & \text{if } t<\frac{1}{k},\\ t & \text{if } t\in[0,T]\setminus\Delta_k,\\ T-\frac{1}{k} & \text{if } t>T-\frac{1}{k},\end{cases} \qquad \delta_k(x)=\begin{cases}|x| & \text{if } |x|\geqslant\frac{1}{k},\\ \frac{1}{k} & \text{if } |x|<\frac{1}{k}.\end{cases}$$

Then we construct the sequence of regular functions

$$f_k(t,x,y) = f\big(\gamma_k(t), \delta_k(x), y\big) \quad \text{for a.e. } t \in [0,T], \text{ each } x, y \in \mathbb{R} \tag{4.30}$$

and the sequence of regular problems

$$u'' + f_k(t,u,u') = 0, \quad u(0) = u(T) = \frac{1}{k}, \tag{4.31}$$

where $f_k \in Car([0,T] \times \mathbb{R}^2)$, $k \in \mathbb{N}$, $k \geqslant \frac{3}{T}$.

THEOREM 4.16 (Existence principle for solutions of (4.28)). *Let* (4.29) *hold. Assume that*

$$\begin{cases} \textit{there exists a bounded set } \Omega \subset C^1[0,T] \textit{ such that} \\ \textit{problem } (4.31) \textit{ has a solution } u_k \in \Omega \textit{ for each } k \in \mathbb{N},\ k \geqslant \frac{3}{T}, \end{cases} \tag{4.32}$$

$$\begin{cases} \textit{there exists a function } \varepsilon \in C[0,T],\ \varepsilon(0) = \varepsilon(T) = 0, \\ \textit{such that } u_k(t) \geqslant \varepsilon(t) > 0 \textit{ for } t \in (0,T) \textit{ and each } k \in \mathbb{N},\ k \geqslant \frac{3}{T} \end{cases} \tag{4.33}$$

and

$$\textit{the sequence } \big\{f_k\big(t, u_k(t), u_k'(t)\big)\big\} \textit{ is uniformly integrable on } [0,T]. \tag{4.34}$$

Then

$$\begin{cases} \textit{there exist a function } u \in \overline{\Omega} \textit{ and a subsequence } \{u_{k_n}\} \subset \{u_k\} \\ \textit{such that } \lim_{n\to\infty} \|u_{k_n} - u\|_{C^1} = 0, \end{cases} \tag{4.35}$$

and

$$u \in AC^1[0,T] \textit{ is a positive solution of problem } (4.28). \tag{4.36}$$

PROOF. Since f can have singularities both at t and at x, we cannot obtain Theorem 4.16 as a direct consequence of some of the theorems in Section 1. Nevertheless, we can use the ideas of their proofs and argue as follows.

Step 1. *Convergence of the sequence of approximating solutions.*

By (4.32) we get that $\{u_k\}$ and $\{u_k'\}$ are bounded in $C[0,T]$. The boundedness of $\{u_k'\}$ implies the equicontinuity of $\{u_k\}$ on $[0,T]$. Condition (4.34) yields the equicontinuity of $\{u_k'\}$ on $[0,T]$ and hence the Arzelà–Ascoli theorem gives the assertion (4.35).

Step 2. *Convergence of the sequence of regular right-hand sides.*

The conditions $u_{k_n}(0) = u_{k_n}(T) = \frac{1}{k_n}$ imply $u(0) = u(T) = 0$. By (4.33) we get $u(t) > 0$ on $(0,T)$. Choose $\xi \in (0,T)$ such that $f(\xi, \cdot, \cdot) : \mathcal{D} \to \mathbb{R}$ is continuous. Then, by virtue of

(4.30) and (4.33), we have

$$u_{k_n}(\xi) \geqslant \varepsilon(\xi) > \frac{1}{k_n}, \quad \xi \in [0,T] \setminus \Delta_{k_n},$$
$$f_{k_n}\big(\xi, u_{k_n}(\xi), u'_{k_n}(\xi)\big) = f\big(\xi, u_{k_n}(\xi), u'_{k_n}(\xi)\big)$$

for a sufficiently large k_n. Therefore

$$\lim_{n\to\infty} f_{k_n}\big(t, u_{k_n}(t), u'_{k_n}(t)\big) = f\big(t, u(t), u'(t)\big) \quad \text{for a.e. } t \in [0,T]. \tag{4.37}$$

Step 3. *Properties of the limit u.*

By (4.34), (4.37) and Vitali's convergence theorem, we get that $f(t,u(t),u'(t))$ belongs to $L_1[0,T]$ and we can pass to the limit in the sequence

$$u'_{k_n}(t) = u'_{k_n}(0) - \int_0^t f_{k_n}\big(s, u_{k_n}(s), u'_{k_n}(s)\big)\,\mathrm{d}s, \quad t \in [0,T],$$

thus obtaining

$$u'(t) = u'(0) - \int_0^t f\big(s, u(s), u'(s)\big)\,\mathrm{d}s, \quad t \in [0,T].$$

Hence (4.36) is true. □

REMARK 4.17. Having in mind the absolute continuity of the Lebesgue integral we see that if there exists $\varphi \in L_1[0,T]$ such that

$$\big|f_k\big(t, u_k(t), u'_k(t)\big)\big| \leqslant \varphi(t) \quad \text{for a.e. } t \in [0,T] \text{ and each } k \in \mathbb{N},\ k \geqslant \tfrac{3}{T}, \tag{4.38}$$

then condition (4.34) is valid.

In Section 4.1, the classical lower and upper functions method has been extended to singular problems (see Theorems 4.6, 4.8, 4.11 and 4.12). Motivated by Agarwal, O'Regan, Lakshmikantham and Leela [13], we will show another approach which consists in the employment of a sequence of lower and upper functions of approximating regular problems.

THEOREM 4.18. *Let* (4.29) *and* (4.30) *hold. Assume that there exists $k_0 \in \mathbb{N}$ such that for each $k \in \mathbb{N}$, $k \geqslant k_0$, the following conditions are satisfied*:

$$\begin{cases} \alpha_k \textit{ and } \beta \textit{ are a lower and an upper function of } (4.31) \textit{ and} \\ \alpha'_k, \beta' \in L_\infty[0,T], \quad \frac{1}{k} \leqslant \alpha_k(t) \leqslant \beta(t) \quad \textit{for } t \in [0,T], \end{cases} \tag{4.39}$$

$$\begin{cases} \big|f_k(t,x,y)\big| \leqslant \psi(t) g(x) \omega(|y|) \quad \textit{for a.e. } t \in [0,T] \\ \textit{and for all } x \in \big(0, \|\beta\|_\infty\big),\ y \in \mathbb{R}, \end{cases} \tag{4.40}$$

with

$$\begin{cases} \text{positive functions } \psi \in L_1[\frac{1}{k}, T - \frac{1}{k}],\ \omega \in C[0,\infty) \text{ and} \\ \text{a positive nonincreasing function } g \in C(0,\infty). \end{cases} \tag{4.41}$$

Further, assume that there is a function $\alpha \in C[0,T]$ *with* $\alpha(0) = \alpha(T) = 0$, $\alpha > 0$ *on* $(0,T)$, $\alpha_k \geqslant \alpha$ *on* $[0,T]$, *such that*

$$\int_0^T \psi(t) g\bigl(\alpha(t)\bigr)\,\mathrm{d}t < \int_0^\infty \frac{\mathrm{d}s}{\omega(s)}. \tag{4.42}$$

Then problem (4.28) *has a positive solution* $u \in AC^1[0,T]$ *satisfying*

$$\alpha(t) \leqslant u(t) \leqslant \beta(t) \quad \textit{for } t \in [0,T].$$

SKETCH OF THE PROOF. Theorem 4.18 can be proved by means of Theorem 4.16 in the following way.

Step 1. *Construction of the sequence of regular problems.*

Condition (4.42) implies that there exists $\rho > 0$ such that

$$\int_0^T \psi(t) g\bigl(\alpha(t)\bigr)\,\mathrm{d}t < \int_0^\rho \frac{\mathrm{d}s}{\omega(s)}. \tag{4.43}$$

For $k \in \mathbb{N}$, $k \geqslant k_0$, put $\rho_k = \max\{\rho, \|\alpha_k'\|_\infty, \|\beta'\|_\infty\}$ and consider a sequence of regular problems

$$u'' + \widetilde{f}_k(t,u,u') = 0, \quad u(0) = u(T) = \frac{1}{k}, \tag{4.44}$$

where for a.e. $t \in [0,T]$ and all $(x,y) \in \mathbb{R}^2$ we set

$$\widetilde{f}_k(t,x,y) = \chi_k(y) f_k(t,x,y)$$

and

$$\chi_k(y) = \begin{cases} 1 & \text{if } |y| \leqslant \rho_k, \\ 2 - \frac{|y|}{\rho_k} & \text{if } \rho_k < |y| < 2\rho_k, \\ 0 & \text{if } |y| \geqslant 2\rho_k. \end{cases}$$

Then α_k and β are respectively a lower and an upper function of (4.44), $\widetilde{f}_k$ satisfies (4.40) and, moreover, there exists $\widetilde{h}_k \in L_1[0,T]$ such that

$$\bigl|\widetilde{f}_k(t,x,y)\bigr| \leqslant \widetilde{h}_k(t) \quad \text{for a.e. } t \in [0,T],\ \text{all } x \in \bigl[\alpha_k(t), \beta(t)\bigr],\ y \in \mathbb{R}.$$

The classical existence result based on the lower and upper functions method for regular problems (see, e.g., [146]) guarantees that for each $k \geqslant k_0$ problem (4.44) has a solution u_k with

$$\alpha \leqslant \alpha_k \leqslant u_k \leqslant \beta \quad \text{on } [0, T]. \tag{4.45}$$

Step 2. *A priori estimates of approximate solutions.*

Using (4.40), (4.41) and (4.43) we deduce that $\|u_k'\|_\infty \leqslant \rho$, which implies that u_k is also a solution of (4.31) for $k \geqslant k_0$. Hence, if we put $\varepsilon(t) = \alpha(t)$ for $t \in [0, T]$ and

$$\Omega = \bigl\{x \in C^1[0, T]\colon\ \alpha \leqslant x \leqslant \beta \text{ on } [0, T],\ \|x'\|_\infty \leqslant \rho\bigr\},$$

we get that (4.32), (4.33) are fulfilled.

Step 3. *Uniform integrability of regular right-hand sides.*

By (4.40), (4.41) and (4.45) we get

$$\bigl|f_k\bigl(t, u_k(t), u_k'(t)\bigr)\bigr| \leqslant \psi(t)g\bigl(u_k(t)\bigr)\omega\bigl(\bigl|u_k'(t)\bigr|\bigr) \leqslant M\psi(t)g\bigl(\alpha(t)\bigr) = \varphi(t)$$

for a.e. $t \in [0, T]$, where $M = \max\{\omega(|s|)\colon\ s \in [-\rho, \rho]\}$. By virtue of (4.43), $\varphi \in L_1[0, T]$ and we conclude by Remark 4.17 that condition (4.34) is valid. Therefore the assertion follows from Theorem 4.16 and condition (4.45). ☐

REMARK 4.19. If f does not depend on y, then (4.42) takes the form

$$\int_0^T \psi(t)g\bigl(\alpha(t)\bigr)\,\mathrm{d}t < \infty.$$

In the rest of this section we present a selection of existence results about w-solutions which can be obtained by theorems from Section 4.1. Consider a singular Dirichlet problem

$$u'' + f(t, u) = 0, \quad u(0) = u(T) = 0, \tag{4.46}$$

where

$$\begin{cases} f \in Car\bigl((0, T) \times (0, \infty)\bigr) \text{ can change its sign and} \\ \text{can have singularities at } t = 0,\ t = T \text{ and at } x = 0. \end{cases} \tag{4.47}$$

The first result is due to Lomtatidze [99].

THEOREM 4.20. *Let* (4.47) *hold. Assume that* f *is nonincreasing as a function of its second argument and that there is* $\varepsilon > 0$ *for which*

$$f(t, \varepsilon) \geqslant 0 \quad \textit{for a.e. } t \in [0, T], \operatorname{meas}\bigl\{t \in (0, T)\colon\ f(t, \varepsilon) > 0\bigr\} > 0.$$

Then the condition

$$\int_0^T t(T-t)\bigl|f(t,\delta)\bigr|\,\mathrm{d}t < \infty \quad \textit{for any } \delta \in (0,\varepsilon]$$

is necessary and sufficient for problem (4.46) *to have a unique positive w-solution* $u \in C[0,T] \cap AC^1_{loc}(0,T)$.

The proof of Theorem 4.20 is based on the lower and upper functions method via Theorem 4.11. More general results were obtained by Lomtatidze and Torres in [100], where a differential equation having moreover the term $g(t,u)u'$ was investigated.

A very similar result for a simpler case of problem (4.46) with f satisfying (4.49) was proved by Habets and Zanolin in [80]. The analogue of their results was proved by Jiang [86] for a singular Dirichlet problem with the p-Laplacian

$$\bigl(\phi_p(u')\bigr)' + f(t,u) = 0, \quad u(0) = u(T) = 0, \tag{4.48}$$

where $\phi(y) = |y|^{p-2}y,\ p > 1$ and

$$\begin{cases} f \in C\bigl((0,T) \times (0,\infty)\bigr) \text{ can change its sign and} \\ \text{can have singularities at } t = 0,\ t = T \text{ and at } x = 0. \end{cases} \tag{4.49}$$

THEOREM 4.21 [86, Theorem 3]. *Let* (4.49) *hold. Assume that*

(i) *there exists a constant* $L > 0$ *such that for any compact set* $K \subset (0,T)$ *there is* $\varepsilon = \varepsilon_K > 0$ *such that*

$$f(t,x) > L \quad \textit{for all } t \in K,\ x \in (0,\varepsilon],$$

(ii) *for any* $\delta > 0$ *there are* $h_\delta \in C(0,T)$ *and* $\mu, \nu \in [0, p-1)$ *such that*

$$\bigl|f(t,x)\bigr| \leqslant h_\delta(t) \quad \textit{for all } t \in (0,T),\ x \geqslant \delta$$

and

$$\int_0^T t^\mu (T-t)^\nu h_\delta(t)\,\mathrm{d}t < \infty.$$

Then problem (4.48) *has a positive w-solution* $u \in C[0,T]$ *with* $\phi_p(u') \in C^1(0,T)$. *If, moreover, for each* $t \in (0,T)$ *the function* $f(t,x)$ *is nonincreasing in* x, *then* u *is a unique w-solution.*

Another existence result for differential equation where the nonlinearity f can depend on u' is due to Jiang in [87]. He studied the singular Dirichlet problem of the form

$$u'' + f(t,u,u') = 0, \quad u(0) = u(1) = 0, \tag{4.50}$$

where

$$\begin{cases} f \in C((0,1)\times(0,\infty)\times\mathbb{R}) \text{ can change its sign} \\ \text{and can have singularities at } t=0,\ t=1 \text{ and } x=0. \end{cases} \tag{4.51}$$

Motivated by the example

$$f(t,x,y)=\frac{\delta}{t^m(1-t)^n}\left(\frac{1}{x^\alpha}+x^\beta+\sin(8\pi t)\right)\bigl(1+t(1-t)|y|^{1/\gamma}\bigr) \tag{4.52}$$

with real numbers $\alpha>0$, $\beta\geqslant 0$, $\gamma>1$, $\delta>0$, $0\leqslant m,n<2$, he proved existence of a positive w-solution of (4.50) for f satisfying (4.51) and having superlinear growth in x for large x and sublinear growth in y for large $|y|$. Particularly, if f in (4.50) has the form (4.52), then an upper bound for δ is found guaranteeing that (4.50) has a positive w-solution. The proof is based on the papers [80] and [4].

Let us turn back to problem (4.17). Using the lower and upper functions method established in Theorem 4.12 we can get sufficient conditions for the existence of positive w-solutions. Specifically, we report the result motivated by Agarwal, Lü and O'Regan [2].

THEOREM 4.22. *Let* (4.18)–(4.21) *hold. Assume that there exist* $n_0\in\mathbb{N}$, $n_0\geqslant 3$ *and* $c_0\in(0,\infty)$ *such that*

$$\psi(t)g(t,x)\geqslant c_0 \quad \textit{for } t\in\left[\frac{1}{2^{n+1}},1-\frac{1}{2^{n+1}}\right],\ x\in(0,\rho_n],\ n\geqslant n_0$$

and

$$\bigl|g(t,x)\bigr|\leqslant \eta(x)+h(x) \quad \textit{for } (t,x)\in[0,1]\times(0,\infty),$$

where $\eta\in AC_{loc}(0,\infty)$ *is positive*, $h\in C[0,\infty)$ *is nonnegative and* $\frac{h}{\eta}$ *is nondecreasing on* $(0,\infty)$. *Further assume that*

$$\eta'<0 \quad \textit{a.e. on } (0,R) \quad \textit{and} \quad \frac{\eta'}{\eta^2}\in L_1[0,R] \quad \textit{for any } R>0$$

and that there exists $r\in(0,\infty)$ *such that*

$$b_0<\left(\phi_p^{-1}\left(1+\frac{h(r)}{\eta(r)}\right)\right)^{-1}\int_0^r\frac{\mathrm{d}u}{\phi_p^{-1}(\eta(u))}.$$

Then problem (4.17) *has a positive w-solution* $u\in C[0,1]$ *with* $\phi_p(u')\in C^1(0,1)$.

4.4. *Sign-changing solutions and w-solutions*

If we consider a singular differential equation with a space singularity at $x = 0$, a question about the existence of sign-changing solutions of such an equation can arise. A single result in this direction was proved by Rachůnková and Staněk in [118], where an equation of the form $(r(u)u')' = \mu q(t) f(t, u)$ has been studied.

Here we will present this result for a simplified equation

$$u'' + \mu f(t, u) = 0 \tag{4.53}$$

with a positive real parameter μ and with boundary conditions

$$u(0) = u(T) = 0, \quad \max\{u(t)\colon\ t \in [0, T]\} \min\{u(t)\colon\ t \in [0, T]\} < 0. \tag{4.54}$$

We assume that

$$\begin{cases} f : [0, T] \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R} \text{ is continuous} \\ \text{and can have a space singularity at } x = 0. \end{cases} \tag{4.55}$$

By a *solution of problem* (4.53), (4.54) we mean a function $u \in C^1[0, T]$ having precisely one zero $t_u \in (0, T)$. Moreover, $u \in C^2((0, T) \setminus \{t_u\})$ fulfills (4.54) and there exists $\mu_u > 0$ such that u satisfies (4.53) for $\mu = \mu_u$ and $t \in (0, T) \setminus \{t_u\}$.

THEOREM 4.23. *Let* (4.55) *hold. Assume that for each* $t \in [0, T]$ *the function* $f(t, x)$ *is nondecreasing with respect to* x *on* $(-\infty, 0)$ *and nonincreasing on* $(0, \infty)$ *and*

$$\varepsilon \leqslant f(t, x) \operatorname{sign} x \leqslant g(x) \quad \textit{for } (t, x) \in [0, T] \times \big(\mathbb{R} \setminus \{0\}\big),$$

where $\varepsilon \in (0, \infty)$ *and* $g \in C(\mathbb{R} \setminus \{0\}) \cap L_1[-1, 1]$. *Then for each* $A \in (0, \infty)$ *and* $B \in (-\infty, 0)$ *there exist solutions* u *and* v *of problem* (4.53), (4.54) *satisfying*

$$\max\{u(t)\colon\ t \in [0, T]\} = A \quad \textit{and} \quad \min\{v(t)\colon\ t \in [0, T]\} = B. \tag{4.56}$$

By virtue of Theorem 4.23, any solution u of problem (4.53), (4.54) vanishes at some point $t_u \in (0, T)$. Since f has a singularity at $x = 0$ and we do not know the position of t_u, we see that t_u is a singular point of type II.

The proof of Theorem 4.23 is based on a combination of four main theorems in [118], where a new method of proofs was developed. It is based on “gluing” the positive and negative parts of solutions and smoothing them.

In accordance with the paper [119] by Rachůnková and Staněk we define *a w-solution of problem* (4.53), (4.54) as a function $u \in C[0, T]$ having precisely one zero $t_u \in (0, T)$. Further, $u \in C^2((0, T) \setminus \{t_u\})$ fulfills (4.54), there exist finite limits

$$\lim_{t \to t_u-} u'(t), \quad \lim_{t \to t_u+} u'(t)$$

and there exists $\mu_u > 0$ such that u satisfies (4.53) for $\mu = \mu_u$ and $t \in (0,T) \setminus \{t_u\}$.

In [119], the following existence result for w-solutions is proved.

THEOREM 4.24. *Let all assumptions of Theorem* 4.23 *be satisfied. Then for each* $t_0 \in (0,T)$ *and each* $A \in (0,\infty)$, $B \in (-\infty,0)$ *problem* (4.53), (4.54) *has just two different w-solutions vanishing at* t_0 *and having their maximum value on* $[0,T]$ *equal to* A, *and just two different w-solutions vanishing at* t_0 *and having their minimum value on* $[0,T]$ *equal to* B.

In the proof of Theorem 4.24, w-solutions are constructed by means of solutions of auxiliary Dirichlet problems on $[0,t_0]$ and $[t_0,T]$.

EXAMPLE. Let $\alpha, \beta \in (0,1)$, $a \in (0,\infty)$, $b \in (-\infty,0)$ and

$$f(x) = \begin{cases} \frac{a}{x^\alpha} & \text{for } x > 0, \\ \frac{b}{(-x)^\beta} & \text{for } x < 0. \end{cases}$$

Consider the differential equation

$$u'' + \mu f(u) = 0. \tag{4.57}$$

By Theorem 4.23, for each $A > 0$ and $B < 0$ there exist solutions u and v of problem (4.57), (4.54) satisfying (4.56). Moreover, by Theorem 4.24, for each $t_0 \in (0,T)$ and for each $A > 0$ there exist just two different w-solutions u_1 and u_2 of problem (4.57), (4.54) satisfying

$$\max\{u_1(t)\colon t \in [0,T]\} = \max\{u_2(t)\colon t \in [0,T]\} = A.$$

Further, for each $t_0 \in (0,T)$ and for each $B < 0$ there exist just two different w-solutions v_1 and v_2 of problem (4.57), (4.54) satisfying

$$\min\{v_1(t)\colon t \in [0,T]\} = \min\{v_2(t)\colon t \in [0,T]\} = B.$$

4.5. *Historical and bibliographical notes*

A systematic study of solvability of Dirichlet problems having both time and space singularities was initiated in 1979 by Taliaferro [141], who found necessary and sufficient conditions for the existence of solutions (w-solutions) of problem (1.3), (1.4). A contribution to the more general problem (4.22) was published by Bobisud, O'Regan and Royalty [37] in 1988. In 1989, in contrast to the shooting method used in [141] and the topological transversality method applied in [37], Gatica, Oliker and Waltman [72] proved a fixed point theorem for decreasing maps on cones and applying it they obtained solvability of (4.14). However, in these works the nonlinearity f had to be bounded in its space variables x and y for large x and large $|y|$.

An extension of the above results permitting linear growth of f in its third variable y for large $|y|$ was treated by Baxley [24] in 1991 and by Tineo [142] in 1992. The condition of boundedness of f in its second variable x for large x was overcome by Agarwal and O'Regan in [3] (1996), where the existence of a positive w-solution was proved even for f having superlinear growth for large x. In 1999, the first multiplicity result for Dirichlet problems with time and space singularities was reached. Particularly, Agarwal and O'Regan [6] proved the existence of two different positive w-solutions. All these results rely on the fact that nonlinearities in the equations considered are positive.

In 1987, this assumption was removed by Lomtatidze [99] for problem (4.14). We can also refer to papers by Janus and Myjak [83] for a nonhomogeneous equation (1.3) and by Habets and Zanolin [80] for the continuous case of (4.14) which appeared in 1994. From papers providing more general existence results for problems with sign-changing nonlinearities we mention the recent papers by Jiang [87] (2002), by Agarwal, Staněk [17] (2003) or by Lomtatidze, Torres [100] (2003). These papers deal with problems of the type (4.2).

Existence results for problems of the type (4.1) with the ϕ-Laplacian and sign-changing nonlinearities were presented by Wang and Gao in [149] (1996), where Taliaferro's results were extended. Existence results in the spirit of Habets and Zanolin which are applicable to problems with the p-Laplacian of the form (4.48) were given by Jiang in [86] (2001). In 2003, Agarwal, Lü and O'Regan [2] published the existence result for the problem with the p-Laplacian of the form (4.17).

Further results and references for positive and for sign-changing nonlinearities can be found in the monographs by Kiguradze [92] (1975), by Kiguradze and Shekhter [93] (1987), by O'Regan [110] (1994), by Agarwal and O'Regan [10] (2003) and in [11] (2004). Note that there exists a large group of papers investigating Dirichlet boundary value problems having only time singularities. These results are not discussed here but some of them can be found in the above cited monographs.

In the study of Dirichlet problems with space singularities and singular points both of type I and of type II the first existence result was reached by Staněk [135] in 2001, and the existence of sign-changing solutions was proved by Rachůnková and Staněk [118] in 2003. Numerical algorithms and computation of solutions and w-solutions of singular Dirichlet problems were given by Baxley [25] (1995) and by Baxley and Thompson [28] (2000).

5. Singular periodic BVPs with ϕ-Laplacian

The aim of this section is to present existence results for singular periodic problems of the form

$$\big(\phi(u')\big)' = f(t, u, u'), \tag{5.1}$$

$$u(0) = u(T), \quad u'(0) = u'(T), \tag{5.2}$$

where $0 < T < \infty$, $\phi : \mathbb{R} \to \mathbb{R}$ is an increasing and odd homeomorphism such that $\phi(\mathbb{R}) = \mathbb{R}$ and

$$\begin{cases} f \in Car\big([0,T] \times \big((0,\infty) \times \mathbb{R}\big)\big) \quad \text{and} \\ f \text{ can have a space singularity at } x = 0. \end{cases} \tag{5.3}$$

In accordance with Section 1.3, this means that

$$\limsup_{x \to 0+} \big| f(t,x,y) \big| = \infty \quad \text{for a.e. } t \in [0,T] \text{ and some } y \in \mathbb{R}$$

may happen.

REMARK 5.1. Physicists say that f has an *attractive* singularity at $x = 0$ if

$$\liminf_{x \to 0+} f(t,x,y) = -\infty \quad \text{for a.e. } t \in [0,T] \text{ and some } y \in \mathbb{R}$$

since near the origin the force is directed inward. Alternatively, f is said to have a *repulsive* singularity at $x = 0$ if

$$\limsup_{x \to 0+} f(t,x,y) = \infty \quad \text{for a.e. } t \in [0,T] \text{ and some } y \in \mathbb{R}$$

holds.

In the setting of Section 1.3, problem (5.1), (5.2) is investigated on the set $[0,T] \times \mathcal{A}$, where $\mathcal{A} = [0,\infty) \times \mathbb{R}$. In contrast to the Dirichlet problem (4.1), where each solution vanishes at $t = 0$ and $t = T$ and hence enters the space singularity $x = 0$ of f, all known existence results for the periodic problem (5.1), (5.2) under the assumption (5.3) concern positive solutions which do not touch the space singularity $x = 0$ of the function f.

DEFINITION 5.2. A function $u : [0,T] \to \mathbb{R}$ is a *positive solution* to problem (5.1), (5.2) if $\phi(u') \in AC[0,T]$, $u > 0$ on $[0,T]$, $(\phi(u'(t)))' = f(t,u(t),u'(t))$ for a.e. $t \in [0,T]$ and (5.2) is satisfied.

The restriction to positive solutions causes that the general existence principle in Theorem 1.8 about a limit of a sequence of approximate solutions need not be employed here. On the other hand, the singular problem (5.1), (5.2) will be also investigated through regular approximating periodic problems having differential equations of the form

$$\big(\phi(u')\big)' = h(t,u,u'), \tag{5.4}$$

where $h \in Car([0,T] \times \mathbb{R}^2)$. As usual, by a *solution of the regular problem* (5.4), (5.2) we understand a function u such that $\phi(u') \in AC[0,T]$, (5.2) is true and $(\phi(u'(t)))' = h(t,u(t),u'(t))$ for a.e. $t \in [0,T]$.

Notice that the requirement $\phi(u') \in AC[0,T]$ implies that $u \in C^1[0,T]$.

We will also discuss various special cases of (5.1) including the classical one with $\phi(y) \equiv y$ or those with f not depending on u' or with f depending on u' linearly.

Let us notice that the assumption that ϕ is an odd function is only technical and it is sufficient to assume (3.3) as in Section 3 in most cases. We employ it just to simplify some formulas occurring in this section.

5.1. *Method of lower and upper functions for regular problems*

First, we will consider problem (5.4), (5.2), where $h \in Car([0,T] \times \mathbb{R}^2)$. We bring some results which will be exploited in the investigation of the singular problem (5.1), (5.2). The lower and upper functions method combined with the topological degree argument is an important tool for proofs of solvability of regular periodic problems. Several rather general definitions of lower and upper functions are available (see, e.g., [47,48,62,93,124, 147]). However, for our purposes the following one seems to be optimal.

DEFINITION 5.3. We say that a function $\sigma \in C[0,T]$ is a *lower function* of problem (5.4), (5.2) if there is a finite set $\Sigma \subset (0,T)$ such that $\phi(\sigma') \in AC_{loc}([0,T] \setminus \Sigma)$, $\sigma'(\tau+)$, $\sigma'(\tau-) \in \mathbb{R}$ for each $\tau \in \Sigma$ and

$$\big(\phi\big(\sigma'(t)\big)\big)' \geqslant h\big(t, \sigma(t), \sigma'(t)\big) \quad \text{for a.e. } t \in [0,T], \tag{5.5}$$

$$\sigma(0) = \sigma(T), \quad \sigma'(0) \geqslant \sigma'(T), \tag{5.6}$$

$$\sigma'(\tau+) > \sigma'(\tau-) \quad \text{for all } \tau \in \Sigma. \tag{5.7}$$

If the inequalities in (5.5)–(5.7) are reversed, σ is called an *upper function* of problem (5.4), (5.2).

The role of lower and upper functions is demonstrated by the following "maximum principle":

LEMMA 5.4. *Let σ_1 and σ_2 be a lower and an upper function of* (5.4), (5.2) *and* $\sigma_1 \leqslant \sigma_2$ *on* $[0,T]$. *Then for each $d \in [\sigma_1(0), \sigma_2(0)]$ and each $\tilde{f} \in Car([0,T] \times \mathbb{R}^2)$ such that*

$$\begin{cases} \tilde{f}(t,x,y) < h\big(t, \sigma_1(t), \sigma_1'(t)\big) \textit{ for a.e. } t \in [0,T], \textit{ all } x \in \big(-\infty, \sigma_1(t)\big) \\ \quad \textit{and all } y \in \mathbb{R} \textit{ such that } \big|y - \sigma_1'(t)\big| \leqslant \dfrac{\sigma_1(t) - x}{\sigma_1(t) - x + 1}, \\ \tilde{f}(t,x,y) > h\big(t, \sigma_2(t), \sigma_2'(t)\big) \textit{ for a.e. } t \in [0,T], \textit{ all } x \in \big(\sigma_2(t), \infty\big) \\ \quad \textit{and all } y \in \mathbb{R} \textit{ such that } \big|y - \sigma_2'(t)\big| \leqslant \dfrac{x - \sigma_2(t)}{x - \sigma_2(t) + 1}, \end{cases} \tag{5.8}$$

any solution u of the problem

$$\big(\phi(u')\big)' = \tilde{f}(t,u,u'), \quad u(0) = u(T) = d$$

satisfies $\sigma_1 \leqslant u \leqslant \sigma_2$ *on* $[0, T]$.

PROOF. Denote $v = u - \sigma_1$ and assume that $v(\alpha) = \min\{v(t)\colon t \in [0, T]\} < 0$. Since $d \in [\sigma_1(0), \sigma_2(0)]$ and thanks to (5.6) and (5.7), we may assume that $\alpha \in (0, T) \setminus \Sigma$, $v'(\alpha) = 0$, and there is $\beta \in (\alpha, T]$ such that $(\alpha, \beta] \cap \Sigma = \emptyset$ and

$$v(t) < 0 \quad \text{and} \quad \left|v'(t)\right| < \frac{-v(t)}{1 - v(t)} \quad \text{for all } t \in [\alpha, \beta].$$

Using (5.5) (where $\sigma = \sigma_1$) and (5.8), we obtain

$$\bigl(\phi\bigl(u'(t)\bigr) - \phi\bigl(\sigma_1'(t)\bigr)\bigr)' < h\bigl(t, \sigma_1(t), \sigma_1'(t)\bigr) - \bigl(\phi\bigl(\sigma_1'(t)\bigr)\bigr)' \leqslant 0$$
$$\text{for a.e. } t \in [\alpha, \beta].$$

Hence

$$0 > \int_\alpha^t \bigl(\phi\bigl(u'(s)\bigr) - \phi\bigl(\sigma_1'(s)\bigr)\bigr)' \,\mathrm{d}s = \phi\bigl(u'(t)\bigr) - \phi\bigl(\sigma_1'(t)\bigr) \quad \text{for all } t \in (\alpha, \beta],$$

which leads to a contradiction with the definition of α, i.e. $u \geqslant \sigma_1$ on $[0, T]$. Similarly we can show that $u \leqslant \sigma_2$ on $[0, T]$. □

Problem (5.4), (5.2) is often transformed to a fixed point problem (see, e.g., [41,101,104, 150]). Here we present one possibility how to find an operator representation of (5.4), (5.2) in the space $C^1[0, T]$. Having in mind that the periodic conditions (5.2) can be equivalently rewritten as

$$u(0) = u(T) = u(0) + u'(0) - u'(T),$$

let us consider the quasilinear Dirichlet problem

$$\bigl(\phi(x')\bigr)' = b(t) \quad \text{a.e. on } [0, T], \quad x(0) = x(T) = d \tag{5.9}$$

with $b \in L_1[0, T]$ and $d \in \mathbb{R}$. A function $x \in C^1[0, T]$ is a solution of (5.9) if and only if there is $a \in \mathbb{R}$ such that

$$x(t) = d + \int_0^t \phi^{-1}\left(a + \int_0^s b(\tau)\,\mathrm{d}\tau\right)\mathrm{d}s \quad \text{for } t \in [0, T]$$

and

$$\int_0^T \phi^{-1}\left(a + \int_0^s b(\tau)\,\mathrm{d}\tau\right)\mathrm{d}s = 0. \tag{5.10}$$

Since ϕ is increasing on $\mathbb{R}$ and $\phi(\mathbb{R}) = \mathbb{R}$, Eq. (5.10) has exactly one solution $a = a(b) \in \mathbb{R}$ for each $b \in L_1[0, T]$. So, we can define an operator $\mathcal{K}: L_1[0, T] \to C^1[0, T]$ by

$$\big(\mathcal{K}(b)\big)(t) = \int_0^t \phi^{-1}\left(a(b) + \int_0^s b(\tau)\,\mathrm{d}\tau\right)\mathrm{d}s \quad \text{for } t \in [0, T]. \tag{5.11}$$

Let $\mathcal{N}: C^1[0, T] \to L_1[0, T]$ and $\mathcal{F}: C^1[0, T] \to C^1[0, T]$ have the form $(\mathcal{N}(u))(t) = h(t, u(t), u'(t))$ for a.e. $t \in [0, T]$ and

$$\big(\mathcal{F}(u)\big)(t) = u(0) + u'(0) - u'(T) + \big(\mathcal{K}\big(\mathcal{N}(u)\big)\big)(t) \quad \text{for } t \in [0, T]. \tag{5.12}$$

In view of the definition of $\mathcal{K}$, a function $x \in C^1[0, T]$ is a solution to (5.9) if and only if $x = d + \mathcal{K}(b)$. Therefore, $u \in C^1[0, T]$ is a solution to (5.4), (5.2) if and only if it is a fixed point of $\mathcal{F}$.

An alternative representation of the operator $\mathcal{F}$ can be obtained by inserting $\alpha(u) = u(0) - d$, $\beta(u) = d - u(T)$ and $d = u(0) + u'(0) - u'(T)$ into the operator $\mathcal{P}(1, u)$ defined in the proof of Theorem 3.2. In this way we get

$$\begin{aligned}\big(\mathcal{F}(u)\big)(t) &= u(0) + u'(0) - u'(T) \\ &\quad + \int_0^t \phi^{-1}\left(\phi\big(u'(T) + u(T) - u(0)\big) + \int_0^s \big(\mathcal{N}(u)\big)(\tau)\,\mathrm{d}\tau\right)\mathrm{d}s\end{aligned}$$

for $t \in [0, T]$ and $u \in C^1[0, T]$.

Taking into account [101, Proposition 2.2] or the proof of Theorem 3.2, we can summarize:

LEMMA 5.5. *Let $\mathcal{F}: C^1[0, T] \to C^1[0, T]$ be defined by* (5.12). *Then $\mathcal{F}$ is completely continuous and $u \in C^1[0, T]$ is a solution to* (5.4), (5.2) *if and only if $\mathcal{F}(u) = u$.*

The next lemma describes the relationship between lower and upper functions and the Leray–Schauder topological degree. We will consider the class of auxiliary problems

$$\big(\phi(v')\big)' = \eta(v')h(t, v, v'), \quad v(0) = v(T), \quad v'(0) = v'(T), \tag{5.13}$$

where $\eta: \mathbb{R} \to [0, 1]$ may be an arbitrary continuous function.

LEMMA 5.6. *Let σ_1 and σ_2 be a lower and an upper function of* (5.4), (5.2) *and $\sigma_1 < \sigma_2$ on* $[0, T]$. *Furthermore, let there exist $r^* > 0$ such that*

$$\begin{cases} \|v'\|_\infty < r^* & \textit{for each continuous } \eta: \mathbb{R} \to [0, 1] \textit{ and for} \\ \textit{each solution } v \textit{ of } (5.13) \textit{ such that } \sigma_1 \leqslant v \leqslant \sigma_2 \textit{ on } [0, T]. \end{cases} \tag{5.14}$$

Finally, assume that $\mathcal{F}: C^1[0, T] \to C^1[0, T]$ is defined by (5.12) *and, for $\rho > 0$, denote*

$$\Omega_\rho = \big\{u \in C^1[0, T]\colon \sigma_1 < u < \sigma_2 \textit{ on } [0, T] \textit{ and } \|u'\|_\infty < \rho\big\}. \tag{5.15}$$

Then

$$\deg(\mathcal{I}-\mathcal{F},\Omega_\rho)=1 \quad \textit{for each } \rho \geqslant r^* \textit{ such that } \mathcal{F}(u)\neq u \textit{ on } \partial\Omega_\rho.$$

PROOF. Put $\Omega=\Omega_{r^*}$, $R^*=r^*+\|\sigma_1'\|_\infty+\|\sigma_2'\|_\infty$,

$$\eta(y)=\begin{cases}1 & \text{if } |y|\leqslant R^*,\\ 2-\frac{|y|}{R^*} & \text{if } R^*<|y|\leqslant 2R^*,\\ 0 & \text{if } |y|>2R^*\end{cases}$$

and assume that $\mathcal{F}(x)\neq x$ for all $x\in\partial\Omega$. Then σ_1 and σ_2 are a lower and an upper function for the modified problem (5.13) and there is a $\psi\in L_1[0,T]$ satisfying

$$\bigl|\eta(y)h(t,x,y)\bigr|\leqslant\psi(t) \quad \text{for a.e. } t\in[0,T] \text{ and all } (x,y)\in[\sigma_1(t),\sigma_2(t)]\times\mathbb{R}.$$

We can construct (see the proof of Theorem 2.1 in [127]) the function $\widetilde{f}\in Car([0,T]\times\mathbb{R}^2)$ so that

$$\begin{aligned}&\widetilde{f}(t,x,y)=\eta(y)h(t,x,y)\\ &\quad\text{for a.e. } t\in[0,T] \text{ and all } (x,y)\in[\sigma_1(t),\sigma_2(t)]\times\mathbb{R},\\ &\bigl|\widetilde{f}(t,x,y)\bigr|\leqslant\widetilde{\psi}(t)\\ &\quad\text{for a.e. } t\in[0,T],\ \text{all } (x,y)\in\mathbb{R}^2 \text{ and some } \widetilde{\psi}\in L_1[0,T]\end{aligned}$$

and $\widetilde{f}$ satisfies the assumptions of Lemma 5.4 with $\eta(y)h(t,x,y)$ in place of $h(t,x,y)$. Define $\widetilde{\mathcal{F}}: C^1[0,T]\to C^1[0,T]$ by

$$\widetilde{\mathcal{F}}(u)=\alpha\bigl(u(0)+u'(0)-u'(T)\bigr)+\mathcal{K}\bigl(\widetilde{\mathcal{N}}(u)\bigr),$$

where

$$\bigl(\widetilde{\mathcal{N}}(u)\bigr)(t)=\widetilde{f}\bigl(t,u(t),u'(t)\bigr) \quad \text{for } u\in C^1[0,T] \text{ and a.e. } t\in[0,T],$$

$$\alpha(x)=\begin{cases}\sigma_1(0) & \text{if } x<\sigma_1(0),\\ x & \text{if } \sigma_1(0)\leqslant x\leqslant\sigma_2(0),\\ \sigma_2(0) & \text{if } x>\sigma_2(0)\end{cases}$$

and let $\mathcal{K}: L_1[0,T]\to C^1[0,T]$ be given in (5.11). By Lemma 5.5, the operator $\widetilde{\mathcal{F}}$ is completely continuous. Moreover, it follows from the definition of the operator $\mathcal{K}$ that the problem

$$\bigl(\phi(u')\bigr)'=\widetilde{f}(t,u,u'), \quad u(0)=u(T)=\alpha\bigl(u(0)+u'(0)-u'(T)\bigr)$$

is equivalent to the operator equation $\widetilde{\mathcal{F}}(u) = u$. We can find $r_0 \in (0, \infty)$ such that for any $\lambda \in [0, 1]$, each fixed point u of the operator $\lambda\widetilde{\mathcal{F}}$ belongs to $\mathcal{B}(r_0) = \{x \in C^1[0, T]\colon \|x\|_\infty + \|x'\|_\infty < r_0\}$. So, $\mathcal{I} - \lambda\widetilde{\mathcal{F}}$ is a homotopy on $\overline{\mathcal{B}}(r_0) \times [0, 1]$ and

$$\deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \mathcal{B}(r_0)\bigr) = \deg\bigl(\mathcal{I}, \mathcal{B}(r_0)\bigr) = 1.$$

Put $\Omega_1 = \{u \in \Omega\colon \sigma_1(0) < u(0) + u'(0) - u'(T) < \sigma_2(0)\}$. Clearly, $\widetilde{\mathcal{F}} = \mathcal{F}$ on $\overline{\Omega}_1$ and $u \in \Omega_1$ whenever $\mathcal{F}(u) = u$ and $u \in \Omega$. Using Lemma 5.4, we can prove that

$$\bigl(\widetilde{\mathcal{F}}(u) = u\bigr) \quad\Longrightarrow\quad u \in \Omega_1$$

which, by the excision property of the degree, yields

$$\deg(\mathcal{I} - \mathcal{F}, \Omega) = \deg(\mathcal{I} - \mathcal{F}, \Omega_1) = \deg(\mathcal{I} - \widetilde{\mathcal{F}}, \Omega_1) = \deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \mathcal{B}(r_0)\bigr) = 1.$$

Finally, according to (5.14) all fixed points u of $\mathcal{F}$ such that $\sigma_1 < u < \sigma_2$ on $[0, T]$ belong to Ω. Thus

$$\deg(\mathcal{I} - \mathcal{F}, \Omega_\rho) = \deg(\mathcal{I} - \mathcal{F}, \Omega) = 1$$

for each $\rho \geqslant r^*$ such that $\mathcal{F}(x) \neq x$ on $\partial\Omega_\rho$. □

Lemma 5.6 offers a possibility to get existence results for problems having a pair of lower and upper functions σ_1 and σ_2 satisfying

$$\sigma_1 \leqslant \sigma_2 \quad \text{on } [0, T]. \tag{5.16}$$

In such a case we say that σ_1 and σ_2 are *well-ordered* and the existence of an a priori estimate r^* with the property (5.14) is usually ensured by conditions of Nagumo type. The most general known version of such conditions is provided by the next lemma which is a modified version of the result by Staněk [133, Lemma 1].

LEMMA 5.7. *Let* $\sigma_1, \sigma_2 \in C[0, T]$ *satisfy* (5.16) *and assume that*

$$\begin{cases} \psi \in L_1[0, T] \text{ is nonnegative}, \ \varepsilon_1, \varepsilon_2 \in \{-1, 1\}, \\ \omega \in C(\mathbb{R}) \text{ is positive and } \displaystyle\int_{-\infty}^{0} \frac{\mathrm{d}s}{\omega(s)} = \int_0^\infty \frac{\mathrm{d}s}{\omega(s)} = \infty. \end{cases} \tag{5.17}$$

Then there is an $r^* > 0$ *such that*

$$\|v'\|_\infty < r^* \tag{5.18}$$

holds for each $v \in C^1[0,T]$ *such that* $\phi(v') \in AC[0,T]$, $v(0) = v(T)$, $v'(0) = v'(T)$, $\sigma_1 \leqslant v \leqslant \sigma_2$ *on* $[0,T]$ *and, for a.e.* $t \in [0,T]$,

$$\begin{cases} \varepsilon_1\bigl(\phi\bigl(v'(t)\bigr)\bigr)' \leqslant \bigl(\psi(t) + v'(t)\bigr)\omega\bigl(\phi\bigl(v'(t)\bigr)\bigr) & \text{if } v'(t) > 0, \\ \varepsilon_2\bigl(\phi\bigl(v'(t)\bigr)\bigr)' \leqslant \bigl(\psi(t) - v'(t)\bigr)\omega\bigl(\phi\bigl(v'(t)\bigr)\bigr) & \text{if } v'(t) < 0. \end{cases}$$

□

Lemma 5.6 provides also a crucial argument for the proof of existence of a solution even in the case that the given problem possesses lower and upper functions σ_1 and σ_2 which do not satisfy (5.16), i.e. if

$$\sigma_1(\tau) > \sigma_2(\tau) \quad \text{for some } \tau \in [0,T]. \tag{5.19}$$

In such a case, the following a priori estimate is available.

LEMMA 5.8. *Let* $\psi \in L_1[0,T]$. *Then there is* $r^* > 0$ *such that* (5.18) *holds for each* $v \in C^1[0,T]$ *fulfilling* $\phi(v') \in AC[0,T]$, $v(0) = v(T)$, $v'(0) = v'(T)$ *and* $(\phi(v'(t)))' > \psi(t)$ (*or* $(\phi(v'(t)))' < \psi(t)$) *for a.e.* $t \in [0,T]$.

PROOF. We will restrict ourselves to the case that $(\phi(v'(t)))' > \psi(t)$ for a.e. $t \in [0,T]$. (The other case can be proved by a similar argument.) By the proof of [129, Lemma 1.1], we can see that $\|w\|_\infty < \|\psi\|_1$ holds for each $w \in AC[0,T]$ such that $w(0) = w(T)$, $w(t_w) = 0$ for some $t_w \in (0,T)$ and $w'(t) > \psi(t)$ for a.e. $t \in [0,T]$. The assertion of the lemma follows by setting $w = \phi(v')$ and

$$r^* = \phi^{-1}\bigl(\|\psi\|_1\bigr). \tag{5.20}$$

□

The next lemma provides an existence principle which will be helpful later:

LEMMA 5.9. *Let* σ_1 *and* σ_2 *be a lower and an upper function of* (5.4), (5.2) *and let* (5.19) *be true. Furthermore, let there be* $m \in L_1[0,T]$ *such that*

$$h(t,x,y) > m(t) \ (\text{or } h(t,x,y) < m(t)) \quad \text{for a.e. } t \in [0,T] \text{ and all } x, y \in \mathbb{R}$$

and let $r^* > 0$ *be given by* (5.20), *where* $\psi = |m| + 2$. *Then problem* (5.4), (5.2) *has a solution u satisfying*

$$\|u'\|_\infty < r^* \tag{5.21}$$

and

$$\min\bigl\{\sigma_1(\tau_u), \sigma_2(\tau_u)\bigr\} \leqslant u(\tau_u) \leqslant \max\bigl\{\sigma_1(\tau_u), \sigma_2(\tau_u)\bigr\} \quad \text{for some } \tau_u \in [0,T]. \tag{5.22}$$

SKETCH OF THE PROOF. We follow the ideas of the proof of Theorem 3.2 in [127]. Assume, e.g., that $h(t,x,y) > m(t)$ for a.e. $t \in [0,T]$ and all $x, y \in \mathbb{R}$.

Step 1. *Construction of an auxiliary problem and the operator representation.*

Define $\psi(t)\colon = -(|m(t)| + 2)$ for a.e. $t \in [0,T]$, find $r^* > 0$ as in Lemma 5.8 and set $c^* = \|\sigma_1\|_\infty + \|\sigma_2\|_\infty + Tr^*$. Consider the auxiliary problem

$$\bigl(\phi(u')\bigr)' = \widetilde{f}(t,u,u'), \quad u(0) = u(T), \quad u'(0) = u'(T), \tag{5.23}$$

where

$$\widetilde{f}(t,x,y) = \begin{cases} -\bigl(|m(t)| + 1\bigr) & \text{if } x \leqslant -(c^*+1), \\ h(t,x,y) + (x + c^*)\bigl(|m(t)| + 1 + h(t,x,y)\bigr) & \\ & \text{if } -(c^*+1) < x < -c^*, \\ h(t,x,y) & \text{if } -c^* \leqslant x \leqslant c^*, \\ h(t,x,y) + (x - c^*)|m(t)| & \text{if } c^* < x < c^* + 1, \\ h(t,x,y) + |m(t)| & \text{if } x \geqslant c^* + 1. \end{cases}$$

We have

$$\begin{cases} \widetilde{f}(t,x,y) < 0 & \text{if } x \leqslant -(c^*+1), \\ \widetilde{f}(t,x,y) > 0 & \text{if } x \geqslant c^* + 1, \\ \widetilde{f}(t,x,y) = h(t,x,y) & \text{if } x \in [-c^*, c^*], \end{cases} \tag{5.24}$$

$$\widetilde{f}(t,x,y) > \psi(t) \quad \text{for a.e. } t \in [0,T] \text{ and all } x, y \in \mathbb{R} \tag{5.25}$$

and σ_1 and σ_2 are a lower and an upper function of (5.23). Moreover, $\sigma_3(t) \equiv -c^* - 2$ and $\sigma_4(t) \equiv c^* + 2$ form another pair of a lower and an upper function for (5.23) and

$$\sigma_3 < \min\{\sigma_1, \sigma_2\} \leqslant \max\{\sigma_1, \sigma_2\} < \sigma_4 \text{ on } [0,T].$$

Denote $\Omega_0 = \{u \in C^1[0,T]\colon \sigma_3 < u < \sigma_4 \text{ on } [0,T],\ \|u'\|_\infty < r^*\}$,

$$\Omega_1 = \bigl\{u \in \Omega_0\colon \sigma_3 < u < \sigma_2 \text{ on } [0,T]\bigr\},$$
$$\Omega_2 = \bigl\{u \in \Omega_0\colon \sigma_1 < u < \sigma_4 \text{ on } [0,T]\bigr\}$$

and $\Omega = \Omega_0 \setminus \overline{\Omega_1 \cup \Omega_2}$. By Lemma 5.5, problem (5.23) is equivalent to the operator equation $\widetilde{\mathcal{F}}(u) = u$ in $C^1[0,T]$, where $\widetilde{\mathcal{F}}(u) = u(0) + u'(0) - u'(T) + \mathcal{K}(\widetilde{\mathcal{N}}(u))$, $(\widetilde{\mathcal{N}}(u))(t) = \widetilde{f}(t, u(t), u'(t))$ and $\mathcal{K}\colon L_1[0,T] \to C^1[0,T]$ is given by (5.11). Clearly, $\widetilde{\mathcal{F}}(u) = \mathcal{F}(u)$ for $u \in C^1[0,T]$ such that $\|u\|_\infty \leqslant c^*$.

Step 2. *A priori estimates*. We show that

$$\|u'\|_\infty < r^* \quad \text{and} \quad \|u\|_\infty < c^*$$

is true for all $u \in \overline{\Omega}$ such that $\widetilde{\mathcal{F}}(u) = u$.

Step 3. *Existence of a solution to* (5.4), (5.2).

Let $\widetilde{\mathcal{F}}(u) = u$ and $u \in \partial\Omega$. By Step 3, we have $\mathcal{F}(u) = \widetilde{\mathcal{F}}(u) = u$ and u solves (5.4), (5.2). Let $\widetilde{\mathcal{F}}(u) \neq u$ on $\partial\Omega$. Then using Lemma 5.6 we get

$$\deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \Omega_0\bigr) = \deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \Omega_1\bigr) = \deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \Omega_2\bigr) = 1.$$

Furthermore, by (5.19), we have $\Omega_1 \cap \Omega_2 = \emptyset$. Therefore, due to the additive property of the degree,

$$\deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \Omega\bigr) = \deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \Omega_0\bigr) - \deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \Omega_1\bigr) - \deg\bigl(\mathcal{I} - \widetilde{\mathcal{F}}, \Omega_2\bigr) = -1$$

which implies that $\widetilde{\mathcal{F}}$ has a fixed point $u \in \Omega$. It follows by Step 3 that $\|u\|_\infty < c^*$ which, by virtue of (5.24), means that u solves (5.4), (5.2).

We can proceed analogously when $h(t, x, y) < m(t)$ for a.e. $t \in [0, T]$ and all $x, y \in \mathbb{R}$. □

5.2. *Method of lower and upper functions for singular problems*

Now, we consider problem (5.1), (5.2) where f satisfies (5.3). We present sufficient conditions in terms of lower and upper functions for the existence of positive solutions to (5.1), (5.2). Lower and upper functions σ_1 and σ_2 are defined in the same way as for the regular problem (5.4), (5.2) (see Definition 5.3). However, since problem (5.1), (5.2) is investigated on $[0, T] \times \mathcal{A}$ where $\mathcal{A} = [0, \infty) \times \mathbb{R}$, only such σ_1 and σ_2 which are positive a.e. on $[0, T]$ make sense.

The first existence result concerns problem (5.1), (5.2) having well-ordered lower and upper functions.

THEOREM 5.10. *Let there exist lower and upper functions* σ_1 *and* σ_2 *of problem* (5.1), (5.2) *such that* (5.16) *is true and* $\sigma_1 > 0$ *on* $[0, T]$. *Furthermore, let for a.e.* $t \in [0, T]$ *and each* $(x, y) \in [\sigma_1(t), \sigma_2(t)] \times \mathbb{R}$ *the inequalities*

$$\begin{cases} \varepsilon_1 f(t, x, y) \leqslant \bigl(\psi(t) + y\bigr)\omega\bigl(\phi(y)\bigr) & \text{if } y > 0, \\ \varepsilon_2 f(t, x, y) \leqslant \bigl(\psi(t) - y\bigr)\omega\bigl(\phi(y)\bigr) & \text{if } y < 0 \end{cases} \tag{5.26}$$

hold with $\varepsilon_1, \varepsilon_2, \omega$ *and* ψ *satisfying* (5.17). *Then problem* (5.1), (5.2) *has a positive solution* u *such that*

$$\sigma_1 \leqslant u \leqslant \sigma_2 \quad \text{on } [0, T]. \tag{5.27}$$

PROOF. *Step* 1. *The case* $\sigma_1 < \sigma_2$.

Assume that $\sigma_1 < \sigma_2$ on $[0, T]$. Consider the auxiliary problem (5.4), (5.2) with $h(t, x, y) = f(t, \max\{\sigma_1(t), \min\{x, \sigma_2(t)\}, y)$ for a.e. $t \in [0, T]$ and $(x, y) \in \mathbb{R}^2$. Clearly, $h \in Car([0, T] \times \mathbb{R}^2)$ and $h(t, x, y) = f(t, x, y)$ if $x \in [\sigma_1(t), \sigma_2(t)]$. Further, σ_1 and σ_2 are a lower and an upper function of (5.4), (5.2). Choose an arbitrary continuous $\eta : \mathbb{R} \to [0, 1]$

and let v be an arbitrary solution of (5.13) fulfilling $\sigma_1 \leqslant v \leqslant \sigma_2$ on $[0, T]$. Since (5.26) is satisfied with h in place of f, we have for a.e. $t \in [0, T]$

$$\begin{aligned}\varepsilon_1\big(\phi\big(v'(t)\big)\big)' &= \varepsilon_1\eta\big(v'(t)\big)h\big(t, v(t), v'(t)\big)\\ &\leqslant \eta\big(v'(t)\big)\big(\psi(t) + v'(t)\big)\omega\big(\phi\big(v'(t)\big)\big)\\ &\leqslant \big(\psi(t) + v'(t)\big)\omega\big(\phi\big(v'(t)\big)\big) \quad \text{if } v'(t) > 0\end{aligned}$$

and

$$\varepsilon_2\big(\phi\big(v'(t)\big)\big)' \leqslant \big(\psi(t) - v'(t)\big)\omega\big(\phi\big(v'(t)\big)\big) \quad \text{if } v'(t) < 0.$$

Hence we can apply Lemma 5.7 to deduce that (5.14) is satisfied. Let $\mathcal{F}: C^1[0, T] \to C^1[0, T]$ and $\Omega = \Omega_{r^*}$ be defined by (5.12) and (5.15), respectively. Then there are two possibilities: either $\mathcal{F}$ has a fixed point $u \in \partial\Omega$ or $\mathcal{F}(u) \neq u$ on $\partial\Omega$.

(a) Let $\mathcal{F}(u) = u$ for some $u \in \partial\Omega$. In view of Lemma 5.5 and of the definition of h, it follows that u is a solution to (5.1), (5.2) fulfilling (5.27).

(b) If $\mathcal{F}(u) \neq u$ on $\partial\Omega$, then by Lemma 5.6 we have $\deg(\mathcal{I} - \mathcal{F}, \Omega) = 1$, which implies that $\mathcal{F}$ has a fixed point $u \in \Omega$. As in (a), this fixed point is a solution to (5.1), (5.2) fulfilling (5.27).

Step 2. *The case* $\sigma_1 \leqslant \sigma_2$.

For each $k \in \mathbb{N}$, the function $\widetilde{\sigma}_k = \sigma_2 + \frac{1}{k}$ is also an upper function of (5.4), (5.2) and $\sigma_1 < \widetilde{\sigma}_k$ on $[0, T]$. Hence, in the general case, when the strict inequality between σ_1 and σ_2 need not hold, we can use Step 1 to show that for each $k \in \mathbb{N}$ there exists a solution u_k to (5.4), (5.2) such that

$$u_k(t) \in \big[\sigma_1(t), \sigma_2(t) + \tfrac{1}{k}\big] \quad \text{for } t \in [0, T] \quad \text{and} \quad \|u_k'\|_\infty < \rho^*,$$

where $\rho^* > 0$ is the a priori estimate given by Lemma 5.7 with $\sigma_2 + 1$ in place of σ_2. Using the Arzelà–Ascoli theorem and the Lebesgue dominated convergence theorem for the sequence $\{u_k\}$ we get a solution u of (5.1), (5.2) as the C^1-limit of a subsequence of $\{u_k\}$. □

REMARK 5.11. Theorem 5.10 provides the existence of a positive solution to problem (5.1), (5.2) with $f(t, x, y) = -h(x)y + g(t, x)$, if $h \in C[0, \infty)$, $g \in Car([0, T] \times (0, \infty))$ and if the existence of well-ordered and positive lower and upper functions is ensured. Indeed, for a.e. $t \in [0, T]$ and each $(x, y) \in [\sigma_1(t), \sigma_2(t)] \times \mathbb{R}$, we have

$$\big|f(t, x, y)\big| \leqslant \big|h(x)\big||y| + \big|g(t, x)\big| \leqslant K\big(\psi(t) + |y|\big)$$

where $K = 1 + \max\{|h(x)|\colon x \in [\delta, \|\sigma_2\|_\infty]\}$, $\psi(t) = \sup\{|g(t, x)|\colon x \in [\delta, \|\sigma_2\|_\infty]\}$ and $\delta = \min\{\sigma_1(t)\colon t \in [0, T]\}$. (By assumption, we have $\delta > 0$.)

Now, we will consider problem (5.1), (5.2) which has lower and upper functions, but no pair of them is well-ordered. We will deal with the periodic problem for the equation

$$\big(\phi(u')\big)' = g(u) + p(t, u, u'). \tag{5.28}$$

THEOREM 5.12. *Assume* $g \in C(0,\infty)$, $p \in Car([0,T]\times\mathbb{R}^2)$ *and*

$$\lim_{x\to 0+}\int_x^1 g(\xi)\,\mathrm{d}\xi = +\infty. \tag{5.29}$$

Let there exist lower and upper functions σ_1 *and* σ_2 *of problem* (5.28), (5.2) *such that* (5.19) *is true and* $\sigma_2 > 0$ *on* $[0,T]$. *Furthermore, let there exist an* $m \in L_1[0,T]$ *such that*

$$g(x) + p(t,x,y) > m(t) \quad \textit{for a.e. } t \in [0,T] \textit{ and all } x > 0,\ y \in \mathbb{R} \tag{5.30}$$

holds and let r^* *be given by* (5.20) *with* $\psi = |m| + 2$. *Then problem* (5.28), (5.2) *has a positive solution* u *satisfying* (5.21) *and* (5.22).

PROOF. We will use the ideas of [130]. Similarly to [130, Lemma 2.5] we can deduce from (5.29) and (5.30) that σ_1 is positive on $[0,T]$. Thus, $\delta := \min\{\{\sigma_1(t),\sigma_2(t)\}\colon\ t \in [0,T]\} > 0$. Put $R = \|\sigma_1\|_\infty + \|\sigma_2\|_\infty$ and $B = R + r^*T$. Furthermore, as $p \in Car([0,T]\times\mathbb{R}^2)$, there is $\widetilde{p} \in L_1[0,T]$ such that $|p(t,x,y)| \leqslant \widetilde{p}(t)$ for a.e. $t \in [0,T]$ and all $(x,y) \in [0,B]\times[-r^*,r^*]$. Put

$$K = \|\widetilde{p}\|_1 r^* + \int_\delta^B |g(\xi)|\,\mathrm{d}\xi.$$

By (5.29) there exists $\varepsilon \in (0,\delta)$ such that $g(\varepsilon) > 0$ and

$$\int_\varepsilon^\delta g(\xi)\,\mathrm{d}\xi > K. \tag{5.31}$$

For a.e. $t \in [0,T]$ and $(x,y) \in \mathbb{R}^2$, define

$$h(t,x,y) = \widetilde{g}(x) + p(t,x,y), \quad \text{where } \widetilde{g}(x) = \begin{cases} g(\varepsilon) & \text{if } x < \varepsilon,\\ g(x) & \text{if } x \geqslant \varepsilon.\end{cases}$$

Then $h \in Car([0,T]\times\mathbb{R}^2)$, σ_1 and σ_2 are lower and upper functions of (5.4), (5.2) and, by (5.30), $h(t,x,y) > m(t)$ for a.e. $t \in [0,T]$ and all $x > 0$, $y \in \mathbb{R}$. By Lemma 5.9, problem (5.4), (5.2) has a solution u satisfying (5.21) and $\delta \leqslant u(t_u) \leqslant R$ for some $t_u \in [0,T]$. In particular, $u \leqslant B$ for all $t \in [0,T]$. It remains to show that $u \geqslant \varepsilon$ on $[0,T]$.

Let $t_0, t_1 \in [0,T]$ be such that $u(t_0) = \min\{u(t)\colon\ t \in [0,T]\}$ and $u(t_1) = \max\{u(t)\colon\ t \in [0,T]\}$. We have $u'(t_0) = u'(t_1) = 0$ and $u(t_1) \in [\delta, B]$. Put $v(t) = \phi(u'(t))$ for $t \in [0,T]$. Then $u'(t) = \phi^{-1}(v(t))$ on $[0,T]$, $v(t_0) = v(t_1) = \phi(0)$ and

$$\int_{t_0}^{t_1} \bigl(\phi\bigl(u'(s)\bigr)\bigr)' u'(s)\,\mathrm{d}s = \int_{t_0}^{t_1} v'(s)\phi^{-1}\bigl(v(s)\bigr)\,\mathrm{d}s = \int_{v(t_0)}^{v(t_1)} \phi^{-1}(\xi)\,\mathrm{d}\xi = 0.$$

Thus, multiplying both sides of the equality $(\phi(u'(t)))' = h(t, u(t), u'(t))$ by $u'(t)$ and integrating from t_0 to t_1, we get

$$\int_{u(t_0)}^{u(t_1)} \widetilde{g}(\xi)\,\mathrm{d}\xi \leqslant \int_{t_0}^{t_1} \bigl|p\bigl(t, u(t), u'(t)\bigr)\bigr|\bigl|u'(t)\bigr|\,\mathrm{d}t \leqslant \|\widetilde{p}\|_1 r^*.$$

Therefore

$$\begin{aligned} g(\varepsilon)\bigl(\varepsilon - u(t_0)\bigr) + \int_{\varepsilon}^{\delta} g(\xi)\,\mathrm{d}\xi &= \int_{u(t_0)}^{\delta} \widetilde{g}(\xi)\,\mathrm{d}\xi \\ \leqslant \int_{u(t_0)}^{u(t_1)} \widetilde{g}(\xi)\,\mathrm{d}\xi + \int_{\delta}^{B} \bigl|g(\xi)\bigr|\,\mathrm{d}\xi &\leqslant \|\widetilde{p}\|_1 r^* + \int_{\delta}^{B} \bigl|g(\xi)\bigr|\,\mathrm{d}\xi = K. \end{aligned}$$

Since $g(\varepsilon) > 0$, this contradicts (5.31) whenever $u(t_0) = \min\{u(t)\colon t \in [0, T]\} < \varepsilon$. Hence, $u(t) \geqslant \varepsilon$ on $[0, T]$ which means that u is a solution to (5.28), (5.2). ☐

REMARK 5.13. Let g and p fulfill the assumptions of Theorem 5.12 and $f(t, x, y) = g(x) + p(t, x, y)$. Then the condition (5.29) implies that

$$\limsup_{x \to 0+} g(x) = +\infty, \tag{5.32}$$

which means by Remark 5.1 that f and g have a space repulsive singularity at $x = 0$. Each repulsive singularity having the property (5.29) is called a *strong singularity* of f and the corresponding function g is usually called a *strong repulsive singular force*. On the contrary, if (5.32) holds together with

$$\lim_{x \to 0+} \int_x^1 g(\xi)\,\mathrm{d}\xi \in \mathbb{R}, \tag{5.33}$$

then the singularity of f at $x = 0$ is called a *weak singularity* and g is called a *weak repulsive singular force.*

5.3. *Attractive singular forces*

This section is devoted to singular problem (5.1), (5.2) where f can have an attractive singularity at $x = 0$. (See Remark 5.1.)

In what follows we use the standard notation for mean values of integrable functions: for $y \in L_1[0, T]$, the symbol $\overline{y}$ stands for

$$\overline{y}\colon = \frac{1}{T} \int_0^T y(t)\,\mathrm{d}t.$$

THEOREM 5.14. *Let there exist* $r > 0$, $A > r$ *and* $b \in L_1[0,T]$ *such that* $\overline{b} \geqslant 0$,

$$f(t,r,0) \leqslant 0 \quad \textit{for a.e. } t \in [0,T], \tag{5.34}$$

$$f(t,x,y) \geqslant b(t)$$
$$\textit{for a.e. } t \in [0,T] \textit{ and all } x \in [A,B] \textit{ and } |y| \leqslant \phi^{-1}\big(\|b\|_1\big), \tag{5.35}$$

where

$$B - A \geqslant 2T\phi^{-1}\big(\|b\|_1\big). \tag{5.36}$$

Furthermore, let for a.e. $t \in [0,T]$ *and each* $(x,y) \in [r,B] \times \mathbb{R}$ *the inequalities* (5.26) *hold with* ε_1, ε_2, ω *and* ψ *satisfying* (5.17).

Then problem (5.1), (5.2) *has a positive solution* u *such that*

$$r \leqslant u \leqslant B \quad \textit{on } [0,T]. \tag{5.37}$$

PROOF. For a given $d \in \mathbb{R}$, let x_d be a solution of (5.9). Then

$$\phi\big(x_d'(t)\big) = \phi\big(x_d'(t_0)\big) + \int_{t_0}^{t} b(s)\,\mathrm{d}s \qquad \text{for all } t, t_0 \in [0,T].$$

Since $\overline{b} \geqslant 0$, it follows that $x_d'(T) \geqslant x_d'(0)$. Since $x_d(0) = x_d(T)$, there is a $t_d \in (0,T)$ such that $x_d'(t_d) = 0$. Thus

$$\phi\big(x_d'(t)\big) = \int_{t_d}^{t} b(s)\,\mathrm{d}s \quad \text{for } t \in [0,T]$$

and so $\|x_d'\|_\infty \leqslant \phi^{-1}(\|b\|_1)$ for each $d \in \mathbb{R}$ and $\|x_0\|_\infty \leqslant T\phi^{-1}(\|b\|_1)$. Put $\sigma_2 = A + T\phi^{-1}(\|b\|_1) + x_0$. Then

$$A \leqslant \sigma_2 \leqslant A + 2T\phi^{-1}\big(\|b\|_1\big) \leqslant B \qquad \text{on } [0,T]. \tag{5.38}$$

Having in mind (5.35) and (5.9), we can see that σ_2 is an upper function of (5.1), (5.2). Furthermore, $\sigma_1 = r$ is a lower function of problem (5.1), (5.2) and $0 < \sigma_1 < \sigma_2$ on $[0,T]$. By Theorem 5.10, problem (5.1), (5.2) has a positive solution u satisfying (5.37). □

Now, let us consider the Liénard periodic problem

$$\big(\phi(u')\big)' + h(u)u' = g(t,u) + e(t), \quad u(0) = u(T), \quad u'(0) = u'(T), \tag{5.39}$$

where g can have an attractive space singularity at $x = 0$.

THEOREM 5.15. *Assume*

$$h \in C[0,\infty),\ e \in L_1[0,T],\ g \in Car\big([0,T]\times(0,\infty)\big), \tag{5.40}$$

$$\textit{there exists } \alpha > 0 \textit{ such that } \liminf_{|y|\to\infty} \frac{|\phi(y)|}{|y|^\alpha} > 0, \tag{5.41}$$

$$\begin{cases} \textit{there exists } r > 0 \textit{ such that} \\ g(t,r) + e(t) \leqslant 0 \quad \textit{for a.e. } t \in [0,T], \end{cases} \tag{5.42}$$

$$\begin{cases} \textit{there exist } A > r \textit{ and } g_0 \in L_1[0,T] \textit{ such that} \\ g(t,x) \geqslant g_0(t) \quad \textit{for a.e. } t \in [0,T] \textit{ and all } x \geqslant A \end{cases} \tag{5.43}$$

and

$$\overline{g}_0 + \overline{e} \geqslant 0. \tag{5.44}$$

Then problem (5.39) *has a positive solution u such that $u \geqslant r$ on* $[0,T]$.

SKETCH OF THE PROOF. We follow the ideas of the paper [128]. Define

$$f(t,x,y) = -h(x)y + g(t,x) + e(t)$$
$$\text{for a.e. } t \in [0,T] \text{ and all } x \in (0,\infty),\ y \in \mathbb{R}.$$

Step 1. First, notice that, due to (5.42), $\sigma_1(t) \equiv r$ is a lower function of (5.39).

Step 2. Thanks to (5.41), (5.43) and (5.44), we can construct an upper function σ_2 of (5.39). To this aim, take an arbitrary $C \in \mathbb{R}$ and consider a parameter auxiliary problem

$$\big(\phi(v')\big)' + \lambda h(v+C)v' = \lambda b(t), \quad v(0) = v(T) = 0, \quad \lambda \in [0,1], \tag{5.45}$$

where $b(t) = g_0(t) + e(t)$ for a.e. $t \in [0,T]$. By (5.41), there are $k > 0$ and $y_0 > 0$ such that

$$\big|\phi(y)\big| > \frac{k}{2}|y|^\alpha \quad \text{for } |y| \geqslant y_0. \tag{5.46}$$

Multiplying (5.45) by $v(t)$ and integrating over $[0,T]$, we obtain

$$-\int_0^T \phi\big(v'(t)\big)v'(t)\,\mathrm{d}t = \lambda\int_0^T b(t)v(t)\,\mathrm{d}t. \tag{5.47}$$

Using (5.46), (5.47) and the Hölder inequality, we can find $\rho \in (0,\infty)$, independent of $C \in \mathbb{R}$, such that $v \in \mathcal{B}(\rho) = \{x \in C^1[0,T]\colon \|x\|_\infty + \|x'\|_\infty < \rho\}$ holds for each $\lambda \in [0,1]$ and each solution v of (5.45). Thus, choosing a proper operator representation of problem (5.45) and using a standard homotopy and topological degree argument we can show that, for each $C \in \mathbb{R}$, problem (5.45) with $\lambda = 1$ has a solution $v_C \in \mathcal{B}(\rho)$. Now, it is already easy to see that if $C > A + \rho$, then $\sigma_2 = v_C + C$ is an upper function of (5.39). Indeed, we have $\sigma_2(0) = \sigma_2(T) = C$ and, due to (5.44),

$$\phi\big(\sigma_2'(T)\big) - \phi\big(\sigma_2'(0)\big) = T\overline{b} = T[\overline{g}_0 + \overline{e}] \geqslant 0.$$

Moreover, $\sigma_2(t) \geqslant C - \rho > A > r$ on $[0, T]$. Hence, by (5.43), we have

$$\begin{aligned}\big(\phi\big(\sigma_2'(t)\big)\big)' &= -h\big(\sigma_2(t)\big)\sigma_2'(t) + g_0(t) + e(t)\\ &\leqslant -h\big(\sigma_2(t)\big)\sigma_2'(t) + g\big(t, \sigma_2(t)\big) + e(t) = f\big(t, \sigma_2(t), \sigma_2'(t)\big)\\ &\qquad \text{for a.e. } t \in [0, T].\end{aligned}$$

Step 3. Finally, similarly as in Remark 5.11, we show that f satisfies (5.26) with $\omega(s) \equiv 1 + \max\{|h(x)|\colon x \in [r, \|\sigma_2\|_\infty]\}$, $\psi(t) = |e(t)| + \sup\{|g(t, x)|\colon x \in [r, \|\sigma_2\|_\infty]\}$ and $\varepsilon_1 = \varepsilon_2 = 1$. Therefore, by Theorem 5.10, problem (5.39) has a positive solution u such that $u \geqslant r$ on $[0, T]$. □

REMARK 5.16. If g does not depend on t, i.e. $g(t, x) \equiv g(x)$ for a.e. $t \in [0, T]$ and all $x \in (0, \infty)$, then the condition (5.42) is satisfied if $\liminf_{x\to 0+}(g(x) + \|e\|_\infty) < 0$ which is true, e.g., if $\liminf_{x\to 0+} g(x) = -\infty$ and $\sup\operatorname{ess}\{e(t)\colon t \in [0, T]\} < \infty$. Similarly, the conditions (5.43) and (5.44) are in such a case satisfied if $\liminf_{x\to\infty}(g(x) + \overline{e}) > 0$. In particular, Theorem 5.15 applies to problem (5.39) if $\phi = \phi_p$, $p > 1$, $\sup\operatorname{ess}\{e(t)\colon t \in [0, T]\} < \infty$, $\overline{e} > 0$, $g(t, x) = -\beta(t)x^{-\lambda}$, where $\beta \in L_1[0, T]$, $\beta \geqslant \varepsilon > 0$ a.e. on $[0, T]$ and $\lambda \geqslant 1$. Notice that the condition (5.41) is satisfied, e.g., by $\phi(y) = (|y|y + y)\ln(1 + \frac{1}{|y|})$ or $\phi(y) = y(\exp(y^2) - 1)$.

5.4. *Repulsive singular forces*

In this section we study the singular problem (5.1), (5.2) with f having a repulsive singularity at $x = 0$. Recall (see Remark 5.1) that this means that the relation

$$\limsup_{x\to 0+} f(t, x, y) = \infty \quad \text{for a.e. } t \in [0, T] \text{ and some } y \in \mathbb{R}$$

is true. In general, for the case of a repulsive singularity, the existence of a pair of associated lower and upper functions having opposite order is typical. This causes that such a case is more difficult and more interesting than that of an attractive singularity. The next assertion deals with equation (5.28) and is a direct corollary of Theorem 5.12.

THEOREM 5.17. *Assume* $g \in C(0, \infty)$, $p \in Car([0, T] \times \mathbb{R}^2)$, (5.29) *and* (5.30) *with some* $m \in L_1[0, T]$. *Furthermore, let there be* $r > 0$, $A > r$, $B \geqslant A$ *and* $b \in L_1[0, T]$ *such that* $\overline{b} \leqslant 0$, (5.36),

$$g(r) + p(t, r, 0) \geqslant 0 \quad \textit{for a.e. } t \in [0, T], \tag{5.48}$$

and

$$g(x) + p(t, x, y) \leqslant b(t) \quad \text{for a.e. } t \in [0, T] \text{ and all } x \in [A, B] \text{ and } |y| \leqslant \phi^{-1}\big(\|b\|_1\big) \tag{5.49}$$

hold. Then problem (5.28), (5.2) *has a positive solution u such that $u(t_u) \in [r, B]$ for some $t_u \in [0, T]$.*

PROOF. By (5.48), $\sigma_2(t) \equiv r$ is an upper function of (5.28), (5.2). Furthermore, let x_0 be a solution of

$$\big(\phi(x')\big)' = b(t), \quad u(0) = u(T) = 0.$$

Using (5.49) and having in mind that $\overline{b} \leqslant 0$, we can show by a reasoning analogous to that applied in the proof of Theorem 5.14 to construct an upper function that the function $\sigma_1 = A + T\phi^{-1}(\|b\|_1) + x_0$ is a lower function of (5.28), (5.2). Using Theorem 5.12 we complete the proof. □

In particular, when restricted to the Duffing equation with the ϕ-Laplacian

$$\big(\phi(u')\big)' = g(u) + e(t), \tag{5.50}$$

Theorem 5.17 has the following corollary.

COROLLARY 5.18. *Let $e \in L_1[0, T]$ with* $\operatorname{inf\,ess}\{e(t)\colon t \in [0, T]\} > -\infty$ *and let $g \in C(0, \infty)$ have a strong repulsive singularity* (5.29). *Further, let*

$$g_* := \inf\big\{g(x)\colon x \in (0, \infty)\big\} > -\infty \tag{5.51}$$

and let there be $A > 0$ such that

$$g(x) + \overline{e} \leqslant 0 \quad \text{for } x \in [A, B], \text{ where } B - A \geqslant 2T\phi^{-1}\big(\|e - \overline{e}\|_1\big).$$

Then problem (5.50), (5.2) *has a positive solution u such that $u(t_u) \leqslant B$ for some $t_u \in [0, T]$.*

PROOF. By (5.29) we have (5.32) and consequently there is an $r \in (0, A)$ such that $g(r) + e(t) \geqslant 0$ for a.e. $t \in [0, T]$. The assertion follows from Theorem 5.17 if we put $b(t) = e(t) - \overline{e}$ and $m(t) = g_* + e(t)$ a.e. on $[0, T]$. □

Consider the periodic problem for the Liénard equation

$$\big(\phi_p(u')\big)' + h(u)u' = g(u) + e(t) \tag{5.52}$$

with the p-Laplacian $\phi_p(y) = |y|^{p-2}y$. To this end, the following easy corollary of the continuation type principle due to Manásevich and Mawhin turned out to be essential.

LEMMA 5.19 [101, Theorem 3.1] and [84, Lemma 3]. *Let* $p > 1$, $h \in C[0,\infty)$, $g \in C(0,\infty)$ *and* $e \in L_1[0,T]$. *Furthermore, assume there exist* $r > 0$, $R > r$ *and* $R' > 0$ *such that*

(i) *the inequalities* $r < v < R$ *on* $[0,T]$ *and* $\|v'\|_\infty < R'$ *hold for each* $\lambda \in (0,1]$ *and for each positive solution* v *of the problem*

$$\begin{cases} \big(\phi_p(v')\big)' = \lambda\big(-h(v)v' + g(v) + e(t)\big), \\ v(0) = v(T), \quad v'(0) = v'(T), \end{cases} \tag{5.53}$$

(ii) $(g(x) + \overline{e} = 0) \Longrightarrow r < x < R$,

(iii) $(g(r) + \overline{e})(g(R) + \overline{e}) < 0$.

Then problem (5.52), (5.2) *has at least one solution* u *such that* $r < u < R$ *on* $[0,T]$.

Under the assumptions ensuring that g is bounded below on $(0,\infty)$, the following result was delivered by Jebelean and Mawhin.

THEOREM 5.20 [84, Theorem 2]. *Let* $p > 1$, $h \in C[0,\infty)$, $e \in L_1[0,T]$ *and let* $g \in C(0,\infty)$ *have a strong repulsive singularity* (5.29). *Furthermore, assume*

$$\liminf_{x\to\infty} g(x) > -\infty \tag{5.54}$$

and

$$\liminf_{x\to 0+}\big[g(x) + \overline{e}\big] > 0 > \limsup_{x\to\infty}\big[g(x) + \overline{e}\big]. \tag{5.55}$$

Then problem (5.52), (5.2) *has a positive solution.*

PROOF. We will verify that the assumptions of Lemma 5.19 are satisfied.

Step 1. First, we will show that

$$\begin{cases} \text{there are } R_0 > 0 \text{ and } R_1 > R_0 \text{ such that} \\ v(t_v) \in (R_0, R_1) \text{ for some } t_v \in [0,T] \\ \text{holds for each } \lambda \in (0,1] \text{ and each positive solution } v \text{ of (5.53).} \end{cases} \tag{5.56}$$

To this aim, assume that $\lambda \in (0,1]$ and that v is a positive solution to (5.53). Integrating the differential equation in (5.53) over $[0,T]$ and having in mind the periodicity of v, we get:

$$\int_0^T \big(g\big(v(t)\big) + e(t)\big)\,\mathrm{d}t = 0. \tag{5.57}$$

By the first inequality in (5.55), there is an $R_0 > 0$ such that

$$g(x) + \overline{e} > 0 \quad \text{whenever } x \in (0, R_0). \tag{5.58}$$

If $g(v(t)) + \overline{e} > 0$ were valid on $[0, T]$, we would have

$$\int_0^T \bigl(g\bigl(v(t)\bigr) + e(t)\bigr)\,\mathrm{d}t = \int_0^T \bigl(g\bigl(v(t)\bigr) + \overline{e}\bigr)\,\mathrm{d}t > 0.$$

Since this contradicts (5.57), we see that $\max\{v(t)\colon t \in [0, T]\} > R_0$. Similarly, by the second inequality in (5.55), there is an $R_1 > R_0$ such that $g(x) + \overline{e} < 0$ for $x \geqslant R_1$ and $v(t_1) < R_1$ for some $t_1 \in [0, T]$. Therefore (5.56) is true.

Step 2. Now we show that

$$\begin{cases} \text{there is } R > 0 \text{ such that } v < R \text{ on } [0, T] \\ \text{for each } \lambda \in (0, 1] \text{ and each positive solution } v \text{ of (5.53).} \end{cases} \tag{5.59}$$

Notice that, due to (5.54) and (5.58), we have $g_* = \inf\{g(x)\colon x \in (0, \infty)\} > -\infty$. Thus, multiplying (5.53) by v and integrating over $[0, T]$, we get

$$\|v'\|_p^p \leqslant \int_0^T \bigl(|g_*| + |e(t)|\bigr) v(t)\,\mathrm{d}t.$$

Furthermore, for R_1 given as in Step 1 and $\frac{1}{p} + \frac{1}{q} = 1$, we deduce that

$$\|v'\|_p^p \leqslant \left(\int_0^T \bigl(|g_*| + |e(t)|\bigr)\,\mathrm{d}t\right)\bigl(R_1 + T^{1/q}\|v'\|_p\bigr).$$

The right-hand side being a linear function of $\|v'\|_p$, this is possible only if there is $C_1 > 0$, independent of v and λ and such that $\|v'\|_p < C_1$. Therefore

$$v(t) = v(t_1) + \int_{t_1}^t v'(s)\,\mathrm{d}s < R_1 + T^{1/q} C_1$$

for all $\lambda \in (0, 1]$ and all positive solutions v of (5.53), i.e. the assertion (5.59) is true with $R := R_1 + T^{1/q} C_1$.

Step 3. Next we show that

$$\begin{cases} \text{there is } R_2 > 0 \text{ such that } |v'| < \lambda^{\frac{1}{p-1}} R_2 \text{ on } [0, T] \\ \text{for each } \lambda \in (0, 1] \text{ and each positive solution } v \text{ of (5.53).} \end{cases} \tag{5.60}$$

Having in mind that v satisfies the periodic conditions, we can see that there is $t' \in [0, T]$ such that $v'(t') = 0$. Integrating the differential equation in (5.53) over $[t', t]$ and taking into account (5.59), we get

$$\left|v'(t)\right|^{p-1} \leqslant \lambda\left(\int_0^R \left|h(x)\right| \mathrm{d}x + \|e\|_1 + \left|\int_{t'}^t \left|g\big(v(s)\big)\right| \mathrm{d}s\right|\right)$$
$$\text{for } t \in [0, T]. \tag{5.61}$$

By (5.58), there is $b > 0$ such that $g(x) \geqslant -b$ for all $x \in (0, R]$. So, by (5.59), $g(v(t)) \geqslant -b$ on $[0, T]$ holds for each possible positive solution v of (5.53). Therefore, $|g(v(t))| \leqslant g(v(t)) + 2b$ for all $t \in [0, T]$ wherefrom, using (5.57), we deduce

$$\left|\int_{t'}^t \left|g\big(v(s)\big)\right| \mathrm{d}s\right| \leqslant 2bT + \|e\|_1,$$

which inserted into (5.61) yields (5.60) with

$$R_2^{p-1} = \int_0^R \left|h(x)\right| \mathrm{d}x + 2\big(b + \|e\|_1\big) > 0.$$

Step 4. We show that

$$\begin{cases} \text{there is } r \in (0, R_0) \text{ such that } v > r \text{ on } [0, T] \\ \text{for each } \lambda \in (0, 1] \text{ and each positive solution } v \text{ of (5.53).} \end{cases} \tag{5.62}$$

Put $h_R := \max\{|h(x)|\colon\ x \in [0, R]\}$, $R^* = \frac{R_2^p}{q} + R_2(h_R R_2 T + \|e\|_1)$ and

$$K^* = R^* + \int_{R_0}^R \left|g(x)\right| \mathrm{d}x. \tag{5.63}$$

By (5.29), there is $r > 0$ such that

$$\int_r^{R_0} g(x)\,\mathrm{d}x > K^*. \tag{5.64}$$

Put $w(t) = \phi_p(v'(t))$ for $t \in [0, T]$. Then $|w(t)|^q = |v'(t)|^p$ for $t \in [0, T]$,

$$v'(t) = \left|w(t)\right|^{q-2} w(t) \quad \text{for } t \in [0, T] \tag{5.65}$$

and

$$w'(t) = \lambda\big(-h\big(v(t)\big)v'(t) + g\big(v(t)\big) + e(t)\big) \quad \text{for a.e. } t \in [0, T]. \tag{5.66}$$

Multiplying (5.65) by $w'(t)$ and (5.66) by $v'(t)$ and subtracting we get

$$\frac{1}{q}\big(|v'(t)|^p\big)' = \lambda\big(-h\big(v(t)\big)\big(v'(t)\big)^2 + g\big(v(t)\big)v'(t) + e(t)v'(t)\big)$$
$$\text{for a.e. } t \in [0,T]. \tag{5.67}$$

Now, suppose that $\min\{v(t)\colon\ t \in [0,T]\} < r$. Let us extend v to a T-periodic function on $\mathbb{R}$ and choose $t' \in [0,T]$ and $t^* \in (t', t'+T]$ so that $v(t') = r$, $v(t^*) = \max\{v(t)\colon\ t \in [0,T]\}$ and $v(t) \geqslant r$ on $[t', t^*]$. Integrating (5.67) from t' to t^*, we get

$$\lambda \int_r^{v(t^*)} g(x)\,\mathrm{d}x$$
$$= -\frac{1}{q}\big|v'(t')\big|^p - \lambda\bigg(\int_{t'}^{t^*} h\big(v(t)\big)\big(v'(t)\big)^2\,\mathrm{d}t + \int_{t'}^{t^*} e(t)v'(t)\,\mathrm{d}t\bigg).$$

Consequently, by (5.59) and (5.60),

$$\int_r^{v(t^*)} g(x)\,\mathrm{d}x \leqslant \frac{\lambda^{\frac{p}{p-1}}}{q} R_2^p + h_R \lambda^{\frac{2}{p-1}} R_2^2 T + \|e\|_1 \lambda^{\frac{1}{p-1}} R_2$$
$$\leqslant \frac{R_2^p}{q} + R_2\big(h_R R_2 T + \|e\|_1\big) = R^*,$$

which, by virtue of (5.63), finally gives

$$\int_r^{R_0} g(x)\,\mathrm{d}x = \int_r^{v(t^*)} g(x)\,\mathrm{d}x - \int_{R_0}^{v(t^*)} g(x)\,\mathrm{d}x \leqslant R^* + \int_{R_0}^{R} \big|g(x)\big|\,\mathrm{d}x = K^*.$$

This being contradictory to (5.64) implies that $v > r$ holds on $[0,T]$, i.e. (5.62) is true.

Step 5. To summarize, there are r, R and R' such that the assumption (i) from Lemma 5.19 is satisfied. Furthermore, since by Step 1 we have

$$g(x) + \overline{e} > 0 \quad \text{if } 0 < x < R_0 \quad \text{and} \quad g(x) + \overline{e} < 0 \quad \text{if } x > R_1$$

and $0 < r < R_0 < R_1 < R$, it is easy to see that also the assumptions (ii) and (iii) of Lemma 5.19 are satisfied. □

Assume that the dissipativity condition

$$h(x) \geqslant h_* > 0 \quad \text{or} \quad h(x) \leqslant -h_* < 0 \quad \text{for all } x \in [0,\infty) \tag{5.68}$$

is fulfilled instead of (5.54) and $e \in L_2[0,T]$. Then the existence of a positive solution to problem (5.52) is ensured by Jebelean and Mawhin.

THEOREM 5.21 [85, Theorem 3]. *Let* $p > 1$, $h \in C[0,\infty)$, $e \in L_2[0,T]$ *and let* $g \in C(0,\infty)$ *have a strong repulsive singularity* (5.29). *Furthermore, assume* (5.55) *and* (5.68). *Then problem* (5.52), (5.2) *has a positive solution.*

PROOF. The proof is analogous to that of Theorem 5.20, just the estimate (5.59) is, thanks to (5.68), obtained more easily. Indeed: let $\lambda \in (0,1]$ and let v be a positive solution of (5.53). Let R_0, R_1 and t_1 be found as in Step 1 of the proof of Theorem 5.20, i.e. R_0 is such that (5.58) is true, $R_1 > R_0$, $g(x) + \overline{e} < 0$ for $x \geqslant R_1$ and $v(t_1) < R_1$. Integrating equality (5.67) over $[0,T]$, we get $h_* \|v'\|_2 \leqslant \|e\|_2$ and, consequently,

$$v(t) = v(t_1) + \int_{t_1}^{t} v'(s)\,\mathrm{d}s < R_1 + \sqrt{T}\frac{\|e\|_2}{h_*} + 1 \quad \text{for all } t \in [0,T].$$

Thus, (5.59) is true with $R = R_1 + \sqrt{T}\frac{\|e\|_2}{h_*} + 1$. Now, we can repeat Steps 3–5 of the proof of Theorem 5.20. □

Lemma 5.19 enables us to prove also the following result concerning both the non-dissipative case and the case where g need not be bounded below on $(0,\infty)$. Recall that the symbol π_p is defined for $p > 1$ by

$$\pi_p = \frac{2\pi(p-1)^{1/p}}{p\sin(\frac{\pi}{p})}$$

and $(\frac{\pi_p}{T})^p$ is the first eigenvalue of the Dirichlet problem

$$\bigl(\phi_p(u')\bigr)' + \lambda\phi_p(u) = 0, \quad u(0) = u(T) = 0$$

(see [57]).

THEOREM 5.22. *Let* $p > 1$, $h \in C[0,\infty)$, $e \in L_1[0,T]$ *and let* $g \in C(0,\infty)$ *have a strong repulsive singularity* (5.29). *Furthermore, assume* (5.55) *and*

$$\begin{cases} \text{there exist } a,\ 0 \leqslant a < \bigl(\frac{\pi_p}{T}\bigr)^p, \text{ and } \gamma \geqslant 0 \text{ such that} \\ g(x)x \geqslant -\bigl(ax^p + \gamma\bigr) \quad \text{for all } x > 0. \end{cases} \tag{5.69}$$

Then problem (5.52), (5.2) *has a positive solution.*

PROOF. Similarly to the proof of Theorem 5.21, it suffices to verify (5.59). Assume that $\lambda \in (0,1]$, v is a positive solution to (5.53) and let R_1 and t_1 have the same meaning as in Step 1 of the proof of Theorem 5.20. Multiplying (5.53) by $v(t)$ and integrating over $[0,T]$, we get

$$\|v'\|_p^p \leqslant a\|v\|_p^p + \|e\|_1\|v\|_\infty + \gamma T. \tag{5.70}$$

Since $v(t_1) \leqslant R_1$, we have

$$0 < v(t) < R_1 + T^{1/q}\|v'\|_p \quad \text{for } t \in [0,T], \tag{5.71}$$

where $\frac{1}{p}+\frac{1}{q}=1$. Now put

$$y(t)=\begin{cases} v(t+t_1)-v(t_1) & \text{if } 0\leqslant t\leqslant T-t_1, \\ v(t+t_1-T)-v(t_1) & \text{if } T-t_1\leqslant t\leqslant T. \end{cases}$$

We have $y\in C^1[0,T]$, $y(0)=y(T)=0$ and $\|y+v(t_1)\|_p^p=\|v\|_p^p$. Therefore, by the generalized Poincaré-Wirtinger inequality (see, e.g., [153, Lemma 3]),

$$\|y\|_p\leqslant\frac{T}{\pi_p}\|y'\|_p=\frac{T}{\pi_p}\|v'\|_p.$$

Hence, for an arbitrary $\varepsilon>0$, there is a $C_1>0$ such that

$$\|v\|_p^p\leqslant\bigl(\|y\|_p+v(t_1)T^{1/p}\bigr)^p\leqslant(1+\varepsilon)\left(\frac{T}{\pi_p}\right)^p\|v'\|_p^p+C_1.$$

Inserting this into (5.70), choosing $\varepsilon\in(0,\frac{1}{a}(\frac{\pi_p}{T})^p-1)$ and having in mind (5.71), we deduce that

$$\alpha\|v'\|_p^p\leqslant T^{1/q}\|e\|_1\|v'\|_p+C_2$$

for some $C_2>0$, where $\alpha=(1-a(1+\varepsilon)(\frac{T}{\pi_p})^p)>0$. However, this is possible only if there is $R_p\in(0,\infty)$, independent of λ and v, such that $\|v'\|_p<R_p$. Therefore $0<v(t)<R_1+T^{1/q}R_p+1$ on $[0,T]$ for all $\lambda\in(0,1]$ and all positive solutions v of (5.53), i.e. the assertion (5.59) is true with $R=R_1+T^{1/q}R_p+1$.

By virtue of (5.55), we can choose $b>0$ so that $\inf\{g(x)\colon x\in(0,R]\}\geqslant -b$. Thus, we can continue by Steps 3–5 of the proof of Theorem 5.20 to verify that the assumptions of Lemma 5.19 are satisfied. □

REMARK 5.23. Theorem 5.22 is a slightly modified scalar version of the result by Liu [98, Theorem 1].

In the undamped case of the Duffing type equation

$$\bigl(\phi_p(u')\bigr)'=g(u)+e(t), \tag{5.72}$$

condition (5.69) can be replaced by a related asymptotic condition. It is shown in the next theorem which has been proved for $p=2$ by Rachůnková and Tvrdý in [125, Theorem 3.1].

THEOREM 5.24. *Let* $p>1$, $\frac{1}{p}+\frac{1}{q}=1$, $g\in C(0,\infty)$ *and* $e\in L_q[0,T]$. *Furthermore, assume* (5.29),

$$\liminf_{x\to 0+} g(x)>-\infty, \tag{5.73}$$

$$\liminf_{x\to\infty} \frac{g(x)}{x^{p-1}} > -\left(\frac{\pi_p}{T}\right)^p. \tag{5.74}$$

Further, assume that there exist $r > 0$ *and* $A > r$ *such that the conditions*

$$g(r) + e(t) \geqslant 0 \quad \textit{for a.e. } t \in [0, T] \tag{5.75}$$

and

$$g(x) + \overline{e} \leqslant 0 \quad \textit{for } x \in [A, B], \tag{5.76}$$

where

$$B - A \leqslant 2T \|e - \overline{e}\|_1^{q-1}, \tag{5.77}$$

are satisfied. Then problem (5.72), (5.2) *has a positive solution* u *such that* $u(t_u) \in [r, B]$ *for some* $t_u \in [0, T]$.

PROOF. *Step* 1. *Lower and upper functions.*

By (5.75), $\sigma_2 \equiv r$ is an upper function of (5.72), (5.2). Let v be a solution of the quasilinear Dirichlet problem (5.9) with $b(t) = e(t) - \overline{e}$ a.e. on $[0, T]$ and $d = 0$ and let $\sigma_1 = A + T\phi_p^{-1}(\|e - \overline{e}\|_1) + v$ on $[0, T]$. Let us recall that $\phi_p^{-1} = \phi_q$. Hence $\phi_p^{-1}(\|e - \overline{e}\|_1) = \|e - \overline{e}\|_1^{q-1}$. Having in mind assumption (5.76), we can see, similarly to the proof of Theorem 5.17 (see also the proof of Theorem 5.14), that σ_1 is a lower function of (5.72), (5.2) and $\sigma_1(t) \in [A, B]$ for $t \in [0, T]$.

Step 2. *Construction of an auxiliary problem having a right-hand side bounded below.*

By (5.29) we have (5.32) and hence there is a sequence $\{\varepsilon_n\} \subset (0, r)$ such that

$$g(\varepsilon_n) > 0 \quad \text{for } n \in \mathbb{N}, \qquad \lim_{n\to\infty} \varepsilon_n = 0 \quad \text{and} \quad \lim_{n\to\infty} g(\varepsilon_n) = \infty. \tag{5.78}$$

For $n \in \mathbb{N}$ and $M \in \mathbb{R}$, $M > r$, define

$$g_{n,M}(x) = \begin{cases} 0 & \text{if } x < 0, \\ \frac{g(\varepsilon_n)}{\varepsilon_n^{p-1}} x^{p-1} & \text{if } x \in [0, \varepsilon_n], \\ g(x) & \text{if } x \in [\varepsilon_n, M], \\ g(M) & \text{if } x > M. \end{cases} \tag{5.79}$$

By (5.74), there are $\eta \in (0, (\frac{\pi_p}{T})^p)$ and $x_0 > 1$ such that

$$\frac{g(x)}{x^{p-1}} \geqslant -\left(\left(\frac{\pi_p}{T}\right)^p - \eta\right) \quad \text{for all } x \geqslant x_0.$$

Put

$$p(x) = \begin{cases} 0 & \text{if } x \leqslant 0, \\ \frac{g(x_0)}{x_0^{p-1}} x^{p-1} & \text{if } x \in (0, x_0), \\ g(x) & \text{if } x \geqslant x_0 \end{cases}$$

and $q_{n,M}(x) = g_{n,M}(x) - p(x)$ for $x \in \mathbb{R}$. By virtue of (5.73), there is $\gamma \geqslant 0$ such that $q_{n,M}(x) \geqslant -\gamma$ for all $x \in \mathbb{R}$, $n \in \mathbb{N}$ and $M > r$. Consequently, each function $\widetilde{g}(x) = g_{n,M}(x)$, $n \in \mathbb{N}$, $M > r$, satisfies the estimate

$$\widetilde{g}(x)x \geqslant -\left(\left(\frac{\pi_p}{T}\right)^p - \eta\right)|x|^p - \gamma|x| \quad \text{for all } x \in \mathbb{R}. \tag{5.80}$$

Step 3. *A priori estimates.*

Now, we will give uniform a priori estimates for solutions of periodic problems associated to the equations

$$\big(\phi_p(u')\big)' = \widetilde{g}(u) + e(t), \tag{5.81}$$

where $\widetilde{g}$ may be an arbitrary function satisfying the estimate (5.80). To this aim, we will prove the following assertion.

CLAIM. *Let* $\gamma \geqslant 0$ *and* $\eta \in (0, (\frac{\pi_p}{T})^p)$. *Then for any* $\delta > 0$, *there are* $R \geqslant \delta$ *and* $R' > 0$ *such that the estimates*

$$u \leqslant R \quad \textit{on } [0, T] \quad \textit{and} \quad \|u'\|_p \leqslant R' \tag{5.82}$$

hold whenever

$$\begin{cases} \widetilde{g} \in C(0,\infty) \textit{ fulfills } (5.80) \textit{ and } u \textit{ is a solution of } (5.81), (5.2) \\ \textit{such that } \min\{u(t)\colon\ t \in [0, T]\} \leqslant \delta. \end{cases} \tag{5.83}$$

PROOF OF CLAIM. We will follow ideas from the proof of [130, Lemma 2.4]. Suppose that for each $k \in \mathbb{N}$ there are $g_k \in C(0, \infty)$ and a solution u_k of

$$\big(\phi_p(u')\big)' = g_k(u) + e(t), \quad u(0) = u(T), \quad u'(0) = u'(T) \tag{5.84}$$

such that

$$g_k(x)x \geqslant -\left(\left(\frac{\pi_p}{T}\right)^p - \eta\right)|x|^p - \gamma|x| \tag{5.85}$$

and

$$u_k(t_k) = \delta \quad \text{for some } t_k \in [0, T] \quad \text{and} \quad \max\{u_k(t)\colon\ t \in [0, T]\} > k. \tag{5.86}$$

In particular, we have

$$\lim_{k\to\infty} \max\bigl\{u_k(t)\colon\ t\in[0,T]\bigr\}=\infty. \tag{5.87}$$

Let us extend u_k and e to functions T-periodic on $\mathbb{R}$. We have

$$\bigl(\phi_p\bigl(u_k'(t)\bigr)\bigr)'=g_k\bigl(u_k(t)\bigr)+e(t)\quad \text{for a.e. } t\in\mathbb{R}.$$

Multiplying this equality by u_k, integrating from t_k to t_k+T and making use of (5.85), we obtain

$$\begin{aligned}\|u_k'\|_p^p&=-\int_{t_k}^{t_k+T} g_k\bigl(u_k(s)\bigr)u_k(s)\,\mathrm{d}s-\int_{t_k}^{t_k+T} e(s)u_k(s)\,\mathrm{d}s\\ &\leqslant\left(\left(\frac{\pi_p}{T}\right)^p-\eta\right)\|u_k\|_p^p+\gamma T^{1/q}\|u_k\|_p+\|e\|_q\|u_k\|_p.\end{aligned}$$

Let us set $v_k=u_k-\delta$. By (5.87) we have $\lim_{k\to\infty}\|v_k\|_\infty=\infty$. Therefore, applying the Hölder inequality we can conclude that

$$\lim_{k\to\infty}\|v_k'\|_p=\infty. \tag{5.88}$$

Furthermore, it is easy to verify that

$$\|v_k'\|_p^p\leqslant\left(\left(\frac{\pi_p}{T}\right)^p-\epsilon\right)\|v_k\|_p^p+a\|v_k\|_p+b \tag{5.89}$$

holds with some $\epsilon\in(0,(\frac{\pi_p}{T})^p)$ and a, $b\geqslant 0$ not depending on v_k. This, together with (5.88), gives

$$\lim_{k\to\infty}\|v_k\|_p=\infty. \tag{5.90}$$

Moreover, as $v_k(t_k)=v_k(t_k+T)=0$, we can apply the generalized Poincaré–Wirtinger inequality (see, e.g., [153, Lemma 3]) to get

$$\|v_k\|_p^p\leqslant\left(\frac{T}{\pi_p}\right)^p\|v_k'\|_p^p\quad \text{for each } k\in\mathbb{N}.$$

Hence the inequality (5.89) can be rewritten as

$$\left(\frac{\pi_p}{T}\right)^p\leqslant\frac{\|v_k'\|_p^p}{\|v_k\|_p^p}\leqslant\left(\frac{\pi_p}{T}\right)^p-\epsilon+\frac{a}{\|v_k\|_p^{p-1}}+\frac{b}{\|v_k\|_p^p}, \tag{5.91}$$

which, in view of (5.90), leads to a contradiction

$$\left(\frac{\pi_p}{T}\right)^p \leqslant \left(\frac{\pi_p}{T}\right)^p - \epsilon.$$

As a consequence, we can conclude that the sequences $\{\|v_k\|_\infty\}$ and $\{\|v_k\|_p\}$ are bounded. By (5.89), this implies that also the sequence $\{\|v_k'\|_p\}$ is bounded. In particular, there are $R \in [\delta, \infty)$ and $R' \in (0, \infty)$ such that $u \leqslant R$ and $\|u'\|_p \leqslant R'$ hold whenever (5.83) is true. This completes the proof of Claim.

Now, let $R > B$ and $R' > 0$ be constants given by Claim for $\delta = B$. Put

$$K = \int_A^R \left|g(x)\right| \mathrm{d}x + \|e\|_q R'.$$

It follows from (5.29) and (5.78) that we can choose $\varepsilon = \varepsilon_{n^*} \in \{\varepsilon_n\}$ such that

$$\int_\varepsilon^A g(x)\,\mathrm{d}x > K \quad \text{and} \quad g(\varepsilon) > 0. \tag{5.92}$$

By (5.73), there is $g_R \in \mathbb{R}$ such that

$$g(x) \geqslant g_R \quad \text{for } x \in (0, R]. \tag{5.93}$$

Define

$$\widetilde{g}(x) = g_{n^*,R}(x) \quad \text{for } x \in \mathbb{R} \tag{5.94}$$

and consider the regular periodic problem for the auxiliary equation

$$\left(\phi_p(u')\right)' = \widetilde{g}(u) + e(t). \tag{5.95}$$

Clearly, σ_1 and σ_2 are lower and upper functions of (5.95), (5.2) and $\widetilde{g}(x) + e(t) \geqslant g_R + e(t)$ for a.e. $t \in [0, T]$ and all $x \in \mathbb{R}$. Thus, by Lemma 5.9, problem (5.95), (5.2) possesses a solution u such that $u(t_u) \in [r, B]$ for some $t_u \in [0, T]$. In particular, $\min\{u(t)\colon t \in [0, T]\} \leqslant B$. Furthermore, by Step 2 it is easy to see that $\widetilde{g}$ satisfies the estimate (5.80). Thus, by Claim and by the definitions of R and R', the estimates (5.82) are true.

Step 4. *Existence of a solution to* (5.72), (5.2).

It remains to show that $u \geqslant \varepsilon$ on $[0, T]$. Let $t_0, t_1 \in [0, T]$ be such that

$$u(t_0) = \min\bigl\{u(t)\colon t \in [0, T]\bigr\} \quad \text{and} \quad u(t_1) = \max\bigl\{u(t)\colon t \in [0, T]\bigr\}.$$

Due to the periodicity of u, we have $u'(t_0) = u'(t_1) = 0$. Multiplying the equality $(\phi_p(u'(t)))' = \widetilde{g}(u(t)) + e(t)$ by $u'(t)$ and integrating, we get

$$\int_{u(t_0)}^{u(t_1)} \widetilde{g}(x)\,\mathrm{d}x \leqslant \|e\|_q R'.$$

Therefore,

$$\int_{u(t_0)}^{A} \widetilde{g}(x)\,\mathrm{d}x \leqslant \int_{A}^{R} \left|g(x)\right|\mathrm{d}x + \|e\|_q R' = K.$$

Let $u(t_0) < \varepsilon$. Then, by (5.92),

$$\begin{aligned}\int_{u(t_0)}^{A} \widetilde{g}(x)\,\mathrm{d}x &= \int_{u(t_0)}^{\varepsilon} \widetilde{g}(x)\,\mathrm{d}x + \int_{\varepsilon}^{A} g(x)\,\mathrm{d}x \\ &= g(\varepsilon)\bigl(\varepsilon - u(t_0)\bigr) + \int_{\varepsilon}^{A} g(x)\,\mathrm{d}x \\ &> \int_{\varepsilon}^{A} g(x)\,\mathrm{d}x > K,\end{aligned}$$

a contradiction. So $u(t) \geqslant \varepsilon$ on $[0, T]$, which together with (5.79), (5.82) and (5.94) yields that u is a solution of (5.72), (5.2). □

EXAMPLES. (i) Let $p > 1$, $h \in C[0, \infty)$, $\beta > 0$, $\alpha \geqslant 1$, $e \in L_1[0, T]$. Then, by Theorem 5.20, the problem

$$\bigl(|u'|^{p-2}u'\bigr)' + h(u)u' = \frac{\beta}{u^{\alpha}} + e(t), \quad u(0) = u(T), \quad u'(0) = u'(T) \tag{5.96}$$

has a positive solution if $\overline{e} < 0$. Integrating both sides of the differential equation in (5.96) over $[0, T]$ and taking into account the positivity of $g(x) = \beta x^{-\alpha}$ on $(0, \infty)$, we can see that the condition $\overline{e} < 0$ is also necessary for the existence of a positive solution to (5.96).

(ii) Let $p > 1$, $c \neq 0$, $a > 1$, $\beta > 0$, $\alpha \geqslant 1$. Then, by Theorem 5.21, the problem

$$\bigl(|u'|^{p-2}u'\bigr)' + cu' = \frac{\beta}{u^{\alpha}} - a\exp(u) + e(t), \quad u(0) = u(T), \quad u'(0) = u'(T)$$

has a solution for each $e \in L_2[0, T]$.

(iii) Let $p > 1$, $h \in C[0, \infty)$, $0 < a < (\frac{\pi_p}{T})^p$, $\beta > 0$ and $\alpha \geqslant 1$. Then, by Theorem 5.22, the problem

$$\bigl(|u'|^{p-2}u'\bigr)' + h(u)u' = -au^{p-1} + \frac{\beta}{u^{\alpha}} + e(t), \quad u(0) = u(T), \quad u'(0) = u'(T)$$

has a positive solution for each $e \in L_1[0, T]$.

For the classical case $p = 2$, the following result due to Omari and Ye is known. Its proof combines the lower and upper functions method, the degree theory and connectedness arguments for some properly chosen truncated equations and a posteriori estimates.

THEOREM 5.25 [108, Theorem 1.2]. *Assume* $e \in L_\infty[0,T]$, (5.29), $\lim_{x\to 0+} g(x) = \infty$ *and*

$$\liminf_{x\to\infty} \frac{g(x)}{x} \geqslant -\left(\frac{\pi}{T}\right)^2 \quad \text{and} \quad \liminf_{x\to\infty} \frac{2G(x)}{x^2} > -\left(\frac{\pi}{T}\right)^2,$$

where

$$G(x) = \int_x^1 g(\xi)\,\mathrm{d}\xi \quad \text{for } x \in (0,\infty).$$

Then the problem $u'' + h(u)u' = g(u) + e(t)$, (5.2) *has a solution if and only if it possesses a lower function* $\sigma_1 \in AC^1[0,T]$.

Hitherto we have assumed the strong singularity condition (5.29). The next existence principle enables us to treat also problems with weak repulsive singularities. We shall restrict ourselves to the case that $\phi(y) \equiv y$ and f does not depend on u', i.e. we consider the equation

$$u'' = f(t,u), \tag{5.97}$$

where $f : [0,T] \times (0,\infty) \to \mathbb{R}$.

THEOREM 5.26. *Let* $f \in Car([0,T] \times (0,\infty))$, $r > 0$, $A \geqslant r$ *and let* $\mu \in L_1[0,T]$ *and* $\beta \in L_1[0,T]$ *be such that* $\mu(t) \geqslant 0$ *a.e. on* $[0,T]$, $\overline{\mu} > 0$,

$$\overline{\beta} \leqslant 0 \quad \text{and} \quad f(t,x) \leqslant \beta(t) \quad \text{for a.e. } t \in [0,T] \text{ and all } x \in [A,B] \tag{5.98}$$

and

$$f(t,x) \geqslant -\mu(t)(x-r) \quad \text{for a.e. } t \in [0,T] \text{ and all } x \in [r,B], \tag{5.99}$$

where

$$B - A \geqslant \frac{T^2}{2}\overline{m},$$

$$m(t) = \max\bigl\{\sup\bigl\{f(t,x)\colon\ x \in [r,A]\bigr\}, \beta(t), 0\bigr\} \quad \text{for a.e. } t \in [0,T]$$

and

$$\begin{cases} v \geqslant 0 \text{ on } [0,T] \text{ holds for each } v \in AC^1[0,T] \text{ such that} \\ v''(t) + \mu(t)v(t) \geqslant 0 \quad \text{for a.e. } t \in [0,T], \\ v(0) = v(T), \quad v'(0) = v'(T). \end{cases} \tag{5.100}$$

Then problem (5.97), (5.2) *has a positive solution* u *such that* $r \leqslant u \leqslant B$ *on* $[0,T]$.

PROOF. The proof follows the ideas of the proof of [129, Theorem 2.5]. First, assume that $\overline{\beta} < 0$.

Step 1. *Existence of a solution u to a certain auxiliary problem.*

Put

$$\widetilde{f}(t,x) = \begin{cases} f(t,r) - \mu(t)(x-r) & \text{if } x \leqslant r, \\ f(t,x) & \text{if } x \in [r,B], \\ f(t,B) & \text{if } x \geqslant B \end{cases} \tag{5.101}$$

and consider the problem

$$u'' = \widetilde{f}(t,u), \quad u(0) = u(T), \quad u'(0) = u'(T). \tag{5.102}$$

We have $\widetilde{f} \in Car([0,T] \times \mathbb{R})$. By (5.98), (5.99) and (5.101), the inequalities

$$\widetilde{f}(t,x) \leqslant \beta(t) \quad \text{if } x \geqslant A \tag{5.103}$$

and

$$\widetilde{f}(t,x) \geqslant -\mu(t)\min\{x-r, B-r\} \tag{5.104}$$

are valid for a.e. $t \in [0,T]$ and all $x \in \mathbb{R}$. In particular, $\widetilde{f}(t,x) \geqslant -\mu(t)(B-r)$. By (5.104), $\sigma_2 \equiv r$ is an upper function of (5.102). Further, let σ_0 be the solution of the Dirichlet problem $v'' = b$, $v(0) = v(T) = 0$, where $b(t) = \beta(t) - \overline{\beta}$ for a.e. $t \in [0,T]$, and let $\sigma_c(t) = c + \sigma_0(t)$ for $t \in [0,T]$ and $c \in \mathbb{R}$. Then $\sigma_c'' = b$ a.e. on $t \in [0,T]$ and $\sigma_c(0) = \sigma_c(T) = c$. Moreover, $\sigma_c'(T) - \sigma_c'(0) = T\overline{b} = 0$. Let us choose $c^* > 0$ so that $\sigma_1 := \sigma_{c^*} \geqslant A$ on $[0,T]$. Due to (5.103), where $\beta < b$ a.e. on $[0,T]$, we can see that σ_1 is a lower function of (5.102). Therefore, by Lemma 5.9, the regular problem (5.102) has a solution u such that $u(t_u) \geqslant r$ for some $t_u \in [0,T]$.

Step 2. *Lower estimate for u.*

We shall show that

$$u \geqslant r \quad \text{on } [0,T]. \tag{5.105}$$

Set $z = u - r$. By virtue of (5.99) and (5.101), we have

$$z''(t) + \mu(t)z(t) = u''(t) + \mu(t)z(t) = \widetilde{f}\big(t,u(t)\big) + \mu(t)\big(u(t) - r\big) \geqslant 0$$

for a.e. $t \in [0,T]$. By (5.100), it follows that $z(t) \geqslant 0$ on $[0,T]$, i.e. (5.105) is true.

Step 3. *Upper estimate for u.*

We shall show that

$$u \leqslant B \quad \text{on } [0,T]. \tag{5.106}$$

By the definition of m and by (5.101) and (5.103) we have

$$\widetilde{f}(t,x) \leqslant m(t) \quad \text{for a.e. } t \in [0,T] \text{ and all } x \geqslant r.$$

Hence, we can use Lemma 5.8 (see (5.20)) to get

$$\|u'\|_\infty \leqslant \|m\|_1 = T\,\overline{m}. \tag{5.107}$$

If $u \geqslant A$ were valid on $[0,T]$, then taking into account the periodicity of u' and (5.103), we would get

$$0 = \int_0^T \widetilde{f}\big(t,u(t)\big)\,\mathrm{d}t \leqslant \int_0^T \beta(t)\,\mathrm{d}t = T\overline{\beta} < 0,$$

a contradiction. Thus, there is $\tau \in [0,T]$ such that $u(\tau) < A$. Now, assume that $u(s) > A$ for some $s \in [0,T]$ and extend u to the function T-periodic on $\mathbb{R}$. There are s_1, s_2 and $s^* \in \mathbb{R}$ such that $s_1 < s^* < s_2$, $s_2 - s_1 < T$, $u(s_1) = u(s_2) = A$ and $u(s^*) = \max\{u(s)\colon s \in [0,T]\} > A$. In particular, due to (5.107),

$$2\big(u(s^*) - A\big) = \int_{s_1}^{s^*} u'(s)\,\mathrm{d}s + \int_{s_2}^{s^*} u'(s)\,\mathrm{d}s \leqslant T^2\overline{m},$$

wherefrom the estimate

$$u(t) - A \leqslant \frac{T^2}{2}\overline{m} \leqslant B - A \quad \text{on } [0,T]$$

follows. Consequently, (5.106) is true.

Step 4. *Conclusion: u is a solution to* (5.97)*,* (5.2).

The estimates (5.105) and (5.106) mean that $r \leqslant u \leqslant B$ holds on $[0,T]$. By (5.101), we conclude that u is a solution to (5.97), (5.2).

If $\overline{\beta} = 0$, we can approximate the solution to (5.1), (5.2) by solutions of the problems

$$u'' = \widetilde{f}_n(t,u), \quad u(0) = u(T), \quad u'(0) = u'(T),$$

where

$$\widetilde{f}_n(t,x) = \begin{cases} f(t,r) & \text{if } x \leqslant r, \\ f(t,x) & \text{if } x \in [r,A], \\ f(t,x) - \mu(t)\frac{1}{n}\frac{x-A}{x-A+1} & \text{if } x \in [A,B], \\ f(t,B) - \mu(t)\frac{1}{n}\frac{B-A}{B-A+1} & \text{if } x \geqslant B. \end{cases}$$

□

Recently, using the Krasnoselskii fixed point theorem, Torres proved for the Hill equation

$$u'' + \mu(t)u = g(t, u) \tag{5.108}$$

the existence result which is related to Theorem 5.26.

THEOREM 5.27 [144, Theorem 4.5]. *Let $\mu \in L_1[0, T]$ be such that the problem*

$$v'' + \mu(t)v = 0, \quad v(0) = v(T), \quad v'(0) = v'(T) \tag{5.109}$$

possesses the Green function $G(t, s)$ which is positive on $[0, T] \times [0, T]$. Moreover, assume that there is an $R > 0$ such that

$$g(t, x) \geqslant 0 \quad \textit{for all } x \in \left(0, \frac{M}{m}R\right] \quad \textit{and}$$

$$g(t, x) \leqslant \frac{1}{TM}x \quad \textit{for all } x \in \left[R, \frac{M}{m}R\right]$$

for a.e. $t \in [0, T]$, where

$$m = \min\bigl\{G(t, s)\colon\ t, s \in [0, T]\bigr\} \quad \textit{and} \quad M = \max\bigl\{G(t, s)\colon\ t, s \in [0, T]\bigr\}.$$

Then problem (5.108), (5.2) *has a positive solution.*

It is easy to check that the function $G(t, s) = \sin(\frac{\pi}{T}|t - s|)$, $t, s \in [0, T]$, is the Green function for $v'' + (\frac{\pi}{T})^2 v = 0$, $v(0) = v(T)$, $v'(0) = v'(T)$ and $G(t, s) \geqslant 0$ on $[0, T] \times [0, T]$. Hence, the statement (5.100) holds if $\mu(t) \equiv \mu_1 = (\frac{\pi}{T})^2$. Notice that $\mu_1 = (\frac{\pi}{T})^2$ is the first eigenvalue of the related Dirichlet problem and it is optimal in the sense that for $\mu(t) = \mu$ a.e. on $[0, T]$ and $\mu \in (\mu_1, 4\mu_1)$ the corresponding Green function of $v'' + \mu v = 0$, $v(0) = v(T)$, $v'(0) = v'(T)$ is not nonnegative on $[0, T] \times [0, T]$.

In particular, when restricted to the Duffing equation

$$u'' = g(u) + e(t), \tag{5.110}$$

Theorem 5.26 has the following corollary.

COROLLARY 5.28 [129, Corollary 3.7]. *Suppose that $g \in C(0, \infty)$, $e \in L_1[0, T]$,*

$$\overline{e} + \limsup_{x \to \infty} g(x) < 0$$

and there is $r > 0$ such that

$$e(t) + g(x) + \left(\frac{\pi}{T}\right)^2 x \geqslant \left(\frac{\pi}{T}\right)^2 r \quad \textit{for a.e. } t \in [0, T] \textit{ and all } x > r.$$

Then problem (5.110), (5.2) *has a positive solution* u *such that* $u \geqslant r$ *on* $[0, T]$.

More detailed information on the sign properties of the associated Green functions is provided by the next proposition which is due to Torres [144] (see also [154, Lemma 2.5]). Before formulating it, let us define the function $K : [0, \infty] \to (0, \infty)$ by

$$K(z) = \begin{cases} \dfrac{2\pi}{zT^{1+2/z}}\left(\dfrac{2}{2+z}\right)^{1-2/z}\left(\dfrac{\Gamma(\frac{1}{z})}{\Gamma(\frac{1}{2}+\frac{1}{z})}\right)^2 & \text{if } 1 \leqslant z < \infty, \\ \dfrac{4}{T} & \text{if } z = \infty. \end{cases} \tag{5.111}$$

Let us recall that for a given z, $1 \leqslant z \leqslant \infty$, $K(z)$ is the best Sobolev constant for the inequality $C\|u\|_z^2 \leqslant \|u'\|_2^2$, i.e.

$$K(z) = \inf\left\{ \frac{\|u'\|_2^2}{\|u\|_z^2} \colon u \in H_0^1[0, T] \setminus \{0\} \right\},$$

where $H_0^1[0, T] = \{u \in AC[0, T] \colon u' \in L_2[0, T],\ u(0) = u(T) = 0\}$.

PROPOSITION 5.29 [144, Corollary 2.3]. *Let* $1 \leqslant q \leqslant \infty$ *and let* $\mu \in L_q[0, T]$. *Then* (5.100) *is true provided*

$$\mu(t) \geqslant 0 \quad \textit{a.e. on } [0, T], \quad \overline{\mu} > 0 \quad \textit{and} \quad \|\mu\|_q \leqslant K(2q^*), \tag{5.112}$$

where

$$\begin{cases} \frac{1}{q} + \frac{1}{q^*} = 1 & \textit{if } 1 < q < \infty, \\ q^* = \infty & \textit{if } q = 1, \\ q^* = 1 & \textit{if } q = \infty \end{cases} \tag{5.113}$$

and the function K *is defined by* (5.111).

Moreover, if $\|\mu\|_q < K(2q^*)$, *then problem* (5.109) *has the Green function which is positive on* $[0, T] \times [0, T]$.

Notice that if $\mu(t) \equiv \mu \in (0, \infty)$ on $[0, T]$, then we can take $q = \infty$, $q^* = 1$ and so we get $K(2q^*) = K(2) = (\frac{\pi}{T})^2$, which confirms the above mentioned fact that in such a case (5.100) is satisfied if $\mu \in (0, (\frac{\pi}{T})^2]$.

EXAMPLE. Consider the Brillouin beam focusing equation

$$u'' + a(1 + \cos t)u = \frac{1}{u} \tag{5.114}$$

on the interval $[0, 2\pi]$, where $a > 0$ is a parameter. (See [34] for a description of the model.) The problem of existence of a positive 2π-periodic solution to (5.114) has been considered by several authors (see, e.g., [55,144,145,152,154]). Put $A = \frac{1}{\sqrt{a}}$. Then for all $x \geqslant A$ and $t \in [0, 2\pi]$ we have

$$f(t, x) := \frac{1}{x} - a(1 + \cos t)x \leqslant \beta(t) := \frac{1}{A} - a(1 + \cos t)A$$

and $\overline{\beta} = \frac{1}{A} - aA = 0$. So, the assumption (5.98) is satisfied with $B = \infty$ and $T = 2\pi$. Furthermore, for $r \in (0, r_0]$ and a.e. $t \in [0, T]$ define

$$m_r(t) := \max\bigl\{\sup\bigl\{f(t, x)\colon\ x \in [r, A]\bigr\}, \beta(t), 0\bigr\}.$$

Let $r_0 = \frac{1}{\sqrt{2a}}$. Then $\frac{1}{r} - a(1 + \cos t)r \geqslant \frac{1}{r_0} - 2ar_0 = 0$ holds for $r \in (0, r_0]$ and $t \in [0, T]$. Consequently,

$$m_r(t) = \frac{1}{r} - a(1 + \cos t)r \quad \text{for a.e. } t \in [0, T] \text{ and all } r \in (0, r_0].$$

For a given $r \in (0, r_0]$, put $B_r = A + \pi\overline{m_r}$. Now, it is easy to check that it is possible to find $r \in (0, r_0]$ such that the assumption (5.99) is satisfied with $B = B_r$ and $\mu(t) = a(1 + \cos t)$ whenever $a < \frac{1}{2\pi} \approx 0.15915$. Finally, notice that by virtue of Proposition 5.29, the assumption (5.100) is satisfied if

$$a \leqslant K_{\max} := \max\left\{\frac{K(2q^*)}{\|1 + \cos t\|_q}\colon\ 1 \leqslant q \leqslant \infty\right\} \approx 0.16488.$$

(The maximum is attained at $q \approx 2.1941$, see [145, Corollary 4.8].) By Theorem 5.26, we can conclude that the equation (5.114) has a positive 2π-periodic solution for $a < \frac{1}{2\pi}$. To compare, notice that for $q = \infty$ and $T = 2\pi$, we get $q^* = 1$ and

$$\frac{K(2q^*)}{\|1 + \cos t\|_\infty} = \frac{1}{8} = 0.125.$$

Finally, let us note that using more sophisticated and involved techniques, Zhang proved (see [154, Theorem 4.5]) that for $a < K_{\max}$, $b > 0$, $\lambda \geqslant 1$, $e \in C[0, T]$ and $h \in C[0, \infty)$ the problem

$$u'' + h(u)u' + a(1 + \cos t)u = \frac{b}{u^\lambda} + e(t), \quad u(0) = u(2\pi), \quad u'(0) = u'(2\pi)$$

has a positive solution.

The hitherto mentioned conditions for the existence of a positive solution of problem (5.97), (5.2) concern the case when $f(t, x)$ asymptotically behaves like $-kx$ with $k \leqslant \mu_1$, $\mu_1 = (\frac{\pi}{T})^2$ being the first eigenvalue of the related Dirichlet problem. The next theorem deals with the case corresponding to k lying between two adjacent higher eigenvalues.

Let us denote by $\{\mu_k\}_{k=1}^{\infty}$ the sequence of eigenvalues of the related linear Dirichlet problem $u'' + \mu x = 0$, $u(0) = u(T) = 0$, that is,

$$\mu_k = \left(\frac{\pi k}{T}\right)^2, \quad k \in \mathbb{N}. \tag{5.115}$$

Furthermore, we set $\mu_0 = 0$.

THEOREM 5.30 [52, Theorem 1.1]. *Assume that $f : [0, T] \times (0, \infty) \to \mathbb{R}$ is continuous and there are positive constants c, c', δ and $\nu \geqslant 1$ such that*

$$\frac{c'}{x^\nu} \leqslant f(t, x) \leqslant \frac{c}{x^\nu} \quad \textit{for all } x \in (0, \delta). \tag{5.116}$$

Moreover, let there exist a nonnegative integer k such that

$$-\mu_{k+1} < \liminf_{x\to\infty} \frac{f(t, x)}{x} \leqslant \limsup_{x\to\infty} \frac{f(t, x)}{x} < -\mu_k$$
$$\textit{uniformly in } t \in [0, T]. \tag{5.117}$$

Then problem (5.97), (5.2) *has a positive solution.*

SKETCH OF THE PROOF. For a given $e \in C[0, T]$ denote by $\mathcal{R}(e)$ the unique solution of the problem

$$u'' + u = e(t), \quad u(0) = u(T), \quad u'(0) = u'(T).$$

It is known that $\mathcal{R}$ defines a compact linear operator on the space $C_T[0, T]$ of continuous T-periodic functions endowed with the sup norm $\|\cdot\|_\infty$. Problem (5.97), (5.2) is thus equivalent to finding a positive solution $u \in C_T[0, T]$ of the fixed point problem $u = \mathcal{T}(u)$, where

$$\mathcal{T}(u) = \mathcal{R}\big(u + f(\cdot, u)\big) \quad \text{for } u \in C_T[0, T].$$

For $0 < \varepsilon < M$ define $\Omega_{\varepsilon,M} = \{u \in C_T[0, T]\colon\ \varepsilon < u < M \text{ on } [0, T]\}$. Then $\mathcal{T} : \overline{\Omega}_{\varepsilon,M} \to C_T[0, T]$ is a completely continuous operator.

The proof of the theorem consists in showing that there are ε, M such that $\deg(\mathcal{I} - \mathcal{T}, \Omega_{\varepsilon,M}) \neq 0$:

Let k be from (5.117) and choose an arbitrary $\gamma \in (\mu_k, \mu_{k+1})$. Further, for $\lambda \in [0, 1]$ and $u \in C_T[0, T]$, define

$$\widetilde{\mathcal{T}}(u) = \mathcal{R}\left(\gamma u - \frac{1}{u^\nu} + u\right) \quad \text{and} \quad \mathcal{T}_\lambda(u) = \lambda\widetilde{\mathcal{T}}(u) + (1 - \lambda)\mathcal{T}(u).$$

Notice that

$$\mathcal{T}_\lambda(u) = \mathcal{R}\left(\lambda\left(\gamma u - \frac{1}{u^\nu}\right) + (1-\lambda) f(\cdot, u) + u\right).$$

Furthermore, $\mathcal{T}_0 = \mathcal{T}$ and $\mathcal{T}_1 = \widetilde{\mathcal{T}}$. It can be shown that there are ε, M such that $0 < \varepsilon < M$ and $u \neq \mathcal{T}_\lambda(u)$ for all $u \in \partial\Omega_{\varepsilon,M}$. By the homotopy property of the degree, it follows that $\deg(\mathcal{I} - \mathcal{T}, \Omega_{\varepsilon,M}) = \deg(\mathcal{I} - \widetilde{\mathcal{T}}, \Omega_{\varepsilon,M})$.

Define

$$\widetilde{\mathcal{S}}(u) = \mathcal{R}\left(\gamma u - \frac{\lambda}{u^3} + u\right) \quad \text{and} \quad \mathcal{S}_\lambda(u) = \lambda \widetilde{\mathcal{T}}(u) + (1-\lambda)\widetilde{\mathcal{S}}(u)$$

for $u \in C_T[0, T]$ and $\lambda \in [0, 1]$. We have

$$\mathcal{S}_\lambda(u) = \mathcal{R}\left(\gamma u + u - (1-\lambda)\frac{1}{u^\nu} - \lambda\frac{1}{u^3}\right),$$

$\mathcal{S}_0 = \widetilde{\mathcal{S}}$ and $\mathcal{S}_1 = \widetilde{\mathcal{T}}$. Now, we prove that $u \neq \mathcal{S}_\lambda(u)$ on $\partial\Omega_{\varepsilon,M}$ for each $\lambda \in [0, 1]$ and for some suitable ε and M. Similarly to Step 1, this yields that $\deg(\mathcal{I} - \widetilde{\mathcal{T}}, \Omega_{\varepsilon,M}) = \deg(\mathcal{I} - \widetilde{\mathcal{S}}, \Omega_{\varepsilon,M})$.

The proof is completed by proving that $\deg(\mathcal{I} - \widetilde{\mathcal{S}}, \Omega_{\varepsilon,M}) \neq 0$. □

REMARK 5.31. Consider the problem

$$u'' + ku = \frac{\beta}{u^\lambda} + e(t), \quad u(0) = u(T), \quad u'(0) = u'(T) \tag{5.118}$$

with $\lambda > 0$, $\beta > 0$ and $k \geqslant 0$. Denote

$$g(x) = \frac{\beta}{x^\lambda} - kx \quad \text{for } x > 0.$$

If $e \in L_1[0, T]$, $k = 0$ and $\lambda \geqslant 1$, i.e. the function g has a strong singularity at $x = 0$, then by [96, Theorem 3.12] problem (5.118) has a positive solution if and only if $\overline{e} < 0$ and, in the case $\lambda \in (0, 1)$, this condition need not ensure the existence of a positive solution to (5.118) (cf. [96, Theorem 4.1]). Further, if $e \in C[0, T]$ and $\lambda \geqslant 1$, then by Theorem 5.30, problem (5.118) has a positive solution whenever the condition

$$k \neq \left(n\frac{\pi}{T}\right)^2 \quad \text{for all } n \in \mathbb{N}$$

is satisfied. It is worth mentioning that the resonance case of $k = (\frac{\pi}{T})^2$ is covered neither by Theorem 5.30 nor by Theorem 5.25 even for the strong singularity $\lambda \geqslant 1$.

In comparison to these results, it should be pointed out that using Corollary 5.28, we can obtain existence results also for the cases $\lambda \in (0,1)$ and $k = (\frac{\pi}{T})^2$. In particular, for problem (5.118), with $e \in L_1[0,T]$ we get the existence of a positive solution in the following cases:

$$k = 0, \quad \bar{e} < 0 \quad \text{and} \quad \inf\operatorname{ess}\{e(t)\colon t \in [0,T]\} > -\left(\frac{\pi^2}{T^2\lambda\beta}\right)^{\frac{\lambda}{\lambda+1}}(\lambda+1)\beta$$

or

$$0 < k < \left(\frac{\pi}{T}\right)^2 \quad \text{and} \quad \inf\operatorname{ess}\{e(t)\colon t \in [0,T]\} > -\left(\frac{\pi^2 - T^2k}{T^2\lambda\beta}\right)^{\frac{\lambda}{\lambda+1}}(\lambda+1)\beta$$

or

$$k = \left(\frac{\pi}{T}\right)^2 \quad \text{and} \quad \inf\operatorname{ess}\{e(t)\colon t \in [0,T]\} > 0.$$

Notice that for the case $0 < k < (\frac{\pi}{T})^2$, Theorem 5.27 provides a complementary existence condition. For details, see [144, Corollary 4.6].

We close this section by mentioning some results concerning the case when the nonlinearity can have both a space singularity at $x = 0$ and superlinear descent for large x. The first is due to del Pino and Manásevich. It was motivated by [156], where an equation governing the nonlinear vibrations of a radially forced thick-walled and incompressible material was derived. The proof makes use of a version of the Poincaré–Birkhoff theorem due to Ding [56] together with an analysis of some oscillatory properties of solutions to the related initial value problems.

THEOREM 5.32 [51, Theorem 2.1]. *Let* $f : \mathbb{R} \times (0,\infty) \to \mathbb{R}$ *be continuous, locally Lipschitz in* x, T*-periodic in* t *and such that for* $s, \beta \in \mathbb{R}$ *and* $\alpha > 0$, *the solution* $u(t)$ *of the local initial value problem*

$$u'' = f(t,u), \quad u(s) = \alpha, \quad u'(s) = \beta$$

is continuable to the whole real line $\mathbb{R}$ *and* $u > 0$ *on* $\mathbb{R}$. *Furthermore, assume that*

$$0 < \liminf_{x\to 0+} xf(t,x) \leqslant \infty \quad \textit{and} \quad \lim_{x\to\infty} \frac{f(t,x)}{x} = -\infty$$

uniformly in t. *Then there is* $n_0 \in \mathbb{N}$ *such that for every* $n \geqslant n_0$ *there exist two distinct* T*-periodic positive solutions* u_{n1}, u_{n2} *of* (5.97), (5.2) *such that both* $u_{n1} - 1$ *and* $u_{n2} - 1$ *have exactly* $2n$ *zeros in* $[0,T)$. *In particular, problem* (5.97), (5.2) *possesses infinitely many* T*-periodic positive solutions.*

The other result is due to Ge and Mawhin. It deals with the equation

$$u'' = g(u) + p(t, u, u'). \tag{5.119}$$

Its proof was obtained by use of some continuation theorems valid in absence of a priori bounds and given by Capietto, Mawhin and Zanolin, e.g., in [46] and [103].

THEOREM 5.33 [73, Theorem 1]. *Let* $g \in C(0, \infty)$ *and* $p \in Car([0, T] \times \mathbb{R}^2)$. *Furthermore, assume that there are constants* $\alpha, \beta \geqslant 1$, $M \geqslant 0$, $L \geqslant 0$ *such that*

$$\lim_{x\to 0+} g(x)x^{\alpha} = \infty, \qquad \lim_{x\to\infty} \frac{g(x)}{x^{\beta}} = -\infty, \tag{5.120}$$

$$\left|p(t, x, y)\right| \leqslant \begin{cases} M\left(|x|^{\frac{1-\alpha}{2}} + |y|\right) + L & \textit{if } 0 < x < 1, \\ M\left(|x|^{\frac{1+\beta}{2}} + |y|\right) + L & \textit{if } x \geqslant 1. \end{cases} \tag{5.121}$$

Then problem (5.119), (5.2) *has a positive solution.*

5.5. *Historical and bibliographical notes*

In 1958 Bevc, Palmer and Süsskind [34] searched for positive 2π-periodic solutions of the Brillouin electron beam focusing system (5.114) which is a singular perturbation of the Mathieu equation. Before, in 1950, Pinney [113] considered the so-called Ermakov–Pinney equation $r'' + a(t)r = \frac{K}{r^3}$, where $a(t)$ is T-periodic and $K > 0$. Another example of the singular problem mentioned in literature are the parametric resonances of certain nonlinear Schrödinger systems (see [71]). As mentioned by Mawhin and Jebelean in their exhaustive historical introduction to the paper [84], second order nonlinear differential equations or systems with singularities appear naturally in the description of particles submitted to Newtonian type forces or to forces caused by compressed gases. Their mathematical study started in the sixties by Forbat and Huaux [69], Huaux [82], Derwidué [54] and Faure [65], who considered positive solutions of equations describing, e.g., the motion of a piston in a cylinder closed at one extremity and submitted to a T-periodic exterior force, to the restoring force of a perfect gas and to a viscosity friction. The equations under their study may be after suitable substitutions transformed to

$$u'' + cu' = \frac{\beta}{u} + e(t),$$

where $c \neq 0$ and $\beta \in \mathbb{R}$ can be either positive or negative. Equations of this form are usually called Forbat's equations and their Liénard type generalizations like

$$u'' + h(u)u' = g(t, u) + e(t) \tag{5.122}$$

are sometimes also referred to as the generalized Forbat's equations. It is worth mentioning that, while Forbat and others relied on the dissipativeness properties, Faure made use of the Leray–Schauder topological method.

Later, in the seventies, techniques of critical point theory were applied for the first time by Gordon [76], who also introduced the *strong force condition* of the type (5.29).

In 1988, Gaete and Manásevich [70], using variational methods, proved the existence of at least two different positive T-periodic solutions of the equation $u'' = p(t)u^2 - u + u^{-5}$ which governs the radial oscillations of an elastic spherical membrane made up of a neo-Hookean material, and subjected to an internal pressure $p : \mathbb{R} \to (0, \infty)$ continuous, T-periodic and non-constant.

In 1987, keeping in mind the model equation $u'' = \beta u^{-\lambda} + e(t)$ with $\lambda > 0$, $\beta \neq 0$ and $e \in L_1[0, T]$, Lazer and Solimini [96] employed topological arguments and the lower and upper functions method to investigate the existence of positive solutions to the Duffing equation $u'' = g(u) + e(t)$. The restoring force g was allowed to have an attractive singularity or a strong repulsive singularity at origin. Starting with this paper, the interest in periodic singular problems considerably increased. The results by Lazer and Solimini have been generalized or extended, e.g., by Habets and Sanchez [78] (1990), Mawhin [102] (1991), del Pino, Manásevich and Montero [52] (1992), del Pino and Manásevich [51] (1993), Fonda [66] (1993), Omari and Ye [108] (1995), Zhang [152] (1996) and [154] (1998) and Ge and Mawhin [73] (1998). Some of these papers (e.g., [78,108,152] or [154]) cover also the Liénard equation (5.122) with $g(t, x)$ having at $x = 0$ an essentially autonomous singularity. However, all of them, when dealing with the repulsive singularity, supposed that the strong force condition of the type (5.29) is satisfied. Furthermore, except for [51,52] and [73], they dealt with restoring forces $g(t, x)$ behaving at ∞ like $-kx$ with $0 < k < \mu_1$, $\mu_1 = (\frac{\pi}{T})^2$ being the first Dirichlet eigenvalue of $x'' + \mu x = 0$. The paper [52] was concerned with the cases corresponding to k lying between two adjacent higher eigenvalues, while the papers [51] and [73] dealt with the superlinear case. Recently, Yan and Zhang [151] (2003) proved an existence result assuming that the nonlinearity grows semilinearly as $x \to \infty$ and fulfill a certain higher-order non-resonance condition in terms of the periodic and antiperiodic eigenvalues. Let us mention also that Martinez-Amores and Torres [105] considered in 1996 stability of periodic solutions of problems with singularities of attractive type. Furthermore, in 1998, Torres [143] delivered results on the existence of bounded solutions to singular equations of repulsive type.

In 2001, Rachůnková, Tvrdý and Vrkoč [129], motivated by results on the existence of positive solutions to regular periodic problems by Nkashama and Santanilla [107] (1990) and Sanchez [131] (1992), made use of the lower and upper functions method to deliver related results in the form applicable also to singular problems. Unlike the above mentioned papers, their results concern also the resonance case $k = \mu_1$ and do not need any strong force condition. Later, in 2002, further step was done by Bonheure, Fonda and Smets [39] who made use of the properties of forced isochronous oscillators. Their results are also valid in the resonance case $k = \mu_1$ with a weak singularity. It turned out that in the resonance case $k = \mu_1$ problem (5.118) with $\lambda \geqslant 0$ has a solution whenever there is a $\delta > 0$ such that

$$\min_{t \in [0,T]} \int_t^{t+T} e(t) \sin\left(\pi \frac{s-t}{T}\right) \mathrm{d}s \geqslant \delta.$$

Analogous results were derived also by Bonheure and De Coster [38] in 2003 by means of the lower and upper functions method. Simultaneously, Torres [144] noticed that having a thorough analysis of the sign properties of the related Green's functions, solvability of the periodic problem with a weak singularity can be ensured also by the Krasnoselskii fixed point theorem. His results turned out to be complementary to those already known when $0 < k < \mu_1$.

For related multiplicity results we refer to the papers by Fonda, Manásevich and Zanolin [68] (1993), Rachůnková [115] and [116] (2000) and Rachůnková, Tvrdý and Vrkoč [130] (2003). In particular, in [115] an extension of the results of Gaete and Manásevich from [70] applicable to the equation modelling radial oscillations of an elastic spherical membrane can be found.

The regular periodic problem with ϕ- or p-Laplacian on the left-hand side was considered by several authors. For example, del Pino, Manásevich and Murúa [53] (1992) and Yan [150] (2003) proved the existence or multiplicity of periodic solutions of the equation $(\phi_p(u'))' = f(t,u)$ under non resonance conditions imposed on f. In 1998, general existence principles for the regular vector problem, based on the homotopy to the averaged nonlinearity, were presented by Manásevich and Mawhin [101] (1998) (see also Mawhin [104]). Multiplicity results were given by Liu in [97] (1998) and by Jiang, Chu and Zhang in [88] (2005). The resonance case was considered by Fabry and Fayyad in [61] (1992).

The first steps to establish the lower/upper functions method for problems with a ϕ-Laplacian operator on the left-hand side were done by Cabada and Pouso in [41] (1997) and by Jiang and Wang in [89] (1997), the latter paper dealing with the p-Laplacian. They assumed the existence of a pair of well-ordered lower and upper functions and both-sided Nagumo conditions. These results were extended by Staněk [133] (2001) to the case when a functional right-hand side fulfills one-sided growth conditions of Nagumo type. The paper by Cabada, Habets and Pouso [42] (1999) was the first to present the lower/upper functions method for periodic problems with a ϕ-Laplacian operator under the assumption that lower/upper functions are in the reverse order, see also [42] (2000). If $\phi = \phi_p$ the authors got the solvability for $1 < p \leqslant 2$ only. The general existence principle valid also when lower/upper functions are non-ordered was presented by Rachůnková and Tvrdý in [127] (2005) and for the case when impulses are admitted also in [126] (2005).

The singular periodic problem for the Liénard equation $(\phi_p(u'))' + h(u)u' = g(u) + e(t)$ with p-Laplacian on the left-hand side was treated by Liu [98] (2002) and Jebelean and Mawhin [84] (2002) and [85] (2004). Their main tool was the existence principle due to Manásevich and Mawhin from [101] (1998). Furthermore, in [84], the significance of the lower/upper functions method was shown.

The only existence results for problem (5.1), (5.2) with f having a singularity for $u' = 0$ were delivered by Staněk in [134] (2001) and [136] (2002).

Extensions to vector systems of the second order were not the subject of this text. We can only refer, e.g., to the papers by Habets and Sanchez [79] (1990), Solimini [132] (1990), Fonda [67] (1995) and Zhang [155] (1999) for the classical case and by Manásevich and Mawhin [101] (1998) and Liu [98] (2002) for systems with p-Laplacian operators on their left-hand sides.

6. Other types of two-point second order BVPs

In Sections 4 and 5, under the assumption that ϕ satisfies (3.3), we have investigated the nonlinear second order differential equation of the form

$$\big(\phi(u')\big)' + f(t,u,u') = 0, \tag{6.1}$$

subjected to Dirichlet and periodic boundary conditions, respectively. In this section we will study solvability of Eq. (6.1) with some other types of two-point boundary conditions on the interval $[0,T] \subset \mathbb{R}$.

We will focus our attention on problems with *space* or with *mixed* (*time and space*) singularities. According to Section 1, *solutions* and *w-solutions* of the problems are defined in the same way as for the Dirichlet problem (see Section 4.1) just replacing the Dirichlet conditions by the boundary conditions under consideration. We can define *lower and upper functions* of the second order boundary value problem in the same way as in Definition 4.3 replacing inequalities (4.6) with inequalities corresponding to the boundary conditions in question.

In the sequel we consider two-point linear boundary conditions arising in the study of physical, chemical or engineering problems and having the form

$$\begin{cases} a_0u(0) - b_0u'(0) = 0, \quad a_1u(T) + b_1u'(T) = 0, \\ a_i, b_i \in \mathbb{R}, \quad a_i^2 + b_i^2 > 0, \quad i = 0, 1. \end{cases} \tag{6.2}$$

Conditions (6.2) include conditions of the Dirichlet type (with $b_0 = b_1 = 0$), of the Neumann type (with $a_0 = a_1 = 0$), of the mixed type (with $a_0 = b_1 = 0$ or $b_0 = a_1 = 0$), of the Robin type (with $a_i > 0$, $b_i > 0$, $i = 0, 1$) and of the standard Sturm–Liouville type (with $a_i, b_i \in [0,\infty)$, $i = 0, 1$). We will also mention problems involving inhomogeneous form of the above boundary conditions, i.e.

$$\begin{cases} a_0u(0) - b_0u'(0) = A, \quad a_1u(T) + b_1u'(T) = B, \\ a_i, b_i \in \mathbb{R}, \quad a_i^2 + b_i^2 > 0, \quad i = 0, 1,\ A, B \in \mathbb{R}. \end{cases} \tag{6.3}$$

However, there is no restriction in assuming just the homogeneous conditions since a change from $u(t)$ to $y(t) = u(t) - q(t)$, where q is a polynomial satisfying (6.3), will reduce (6.3) to (6.2).

Consider a class of nonlinear singular boundary value problems whose importance is derived, in part, from the fact that they arise when searching for *positive, radially symmetric solutions* to the nonlinear elliptic partial differential equation

$$\Delta u + g(r,u) = 0 \quad \text{on } \Omega, \quad u\,|_\Gamma = 0,$$

where Δ is the Laplace operator, Ω is the open unit disk in $\mathbb{R}^n$ (centered at the origin), Γ is its boundary, and r is the radial distance from the origin. Radially symmetric solutions to this problem are solutions of the ordinary differential equation with the mixed boundary conditions

$$u'' + \frac{n-1}{t}u' + g(t,u) = 0, \quad u'(0) = 0, \quad u(1) = 0.$$

See, e.g., [29] or [74]. Particularly, Gatica, Oliker and Waltman [72] investigated the singular problem

$$u'' + \frac{n-1}{t}u' + \psi(t)\frac{1}{u^{\alpha}} = 0, \quad u'(0) = 0, \quad u(1) = 0, \tag{6.4}$$

where

$$\begin{cases} n \geqslant 2,\ \alpha \in (0,1),\ \psi \in C[0,1) \text{ is non-negative,} \\ \psi \text{ can have a time singularity at } t = 1. \end{cases} \tag{6.5}$$

THEOREM 6.1 [72, Theorem 4.1]. *Let* (6.5) *hold. Assume that*

$$0 < \int_0^1 (1-t)^{-\alpha}\psi(t)\,\mathrm{d}t < \infty.$$

Then problem (6.4) *has a solution that is positive on* [0, 1).

The technical arguments in the proof involve concavity of solutions and the use of iterative techniques. The main tool is a fixed point theorem for decreasing mappings on cones.

In the theory of diffusion and reaction a class of differential equations

$$u'' - \eta^2 g_\kappa(u) = 0 \tag{6.6}$$

appears. Here η^2 is the (positive) Thiele modulus, $u \geqslant 0$ is the concentration of one of the reactants and κ is a positive parameter. The functions g_κ are continuous on $[0,\infty)$,

$$\lim_{\kappa \to 0+} g_\kappa(x) = g(x) \quad \text{for } x \in (0,\infty)$$

and g can have a space singularity at $x = 0$. The model functions are

$$g_\kappa(x) = \frac{x}{\kappa + x^{1+\gamma}}, \tag{6.7}$$

where γ is a positive parameter. Aris [21] proposed such equations as descriptions of the steady state for chemicals reacting and diffusing according to the Langmuir–Hinshelwood kinetics. Bobisud [35] studied a class of equations (6.6) on $[-1,1]$ subjected to the inhomogeneous Robin boundary conditions

$$\alpha u(-1) - u'(-1) = A, \quad \alpha u(1) + u'(1) = A, \quad \alpha, A > 0, \tag{6.8}$$

and with functions g_κ behaving qualitatively very much like the model functions in (6.7). He proved that for η^2 in (6.6) sufficiently small the limit problem with $\kappa = 0$ has a positive solution which can be approximated uniformly on $[-1, 1]$ by solutions of (6.6), (6.8) with small κ.

Motivated by problem (6.6), (6.8) as well as by problem (6.4), Baxley and Gersdorff [27] studied a singular equation in which u' can appear nonlinearly,

$$u'' + h(t, u') - \eta^2 g(t, u) = 0, \quad \eta^2 > 0, \tag{6.9}$$

with inhomogeneous Sturm–Liouville boundary conditions

$$u'(0) = 0, \quad \alpha u(T) + \beta u'(T) = A, \quad \alpha, A > 0, \ \beta \geqslant 0, \tag{6.10}$$

where

$$\begin{cases} h \in C\big((0, T] \times [0, \infty)\big) \text{ is non-negative} \\ \text{and can have a time singularity at } t = 0, \end{cases} \tag{6.11}$$

$$\begin{cases} g \in C\big([0, T] \times \big(0, \frac{A}{\alpha}\big]\big) \text{ is positive} \\ \text{and can have a space singularity at } x = 0. \end{cases} \tag{6.12}$$

In contrast to Theorem 6.1 where positive singular nonlinearity $\psi(t)x^{-\alpha}$ appears, the next theorem applies to equations involving a negative singular term $-\eta^2 g(t, x)$.

THEOREM 6.2 [27, Theorem 17]. *Let* (6.11) *and* (6.12) *hold. Assume that*

$$h(t, 0) = 0 \quad \text{for } t \in (0, T]$$

and that there exists $G \in L[0, \frac{A}{\alpha}]$ *satisfying*

$$g(t, x) \leqslant G(x) \quad \text{on } [0, T] \times \big(0, \tfrac{A}{\alpha}\big]. \tag{6.13}$$

Then for η^2 *sufficiently small problem* (6.9), (6.10) *has a solution that is positive on* $[0, T]$.

Moreover, if η^2 is sufficiently large, Baxley and Gersdorff guarantee the existence of the so called *dead core* solutions which are defined as functions belonging for some $t_0 \in (0, T)$ to $C^1[0, T] \cap C^2(t_0, T]$, satisfying equation (6.9) on $(t_0, T]$, vanishing on $[0, t_0]$ and fulfilling (6.10). The proof is based on a priori estimates of approximate solutions of auxiliary regular problems and on the Arzelà–Ascoli theorem.

Agarwal, O'Regan and Staněk [15] considered a singular equation with a ϕ-Laplacian generalizing (6.9) and subjected to inhomogeneous mixed conditions

$$\big(\phi(u')\big)' - \mu f(t, u, u') = 0, \quad u'(0) = 0, \quad u(T) = b, \quad b > 0, \tag{6.14}$$

where μ is a real positive parameter and

$$\begin{cases} \phi(0)=0, \quad f \in Car\big([0,T]\times(\mathbb{R}\setminus\{b\})\times(\mathbb{R}\setminus\{0\})\big), \\ f \text{ can have space singularities at } x=b \text{ and } y=0. \end{cases} \tag{6.15}$$

THEOREM 6.3 [15, Theorem 3.1]. *Let* (6.15) *hold. Assume that there exist* $\varepsilon > 0$, $\nu \in (0,T]$ *and a positive non-decreasing function* $\rho \in C[0,T]$ *such that*

$$f\big(t,x,\rho(t)\big)=0 \quad \textit{for a.e. } t \in [0,T] \textit{ and all } x \in [0,b),$$

$$\varepsilon \leqslant f(t,x,y) \quad \textit{for a.e. } t \in [0,\nu] \textit{ and all } x \in [0,b),\ y \in [0,\nu].$$

Further, let for a.e. $t \in [0,T]$ *and for all* $x \in [0,b)$, $y \in (0,\rho(t)]$

$$0 \leqslant f(t,x,y) \leqslant \big(h_1(x)+h_2(x)\big)\big(\omega_1(y)+\omega_2(y)\big),$$

where $h_1 \in C[0,b]$, $\omega_1 \in C[0,\infty)$ *are non-negative*, $h_2 \in C[0,b)$, $\omega_2 \in C(0,\infty)$ *are positive*, h_1 *and* ω_2 *are non-increasing*, h_2 *and* ω_1 *are non-decreasing and* $\omega_1+\omega_2$ *is non-increasing on a right neighbourhood of* 0. *Moreover, let*

$$\int_0^b h_2(s)\,\mathrm{d}s < \infty, \qquad \int_0^1 \omega_2\big(\phi^{-1}(s)\big)\,\mathrm{d}s < \infty.$$

Finally, let there exist $\mu_* > 0$ *such that*

$$\int_0^b \frac{\mathrm{d}s}{\Omega^{-1}(\mu_* H(s))} = T, \tag{6.16}$$

where

$$H(u) = \int_0^u \big(h_1(s)+h_2(s)\big)\,\mathrm{d}s, \qquad \Omega(u) = \int_0^{\phi(u)} \frac{\phi^{-1}(s)\,\mathrm{d}s}{\omega_1(\phi^{-1}(s))+\omega_2(\phi^{-1}(s))}.$$

Then for each $\mu \in (0,\mu_*)$ *problem* (6.14) *has a solution u satisfying*

$$0 < u(t) \leqslant b, \quad 0 \leqslant u'(t) \leqslant \rho(t) \quad \textit{for } t \in [0,T].$$

To prove this existence result the authors used regularization and sequential techniques. First, they defined a family of auxiliary regular differential equations depending on $n \in \mathbb{N}$ and then, using the topological transversality theorem, they obtained a sequence of positive approximate solutions. Applying the Arzelà–Ascoli theorem and the Lebesgue convergence theorem they showed that its limit is a solution of problem (6.14).

REMARK 6.4. Note that if there exists $\mu_0 \in (0, \infty)$ such that

$$\int_0^b \frac{\mathrm{d}s}{\Omega^{-1}(\mu_0 H(s))} \in (T, \infty),$$

then $\mu_* \in (0, \mu_0)$ satisfying (6.16) can be always found.

Comparing problem (6.9), (6.10) and problem (6.14), they seem to be in some sense close, because both of them have negative singular nonlinearities in differential equations and the boundary conditions of (6.14) are contained in (6.10). However, there is a large difference between them. For example, positive solutions of (6.9), (6.10) do not touch the space singularity of f at $x = 0$. On the other hand, each solution u of (6.14) satisfies $u(T) = b$ and hence enters the space singularity of f at $x = b$. Another difference between them consists in the fact that f in (6.14) can have also a space singularity at $y = 0$ and hence Theorem 6.3 can be used in the following example whereas Theorem 6.2 cannot.

EXAMPLE. Let $\alpha \in (0, \infty)$ and $\beta \in (0, 1)$. By Theorem 6.3 there exists a positive number μ_* depending on β only such that for any $\mu \in (0, \mu_*)$ the problem

$$u'' - \mu\left(1 - |u|^\alpha\right)\left(\frac{1}{|u'|^\beta} - 1\right) = 0, \quad u'(0) = 0, \quad u(1) = \frac{1}{2},$$

has a solution u such that $0 < u(t) \leqslant \frac{1}{2}$, $0 \leqslant u'(t) \leqslant 1$ for $t \in [0, 1]$. An explicit formula for μ_* can be found in [15].

Assumption (6.13) in Theorem 6.2 means that the space singularity of g at $x = 0$ is a *weak singularity*. See Remark 5.1 for more detail. Note that the assumption (6.13) is not satisfied for the problem

$$u'' + \frac{t^2}{32u^2} - \frac{\lambda^2}{8} = 0, \quad u(0) = 0, \quad 2u'(1) - (1 - \nu)u(1) = 0, \tag{6.17}$$

where $\lambda \in (0, \infty)$, $\nu \in (0, 1)$, which models the large deflection membrane response of a spherical cap. This problem has been solved numerically by various techniques in engineering literature [75,106,111]. Baxley in [26] proved existence and uniqueness of a solution of this problem, gave qualitative information about the solution, and used this information to suggest an approach to numerical computation. In the proof the maximum principle plays a fundamental role. In contrast to (6.8), (6.11) and (6.14), problem (6.17) has boundary conditions which are not included in the Sturm–Liouville ones, because $\nu < 1$ and so the coefficient $-(1 - \nu)$ at $u(1)$ is negative.

Existence results for equations whose nonlinearities have a singularity at $x = 0$ and can be increasing for $x \to \infty$ are proved in [4], where Agarwal and O'Regan obtained the existence of a w-solution $u > 0$ on $(0, 1]$ of such equations with mixed boundary conditions. Their theorem can be applied for example to the problem

$$u'' + \left(\frac{1}{u^\alpha} + u^\beta + 1\right)\left(1 + (u')^3\right) = 0, \quad u'(0) = 0, \quad u(1) = 0, \tag{6.18}$$

with $\alpha \in (0,1)$, $\beta \geqslant 0$. We see that the nonlinearity

$$f(t,x,y) = \left(\frac{1}{x^\alpha} + x^\beta + 1\right)\bigl(1 + y^3\bigr) \tag{6.19}$$

has a weak space singularity at $x = 0$ and can be increasing for large x. If $\beta \in (0,1)$ the growth of f is sublinear. For $\beta = 1$ or $\beta > 1$ the growth of f is linear or superlinear, respectively.

In the investigation of singular problems (6.1), (6.2) or (6.1), (6.3), lower and upper functions of the corresponding regular problems can be a fruitful tool. See for example papers by Kannan and O'Regan in [90] or by Agarwal and Staněk in [17]. We will demonstrate the role of lower and upper functions on the following singular problem with mixed boundary conditions

$$u'' + f(t,u,u') = 0, \quad u'(0) = 0, \quad u(T) = 0, \tag{6.20}$$

where

$$\begin{cases} \mathcal{D} = (0,\infty) \times (-\infty,0), \quad f \in Car\bigl((0,T) \times \mathcal{D}\bigr), \\ f \text{ can have time singularities at } t = 0,\ t = T \\ \text{and space singularities at } x = 0,\ y = 0. \end{cases} \tag{6.21}$$

First, consider an auxiliary regular problem

$$u'' + h(t,u,u') = 0, \quad u'(0) = 0, \quad u(T) = 0, \tag{6.22}$$

where $h \in Car([0,T] \times \mathbb{R}^2)$.

DEFINITION 6.5. A function $\sigma \in C[0,T]$ is called *a lower function of* (6.22) if there exists a finite set $\Sigma \subset (0,T)$ such that $\sigma \in AC^1_{loc}([0,T] \setminus \Sigma)$, $\sigma'(\tau+), \sigma'(\tau-) \in \mathbb{R}$ for each $\tau \in \Sigma$,

$$\sigma''(t) + f\bigl(t,\sigma(t),\sigma'(t)\bigr) \geqslant 0 \quad \text{for a.e. } t \in [0,T], \tag{6.23}$$

$$\sigma'(0) \geqslant 0, \quad \sigma(T) \leqslant 0, \quad \sigma'(\tau-) < \sigma'(\tau+) \quad \text{for each } \tau \in \Sigma. \tag{6.24}$$

If the inequalities in (6.23) and (6.24) are reversed, then σ is called *an upper function of* (6.22).

In what follows we will need the classical lower and upper functions result for the mixed problem (6.22).

LEMMA 6.6 [93, Lemma 3.7]. *Let σ_1 and σ_2 be a lower and an upper function for problem* (6.22) *such that $\sigma_1 \leqslant \sigma_2$ on* $[0,T]$. *Assume also that there is a function $\psi \in L_1[0,T]$ such that*

$$\left|h(t,x,y)\right| \leqslant \psi(t) \quad \text{for a.e. } t \in [0,T],\ \text{all } x \in \left[\sigma_1(t), \sigma_2(t)\right],\ y \in \mathbb{R}. \tag{6.25}$$

Then problem (6.22) *has a solution $u \in AC^1[0,T]$ satisfying*

$$\sigma_1(t) \leqslant u(t) \leqslant \sigma_2(t) \quad \text{for } t \in [0,T]. \tag{6.26}$$

We will apply Lemma 6.6 to the singular mixed problem (6.20).

THEOREM 6.7 [117, Theorem 3.1]. *Let* (6.21) *hold. Assume that there exist $\varepsilon \in (0,1)$, $\nu \in (0,T)$, $c \in (\nu,\infty)$ such that*

$$f\bigl(t, c(T-t), -c\bigr) = 0 \quad \text{for a.e. } t \in [0,T], \tag{6.27}$$

$$0 \leqslant f(t,x,y)$$
$$\text{for a.e. } t \in [0,T],\ \text{and all } x \in \bigl(0, c(T-t)\bigr],\ y \in [-c,0), \tag{6.28}$$

$$\varepsilon \leqslant f(t,x,y)$$
$$\text{for a.e. } t \in [0,\nu],\ \text{and all } x \in \bigl(0, c(T-t)\bigr],\ y \in [-\nu,0). \tag{6.29}$$

Then problem (6.20) *has a solution $u \in AC^1[0,T]$ satisfying*

$$0 < u(t) \leqslant c(T-t), \quad -c \leqslant u'(t) < 0 \quad \text{for } t \in (0,T). \tag{6.30}$$

PROOF. Let $k \in \mathbb{N}$, $k \geqslant \frac{3}{T}$.

Step 1. *Approximate solutions.*

For $t \in [\frac{1}{k}, T - \frac{1}{k}]$, $x \in \mathbb{R}$ and $y \in \mathbb{R}$ put

$$\alpha_k(t,x) = \begin{cases} c(T-t) & \text{if } x > c(T-t), \\ x & \text{if } \frac{c}{k} \leqslant x \leqslant c(T-t), \\ \frac{c}{k} & \text{if } x < \frac{c}{k}, \end{cases}$$

$$\beta_k(y) = \begin{cases} -\frac{\varepsilon}{k} & \text{if } y > -\frac{\varepsilon}{k}, \\ y & \text{if } -c \leqslant y \leqslant -\frac{\varepsilon}{k}, \\ -c & \text{if } y < -c, \end{cases} \qquad \gamma(y) = \begin{cases} \varepsilon & \text{if } y \geqslant -\nu, \\ \varepsilon\frac{c+y}{c-\nu} & \text{if } -c < y < -\nu, \\ 0 & \text{if } y \leqslant -c. \end{cases}$$

For a.e. $t \in [0,T]$ and $x, y \in \mathbb{R}$ define

$$f_k(t,x,y) = \begin{cases} \gamma(y) & \text{if } t \in [0, \frac{1}{k}), \\ f\bigl(t, \alpha_k(t,x), \beta_k(y)\bigr) & \text{if } t \in [\frac{1}{k}, T - \frac{1}{k}], \\ 0 & \text{if } t \in (T - \frac{1}{k}, T]. \end{cases}$$

Then $f_k \in Car([0,T] \times \mathbb{R}^2)$ and there is $\psi_k \in L_1[0,T]$ such that

$$\bigl|f_k(t,x,y)\bigr| \leqslant \psi_k(t) \quad \text{for a.e. } t \in [0,T], \text{ all } x, y \in \mathbb{R}. \tag{6.31}$$

We have got an auxiliary regular problem

$$u'' + f_k(t,u,u') = 0, \quad u'(0) = 0, \quad u(T) = 0. \tag{6.32}$$

Conditions (6.27) and (6.28) yield

$$f_k\bigl(t, c(T-t), -c\bigr) = 0 \quad \text{and} \quad f_k(t,0,0) \geqslant 0 \quad \text{for a.e. } t \in [0,T].$$

Put $\sigma_1(t) = 0$, $\sigma_2(t) = c(T-t)$ for $t \in [0,T]$. Then σ_1 and σ_2 are a lower and an upper function of (6.32). Hence, by Lemma 6.6, problem (6.32) has a solution u_k and

$$0 \leqslant u_k(t) \leqslant c(T-t) \quad \text{on } [0,T]. \tag{6.33}$$

Step 2. *A priori estimates of approximate solutions.*

Since $u_k'(0) = 0$ and $f_k(t,x,y) \geqslant 0$ for a.e. $t \in [0,T]$ and all $x, y \in \mathbb{R}$, we get $u_k'(t) \leqslant 0$ on $[0,T]$. Condition (6.33) and $u_k(T) = 0$ give $u_k(T) - u_k(t) \geqslant -c(T-t)$, which yields $u_k'(T) \geqslant -c$. Since u_k' is nonincreasing on $[0,T]$, we have proved

$$-c \leqslant u_k'(t) \leqslant 0 \quad \text{on } [0,T]. \tag{6.34}$$

Due to $u_k'(0) = 0$, there is $t_k \in (0,T]$ such that

$$-\nu \leqslant u_k'(t) \leqslant 0 \quad \text{for } t \in [0,t_k].$$

If $t_k \geqslant \nu$, we get by (6.29)

$$u_k'(t) \leqslant -\varepsilon t \quad \text{for } t \in [0,\nu]. \tag{6.35}$$

Assume that $t_k < \nu$ and $u_k'(t) < -\nu$ for $t \in (t_k, \nu]$. Then

$$u_k'(t) \leqslant -\varepsilon t \quad \text{for } t \in [0,t_k].$$

Since $-\nu < -\varepsilon t$ for $t \in (t_k, \nu]$, we get (6.35) again. Integrating (6.35) over $[0,\nu]$ and using the concavity of u_k on $[0,T]$ we deduce that

$$\frac{\varepsilon \nu^2}{2T}(T-t) \leqslant u_k(t) \quad \text{on } [0,T]. \tag{6.36}$$

Step 3. *Convergence of a sequence of approximate solutions.*

Consider the sequence $\{u_k\}$. Choose an arbitrary compact interval $J \subset (0,T)$. By virtue of (6.33)–(6.36) there is $k_0 \in \mathbb{N}$ such that for each $k \in \mathbb{N}$, $k \geqslant k_0$,

$$\frac{c}{k_0} \leqslant u_k(t) \leqslant c(T-t), \quad -c \leqslant u_k'(t) \leqslant -\frac{\varepsilon}{k_0} \quad \text{on } J, \tag{6.37}$$

and hence there is $\psi \in L_1(J)$ such that

$$\bigl|f_k\bigl(t, u_k(t), u_k'(t)\bigr)\bigr| \leqslant \psi(t) \quad \text{for a.e. } t \in J. \tag{6.38}$$

Using conditions (6.33), (6.34), (6.38), the Arzelà–Ascoli theorem and the diagonalization principle, we can choose $u \in C[0,T] \cap C^1(0,T)$ and a subsequence of $\{u_k\}$ which we denote for the sake of simplicity in the same way such that

$$\begin{cases} \lim\limits_{k\to\infty} u_k = u \quad \text{uniformly on } [0,T], \\ \lim\limits_{k\to\infty} u_k' = u' \quad \text{locally uniformly on } (0,T). \end{cases} \tag{6.39}$$

Therefore we have $u(T) = 0$.

Step 4. *Convergence of the sequence of approximate problems.*

Choose an arbitrary $\xi \in (0,T)$ such that

$$f(\xi, \cdot, \cdot) \text{ is continuous on } (0,\infty) \times (-\infty, 0).$$

By (6.37) there exist a compact interval $J^* \subset (0,T)$ and $k_* \in \mathbb{N}$ such that $\xi \in J^*$ and for each $k \geqslant k_*$

$$u_k(\xi) > \frac{c}{k_*}, \quad u_k'(\xi) < -\frac{\varepsilon}{k_*}, \quad J^* \subset \left[\frac{1}{k}, T - \frac{1}{k}\right].$$

Therefore

$$f_k\bigl(\xi, u_k(\xi), u_k'(\xi)\bigr) = f\bigl(\xi, u_k(\xi), u_k'(\xi)\bigr)$$

and, due to (6.39),

$$\lim_{k\to\infty} f_k\bigl(t, u_k(t), u_k'(t)\bigr) = f\bigl(t, u(t), u'(t)\bigr) \quad \text{for a.e. } t \in (0,T). \tag{6.40}$$

Choose an arbitrary $t \in (0,T)$. Then there exists a compact interval $J \subset (0,T)$ such that (6.38) holds for all sufficiently large k. By virtue of (6.32) we get

$$u_k'\left(\frac{T}{2}\right) - u_k'(t) = \int_{T/2}^{t} f_k\bigl(s, u_k(s), u_k'(s)\bigr)\,\mathrm{d}s.$$

Letting $k \to \infty$ and using (6.38), (6.39), (6.40) and the Lebesgue convergence theorem on J, we get

$$u'\left(\frac{T}{2}\right) - u'(t) = \int_{T/2}^{t} f\big(s, u(s), u'(s)\big)\,\mathrm{d}s \quad \text{for each } t \in (0, T). \tag{6.41}$$

Therefore $u \in AC^1_{loc}(0, T)$ satisfies

$$u''(t) + f\big(t, u(t), u'(t)\big) = 0 \quad \text{for a.e. } t \in (0, T). \tag{6.42}$$

Further, according to (6.32) and (6.34) we have for each $k \geqslant \frac{3}{T}$

$$\int_0^T f_k\big(s, u_k(s), u_k'(s)\big)\,\mathrm{d}s = -u_k'(T) \leqslant c,$$

which together with (6.28), (6.33), (6.34) and (6.40) yields, by the Fatou lemma, that $f(t, u(t), u'(t)) \in L_1[0, T]$. Therefore, by (6.42), $u \in AC^1[0, T]$. Moreover, for each $k \geqslant \frac{3}{T}$ and $t \in (0, T)$

$$\begin{aligned} \left|u_k'(t)\right| \leqslant{}& \int_0^t \left|f_k\big(s, u_k(s), u_k'(s)\big) - f\big(s, u(s), u'(s)\big)\right|\mathrm{d}s \\ &+ \int_0^t \left|f\big(s, u(s), u'(s)\big)\right|\mathrm{d}s. \end{aligned}$$

Hence by (6.39) and (6.40) for each $\varepsilon > 0$ there exists $\delta > 0$ and for each $t \in (0, \delta)$ there exists $k_0 = k_0(\varepsilon, t) \in \mathbb{N}$ such that

$$\left|u'(t)\right| \leqslant \left|u'(t) - u_{k_0}'(t)\right| + \left|u_{k_0}'(t)\right| < \varepsilon.$$

It means that $u'(0) = \lim_{t\to 0+} u'(t) = 0$. We have proved that u is a solution of problem (6.20). ☐

EXAMPLE. Let $\alpha > 0$, $\beta \geqslant 0$ be arbitrary numbers. By Theorem 6.7 problem (6.18) has a solution $u \in AC^1[0, 1]$ satisfying

$$0 < u(t) \leqslant 1 - t, \quad -1 \leqslant u'(t) < 0 \quad \text{for } t \in (0, 1).$$

Note that Theorem 6.7 guarantees solvability of problem (6.18) even for the nonlinearity (6.19) having a strong space singularity ($\alpha \geqslant 1$) at $x = 0$.

References

[1] R.P. Agarwal, *Boundary Value Problems for Higher Order Differential Equations*, World Scientific, Singapore (1986).

[2] R.P. Agarwal, H. Lü and D. O'Regan, *An upper and lower solution method for the one-dimensional singular p-Laplacian*, Memoirs on Differential Equations and Math. Phys. **28** (2003), 13–31.

[3] R.P. Agarwal and D. O'Regan, *Singular boundary value problems for superlinear second order ordinary and delay differential equations*, J. Differential Equations **130** (1996), 333–355.

[4] R.P. Agarwal and D. O'Regan, *Nonlinear superlinear singular and nonsingular second order boundary value problems*, J. Differential Equations **143** (1998), 60–95.

[5] R.P. Agarwal and D. O'Regan, *Positive solutions for $(p, n-p)$ conjugate boundary value problems*, J. Differential Equations **150** (1998), 462–473.

[6] R.P. Agarwal and D. O'Regan, *Twin solutions to singular Dirichlet problems*, J. Math. Anal. Appl. **240** (1999), 433–445.

[7] R.P. Agarwal and D. O'Regan, *Right focal singular boundary value problems*, Z. Angew. Math. Mech. **79** (1999), 363–373.

[8] R.P. Agarwal and D. O'Regan, *Twin solutions to singular boundary value problems*, Proc. Amer. Math. Soc. **128** (2000), 2085–2094.

[9] R.P. Agarwal and D. O'Regan, *Multiplicity results for singular conjugate, focal, and (n, p) problems*, J. Differential Equations **170** (2001), 142–156.

[10] R.P. Agarwal and D. O'Regan, *Singular Differential and Integral Equations with Applications*, Kluwer, Dordrecht (2003).

[11] R.P. Agarwal and D. O'Regan, *A survey of recent results for initial and boundary value problems singular in the dependent variable*, Handbook of Differential Equations, Ordinary Differential Equations, Vol. 1, A. Cañada, P. Drábek and A. Fonda, eds., Elsevier/North-Holland, Amsterdam (2004), 1–68.

[12] R.P. Agarwal, D. O'Regan and V. Lakshmikantham, *Singular $(p, n-p)$ focal and (n, p) higher order boundary value problems*, Nonlinear Anal. Theory Methods Appl. **42** (2000), 215–228.

[13] R.P. Agarwal, D. O'Regan, V. Lakshmikantham and S. Leela, *Existence of positive solutions for singular initial and boundary value problems via the classical upper and lower solution approach*, Nonlinear Anal. Theory Methods Appl. **50** (2002), 215–222.

[14] R.P. Agarwal, D. O'Regan, I. Rachůnková and S. Staněk, *Two-point higher order BVPs with singularities in phase variables*, Comput. Math. Appl. **46** (2003), 1799–1826.

[15] R.P. Agarwal, D. O'Regan and S. Staněk, *Existence of positive solutions for boundary-value problems with singularities in phase variables*, Proc. Edinburgh Math. Soc. **47** (2004), 1–13.

[16] R.P. Agarwal, D. O'Regan and S. Staněk, *General existence principles for nonlocal boundary value problems with ϕ-Laplacian and their applications*, Abstr. Anal. Appl. **2006** (2006), article ID 96826, 30 pp.

[17] R.P. Agarwal and S. Staněk, *Nonnegative solutions of singular boundary value problems with sign changing nonlinearities*, Comput. Math. Appl. **46** (2003), 1827–1837.

[18] R.P. Agarwal and R.A. Usmani, *Iterative methods for solving right focal point boundary value problems*, J. Comput. Appl. Math. **14** (1986), 371–390.

[19] R.P. Agarwal and R.A. Usmani, *On the right focal point boundary value problems for integro-differential equations*, J. Math. Anal. Appl. **126** (1987), 51–69.

[20] R.P. Agarwal and P.J.Y. Wong, *Existence of solutions for singular boundary value problems for higher order differential equations*, Rend. Semi. Mat. Fis. Milano **65** (1995), 249–264.

[21] R. Aris, *The Mathematical Theory of Diffusion and Reaction in Permeable Catalysts*, Clarendon Press, Oxford (1975).

[22] C. Atkinson and J.E. Bouillet, *Some quantitative properties of solutions of a generalized diffusion equation*, Proc. Cambridge Phil. Soc. **86** (1979), 495–510.

[23] R.G. Bartle, *A Modern Theory of Integration*, Amer. Math. Soc., Providence, RI (2001).

[24] J.V. Baxley, *Some singular nonlinear boundary value problems*, SIAM J. Math. Anal. **22** (1991), 463–479.

[25] J.V. Baxley, *Numerical solution of singular nonlinear boundary value problems*, Proceedings of the Third International Colloquium on Numerical Analysis, D. Bainov and V. Covachev, eds., VSP, Utrecht (1995), 15–24.

[26] J.V. Baxley, *A singular nonlinear boundary value problem: membrane response of a spherical cap*, SIAM J. Appl. Math. **48** (1988), 497–505.

[27] J.V. Baxley and G.S. Gersdorff, *Singular reaction–diffusion boundary value problem*, J. Differential Equations **115** (1995), 441–457.

[28] J.V. Baxley and H.B. Thompson, *Boundary behavior and computation of solutions of singular nonlinear boundary value problems*, Comm. Appl. Anal. **4** (2000), 207–226.

[29] H. Berestycki, P.L. Lions and L.A. Peletier, *An ODE approach to the existence of positive solutions for semilinear problems in* $\mathbb{R}^N$, Indiana Univ. Math. J. **30** (1981), 141–157.

[30] F. Bernis, *Nonlinear parabolic equations arising in semiconductor and viscous droplets models*, Nonlinear Diffusion Equations and their Equilibrium States, N.G. Lloyd et al., eds., Birkhäuser, Boston (1992), 77–88.

[31] F. Bernis, *On some nonlinear singular boundary value problems of higher order*, Nonlinear Anal. Theory Methods Appl. **26** (1996), 1061–1078.

[32] F. Bernis, L.A. Peletier and S.M. Williams, *Source-type solutions of a fourth order nonlinear degenerate parabolic equation*, Nonlinear Anal. Theory Methods Appl. **18** (1992), 217–234.

[33] A.L. Bertozzi, M.P. Brenner, T.F. Dupont and L.P. Kadanoff, *Singularities and Similarities in Interface Flows*, Trends and Perspectives in Applied Mathematics, L. Sirovich, ed., Appl. Math. Sci., Vol. 100, Springer, Berlin (1994), 155–208.

[34] V. Bevc, J.L. Palmer and C. Süskind, *On the design of the transition region of axi-symmetric magnetically beam valves*, J. Br. Inst. Radio Eng. **18** (1958), 697–708.

[35] L.E. Bobisud, *Behavior of solutions for a Robin problem*, J. Differential Equations **85** (1990), 91–104.

[36] L.E. Bobisud, *Steady-state turbulent flow with reaction*, Rocky Mountain J. Math. **21** (1991), 993-1007.

[37] L.E. Bobisud, D. O'Regan and W.D. Royalty, *Solvability of some nonlinear boundary value problems*, Nonlinear Anal. Theory Methods Appl. **12** (1988), 855–869.

[38] D. Bonheure, C. De Coster, *Forced singular oscillators and the method of lower and upper functions*, Topol. Methods Nonlinear Anal. **22** (2003), 297–317.

[39] D. Bonheure, C. Fabry and D. Smets, *Periodic solutions of forced isochronous oscillators*, Discrete Contin. Dyn. Syst. **8** (2002), 907–930.

[40] J.E. Bouillet and C. Atkinson, *A generalized diffusion equation: Radial symmetries and comparison theorems*, J. Math. Anal. Appl. **95** (1983), 37–68.

[41] A. Cabada and R. Pouso, *Existence result for the problem* $(\phi(u'))' = f(t, u, u')$ *with periodic and Neumann boundary conditions*, Nonlinear Anal. Theory Methods Appl. **30** (1997), 1733–1742.

[42] A. Cabada, P. Habets and R. Pouso, *Lower and upper solutions for the periodic problem associated with a* ϕ*-Laplacian equation*, EQUADIFF 1999—International Conference on Differential Equations, Vols. 1, 2 (Berlin, 1999), World Sci. Publishing, River Edge, NJ (2000), 491–493.

[43] A. Callegari and M. Friedman, *An analytical solution of a nonlinear singular boundary value problem in the theory of viscous fluids*, J. Math. Anal. Appl. **21** (1968), 510–529.

[44] A. Callegari and A. Nachman, *Some singular, nonlinear differential equations arising in boundary layer theory*, J. Math. Anal. Appl. **64** (1978), 96–105.

[45] A. Callegari and A. Nachman, *A nonlinear singular boundary value problem in the theory of pseudoplastic fluids*, SIAM J. Appl. Math. **38** (1980), 275–281.

[46] A. Capietto, J. Mawhin and F. Zanolin, *A continuation approach to superlinear periodic boundary value problems*, J. Differential Equations **88** (1990), 347–395.

[47] C. De Coster and P. Habets, *Lower and upper solutions in the theory of ode boundary value problems: Classical and recent results*, Nonlinear Analysis and Boundary Value Problems for Ordinary Differential Equations, F. Zanolin, ed., CISM Courses and Lectures, Vol. 371, Springer, New York (1996), 1–79.

[48] C. De Coster and P. Habets, *The lower and upper solutions method for boundary value problems*, Handbook of Differential Equations, Ordinary Differential Equations, Vol. 1, A. Cañada, P. Drábek and A. Fonda, eds., Elsevier/North-Holland, Amsterdam (2004), 69–161.

[49] C. De Coster and P. Habets, *An overview of the method of lower and upper solutions for ODEs*, Nonlinear Analysis and its Applications to Differential Equations, M.R. Grossinho et al., eds., Progress in Nonlinear Differential Equations and their Applications, Vol. 43, Birkhäuser, Boston (2001), 3–22.

[50] K. Deimling, *Nonlinear Functional Analysis*, Springer, Berlin, Heidelberg (1985).

[51] M. del Pino and R. Manásevich, *Infinitely many* T*-periodic solutions for a problem arising in nonlinear elasticity*, J. Differential Equations **103** (1993), 260–277.

[52] M. del Pino, R. Manásevich and A. Montero, T*-periodic solutions for some second order differential equations with singularities*, Proc. Roy. Soc. Edinburgh Sect. A **120** (1992), 231–243.

[53] M. del Pino, R. Manásevich and A. Murúa, *Existence and multiplicity of solutions with prescribed period for a second order quasilinear O.D.E.*, Nonlinear Anal. Theory Methods Appl. **18** (1992), 79–92.

[54] L. Derwidué, *Systèmes differérentiels non lineáires ayant des solutions périodiques*, Acad. Roy. Belg. Bull. Cl. Sci. (5) **49** (1963), 11–32; **50** (1964), 928–942; **50** (1964), 1130–1142.

[55] T. Ding, *A boundary value problem for the periodic Brillouin focusing system*, Acta Sci. Natur. Univ. Pekinensis **11** (1965), 31–38 (in Chinese).

[56] W.Y. Ding, *A generalization of the Poincaré–Birkhoff theorem*, Proc. Amer. Math. Soc. **88** (1983), 341–346.

[57] P. Drábek, *Solvability and Bifurcation of Nonlinear Equations*, Pitman Res. Notes Math Ser., Vol. 264, Pitman, Boston (1992).

[58] P.W. Eloe and J. Henderson, *Singular nonlinear* $(n-1,1)$ *conjugate boundary value problems*, Georgian Math. J. **4** (1997), 401–412.

[59] P.W. Eloe and J. Henderson, *Singular nonlinear* $(k,n-k)$ *conjugate boundary value problems*, J. Differential Equations **133** (1997), 136–151.

[60] J.R. Esteban and J.L. Vazques, *On the equation of turbulent filtration in one-dimensional porous media*, Nonlinear Anal. Theory Methods Appl. **10** (1986), 1303–1325.

[61] C. Fabry and D. Fayyad, *Periodic solutions of second order differential equation with a* p*-Laplacian and asymmetric nonlinearities*, Rend. Istit. Mat. Univ. Trieste **24** (1992), 207–227.

[62] C. Fabry and P. Habets, *Lower and upper solutions for second-order boundary value problems with nonlinear boundary conditions*, Nonlinear Anal. Theory Methods Appl. **10** (1986), 985–1007.

[63] X.L. Fan and X. Fan, *A Knobloch-type result for* $p(t)$*-Laplacian systems*, J. Math. Anal. Appl. **282** (2003), 453–464.

[64] X.L. Fan, H.Q. Wu and F.Z. Wang, *Hartman-type results for* $p(t)$*-Laplacian systems*, Nonlinear Anal. Theory Methods Appl. **52** (2003), 585–594.

[65] R. Faure, *Solutions périodiques d'equations différentielles et mèthode de Laray–Schauder* (*Cas des vibrations forceés*), Ann. Inst. Fourier (Grenoble) **14** (1964), 195–204.

[66] A. Fonda, *Periodic solution of scalar second order differential equations with a singularity*, Acad Roy. Belg. Cl. Sci. Mém. (3) **4** (1993), 1–39.

[67] A. Fonda, *Periodic solution for a conservative system of differential equations with a singularity of repulsive type*, Nonlinear Anal. Theory Methods Appl. **24** (1995), 667–676.

[68] A. Fonda, R. Manásevich and F. Zanolin, *Subharmonic solutions for some second-order differential equations with singularities*, SIAM J. Math. Anal. **24** (1993), 1294–1311.

[69] N. Forbat and A. Huaux, *Détermination approachée et stabilité locale de la solution périodique d'une equation différentielle non linéaire*, Mém. Public. Soc. Sci. Arts Letters Hainaut **76** (1962), 3–13.

[70] S. Gaete and R.F. Manásevich, *Existence of a pair of periodic solutions of an O.D.E. generalizing a problem in nonlinear elasticity via variational methods*, J. Math. Anal. Appl. **134** (1988), 257–271.

[71] J.J. Garcia-Ripoll, V.M. Pérez-Garcia and P. Torres, *Extended parametric resonances in nonlinear Schrödinger systems*, Phys. Rev. Lett. **83** (1999), 1715–1718.

[72] J.A. Gatica, V. Oliker and P. Waltman, *Singular nonlinear boundary value problems for second-order ordinary differential equations*, J. Differential Equations **79** (1989), 62–78.

[73] W.G. Ge and J. Mawhin, *Positive solutions to boundary value problems for second order ordinary differential equations with singular nonlinearities*, Results. Math. **34** (1998), 108–119.

[74] B. Gidas, W. Ni and L. Nirenberg, *Symmetry of positive solutions of nonlinear elliptic equations in* $\mathbb{R}^N$, Adv. Math. Suppl. Studies **7A** (1981), 369–402.

[75] M.A. Goldberg, *An iterative solution for rotationally symmetric nonlinear membrane problems*, Internat. J. Non-Linear Mech. **1** (1966), 169–178.

[76] W.B. Gordon, *Conservative dynamical systems involving strong forces*, Trans. Amer. Math. Soc. **1975** (1975), 113–135.

[77] A. Granas, R.B. Guenther and J.W. Lee, *Some general existence principles in the Carathéodory theory of nonlinear systems*, J. Math. Pure Appl. **70** (1991), 153–196.

[78] P. Habets and L. Sanchez, *Periodic solutions of some Liénard equations with singularities*, Proc. Amer. Math. Soc. **109** (1990), 1035–1044.

[79] P. Habets and L. Sanchez, *Periodic solutions of dissipative dynamical systems with singular potentials*, Differential Integral Equations **3** (1990), 1039–1149.

[80] P. Habets and F. Zanolin, *Upper and lower solutions for a generalized Emden–Fowler equation*, J. Math. Anal. Appl. **181** (1994), 684–700.

[81] M.A. Herrero and J.L. Vazquez, *On the propagation properties of a nonlinear degenerate parabolic equation*, Comm. Partial Differential Equations **7** (1982), 1381–1402.

[82] A. Huaux, *Sur l'existence d'une solution périodique de l'equation différentielle non linéaire* $u'' + 0.2u' + u/(1-u) = 0.5\cos\omega t$, Acad. Roy. Belg. Bull. Cl. Sci. (5) **48** (1962), 494–504.

[83] J. Janus and J. Myjak, *A generalized Emden–Fowler equation with a negative exponent*, Nonlinear Anal. Theory Methods Appl. **23** (1994), 953–970.

[84] P. Jebelean and J. Mawhin, *Periodic solutions of singular nonlinear perturbations of the ordinary* p*-Laplacian*, Adv. Nonlinear Stud. **2** (2002), 299–312.

[85] P. Jebelean and J. Mawhin, *Periodic solutions of forced dissipative* p*-Liénard equations with singularities*, Vietnam J. Math. **32** (2004), 97–103.

[86] D.Q. Jiang, *Upper and lower solutions method and a singular superlinear boundary value problem for the one-dimensional* p*-Laplacian*, Comp. Math. Appl. **42** (2001), 927–940.

[87] D.Q. Jiang, *Upper and lower solutions method and a superlinear singular boundary value problems*, Comp. Math. Appl. **44** (2002), 323–337.

[88] D.Q. Jiang, J.F. Chu and M.R. Zhang, *Multiplicity of positive solutions to superlinear repulsive singular equations*, J. Differential Equations **21** (2005), 282–202.

[89] D.Q. Jiang and J.Y. Wang, *A generalized periodic boundary value problem for the one-dimensional* p*-Laplacian*, Ann. Polon. Math. **65** (1997), 265–270.

[90] R. Kannan and D. O'Regan, *Singular and nonsingular boundary value problems with sign changing nonlinearities*, J. Inequal. Appl. **5** (2000), 621–637.

[91] H.G. Kaper, M. Knaap and M.K. Kwong, *Existence theorems for second order boundary value problems*, Differential Integral Equations **4** (1991), 543–554.

[92] I.T. Kiguradze, *On Some Singular Boundary Value Problems for Ordinary Differential Equations*, Tbilisi Univ. Press, Tbilisi (1975) (in Russian).

[93] I.T. Kiguradze and B.L. Shekhter, *Singular boundary value problems for second order ordinary differential equations*, Itogi Nauki i Tekhniki Ser. Sovrem. Probl. Mat. Nov. Dost. **30** (1987), 105–201 (in Russian), translated in J. Soviet Math. **43** (1988), 2340–2417.

[94] G.S. Ladde, V. Lakshmikantham and A.S. Vatsala, *Monotone Iterative Techniques for Nonlinear Differential Equations*, Pitman, Boston (1995).

[95] A. Lasota, *Sur les problèmes linéaires aux limites pour un système d'équations différentielles ordinaires*, Bull. Acad. Polon. Sci. Sér. Math. Astron. Phys. **10** (1962), 565–570.

[96] A.C. Lazer and S. Solimini, *On periodic solutions of nonlinear differential equations with singularities*, Proc. Amer. Math. Soc. **99** (1987), 109–114.

[97] Bin Liu, *Multiplicity results for periodic solutions of a second order quasilinear ODE with asymmetric nonlinearities*, Nonlinear Anal. Theory Methods Appl. **33** (1998), 139–160.

[98] Bing Liu, *Periodic solutions of dissipative dynamical systems with singular potential and* p*-Laplacian*, Ann. Pol. Math. **79** (2002), 109–120.

[99] A. Lomtatidze, *Positive solutions of boundary value problems for second order differential equations with singular points*, Differentsial'nye Uravneniya **23** (1987), 1685–1692 (in Russian), translated in Differential Equations **23** (1987), 1146–1152.

[100] A. Lomtatidze and P. Torres, *On a two-point boundary value problem for second order singular equations*, Czechoslovak Math. J. **53** (2003), 19–43.

[101] R. Manásevich and J. Mawhin, *Periodic solutions for nonlinear systems with p-Laplacian-like operators*, J. Differential Equations **145** (1998), 367–393.

[102] J. Mawhin, *Topological degree and boundary value problems for nonlinear differential equations*, Topological Methods for Ordinary Differential Equations, M. Furi and P. Zecca, eds., Lecture Notes in Math., Vol. 1537, Springer, Berlin (1993), 74–142.

[103] J. Mawhin, *Leray–Schauder continuation theorems in the absence of a priori bounds*, Topol. Methods Nonlinear Anal. **9** (1997), 179–200.

[104] J. Mawhin, *Periodic solutions of systems with* p*-Laplacian-like operators*, Nonlinear Analysis and its Applications to Differential Equations, M.R. Grossinho et al., eds., Progress in Nonlinear Differential Equations and their Applications, Vol. 43, Birkhäuser, Boston (2001), 37–63.

[105] P. Martinez-Amores and P.J. Torres, *Dynamics of a periodic differential equation with a singular nonlinearity of attractive type*, J. Math. Anal. Appl. **202** (1996), 1027–1039.

[106] T.Y. Na, *Computational Methods in Engineering Boundary Value Problems*, Academic Press, New York (1979).

[107] M.N. Nkashama and J. Santanilla, *Existence of multiple solutions for some nonlinear boundary value problems*, J. Differential Equations **84** (1990), 148–164.

[108] P. Omari and W.Y. Ye, *Necessary and sufficient conditions for the existence of periodic solutions of second order ordinary differential equations with singular nonlinearities*, Differential Integral Equations **8** (1995), 1843–1858.

[109] D. O'Regan, *Some general existence principles and results for* $(\phi(y'))' = qf(t, y, y'),\ 0 < t < 1,$ SIAM J. Math. Anal. **24** (1993), 648–668.

[110] D. O'Regan, *Theory of Singular Boundary Value Problems*, World Scientific, Singapore (1994).

[111] N. Perrone and R. Kao, *A general nonlinear relaxation iteration technique for solving nonlinear problems in mechanics*, J. Appl. Mech. **38** (1971), 371–378.

[112] N. Phan-Thien, *A method to obtain some similarity solutions to the generalized Newtonian fluid*, Z. Angew. Math. Phys. **32** (1981), 609–615.

[113] E. Pinney, *The nonlinear differential equation* $y''(x) + p(x)y + cy^{-3} = 0$, Proc. Amer. Math. Soc. **1** (1950), 681.

[114] I. Rachůnková, *Lower and Upper Solutions and Topological Degree*, J. Math. Anal. Appl. **234** (1999), 311–327.

[115] I. Rachůnková, *Existence of two positive solutions of a singular nonlinear periodic boundary value problems*, J. Comput. Appl. Math. **113** (2000), 27–34.

[116] I. Rachůnková, *On the existence of more positive solutions of periodic BVPs with singularity*, Appl. Anal. **79** (2000), 257–275.

[117] I. Rachůnková, *Singular mixed boundary value problems*, J. Math. Anal. Appl., to appear.

[118] I. Rachůnková and S. Staněk, *Sign-changing solutions of singular Dirichlet boundary value problems*, Arch. Inequal. Appl. **1** (2003), 11–30.

[119] I. Rachůnková and S. Staněk, *Connections between types of singularities in differential equations and smoothness of solutions of Dirichlet BVPs*, Dynamics Continous Discrete Impulsive Systems **10** (2003), 209–222.

[120] I. Rachůnková and S. Staněk, *Sturm–Liouville and focal higher order BVPs with singularities in phase variables*, Georgian Math. J. **10** (2003), 165–191.

[121] I. Rachůnková and S. Staněk, *General existence principle for singular BVPs and its applications*, Georgian Math. J. **11** (2004), 549–565.

[122] I. Rachůnková and S. Staněk, *A singular boundary value problem for odd order differential equations*, J. Math. Anal. Appl. **291** (2004), 741–756.

[123] I. Rachůnková and S. Staněk, *Zeros of derivatives of solutions to singular* $(p, n-p)$ *conjugate BVPs*, Acta Univ. Palack. Olomuc. Fac. Rerum Natur. Math. **43** (2004), 137–141.

[124] I. Rachůnková and M. Tvrdý, *Nonlinear systems of differential inequalities and solvability of certain nonlinear second order boundary value problems*, J. Inequal. Appl. **6** (2001), 199–226.

[125] I. Rachůnková and M. Tvrdý, *Construction of lower and upper functions and their application to regular and singular boundary value problems*, Nonlinear Anal. Theory Methods Appl. **47** (2001), 3937–3948.

[126] I. Rachůnková and M. Tvrdý, *Second-order periodic problem with* ϕ*-Laplacian and impulses*, Nonlinear Anal. Theory Methods Appl., **63** (2005), e257–e266.

[127] I. Rachůnková and M. Tvrdý, *Periodic problems with* ϕ*-Laplacian involving non-ordered lower and upper functions*, Fixed Point Theory **6** (2005), 99–112.

[128] I. Rachůnková and M. Tvrdý, *Periodic singular problem with quasilinear differential operator*, Math. Bohem., to appear.

[129] I. Rachůnková, M. Tvrdý and I. Vrkoč, *Existence of nonnegative and nonpositive solutions for second order periodic boundary value problems*, J. Differential Equations **176** (2001), 445–469.

[130] I. Rachůnková, M. Tvrdý and I. Vrkoč, *Resonance and multiplicity in periodic BVPs with singularity*, Math. Bohem. **128** (2003), 45–70.

[131] L. Sanchez, *Positive solutions for a class of semilinear two-point boundary value problems*, Bull. Austral. Math. Soc. **45** (1992), 439–451.

[132] S. Solimini, *On forced dynamical systems with a singularity of repulsive type*, Nonlinear Anal. Theory Methods Appl. **14** (1990), 489–500.

[133] S. Staněk, *Periodic boundary value problem for second order functional differential equations*, Math. Notes (Miskolc) **1** (2000), 63–81.

[134] S. Staněk, *On solvability of singular periodic boundary value problems*, Nonlinear Oscil. **4** (2001), 529–538.

[135] S. Staněk, *Positive solutions of singular positone Dirichlet boundary value problems*, Math. Comput. Modelling **33** (2001), 341–351.

[136] S. Staněk, *Positive solutions to Dirichlet and periodic boundary value problems*, Comput. Math. Appl. **43** (2002), 681–692.

[137] S. Staněk, *A nonlocal singular boundary value problem for second-order differential equations*, Math. Notes (Miskolc) **5** (2004), 91–104.

[138] S. Staněk, *A nonlocal boundary value problem with singularities in phase variables*, Math. Comput. Modelling **40** (2004), 101–116.

[139] S. Staněk, *Positive solutions of the Dirichlet problem with state-dependent functional differential equations*, Funct. Differential Equations **11** (2004), 563–586.

[140] S. Staněk, *Singular nonlocal boundary value problems*, Nonlinear Anal. Theory Methods Appl., in press.

[141] S.D. Taliaferro, *A nonlinear singular boundary value problem*, Nonlinear Anal. Theory Methods Appl. **3** (1979), 897–904.

[142] A. Tineo, *Existence theorems for a singular two-point Dirichlet problem*, Nonlinear Anal. Theory Methods Appl. **19** (1992), 323–333.

[143] P.J. Torres, *Bounded solutions in singular equations of repulsive type*, Nonlinear Anal. Theory Methods Appl. **32** (1998), 117–125.

[144] P.J. Torres, *Existence of one-signed periodic solutions of some second-order differential equations via a Krasnoselskii fixed point theorem*, J. Differential Equations **190** (2003), 643–662.

[145] P.J. Torres, *Existence and uniqueness of elliptic periodic solutions of the Brillouin electron beam focusing system*, Math. Methods Appl. Sci. **23** (2000), 1139–1143.

[146] N.I. Vasiliev and Ju.A. Klokov, *Foundation of the Theory of Boundary Value Problems for Ordinary Differential Equations*, Zinatne, Riga (1978) (in Russian).

[147] I. Vrkoč, *Comparison of two definitions of lower and upper functions of nonlinear second order differential equations*, J. Inequal. Appl. **6** (2001), 191–198.

[148] J.Y. Wang, *Solvability of singular nonlinear two-point boundary value problems*, Nonlinear Anal. Theory Methods Appl. **24** (1995), 555–561.

[149] J.Y. Wang and W. Gao, *A singular boundary value problem for the one-dimensional p-Laplacian*, J. Math. Anal. Appl. **201** (1996), 851–866.

[150] P. Yan, *Nonresonance for one-dimensional p-Laplacian with regular restoring*, J. Math. Anal. Appl. **285** (2003), 141–154.

[151] P. Yan and M.R. Zhang, *Higher order non-resonance for differential equations with singularities*, Math. Methods Appl. Sci. **26** (2003), 1067–1074.

[152] M.R. Zhang, *Periodic solutions of Liénard equations with singular forces of repulsive type*, J. Math. Anal. Appl. **203** (1996), 254–269.

[153] M.R. Zhang, *Nonuniform nonresonance at the first eigenvalue of the p-Laplacian*, Nonlinear Anal. Theory Methods Appl. **29** (1997), 41–51.

[154] M.R. Zhang, *A relationship between the periodic and the Dirichlet BVPs of singular differential equations*, Proc. Roy. Soc. Edinburgh Sect. A **128** (1998), 1099–1114.

[155] M.R. Zhang, *Periodic solutions of damped differential systems with repulsive singular forces*, Proc. Amer. Math. Soc. **127** (1999), 401–407.

[156] G. Zhong-Heng and R. Solecki, *Free and forced finite amplitude oscillations of an elastic thick walled hollow sphere made in incompressible material*, Arch. Mech. Stos. **3.25** (1963), 427–433.

Author Index

Roman numbers refer to pages on which the author (or his/her work) is mentioned. Italic numbers refer to reference pages. Numbers between brackets are the reference numbers. No distinction is made between the first author and co-author(s).

Adams, R.A. 172, 188, *199* [1]; 266, *334* [1]
Addou, I. 594, *604* [1]; *604* [2]; *604* [3]
Adje, A. 207, *334* [2]
Agarwal, R.P. 3, *97* [1]; 611, 620, 623, 624, 627, 628, 631, 633, 647, 654, 655, 657, 661, 665, 668, 710–713, *717* [1]; *718* [2]; *718* [3]; *718* [4]; *718* [5]; *718* [6]; *718* [7]; *718* [8]; *718* [9]; *718* [10]; *718* [11]; *718* [12]; *718* [13]; *718* [14]; *718* [15]; *718* [16]; *718* [17]; *718* [18]; *718* [19]; *718* [20]
Agmon-Snir, H. 450, *536* [1]
Ahmad, S. 144, *199* [2]
Aiello, W.G. 460, 525, *536* [2]
Akhmet, M.U. 360, *428* [1]
Al-Omari, J.F.M. 457, *536* [3]
Alessio, F. 129–131, *199* [3]; *199* [4]; *199* [5]; *199* [6]
Alexander, J.C. 360, 399, *429* [28]; *429* [29]
Alikakos, N.D. 321, *334* [3]
Allgower, E.L. 604, *604* [4]
Alt, W. 465, 504, 507, *536* [4]
Alwash, M.A.M. 206, *334* [4]
Amann, H. 252–256, 261, 267, 294, 296, *334* [5]; *334* [6]; *334* [7]; *334* [8]; *334* [9]
Ambrosetti, A. 105, *199* [7]; *199* [8]; *199* [9]; 254, *334* [9]; 552, *604* [5]
an der Heiden, U. 450, *536* [5]
Andersen, K.M. 207, *334* [10]; *334* [11]
Andres, J. 4–7, 9, 12–19, 21, 23, 24, 26, 27, 30, 48, 49, 51, 52, 57, 58, 60, 61, 75, 76, 80, 86, 87, 89, 91, 95–97, *97* [2]; *97* [3]; *97* [4]; *97* [5]; *97* [6]; *97* [7]; *97* [8]; *97* [9]; *98* [10]; *98* [11]; *98* [12]; *98* [13]; *98* [14]; *98* [15]; *98* [16]; *98* [17]; *98* [18]; *98* [19]; *98* [20]; *98* [21]; *98* [22]; *98* [23]; *98* [24]; *98* [25]; *98* [26]; *98* [27]; *98* [28]; *98* [29]; *98* [30]; *98* [31]; *98* [32]; *98* [33]; *98* [34]; *99* [35]; *99* [36]; *99* [37]; *99* [38]
Andrewartha, H.G. 460, *536* [6]
Andronov, A.A. 343, 360, *428* [2]
Anichini, G. 95, *99* [39]
Anuradha, V. 576, *604* [6]
Appell, J. 13, *99* [40]
Arias, M. 107, 133, 137, *199* [10]
Arino, O. 457–461, 465, 470, 472, 482, 504, 507–513, 515, 522, *536* [7]; *536* [8]; *536* [9]; *536* [10]; *536* [11]; *539* [76]; *541* [132]; *541* [144]; *542* [148]
Aris, R. 709, *718* [21]
Arnold, V.I. 399, *428* [3]; *428* [4]
Aronson, D.G. 106, 132, *199* [11]
Aronszajn, N. 3, *99* [41]
Atiya, A.F. 450, *536* [19]
Atkinson, C. 647, 648, *718* [22]; *719* [40]
Attalah, S.H. 460, *545* [224]
Aubin, J.-P. 4, 7, 29, 39, 96, *99* [42]
Augustynowicz, A. 52, 96, *99* [43]

Bader, R. 7, 24, 95, 96, *98* [14]; *99* [44]; *99* [45]
Badiale, M. 105, *199* [7]; *199* [8]
Baĭkuzev, K.B. 251, *334* [12]
Baker, C.T.H. 526, 532, 535, *536* [12]; *536* [13]; *536* [14]; *536* [15]; *536* [16]; *536* [17]; *544* [211]; *544* [212]; *544* [213]; *544* [214]
Bakke, V.L. 526, 535, *536* [18]
Baldi, P. 450, *536* [19]
Bange, D.W. 255, *334* [13]
Bardi, M. 206, *334* [14]
Bartha, M. 515, 518, 522, 523, 525, *537* [20]; *537* [21]; *537* [22]
Bartle, R.G. 618, *718* [23]
Bautin, N.N. 399, *428* [5]

Baxley, J.V. 668, 710, 712, *718* [24]; *718* [25]; *718* [26]; *718* [27]; *719* [28]
Bebernes, J.W. 316, 320, *334* [15]; *334* [16]; 549, 562, *604* [7]
Bejarano, J.D. 360, *428* [6]
Bélair, J. 460, 462, 525, *537* [23]; *537* [24]; *537* [25]; *537* [26]; *542* [149]; *543* [185]
Bellen, A. 526, 528–530, 532, 533, 535, *537* [27]
Bellman, R. 251, 311, 320, *334* [17]; 526, *537* [28]
Beltramo, A. 332, *334* [18]
Benguria, R. 601, *604* [8]
Berestycki, H. 142, *200* [12]; 709, *719* [29]
Bernis, F. 610, 611, *719* [30]; *719* [31]; *719* [32]
Bersani, A.M. 96, *98* [15]
Bertotti, M.L. 129–131, *199* [3]; *200* [13]; *200* [14]
Bertozzi, A.L. 610, *719* [33]
Bevc, V. 700, 705, *719* [34]
Bi, P. 399, 426, 428, *428* [7]; *430* [53]; *430* [54]; *430* [55]; *430* [66]; *430* [67]
Bielawski, D. 53, 96, *99* [46]
Bin Liu 707, *721* [97]
Bing Liu 690, 707, *721* [98]
Birch, L.C. 460, *536* [6]
Birkhoff, G.D. 105, *200* [15]
Blows, T.R. 360, 399, *428* [8]
Bobisud, L.E. 647, 667, 709, *719* [35]; *719* [36]; *719* [37]
Bocharov, G.A. 526, 535, *536* [14]
Bochner, S. 38, *99* [48]
Bolotin, S.V. 105, *200* [16]; *200* [17]
Bonheure, D. 108, 159, 160, 199, *200* [18]; *200* [19]; *200* [20]; 706, 707, *719* [38]; *719* [39]
Borsuk, K. 9, *99* [47]
Bosetto, E. 130, *200* [21]
Bouillet, J.E. 647, 648, *718* [22]; *719* [40]
Brauer, F. 206, *334* [19]
Bravo, J.L. 207, *334* [20]; *334* [21]
Bravo de la Parra, R. 457, *536* [8]
Brenner, M.P. 610, *719* [33]
Bressan, A. 48, *99* [49]; *99* [50]
Brezis, H. 151, *200* [22]; 279, *334* [22]
Brokate, M. 465, 502, 503, *537* [29]
Browder, F.E. 251, *334* [23]; 504, *537* [30]
Brown, K.J. 550, 552, 594, *604* [9]
Brown, R.F. 15, *99* [51]
Brunner, H. 526, *537* [31]
Bryuno, A.D. 378, *428* [9]
Buffoni, B. 155, *200* [23]
Büger, M. 445, 446, *537* [32]; *537* [33]
Butcher, J.C. 526, 535, *537* [34]

Cabada, A. 671, 707, *719* [41]; *719* [42]
Cahlon, B. 526, *537* [35]; *537* [36]
Cai, S. 399, *428* [10]
Calanchi, M. 130, *200* [24]; 206, *335* [24]
Caldiroli, P. 131, *200* [25]
Callegari, A. 610, *719* [43]; *719* [44]; *719* [45]
Campos, J. 107, 133, 137, *199* [10]
Cantrell, R.S. 251, *335* [25]
Cao, J. 440, *544* [221]
Cao, Y. 452, 525, *537* [37]; *537* [38]; *537* [39]
Capietto, A. 705, *719* [46]
Carl, S. 252, *335* [26]
Carminati, C. 130, *199* [4]
Carr, C.E. 450, *537* [40]; *537* [41]
Carr, J. 144, *200* [32]
Castleton, R.N. 526, *537* [42]
Castro, A. 332, *335* [27]
Cecchi, M. 33, 95, *99* [52]
Cellina, A. 4, 7, 29, 39, 96, *99* [42]
Chan, H. 399, *431* [86]; *431* [87]
Chavarriga, J. 360, 399, *428* [11]
Chen, C.-N. 130, *200* [26]
Chen, F. 440, *537* [43]
Chen, G. 399, 428, *430* [56]; *431* [94]
Chen, J. 399, 424, 426, *430* [57]
Chen, L. 399, *428* [12]
Chen, W. 399, 426, *432* [122]; *432* [123]
Cheng, J. 552, 594, 598, *605* [10]; *605* [11]
Cheng, S.S. 438, *543* [188]
Cheng Fu, J. 207, *338* [133]
Cherkas, L.A. 399, 428, *428* [13]
Cherpion, M. 206, 207, 210, 214, *335* [28]
Chicone, C. 360, 376, 399, 425, 428, *429* [14]; *429* [15]
Chow, S.N. 3, 53, *99* [53]; 360, 399, 428, *429* [16]; *429* [17]; *429* [18]; 476, *537* [44]
Christopher, C.J. 360, 399, *429* [19]; *429* [20]; *429* [21]; *432* [103]
Chu, J.F. 707, *721* [88]
Chua, L.O. 398, 428, *432* [111]
Chung, K. 399, *431* [86]; *431* [87]
Churakova, S.V. 438, *539* [80]; *539* [81]; *539* [82]
Cid, J.Á. 243, *335* [29]
Coddington, E.A. 3, *99* [54]; 206, *339* [146]
Coleman, B.D. 206, *335* [30]
Coll, B. 360, 399, *429* [22]
Colonius, F. 465, 502, 503, *537* [29]
Conti, G. 48, 95, 96, *99* [39]; *99* [55]
Cooke, K.L. 438, 483, 524–526, *537* [28]; *537* [45]; *538* [46]; *538* [47]; *538* [48]
Corwin, S.P. 526, 529, 532, 535, *538* [49]
Cosner, C. 251, *335* [25]
Coti Zelati, V. 105, 129, 131, 155, *199* [9]; *200* [27]; *200* [28]; *200* [29]; *202* [72]
Cowan, J.D. 451, 452, *539* [72]; *539* [73]; *539* [74]

Crandall, M.G. 552, 553, *605* [12]; *605* [13]
Cushing, J.M. 206, *335* [31]

Daleckiĭ, J.L. 62, 63, *99* [56]
Dancer, E.N. 206, 252, 254, 256, 257, 290, 294, 298, 305, 310, 311, 318, 321, 323, *335* [32]; *335* [33]; *335* [34]; *335* [35]; *335* [36]
Daners, D. 254, 262, 318, *335* [37]
Davenport, M.R. 451, *538* [51]
Day, S.P. 451, *538* [51]
De Coster, C. 205–207, 210, 214, 251, 254–256, 321, 326, 327, *335* [28]; *335* [38]; *335* [39]; *335* [40]; *335* [41]; *335* [42]; 649, 670, 707, *719* [38]; *719* [47]; *719* [48]; *719* [49]
De Mottoni, P. 206, *335* [43]
de Oliveira, J.F.C. 497, *539* [95]
De Pascale, E. 13, *99* [40]
de Roos, A.M. 459, *538* [52]
DeBlasi, F.S. 45, 96, *99* [57]
Dee, G.T. 145, *200* [30]
Deimling, K. 7, 12, 13, 44, 49, 61, 96, *99* [58]; 633, *719* [50]
del Pino, M. 702, 704, 706, 707, *719* [51]; *719* [52]; *720* [53]
Depassier, M.C. 601, *604* [8]
Derwidué, L. 705, *720* [54]
Desch, W. 525, *538* [53]
Deuel, J. 252, *335* [44]
Devaney, R.L. 155, *200* [31]
Diekmann, O. 438, 457, 462, 466, 467, 484, 485, 488, 494, 505, 506, 514, *538* [54]; *542* [162]
Ding, T. 360, 428, *433* [127]; 701, *720* [55]
Ding, W.Y. 704, *720* [56]
Domoshnitsky, A. 440, 525, *538* [55]; *538* [56]
Dong, G.C. 262, *335* [45]
Dong, Z. 360, 428, *433* [127]
Drábek, P. 689, *720* [57]
Dragoni, R. 96, *99* [59]
Drakhlin, M. 440, 525, *538* [55]; *538* [56]
Driver, R.D. 441–444, 465, 499, *538* [57]; *538* [58]; *538* [59]; *538* [60]; *538* [61]; *538* [62]; *538* [63]; *538* [64]; *540* [111]
Du, Y. 550, *605* [14]
Dugundji, J. 7, 9, 81, *100* [69]
Dumortier, F. 360, 399, 428, *429* [23]; *429* [24]; *429* [25]; *429* [26]
Dupont, T.F. 610, *719* [33]
Dzedzej, Z. 52, 96, *99* [43]

Eberly, D. 549, 562, *604* [7]
Eckmiller, R. 451, *542* [166]
Eichmann, M. 439, 495–497, *538* [65]
Ekeland, I. 105, *200* [27]
El Hachimi, A. 253, *335* [46]
El Hadi Ait Dads 440, *538* [50]
Eloe, P.W. 633, *720* [58]; *720* [59]
Engelborghs, K. 526, 535, *538* [66]; *541* [145]
Engelking, R. 217, 243, 281, *335* [47]
Enright, W.H. 526, 528, 529, 532, 535, *538* [67]; *538* [68]; *538* [69]; *538* [70]; *539* [71]
Ernst, K. 451, 452, *539* [74]
Ernst, U. 451, 452, *539* [73]
Esteban, J.R. 647, 648, *720* [60]
Eurich, C.W. 451, 452, *539* [72]; *539* [73]; *539* [74]; *539* [75]
Ezzinbi, K. 440, *538* [50]

Fabry, C. 670, 706, 707, *719* [39]; *720* [61]; *720* [62]
Fan, J. 525, *537* [37]
Fan, X. 647, *720* [63]
Fan, X.L. 647, *720* [63]; *720* [64]
Fathallah, A. 459–461, *536* [11]; *539* [76]
Faure, R. 705, *720* [65]
Fayyad, D. 707, *720* [61]
Fei, N. 144, *200* [32]
Feldstein, A. 526, 532–535, *539* [77]; *539* [78]; *539* [79]; *542* [168]
Feng, B. 399, 424, *429* [27]
Fernández, M. 207, *334* [20]; *334* [21]
Fiedler, B. 360, 399, *429* [28]; *429* [29]; *429* [30]
Fields, R.D. 451, *543* [193]
Fife, P.C. 142, *200* [33]; 251, *335* [48]; *335* [49]
Filippov, A.F. 6, 46, *99* [60]; 208, 229, 235, *335* [50]
Fišer, J. 4, *98* [16]
Fisher, R.A. 106, *200* [34]
Fonda, A. 706, 707, *720* [66]; *720* [67]; *720* [68]
Forbat, N. 705, *720* [69]
Freedman, H.I. 460, 525, *536* [2]
Friedman, A. 269, *335* [51]
Friedman, M. 610, *719* [43]
Frigon, M. 7, 96, *99* [61]; 207, 210, *336* [52]
Fučík, S. 7, 96, 97, *99* [62]
Furi, M. 7, 33, 95, 96, *99* [52]; *99* [63]; *100* [64]
Fürst, T. 89, 91, 97, *98* [17]; *98* [18]

Gabasov, R. 438, *539* [80]; *539* [81]; *539* [82]
Gabor, G. 7, 14, 26, 95, 96, *98* [19]; *98* [20]; *98* [21]
Gaete, S. 706, 707, *720* [70]
Gaines, R.G. 7, 96, 97, *100* [65]
Gallardo, J. 525, 526, *539* [83]; *539* [84]
Gambell, R. 460, *539* [85]
Gao, W. 668, *723* [149]
Garcia-Margallo, J. 360, *428* [6]
Garcia-Ripoll, J.J. 705, *720* [71]

Gard, T.C. 525, *537* [37]
Gasull, A. 360, 368, 399, *429* [22]; *429* [31]; *429* [32]; *429* [33]
Gatica, J.A. 667, 709, *720* [72]
Gatica, J.G. 456, 465, 525, *539* [86]; *539* [87]; *539* [88]; *539* [89]
Gause, G.F. 207, *336* [53]
Ge, W.G. 705, 706, *720* [73]
Gel'man, B.D. 7, 49, 52, 96, *99* [43]; *99* [45]; *100* [66]
Gersdorff, G.S. 710, *718* [27]
Giacomini, H. 360, 399, *429* [34]
Gidas, B. 551, 554, 585, 588, *605* [15]; 709, *720* [74]
Gilding, B.H. 144, *200* [35]
Glünder, H. 451, *540* [113]
Goldberg, M.A. 712, *720* [75]
Gompper, G. 108, 183, 184, 194, *200* [36]
Gordon, I.I. 343, 360, *428* [2]
Gordon, W.B. 706, *720* [76]
Górniewicz, L. 5–7, 9, 12–19, 21, 23, 24, 26, 30, 47–49, 51, 52, 57, 58, 60, 61, 75, 76, 80, 86, 87, 95–97, *98* [19]; *98* [20]; *98* [21]; *98* [22]; *98* [23]; *98* [24]; *98* [25]; *98* [26]; *98* [27]; *100* [67]; *100* [68]
Gossez, J.P. 254, *336* [54]
Gourley, S.A. 457, *536* [3]
Granas, A. 7, 9, 81, 96, 97, *99* [61]; *100* [69]; *100* [70]; 624, *720* [77]
Grimm, L.J. 526, *537* [42]
Guenther, R.B. 96, 97, *100* [70]; 624, *720* [77]
Guglielmi, N. 526, 535, *539* [90]
Guillamon, A. 360, 368, 399, *429* [31]
Guo, G. 399, *428* [10]
Gurney, W.S.C. 457, *542* [170]
Győri, I. 483, 524, 526, 532, *539* [91]; *539* [92]; *539* [93]

Habets, P. 108, 159, 160, *200* [19]; *200* [37]; 205, 254, 261, *335* [39]; *336* [55]; 649, 651, 653, 664, 665, 668, 670, 706, 707, *719* [42]; *719* [47]; *719* [48]; *719* [49]; *720* [62]; *720* [78]; *720* [79]; *721* [80]
Hadeler, K.P. 458, 465, 504, 507–513, *536* [7]
Hairer, E. 526, 535, *539* [90]
Halanay, H. 437, *539* [94]
Hale, J.K. 105, *200* [38]; 207, 290, *336* [56]; *336* [57]; 360, 398, 399, 428, *429* [16]; *429* [35]; *429* [36]; *429* [37]; 438, 466, 467, 476, 497, 501, 503, *539* [95]; *539* [96]; *540* [97]; *540* [98]
Hamel, F. 144, *200* [39]
Han, M. 360, 398, 399, 412, 424–428, *428* [7]; *429* [38]; *429* [39]; *429* [40]; *429* [41]; *429* [42]; *430* [43]; *430* [44]; *430* [45]; *430* [46]; *430* [47]; *430* [48]; *430* [49]; *430* [50]; *430* [51]; *430* [52]; *430* [53]; *430* [54]; *430* [55]; *430* [56]; *430* [57]; *430* [58]; *430* [59]; *430* [60]; *430* [61]; *430* [62]; *430* [63]; *430* [64]; *430* [65]; *430* [66]; *430* [67]; *430* [68]; *430* [69]; *430* [70]; *430* [71]; *431* [72]; *431* [80]; *432* [100]; *432* [101]; *432* [113]; *432* [114]; *432* [120]; *432* [121]; *432* [123]; *432* [124]; *432* [125]; *433* [126]
Hardy, G.H. 138, *201* [40]
Hartman, P. 3, 48, 58–60, 95, *100* [71]; 146, *201* [41]
Hartung, F. 438, 448, 450, 465, 483, 499, 500, 502–504, 524–526, 532, 535, *539* [91]; *539* [92]; *539* [93]; *540* [99]; *540* [100]; *540* [101]; *540* [102]; *540* [103]; *540* [104]; *540* [105]; *540* [106]; *540* [107]; *540* [108]; *540* [116]
Hauber, R. 526, *540* [109]
Hayashi, H. 526, 532, 535, *538* [68]; *538* [69]
Hbid, M.L. 457, 458, 465, 470, 472, 504, 507–513, 515, 521, 522, *536* [7]; *536* [8]; *541* [144]; *543* [178]
Henderson, J. 633, *720* [58]; *720* [59]
Henrard, M. 254, *335* [40]
Henry, D. 43, *100* [72]; 556, *605* [16]
Herdman, T.L. 526, 535, *540* [102]; *540* [103]
Herrero, M.A. 647, *721* [81]
Hess, P. 208, 229, 251, 252, 254, 256, 257, 290, 294, 298, 305, 310, 311, 318, 321, 323, 332, *334* [3]; *334* [18]; *335* [33]; *335* [34]; *335* [35]; *335* [44]; *336* [58]
Higham, D.J. 528, 532, *540* [110]
Hilbert, D. 343, 399, *431* [73]
Hirsch, M.W. 205, 208, 220, 229, 251, 252, 256, 257, 284, 305, 310, 318, *336* [59]; *336* [60]; *336* [61]
Hoag, J.T. 443, 444, 465, *540* [111]
Hohenberg, P.C. 172, *202* [84]
Hopfield, J.J. 450, *543* [199]
Hoppensteadt, F.C. 456, *540* [112]
Horozov, E. 399, 428, *431* [74]
Hou, X. 577, *605* [17]
Hu, B. 399, *431* [95]
Hu, M. 526, 529, 532, *538* [70]
Hu, S. 7, 13, 50, 87, 96, *100* [73]; *100* [74]; 399, 424, 426, *430* [58]
Huang, Q. 399, *431* [88]
Huang, W. 360, 399, 428, *431* [75]; *433* [127]; 438, 483, 524, *538* [47]; *538* [48]
Huaux, A. 705, *720* [69]; *721* [82]
Hukuhara, M. 3, *100* [75]

Hulshof, J. 158, *202* [90]
Hüning, H. 451, *540* [113]
Hunter, J. 452, 454, *540* [114]

Ibrahim, M.M.A. 550, 552, 594, *604* [9]
Iliev, I.D. 399, 428, *431* [74]; *431* [76]; *431* [77]
Ilyashenko, Yu.S. 206, *336* [62]; *336* [63]; 399, *431* [78]
Innocenti, G.M. 451, *540* [115]
Insperger, T. 447, 448, 450, *540* [116]; *540* [117]
Ito, K. 465, 532, *540* [118]

Jackiewicz, J. 526, *540* [120]
Jackiewicz, Z. 465, 526, 535, *536* [18]; *540* [119]; *540* [121]; *541* [122]
Jackson, K.R. 528, 529, *539* [71]
Jacobs, M. 360, 399, 428, *429* [15]
James, E.M. 399, *431* [79]
Jansen, M. 451, *542* [166]
Janus, J. 668, *721* [83]
Jarník, J. 47, *100* [76]
Jeanjean, L. 131, *199* [5]; *199* [6]; *200* [13]; *200* [25]
Jebelean, P. 685, 688, 705, 707, *721* [84]; *721* [85]
Jezierski, J. 18, 19, 21, 95, 97, *98* [23]; *98* [24]; *98* [25]; *98* [26]
Jiang, B. 18, 21, *100* [77]
Jiang, D.Q. 652, 664, 668, 707, *721* [86]; *721* [87]; *721* [88]; *721* [89]
Jiang, K. 399, *430* [59]
Jiang, Q. 398, 399, *431* [80]
Jiang, T. 399, *431* [96]
Johnson, R.A. 447, *541* [123]
Joyal, P. 399, *431* [81]
Jüttner, L. 4, 96, *98* [16]; *98* [28]; *98* [29]

Kadanoff, L.P. 610, *719* [33]
Kalies, W.D. 107, 152, 155, 156, 158, 186, 189, 190, 198, 199, *201* [42]; *201* [43]
Kamenskii, M.I. 7, 12, 13, 41, 48, 61, 62, 96, *99* [45]; *100* [78]
Kannan, R. 713, *721* [90]
Kao, R. 712, *722* [111]
Kaper, H.G. 647, *721* [91]
Kappel, F. 465, 532, *540* [118]
Karimov, D.H. 251, *334* [12]; *336* [64]; *336* [65]; *336* [66]; *336* [67]
Karoui, A. 526, 532, 535, *541* [124]; *541* [125]
Kazarinoff, N.D. *606* [62]
Kazdan, J.L. 253, *336* [68]
Kelley, W. 144, *201* [44]
Kersner, R.A. 144, *200* [35]
Kiguradze, I.T. 611, 649, 668, 670, 714, *721* [92]; *721* [93]
Kisielewicz, M. 6, 7, *100* [79]
Klokov, Ju.A. 619, 649, 663, *723* [146]
Knaap, M. 647, *721* [91]
Kneser, H. 3, *100* [80]
Koçak, H. 105, *200* [38]; 207, *336* [57]; 398, *429* [36]
Koch Medina, P. 254, 262, 318, *335* [37]
Kolesov, Ju.S. 206, 251, 252, 256, 294, 296, 318, *336* [69]; *336* [70]; *336* [71]; *336* [72]; *336* [73]
Kolmogorov, A. 106, 132, *201* [45]
Konishi, M. 450, *537* [41]
Korman, P. 130, *201* [46]; 550–552, 557, 561–566, 569–573, 575–579, 581, 584–593, 596, 598–600, 602, 604, *605* [17]; *605* [18]; *605* [19]; *605* [20]; *605* [21]; *605* [22]; *605* [23]; *605* [24]; *605* [25]; *605* [26]; *605* [27]; *605* [28]; *605* [29]; *605* [30]; *605* [31]; *605* [32]; *605* [33]; *605* [34]; *605* [35]; *605* [36]; *606* [37]; *606* [38]; *606* [39]; *606* [40]
Kostitzin, V.A. 206, *336* [74]
Krajc, B. 97, *98* [30]; *98* [31]
Krasnosel’skiĭ, M.A. 7, 59, 96, 97, *100* [81]; *100* [82]; 251, 255, 273, 277, *336* [75]; *336* [76]; *336* [77]
Krawcewicz, W. 7, 96, *100* [83]
Krein, M.G. 62, 63, *99* [56]
Krishnan, H.P. 482, *541* [126]; *541* [127]
Krisztin, T. 439, 440, 465, 470, 476, 477, 482, 485, 489, 494, 495, 515, 518, 524, 525, *541* [128]; *541* [129]; *541* [130]; *541* [131]; *541* [132]; *541* [133]; *541* [134]
Kuang, Y. 440, 456, 465, 504, 507, 512, 524, *541* [135]; *541* [136]; *541* [137]; *541* [138]
Kunze, M. 7, *100* [84]
Kurzweil, J. 47, *100* [76]
Kusano, T. 251, *336* [78]
Kwapisz, J. 107, 155–158, 186, 189, 190, 198, 199, *201* [42]; *201* [47]
Kwong, M.K. 647, *721* [91]

Ladde, G.S. 207, *336* [79]; 649, *721* [94]
Ladeira, L.A.C. 501, 503, *539* [96]
Ladyzhenskaya, O.A. 259, 266, 271, *336* [80]
Laetsch, T. 594, *606* [41]
Lakshmikantham, V. 3, *97* [1]; 207, *336* [79]; 623, 624, 628, 649, 661, *718* [12]; *718* [13]; *721* [94]
Lamrani Alaoui, A. 253, *335* [46]
Lasota, A. 3, 53, 56, *99* [53]; *100* [85]; 619, *721* [95]
Lasry, J.M. 255, 273, *337* [81]
Lavrent’ev, M.M., Jr. 254, *339* [154]

Lazer, A.C. 130, 144, *199* [2]; *201* [46]; 332, *335* [27]; *337* [82]; 703, 706, *721* [96]
Lee, J. 96, 97, *100* [70]
Lee, J.W. 624, *720* [77]
Leela, S. 661, *718* [13]
Lehmann, P. 451, *540* [115]
Leontovich, E.A. 343, 360, *428* [2]
Levinson, N. 3, *99* [54]
Li, C. 360, 399, 412, 428, *429* [17]; *429* [18]; *429* [23]; *429* [24]; *429* [25]; *429* [26]; *431* [82]; *431* [83]; *431* [84]; *431* [89]; *431* [94]; *432* [118]
Li, J. 399, *431* [85]; *431* [86]; *431* [87]; *431* [88]; *431* [89]; *431* [90]; *431* [91]; *431* [92]; *431* [93]; *431* [95]
Li, W. 399, *431* [82]
Li, Y. 130, *201* [46]; 440, *541* [138]; 550–552, 562, 564–566, 569–573, 575, 577, 585, 590, 596, 599, 600, *605* [17]; *605* [28]; *605* [29]; *605* [30]; *605* [31]; *605* [32]; *605* [33]
Lieberman, G.M. 256, 262, 291, *337* [83]; *337* [84]
Liebscher, S. 360, 399, *429* [28]; *429* [29]; *429* [30]
Lin, C.S. 557, *606* [42]
Lin, X.B. 476, *540* [97]
Lin, Y. 399, *430* [60]; *431* [90]
Lins Neto, A. 206, *337* [85]
Lions, P.L. 709, *719* [29]
Litsyn, E. 525, *538* [56]
Littlewood, J.E. 138, *201* [40]
Liu, C. 399, *431* [94]
Liu, W. 440, *541* [139]
Liu, X. 399, 424, 426, *430* [58]
Liu, Y. 399, *431* [75]; 526, 535, *541* [140]
Liu, Z. 399, *431* [91]; *431* [92]; *431* [93]; *431* [95]; *431* [96]
Liz, E. 210, *337* [86]
Llibre, J. 360, 399, *428* [11]; *429* [34]; *431* [82]; *431* [97]
Lloyd, N.G. 206, *334* [4]; *337* [87]; *337* [88]; 360, 399, *429* [19]; *429* [20]; *431* [79]; *432* [98]; *432* [99]
Lo, E. 526, 535, *541* [122]
Lomtatidze, A. 651, 653, 663, 664, 668, *721* [99]; *721* [100]
Longtin, A. 450, *541* [141]; *541* [142]; *541* [143]
Lou, Y. 550, *605* [14]
Loud, W.S. 552, 594, *606* [43]
Louihi, M. 465, 470, 472, 522, *541* [144]
Lü, H. 654, 665, 668, *718* [2]
Lu, K. 476, *537* [44]
Luo, D. 360, 399, 428, *430* [61]; *430* [62]; *430* [63]; *432* [100]; *432* [101]
Luzyanina, T. 526, 535, *538* [66]; *541* [145]
Lynch, S. 399, *429* [21]; *432* [102]; *432* [103]

Ma, Z. 399, *432* [104]
Mackey, M.C. 450, 452, 460, 462, *537* [25]; *539* [75]; *541* [146]; *542* [147]; *542* [149]; *543* [185]
Macki, J.W. 96, *99* [59]
Magal, P. 458, 465, 504, 507, 512, 513, *542* [148]
Mahaffy, J. 462, *537* [26]; *542* [149]; *543* [185]
Maier, A.G. 343, 360, *428* [2]
Malaguti, L. 75, 76, 95, 97, *98* [32]; *98* [33]; *98* [34]; 107, 132, 133, 139, 144, *201* [48]; *201* [49]; *201* [50]; *201* [51]
Mallet-Paret, J. 142, *201* [52]; 465, 470, 504, 505, 507, 515, 517–521, *542* [150]; *542* [151]; *542* [152]; *542* [153]; *542* [154]; *542* [155]; *542* [156]; *542* [157]; *542* [158]; *542* [159]
Manásevich, R. 671, 672, 685, 702, 704, 706, 707, *719* [51]; *719* [52]; *720* [53]; *720* [68]; *720* [70]; *721* [101]
Mancini, G. 254, *334* [9]
Manitius, A. 438, 465, 503, *542* [160]
Manosa, V. 360, 368, 399, *429* [31]
Marcelli, C. 107, 132, 133, 139, 144, *201* [48]; *201* [49]; *201* [50]; *201* [51]; 207, 217, *337* [89]; *337* [90]
Marini, M. 33, 95, *99* [52]
Martin, M.R.W. 445, 446, *537* [32]; *537* [33]
Martinez-Amores, P. 706, *722* [105]
Massera, J.L. 63, *100* [86]; 206, 208, 209, 219, *337* [91]
Matano, H. 229, 252, 257, 310–312, 321, *334* [3]; *337* [92]; *337* [93]
Matsuoka, T. 5, *100* [87]; *100* [88]; *100* [89]
Mawhin, J. 7, 96, 97, *100* [65]; *100* [90]; 205–207, *337* [94]; *337* [95]; *337* [96]; *337* [97]; *337* [98]; 671, 672, 685, 688, 705–707, *719* [46]; *720* [73]; *721* [84]; *721* [85]; *721* [101]; *721* [102]; *721* [103]; *721* [104]
Maxwell, T.O. 130, *201* [53]
McKean, H.P. 206, *337* [99]; 576, 577, *606* [44]
McLeod, J.B. 142, *200* [33]
Mel'nikov, V.K. 105, *201* [54]; 399, *432* [105]
Meng, X. 399, 426, *432* [125]
Messer, J.A. 445, *542* [161]
Metz, J.A.J. 457, 462, *542* [162]
Milton, J.G. 450, 451, 452, 454, *539* [72]; *539* [73]; *539* [74]; *540* [114]; *541* [141]; *541* [142]; *541* [143]; *542* [147]; *542* [163]
Mizel, V.J. 180, *201* [55]
Mlak, W. 255, *337* [100]

Montecchiari, P. 105, 129–131, *199* [3]; *199* [4]; *199* [5]; *199* [6]; *200* [14]; *201* [56]
Monteiro Marques, M.D.P. 7, *100* [91]
Montero, A. 702, 706, *719* [52]
Moretto, S. 207, *337* [101]
Moser, J. 105, *201* [57]
Müller, M. 207, *337* [102]
Murphy, K. 535, *542* [164]
Murray, J.D. 251, *337* [103]
Murthy, P. 208, *337* [104]
Murúa, A. 707, *720* [53]
Myjak, J. 45, 96, *99* [57]; 668, *721* [83]
Myshkis, A.D. 437, *542* [165]

Na, T.Y. 712, *722* [106]
Nachman, A. 610, *719* [44]; *719* [45]
Nadirashvili, N. 144, *200* [39]
Nagumo, M. 208, 229, 251, *337* [105]; *337* [106]
Napp-Zinn, H. 451, *542* [166]
Narasimhan, R. 251, *337* [107]
Neugebauer, A. 476, *542* [171]
Neves, K.W. 526, 532–535, *539* [78]; *539* [79]; *542* [167]; *542* [168]; *542* [169]
Ni, W.-M. 551, 554, 557, 585, 588, *605* [15]; *606* [42]; *606* [46]; 709, *720* [74]
Nicolaenko, B. 142, *200* [12]
Nieto, J.J. 206, *338* [108]
Nirenberg, L. 551, 552, 554, 585, 588, *605* [15]; *606* [45]; 709, *720* [74]
Niri, K. 526, *542* [172]
Nisbet, R.M. 457, *542* [170]
Nishihira, S. 526, *543* [173]
Nistri, P. 6, 96, *98* [27]; *99* [59]
Njoku, F.I. 206, *338* [109]
Nkashama, M.N. 207, *338* [110]; 706, *722* [107]
Nolasco, M. 105, *201* [56]
Norris, M. 442, 465, *538* [64]
Nørsett, S.P. 528, 529, *539* [71]
Nouzel, J.-C. 451, *540* [115]
Núñez, D. 206, *338* [111]
Nussbaum, R.D. 445, 465, 470, 504–507, 515, 518–521, *542* [151]; *542* [152]; *542* [153]; *542* [154]; *542* [155]; *542* [156]; *543* [174]; *543* [175]; *543* [176]; *543* [177]; 551, *606* [46]

Obersnel, F. 205–207, 209, 210, 243, *338* [112]; *338* [113]; *338* [114]
Obukhovskii, V.V. 7, 12, 13, 41, 48, 61, 62, 96, *99* [45]; PR9955*100* [78]
Okubo, A. 251, *338* [115]
Oliker, V. 667, 709, *720* [72]
Omari, P. 205–207, 209, 210, 243, 251, 254–256, 261, 321, 326, 327, *335* [41]; *335* [42]; *336* [54]; *336* [55]; *338* [109]; *338* [112]; *338* [113]; *338* [114]; 696, 706, *722* [108]
Opial, Z. 579, 580, *606* [47]
O'Regan, D. 207, 210, *336* [52]; 611, 623, 624, 627, 628, 631, 633, 647, 654, 655, 657, 661, 665, 667, 668, 710–713, *718* [2]; *718* [3]; *718* [4]; *718* [5]; *718* [6]; *718* [7]; *718* [8]; *718* [9]; *718* [10]; *718* [11]; *718* [12]; *718* [13]; *718* [14]; *718* [15]; *718* [16]; *719* [37]; *721* [90]; *722* [109]; *722* [110]
Orlicz, W. 3, *100* [92]
Ortega, R. 205, 206, *335* [36]; *337* [98]; *338* [116]; *338* [117]
Ouifki, R. 465, 515, 521, 522, *543* [178]
Ouyang, T. 550, 551, 561, 562, 564–566, 569–573, 575, 578, 579, 585, 586, 588–591, 593, *605* [30]; *605* [31]; *605* [32]; *605* [33]; *605* [34]; *605* [35]; *605* [36]; *606* [37]; *606* [38]; *606* [39]; *606* [48]; *606* [49]; *606* [50]

Palm, G. 451, *540* [113]
Palmer, J.L. 700, 705, *719* [34]
Palmer, K.J. 66, *101* [93]
Pao, C.V. 205, *338* [118]
Papageorgiou, N.S. 7, 13, 50, 87, 96, *100* [73]; *100* [74]
Paraskevopoulos, P. 465, 470, 504, 505, 507, *542* [156]
Pastor, K. 4, 6, *98* [29]; *99* [35]; *99* [36]
Paul, C.A.H. 526, 532, 535, *536* [15]; *536* [16]; *543* [179]
Pawelzik, K. 451, 452, *539* [73]; *539* [74]
Peano, G. 207, *338* [119]
Pearson, J.M. 360, 399, *432* [98]; *432* [99]
Peletier, L.A. 107, 147, 149–152, 155, 157, 158, 180, *201* [55]; *201* [58]; *201* [59]; *201* [60]; *201* [61]; *201* [62]; *201* [63]; 610, 709, *719* [29]; *719* [32]
Pera, P. 95, *99* [63]
Pérez-Garcia, V.M. 705, *720* [71]
Perko, L.M. 399, *431* [77]; *432* [106]
Perron, O. 207, *338* [120]
Perrone, N. 712, *722* [111]
Persson, L. 459, *538* [52]
Petrovski, I. 106, 132, *201* [45]
Phan-Thien, N. 648, *722* [112]
Piccinini, L.C. 207, 234, *338* [121]
Pinney, E. 705, *722* [113]
Pinto, M. 525, 526, *539* [83]; *539* [84]; *543* [180]
Piscounov, N. 106, 132, *201* [45]
Pliss, V.A. 4, *101* [94]; 206, *338* [122]
Poincaré, H. 105, *201* [64]
Poisson, S.D. 437, *543* [181]
Poláčik, P. 252, 254, 256, 257, 290, *338* [123]

Pólya, G. 138, *201* [40]
Pouso, R.L. 210, *337* [86]; 671, 707, *719* [41]; *719* [42]
Prodi, G. 251, 318, *338* [124]; *338* [125]; *338* [126]; *338* [127]; 552, *604* [5]
Prohens, R. 360, 399, *429* [22]
Protter, M.H. 149, *201* [65]; 262, *338* [128]
Pruszko, T. 53, 96, *99* [46]

Qian, M. 399, 424, *429* [27]

Rabinowitz, P.H. 105, 117, 124, 125, 128–131, 155, *200* [28]; *200* [29]; *202* [66]; *202* [67]; *202* [68]; *202* [69]; *202* [70]; *202* [71]; *202* [72]; *202* [73]; *202* [74]; 552, 553, *605* [12]; *605* [13]
Rachůnková, I. 617, 620, 623, 625–627, 629, 631, 649, 666–668, 670, 673, 675, 676, 679, 682, 690, 692, 697, 699, 706, 707, 714, *718* [14]; *722* [114]; *722* [115]; *722* [116]; *722* [117]; *722* [118]; *722* [119]; *722* [120]; *722* [121]; *722* [122]; *722* [123]; *722* [124]; *722* [125]; *722* [126]; *722* [127]; *722* [128]; *722* [129]; *722* [130]
Rai, S. 525, *543* [182]; *543* [183]
Ramaswamy, M. 577, 584, *606* [51]
Renardi, M. 555, *606* [52]
Rezounenko, A.V. 461, *543* [184]
Rihan, F.A. 526, 535, *536* [14]
Rivero, J. 525, *539* [86]
Robertson, R.L. 525, *543* [182]; *543* [183]
Robles-Pérez, A.M. 107, 133, 137, *199* [10]; 205, *337* [98]; *338* [117]
Rodriguez, G. 360, 399, *431* [97]
Rogers, R.C. 555, *606* [52]
Roose, D. 526, 535, *541* [145]
Rosenblat, S. 206, *338* [129]
Rottschäfer, V. 158, *202* [75]
Roussarie, R. 399, 428, *431* [83]; *432* [107]
Rousseau, C. 360, 399, *428* [8]; *431* [81]; *432* [108]
Roxin, E. 208, 229, *338* [130]
Royalty, W.D. 667, *719* [37]
Rubbioni, P. 207, 217, *337* [89]; *337* [90]
Ruf, B. 206, *335* [24]

Sachdev, P.L. 133, *202* [76]
Samaey, G. 535, *538* [66]
Sánchez, D.A. 206, *334* [19]; *338* [131]; *338* [132]
Sanchez, E. 459, 461, 482, *536* [9]; *536* [10]; *536* [11]
Sanchez, L. 107, 108, 133, 137, 144, 159, 160, *199* [10]; *200* [19]; *200* [20]; *200* [37]; *202* [77]; *202* [78]; 706, 707, *720* [78]; *720* [79]; *722* [131]
Sandqvist, A. 207, *334* [10]; *334* [11]; *338* [133]
Santanilla, J. 706, *722* [107]
Santillan, M. 462, *543* [185]
Sarafyan, D. 526, 529, 532, 535, *538* [49]
Sattinger, D.H. 252, 253, *338* [134]
Saz, E. 360, 399, *432* [99]
Schaaf, R. 552, 557, 579–581, 585, 596–600, *606* [53]
Schaeffer, J.J. 63, *100* [86]
Scheurer, B. 142, *200* [12]
Schiaffino, A. 206, *335* [43]
Schick, M. 108, 183, 184, 194, *200* [36]
Schlomiuk, D. 399, 428, *432* [108]; *432* [109]
Schmitt, K. 316, 320, *334* [15]; *334* [16]
Schwegler, H.C. 452, *539* [75]
Scovel, J.C. 206, *337* [99]; 576, 577, *606* [44]; *606* [54]
Segev, I. 450, *536* [1]
Sell, G. 517, *542* [157]; *542* [158]
Séré, E. 105, 155, *200* [23]; *200* [27]; *202* [79]
Serra, E. 130, 131, *200* [21]; *200* [24]; *202* [80]
Shahshahani, S. 206, *338* [135]
Shampine, L.F. 526, 535, *543* [197]
Sharkovskii, A.N. 4, *101* [95]
Shekhter, B.L. 649, 668, 670, 714, *721* [93]
Shi, J. 440, *537* [43]; 551, 562, 563, 566, 579, 585–588, *606* [40]; *606* [50]; *606* [55]; *606* [56]
Shi, S. 399, *432* [110]
Shilnikov, A.L. 398, 428, *432* [111]
Shilnikov, L.P. 398, 428, *432* [111]
Shivaji, R. 550, 552, 576, 594, *604* [6]; *604* [9]
Si, J.-G. 438, *543* [186]; *543* [187]; *543* [188]
Simoda, S. 251, *337* [106]
Smale, S. 105, *202* [81]; *202* [82]
Smets, D. 199, *202* [83]; 706, *719* [39]
Smirnov, G.V. 7, *101* [96]
Smith, H.L. 205, 208, 220, 229, 251, 252, 256, 257, 284, 305, 310, 318, *336* [60]; *336* [61]; *339* [136]; 456, 465, 497, 504, 507, 512, 523, *541* [136]; *541* [137]; *543* [189]; *543* [190]; *543* [191]
Smoller, J. 550, 569, 594, *606* [57]
Šmulev, I.I. 251, *339* [137]; *339* [138]
Sobolevskiĭ, P.E. 251, 255, 273, *336* [75]; *336* [76]
Solecki, R. 704, *723* [156]
Solimini, S. 703, 706, 707, *721* [96]; *723* [132]
Solonnikov, V.A. 259, 266, 271, *336* [80]
Sorolla, J. 360, 399, *428* [11]
Stampacchia, G. 207, 234, 255, 273, *338* [121]; *339* [139]
Staněk, S. 617, 620, 623, 625–627, 629, 631, 642, 644, 647, 655, 657, 658, 666–668, 674, 707,

710–713, *718* [14]; *718* [15]; *718* [16]; *718* [17]; *722* [118]; *722* [119]; *722* [120]; *722* [121]; *722* [122]; *722* [123]; *723* [133]; *723* [134]; *723* [135]; *723* [136]; *723* [137]; *723* [138]; *723* [139]; *723* [140]
Stanford, L.R. 451, *543* [192]
Stépán, G. 447, 448, 450, *540* [116]; *540* [117]
Stevens, B. 451, *543* [193]
Stokes, A. 526, *543* [194]
Stredulinsky, E. 131, *202* [73]; *202* [74]
Stuart, G.J. 450, *544* [215]
Sugie, J. 526, *543* [195]
Sun, C. 399, 428, *430* [56]
Süskind, C. 700, 705, *719* [34]
Swift, J.B. 172, *202* [84]
Szegö, G.P. 208, *339* [140]
Sznt, I. 360, 399, *432* [99]

Taddei, V. 75, 76, 95, 97, *98* [32]; *98* [33]; *98* [34]
Takáč, P. 256, 290, *339* [141]
Taliaferro, S.D. 610, 655, 667, *723* [141]
Tanaka, K. 131, *202* [85]
Tang, M. 593, *606* [58]
Tank, D.W. 450, *543* [199]
Tanner, S. 451, *543* [193]
Tarallo, M. 108, 159, *200* [20]; *200* [37]
Tavernini, L. 526, *543* [196]
Terracini, S. 105, 108, 159, *200* [20]; *200* [37]; *201* [56]; *202* [86]
Thái, N.H. 13, *99* [40]
Thibaudeau, P. 399, *432* [108]
Thiel, U. 451, 452, *539* [74]
Thompson, H.B. 668, *719* [28]
Thompson, S. 526, 529, 532–535, *538* [49]; *539* [79]; *542* [169]; *543* [197]
Thomsen, P.G. 528, 529, *539* [71]
Tineo, A. 206, 207, *334* [20]; *334* [21]; *339* [142]; 656, 668, *723* [142]
Tolstonogov, A.A. 7, 96, *101* [97]
Torregrosa, J. 360, 399, *429* [32]; *429* [33]
Torres, P.J. 651, 653, 664, 668, 699–701, 704, 705, 706, 707, *720* [71]; *721* [100]; *722* [105]; *723* [143]; *723* [144]; *723* [145]
Travis, S.P. 465, *543* [198]
Treccani, G. 208, *339* [140]
Tricomi, F.G. 207, *339* [143]
Troianiello, G.M. 255, 273, 332, *337* [81]; *339* [144]
Troy, W.C. 107, 147, 149–152, 155, 157, 158, 180, *201* [55]; *201* [58]; *201* [59]; *201* [60]; *201* [61]; *201* [62]; *201* [63]
Tso, K. 262, *339* [145]
Turaev, D.V. 398, 428, *432* [111]
Turi, J. 438, 447, 448, 450, 465, 483, 503, 525, 526, 535, *538* [53]; *540* [102]; *540* [103]; *540* [104]; *540* [105]; *540* [106]; *540* [107]; *540* [108]; *540* [116]; *540* [117]
Tvrdý, M. 649, 670, 673, 675, 676, 679, 682, 690, 692, 697, 699, 706, 707, *722* [124]; *722* [125]; *722* [126]; *722* [127]; *722* [128]; *722* [129]; *722* [130]
Tzeng, S.-Y. 130, *200* [26]

Umanskii, Ya.I. 48, *101* [98]
Unnikrishnan, K.P. 450, *543* [199]
Ural’tseva, N.N. 259, 266, 271, *336* [80]
Usmani, R.A. 620, *718* [18]; *718* [19]

Vaillancourt, R. 526, 532, 535, *541* [124]; *541* [125]
van den Berg, J.B. 107, 147, 157, 158, 180, 181, 199, *202* [83]; *202* [87]; *202* [88]; *202* [89]; *202* [90]
van Gils, S.A. 438, 466, 467, 484, 485, 488, 494, 505, 506, 514, *538* [54]
van Saarloos, W. 145, *200* [30]
Vance, R.R. 206, *339* [146]
VanderVorst, R.C.A.M. 107, 152, 155, 156, 158, 186, 189, 190, 198, 199, *201* [42]; *201* [43]; *201* [58]; *202* [90]
Vasiliev, N.I. 619, 649, 663, *723* [146]
Väth, M. 19, 27, 43, 97, *99* [37]; *99* [38]; *101* [99]
Vatsala, A.S. 207, *336* [79]; 649, *721* [94]
Vazques, J.L. 647, 648, *720* [60]; *721* [81]
Verduyn Lunel, S.M. 438, 466, 467, 484, 485, 488, 494, 505, 506, 514, *538* [54]; *540* [98]
Verriest, E.I. 526, *544* [200]
Viano, M. 360, 399, *429* [34]
Vidossich, G. 206–208, 234, *338* [121]; *339* [147]; *339* [148]
Vishnevskiĭ, M.P. 254, *339* [154]
von Wahl, W. 267, *339* [149]
Vrabie, I.I. 39, *101* [100]
Vrkoč, I. 670, 675, 679, 692, 697, 699, 706, 707, *722* [129]; *722* [130]; *723* [147]

Walter, W. 135, *202* [91]; 205, 207, *339* [150]
Walther, H.O. 438, 439, 441, 444, 445, 465–470, 476, 477, 484, 485, 488, 489, 494, 505, 506, 513, 514, 516, 518, *538* [54]; *541* [133]; *542* [159]; *544* [202]; *544* [203]; *544* [204]; *544* [205]; *544* [206]; *544* [207]; *544* [208]; *544* [209]; *544* [210]
Waltman, P. 454, 456, 465, *539* [87]; *539* [88]; *539* [89]; *540* [112]; *544* [201]; 667, 709, *720* [72]
Wang, D. 399, 428, *429* [17]

Wang, E. 399, *432* [104]
Wang, F.Z. 647, *720* [64]
Wang, J.Y. 579, *606* [56]; 657, 668, 707, *721* [89]; *723* [148]; *723* [149]
Wang, M. 399, *428* [12]
Wang, S.-H. 550, 551, 562, 569, 594, *604* [3]; *606* [59]; *606* [60]; *606* [61]; *606* [62]; *606* [63]
Wang, X.-P. 360, 399, *432* [101]; 438, *543* [186]; *543* [187]; *543* [188]
Wang, Z. 398, 399, *430* [64]; 440, *545* [225]
Warner, F.W. 253, *336* [68]
Wasserman, A. 550, 569, 594, *606* [57]
Wayne, C.E. 158, *202* [75]
Weedermann, M. 399, 428, *429* [37]
Weinberger, H.F. 106, 132, 149, *199* [11]; *201* [65]; 262, *338* [128]
Westphal, H. 251, *339* [151]
Wiggins, S. 105, *202* [92]; 425, 428, *432* [112]
Willé, D.R. 526, 535, *536* [16]; *536* [17]; *544* [211]; *544* [212]; *544* [213]; *544* [214]
Williams, S.M. 610, *719* [32]
Williams, S.R. 450, *544* [215]
Winston, E. 465, 526, *544* [216]; *544* [217]; *544* [218]
Wong, P.J.Y. 633, *718* [20]
Wu, H.Q. 647, *720* [64]
Wu, J. 7, 96, *100* [83]; 450, 452, 454, 460, 461, 470, 476, 477, 489, 494, 518, 525, *536* [2]; *537* [38]; *537* [39]; *540* [114]; *541* [133]; *541* [134]; *543* [184]; *544* [219]
Wu, Y. 399, 426, *428* [7]; *430* [65]; *430* [66]; *430* [67]; *432* [113]

Xia, Z. 428, *433* [131]
Xiang, G. 399, *432* [114]
Xiao, D. 399, 428, *430* [55]; *432* [115]

Yan, J. 526, *544* [220]
Yan, P. 671, 706, 707, *723* [150]; *723* [151]
Yan, W. 526, *544* [220]
Yang, C. 399, 426, *430* [68]
Yang, Z. 399, *431* [96]; 440, *544* [221]
Ye, W.Y. 696, 706, *722* [108]
Ye, Y. 360, 399, 422, 423, 426, 428, *430* [69]; *430* [70]; *432* [116]; *432* [117]
Yebdri, M. 518, *544* [210]
Yeh, T.S. 594, *606* [63]
Yi, Y. 360, 399, 428, *429* [18]
Yoneyama, T. 524, 526, *543* [173]; *544* [222]; *544* [223]
Yorke, J.A. 208, 229, *339* [152]; 437, *539* [94]
Yu, J. 399, 428, *432* [118]
Yu, P. 373, 399, *430* [60]; *432* [119]; *432* [120]; *432* [121]

Zabreĭko, P.P. 7, 13, 96, 97, *99* [40]; *100* [82]; 255, 277, *336* [77]
Zaghrout, A.A.S. 460, *545* [224]
Zang, H. 399, 426, *432* [122]; *432* [123]; *432* [124]; *432* [125]; *433* [126]
Zanolin, F. 651, 653, 664, 665, 668, 705, 707, *719* [46]; *720* [68]; *721* [80]
Zecca, P. 7, 12, 13, 41, 48, 61, 62, 95, 96, *99* [39]; *99* [55]; *99* [59]; *100* [64]; *100* [78]
Zeidler, E. 255, 277, 332, *339* [153]
Zelenyak, T.I. 254, *339* [154]
Zennaro, M. 526, 528–530, 532, 533, 535, *537* [27]
Zhang, M.R. 690, 693, 700, 701, 706, 707, *721* [88]; *723* [151]; *723* [152]; *723* [153]; *723* [154]; *723* [155]
Zhang, P. 399, 428, *433* [128]
Zhang, T. 399, 426, *430* [71]; *432* [122]; *432* [123]; *432* [124]; *432* [125]; *433* [126]
Zhang, Z. 360, 399, 412, 426, 428, *431* [72]; *431* [82]; *431* [84]; *432* [115]; *433* [127]; 440, *545* [225]
Zhao, A. 526, *544* [220]
Zhao, S. 399, 428, *433* [128]
Zhao, X.Q. 251, *339* [155]
Zheng, X. 440, *545* [225]
Zhong-Heng, G. 704, *723* [156]
Zhu, D. 360, 399, 426, 428, *430* [62]; *430* [63]; *430* [70]; *432* [100]; *432* [101]; *433* [129]; *433* [130]; *433* [131]

Subject Index

μ-condensing map, 12
ϕ-Laplacian, 634
π_p, 689
σ-Ca-selectionable map, 47
σ-mLL-selectionable map, 47

absolute neighbourhood retract (ANR), 8, 77, 81
absolute retract (AR), 9
acyclic map, 11
acyclic set, 9, 14, 59
admissible homotopy, 10, 60
admissible map, 3, 10, 60
admissible pair, 10
age-structured models, 462
almost upper semicontinuous (a.u.s.c.) map, 47
amphiphilic, 183
ANR-space, 23
anti-periodic problem, 74, 76
anti-periodic solution, 73, 93, 95
attractive singularity, 669
attractors, 515
autonomous problem, 553

bifurcation diagram, 562, 571, 573, 578
bounded solution, 61, 69, 79
Brillouin equation, 700

C^k equivalent, 352, 353
Carathéodory conditions, 609
– L^1-, 209
– L^1-, 232
– L^p-, 258, 263
Carathéodory map, 12
Carathéodory selection, 13, 47, 87
Carathéodory solution, 3
Cauchy problem, 52
center, 362, 378–380, 385, 386, 392, 395, 397, 406, 407
center manifolds, 476, 483
classical solution, 3, 77
clipping, 189
coincidence point, 14
compact map, 11
compactly admissible homotopy, 11
comparison principles, 313, 314, 316, 320
computing the direction of bifurcation, 575
computing the location of bifurcation, 573
condensing admissible map, 11
condensing composition, 27
condensing homotopy in J_A, 29
condensing map, 12
conjugate problem, 624
continuation principle, 26, 28, 29, 685
continuous dependence, 465, 468, 499
continuous map, 10
continuum, 215, 217, 239, 241, 244, 245, 273, 280, 323, 324
contractible set, 9
control problem, 6
convex nonlinearities, 559
covering (topological) dimension, 49, 52
Crandall–Rabinowitz bifurcation theorem, 553
cubic-like nonlinearities, 568
curves of solutions, 549
cyclicity, 397, 400, 401, 405, 406, 409, 410
– at a periodic orbit, 401
– homoclinic, 400
– Hopf, 397, 400, 407, 422, 423
– Poincaré, 400
cyclicity k, 400

dead core solution, 710
delay adaptation, 450
derivative discontinuities, 533
detection curve, 412
detection function, 412
differentiability
– with respect to parameters, 499
– with respect to parameters in norm, 501
Dirichlet problem, 648
disease transmission, 454
Duffing equation, 684, 690, 699

ejectivity, 504
Emden–Fowler equation, 610
ENR-space, 61

entirely bounded solution, 93, 95
equation
– Fisher, 132
– Fisher–Kolmogorov, 106
– – extended, 145
– Swift–Hohenberg, 179
essential Nielsen class, 18, 22
Euler–Poincaré characteristic, 94
exact multiplicity for a nonlocal problem, 601
exact multiplicity results, 566, 572, 589, 593
existence, 465, 468, 499
exponential dichotomy, 62

Filippov regularization, 6
Filippov solution, 6
first eigenvalue of Dirichlet problem, 689, 699
first order ODE, 206
– Cauchy problem, 206
– periodic problem, 206
fixed point, 3, 14
fixed point index, 23, 24
Floquet problem, 53, 73
flow, 343, 350, 352, 357
focal problem, 620
focus, 360, 362–364, 373, 375, 376, 379, 381–383
– hyperbolic, 364
– of order m, 364
– stable, 362, 363
– unstable, 362, 363
– weak, 364
focus value, 364, 377, 382
Fréchet space, 8
function of type A, 132
fundamental set, 14, 27

generalized Lefschetz number, 15–17, 61, 86
generic convergence, 522
Green function, 62, 72, 73, 90, 559, 620, 627, 700
Green operator, 89

H-Nielsen class, 22
Hamiltonian function, 379
Hamiltonian system, 379, 380, 398
– near-, 399
Hausdorff measure of noncompactness, 43, 44
Hausdorff-continuous map, 10
Hausdorff-metric, 10
heteroclinic, 105
– chain, 124
heteroclinic loop, 400
heteroclinics
– basic, 124
– with endpoint at infinity, 130
homoclinic, 105
homoclinic bifurcation, 423
homoclinic loop, 400, 412, 423, 425, 426
Hopf bifurcation, 360, 385, 386, 391, 399, 401, 495
– degenerate, 386
Hukuhara–Kneser property, 212, 214, 273

implicit function theorem in Banach spaces, 552
implicit ODEs, 7
initial value problem, 44, 49
initial value problem of first order ODE, 211
– lower solution, 212, 215, 217
– solution, 212
– upper solution, 212, 215, 217
initial value problem of second order parabolic PDE, 265
– lower solution, 266, 273, 280
– solution, 266
– upper solution, 266, 273, 280
instability
– Lyapunov, 230, 238, 239, 241, 311, 325, 326, 329
– order, 312, 325, 329
integrably bounded map, 45

J-map, 25, 27
J_A-class, 25, 27

k-set contraction, 12
Kamke function, 43, 44
Kuratowski measure of noncompactness, 44

l-Carathéodory map, 12
l-Carathéodory multivalued function, 87
Lefschetz number, 15, 16, 94
Lefschetz set, 17
Lefschetz theorem, 15, 17
Lidstone problem, 631
Liénard equation, 681, 684
Liénard system, 379, 382, 397, 409
limit cycle, 343–345, 349, 351, 353, 355, 357–359, 381, 383, 385, 390, 400, 401, 404, 409, 412, 426
– hyperbolic, 355–358
– inner stable, 345
– inner unstable, 345
– multiplicity, 345, 349, 351–354, 357, 359
– nonhyperbolic, 357
– outer stable, 345
– outer unstable, 345
– stability, 345, 349, 351, 353, 354, 402, 423, 426
– stable, 345, 348, 404, 423, 424, 426
– unstable, 345, 348, 349, 355, 423, 426
linearization, 472, 502, 506, 510
linearized problem, 554
linearized stability, 448, 450, 482

local invariant manifolds, 476
locally contractible set, 9
locally integrably bounded map, 45
lower and upper functions, 650, 670, 677, 713
lower semicontinuous (l.s.c.) map, 10
lower solution of first order ODE, 210
lower solution of second order parabolic PDE, 259
Lyapunov constant, 364, 381, 395
Lyapunov quantity, 377

Massera convergence and existence theorem, 219, 225, 246
Mathematica program for solving nonlinear boundary value problems, 603
max-plus operators, 521
measurable map, 12
measurable selection, 12
measurable space, 12
measurable-locally Lipschitz map, 46
measure of noncompactness, 11
Melnikov function, 386, 401, 404
milling process, 448
mixed conditions, 710, 712, 713
mixed singularities, 655
Monotone Convergence Criterion, 220, 284, 290
Morse index of solution, 579
multiple solutions, 329
multivalued map, 10
multivalued ODEs, 6, 7

(n, p) problem, 627
N-admissible homotopy, 23
N-admissible map, 22, 23
N-admissible pair, 22
N-Nielsen class, 22
Nagumo condition, 259, 267, 273, 280, 285, 674
Nemytskiĭ operator, 39, 63
Neumann problem, 579
neural networks, 450
Nielsen class, 21, 85
Nielsen number, 18, 79, 95
– modulo H, 22
Nielsen relation, 18
Nielsen theorem, 18
Nielsen-type relation, 19
nonlinear Fredholm alternative, 620
nonlocal problem, 635, 641
normal form, 371
numerical computation of solutions, 602

orbit, 344
– negative, 344
– periodic, 344, 359, 379, 386, 399, 400
– positive, 344
order interval trichotomy, 298, 305
order norm, 261
ordinary Lefschetz number, 86
oscillation, 524

p-Laplacian, 557, 634
$p(t)$-Laplacian, 647
Parabolic Maximum Principle, 262
parabolic PDE
– initial value problem, 265
– periodic problem, 284
pendulum, 109
period, 344, 403, 415
periodic point, 4
periodic problem, 74, 668
periodic problem of first order ODE, 217
– lower solution, 218, 222, 236, 238, 239, 241
– solution, 218, 239, 244
– upper solution, 218, 222, 236, 238, 239, 241
periodic problem of second order parabolic PDE, 284
– lower solution, 285, 294, 297, 318, 320, 324–326
– solution, 285, 323
– upper solution, 294, 297, 318, 320, 324–326
periodic solution, 73–75, 79, 92, 95, 344, 378, 504
perturbed Gelfand problem, 562
phase transition, 183
pitchfork bifurcation, 575
Poincaré map, 93, 345, 348, 350, 353, 354, 356, 361–363, 371, 376, 381, 424, 426
Poincaré translation operator, 3, 59
Poincaré–Wirtinger inequality, 693
polynomial system, 420
population models, 457
position control, 444
positivity for the linearized problem, 591
positone problem, 655
principal eigenvalue, 321, 326, 332
product-measurable map, 13
proper map, 10

quasi-compact map, 41, 42

R_δ-map, 11, 24, 60
R_δ-set, 3, 9, 32, 45, 48, 53, 76
regular point, 343
regular problem, 610
Reidemeister class, 21
Reidemeister set, 20, 21
repulsive singularity, 669, 683
retractible map, 37

S-shaped solution curve, 561, 568
saddle value, 424

saddle-focus, 146, 160, 176, 186
saddle-node, 146
saddle-node bifurcation, 385
Schwarzian derivative, 592
Scorza–Dragoni type result, 47
selected pair, 10
semiflow, 467, 472, 476, 483
Sharkovskii theorem, 4
sign-changing solution, 577
similarity of the solution branches, 582
singular perturbation, 518
singular point, 343, 380, 421, 612
– of type I, 613, 657
– of type II, 613, 657
singular problem, 610
slowly oscillatory solutions, 516
small delay, 521
solution
– fast, 137
– multi-bump, 155
– multi-transition, 150, 194
– sharp type, 144
solution manifold, 467
solution of first order ODE, 210
– almost periodic, 241, 242
– bounded (asymptotic behaviour), 226, 244; 246, 248
– heteroclinic, 227
– maximal, 211, 227
– maximum, 211, 215, 217, 222, 228, 239
– minimal, 211, 227
– minimum, 211, 215, 217, 222, 228, 239
– nonextendible, 210, 245
– subharmonic, 241, 243
solution of singular problem, 611
solution set, 29, 32, 39, 45, 59
solution (strong) of second order parabolic PDE, 259
– bounded (asymptotic behaviour), 288, 289
– heteroclinic, 298, 305
– maximal, 261, 297, 308, 326
– maximum, 261, 273, 280, 294, 320, 324
– minimal, 261, 297, 308, 326
– minimum, 261, 273, 280, 294, 320, 324
– nonextendible, 259, 283
space singularity, 612, 613, 617, 652, 657
stability, 524
– $\mathcal{L}$-asymptotic, 230, 231, 311, 320
– Lyapunov, 230, 232, 233, 236, 311, 312, 316, 318, 323, 329
– of solution, 556, 587
– order, 231–233, 236, 239, 311, 312, 316, 324, 329
– $\mathcal{W}$-asymptotic, 235, 238
– weak, 235, 236, 239
Stable Manifold Theorem, 477
stable manifolds, 476
state-dependent delay, 437
Strong Maximum Principle, 149
strong singularity, 680
strongly measurable map, 12, 39
strongly measurable selection, 12
Sturm–Liouville problem, 628
superposition operator, 39
superpositionally measurable map, 13
symmetric nonlinearities, 585
symmetry breaking, 575

T-monotonicity, 219, 223, 244, 288, 289
target problem, 69
target (terminal) problem, 50
threshold condition, 455, 457, 507, 524
threshold speed, 106
time map formula through global linearization, 595
time maps, 594
time singularity, 611, 613, 651
topological (covering) dimension, 7
topological structure, 14, 66
translation operator, 61
travelling wave, 106, 132, 158
triple-well potential, 108
turning point, 549
turning process, 447
two-body problem, 441
two-point conditions, 708

u-Carathéodory map, 12, 29, 37, 39, 62, 71
u-Carathéodory multivalued function, 87
uniform integrability, 616
uniqueness, 465, 468, 499
– of solutions, 588
unstable manifolds, 476
upper semicontinuous (u.s.c.) map, 10
upper solution of first order ODE, 210
upper solution of second order parabolic PDE, 259, 260

van der Pol equation, 403
Vietoris map, 10, 19, 22

w-solution of singular problem, 611
weak singularity, 680
weakly measurable map, 12
weakly superpositionally measurable map, 13

HANDBOOK
OF DIFFERENTIAL EQUATIONS

ORDINARY DIFFERENTIAL EQUATIONS

VOLUME IV

Contents

Preface v
List of Contributors vii
Contents of Volume 1 xi
Contents of Volume 2 xiii
Contents of Volume 3 xv

1. Symmetric Hopf bifurcation: Twisted degree approach 1
   Z. Balanov and W. Krawcewicz
2. Nonautonomous differential systems in two dimensions 133
   R. Fabbri, R. Johnson and L. Zampogni
3. Complex differential equations 269
   I. Laine
4. Transversal periodic-to-periodic homoclinic orbits 365
   K.J. Palmer
5. Successive approximation techniques in non-linear boundary value problems for ordinary differential equations 441
   A. Rontó and M. Rontó
6. Analytic ordinary differential equations and their local classification 593
   H. Żołądek

Author index 689
Subject index 697

This page intentionally left blank

Contents of Volume 1

Preface v
List of Contributors vii

1. A survey of recent results for initial and boundary value problems singular in the dependent variable 1
   R.P. Agarwal and D. O'Regan
2. The lower and upper solutions method for boundary value problems 69
   C. De Coster and P. Habets
3. Half-linear differential equations 161
   O. Došlý
4. Radial solutions of quasilinear elliptic differential equations 359
   J. Jacobsen and K. Schmitt
5. Integrability of polynomial differential systems 437
   J. Llibre
6. Global results for the forced pendulum equation 533
   J. Mawhin
7. Ważewski method and Conley index 591
   R. Srzednicki

Author index 685
Subject index 693

This page intentionally left blank

Contents of Volume 2

Preface v
List of Contributors vii
Contents of Volume 1 xi

1. Optimal control of ordinary differential equations 1
   V. Barbu and C. Lefter
2. Hamiltonian systems: periodic and homoclinic solutions by variational methods 77
   T. Bartsch and A. Szulkin
3. Differential equations on closed sets 147
   O. Cârjă and I.I. Vrabie
4. Monotone dynamical systems 239
   M.W. Hirsch and H. Smith
5. Planar periodic systems of population dynamics 359
   J. López-Gómez
6. Nonlocal initial and boundary value problems: a survey 461
   S.K. Ntouyas

Author index 559
Subject index 565

This page intentionally left blank

Contents of Volume 3

Preface v
List of Contributors vii
Contents of Volume 1 xi
Contents of Volume 2 xiii

1. Topological principles for ordinary differential equations 1
   J. Andres
2. Heteroclinic orbits for some classes of second and fourth order differential equations 103
   D. Bonheure and L. Sanchez
3. A qualitative analysis, via lower and upper solutions, of first order periodic evolutionary equations with lack of uniqueness 203
   C. De Coster, F. Obersnel and P. Omari
4. Bifurcation theory of limit cycles of planar systems 341
   M. Han
5. Functional differential equations with state-dependent delays: Theory and applications 435
   F. Hartung, T. Krisztin, H.-O. Walther and J. Wu
6. Global solution branches and exact multiplicity of solutions for two point boundary value problems 547
   P. Korman
7. Singularities and Laplacians in boundary value problems for nonlinear ordinary differential equations 607
   I. Rachůnková, S. Staněk and M. Tvrdý

Author index 725
Subject index 735

This page intentionally left blank

CHAPTER 1

Symmetric Hopf Bifurcation: Twisted Degree Approach

Zalman Balanov

Department of Mathematics and Computer Sciences, Netanya Academic College, 1, University str., Netanya 42365, Israel
E-mail: *balanov@mail.netanya.ac.il*

Wieslaw Krawcewicz

Department of Mathematical and Statistical Sciences, University of Alberta, T6G 2G1 Edmonton, Canada
E-mail: *wkrawcew@math.ualberta.ca*

Contents

1. Introduction . . . 3
   Subject and goal . . . 3
   Topological degree approach . . . 5
   Overview . . . 8
2. Auxiliary information . . . 10
   2.1. Basic definitions and notations . . . 10
   2.2. Elements of representation theory . . . 12
   2.3. G-vector bundles and G-manifolds . . . 19
   2.4. Fredholm operators . . . 20
   2.5. Bibliographical remarks . . . 22
3. Twisted equivariant degree: Construction and basic properties . . . 23
   3.1. Topology behind the construction: Equivariant extensions and fundamental domains . . . 23
   3.2. Analysis behind the construction: Regular normal approximations . . . 26
   3.3. Algebra behind the construction: Twisted groups and Burnside modules . . . 27
   3.4. Construction . . . 33
   3.5. Axiomatic approach to twisted degree . . . 34
   3.6. S^1-degree . . . 37
   3.7. Computational techniques for twisted degree . . . 39
   3.8. General concept of basic maps . . . 44
   3.9. Multiplicativity property . . . 47
   3.10. Infinite dimensional twisted degree . . . 48

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 4
Edited by F. Battelli and M. Fečkan

© 2008 Elsevier B.V. All rights reserved

3.11. Bibliographical remarks . . . 52
4. Hopf bifurcation problem for ODEs without symmetries . . . 53
  4.1. Statement of the problem . . . 54
  4.2. S^1-equivariant reformulation of the problem . . . 55
  4.3. S^1-degree method for Hopf bifurcation problem . . . 57
  4.4. Deformation of the map $\mathfrak{F}_\varsigma$: Reduction to a product map . . . 60
  4.5. Crossing numbers . . . 63
  4.6. Conclusions . . . 64
  4.7. Bibliographical remarks . . . 65
5. Hopf bifurcation problem for ODEs with symmetries . . . 66
  5.1. Symmetric Hopf bifurcation and local bifurcation invariant . . . 66
  5.2. Computation of local bifurcation invariant: Reduction to product formula . . . 69
  5.3. Computation of local bifurcation invariant: Reduction to crossing numbers and basic degrees . . . 70
  5.4. Summary of the equivariant degree method . . . 72
  5.5. Usage of Maple© routines . . . 74
  5.6. Bibliographical remarks . . . 75
6. Symmetric Hopf bifurcation for FDEs . . . 75
  6.1. Symmetric Hopf bifurcation for FDEs with delay: General framework . . . 76
  6.2. Symmetric Hopf bifurcation for neutral FDEs . . . 80
  6.3. Global bifurcation problems . . . 81
  6.4. Bibliographical remarks . . . 84
7. Symmetric Hopf bifurcation problems for functional parabolic systems of equations . . . 84
  7.1. Symmetric bifurcation in parameterized equivariant coincidence problems . . . 84
  7.2. Hopf bifurcation for FPDEs with symmetries: Reduction to local bifurcation invariant . . . 89
  7.3. Computation of local bifurcation invariant . . . 95
  7.4. Bibliographical remarks . . . 98
8. Applications . . . 98
  8.1. Γ-symmetric FDEs describing configurations of identical oscillators . . . 98
  8.2. Hopf bifurcation in symmetric configuration of transmission lines . . . 102
  8.3. Global continuation of bifurcating branches . . . 109
  8.4. Symmetric system of Hutchinson model in population dynamics . . . 112
  8.5. Bibliographical remarks . . . 120
Appendix A . . . 122
  Dihedral group D_N . . . 122
  Irreducible representations of dihedral groups . . . 123
  Icosahedral group A_5 . . . 124
  Irreducible representations of A_5 . . . 125
Acknowledgment . . . 126
References . . . 126

1. Introduction

Subject and goal

As it is clear from the title,

(i) the *subject* of this paper is the Hopf bifurcation occurring in dynamical systems admitting a certain group of symmetries;

(ii) the *method* to study the above phenomenon presented in this paper is based on the usage of the so-called twisted equivariant degree.

The *goal* of this paper is to explain why "(ii)" is an appropriate tool to attack "(i)".

To get an idea of what the Hopf bifurcation is about, consider the simplest system of ODEs

$$\dot{x} = Ax \quad (x \in \mathbb{R}^2), \tag{1}$$

where $A : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear operator. Obviously, the origin is a stationary solution to (1), and it is a standard fact of any undergraduate course of ODEs that this system admits a non-constant periodic solution iff the characteristic equation $\det \Delta(\lambda) = 0$, where $\Delta(\lambda) := \lambda \,\mathrm{Id} - A$, has a pair of (conjugate) purely imaginary complex roots (in this case the origin is called a *center* for (1)).

Assume now that the system (1) is included into a one-parameter family of systems

$$\dot{x} = A(\alpha)x \quad (\alpha \in \mathbb{R},\ x \in \mathbb{R}^2), \tag{2}$$

where $A(\cdot) : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear operator (smoothly) depending on α, $A(\alpha_o) = A$ for some 'critical' value α_o and $(\alpha_o, 0)$ is an *isolated* center for (2) (i.e. it is the only center for α close to α_o). If a pair of complex roots λ of the characteristic equation crosses (for $\alpha = \alpha_o$) the imaginary axis, then the stationary solution $(\alpha, 0)$ changes its stability which results in appearance of non-constant periodic solutions. This phenomenon is called the Hopf bifurcation in (2). Similarly, one can speak about the occurrence of the Hopf bifurcation in a one-parameter family of n-dimensional linear systems of ODEs for $n > 2$ (does not matter that the corresponding purely imaginary roots may have multiplicities greater than one).

Next, consider a (nonlinear) autonomous system of ODEs of the type

$$\dot{x} = f(\alpha, x) \quad (\alpha \in \mathbb{R},\ x \in V := \mathbb{R}^n), \tag{3}$$

where $f : \mathbb{R} \oplus V \to V$ is a continuously differentiable function satisfying the condition that $f(\alpha, 0) = 0$ for all $\alpha \in \mathbb{R}$. Clearly, $(\alpha, 0)$ is a stationary solution to (3) for all α. In this situation, a change of stability simply means that some of the complex roots λ of the characteristic equation $\det \Delta_{(\alpha,0)}(\lambda) = 0$, where $\Delta_{(\alpha,0)}(\lambda) := \lambda \,\mathrm{Id} - D_x f(\alpha, 0)$, cross (for $\alpha = \alpha_o$) the imaginary axis. In particular, this means the existence of a purely imaginary characteristic root $i\beta_o$ for $\alpha = \alpha_o$, which (by the implicit function theorem) is a *necessary* condition for the Hopf bifurcation (i.e. for the appearance of non-constant small amplitude periodic solutions). However, simple examples show that, in contrast to the linear case, this algebraic condition on the linearization is not enough, in general, for the occurrence of the Hopf bifurcation. Of course, this is not surprising: the classical Grobman–Hartman Theorem provides the local topological equivalence of an autonomous system to its linearization (near the origin) only under the assumption that the linearization matrix does not have eigenvalues on the imaginary axis. In particular, this means that studying the Hopf bifurcation phenomenon in parametrized families of *nonlinear* systems requires an additional *topological* argument.

A standard way of studying the Hopf bifurcation is the application of the Central Manifold theorem (allowing a two-dimensional reduction) and usage of the Poincaré section associated with the induced system (see [113] for a detailed exposition of this stream of ideas, see also [8,68]). However, this approach meets serious technical difficulties if the multiplicity of a purely imaginary characteristic root is greater than one. To overcome these and other technical difficulties, alternative methods were developed based on Lyapunov–Schmidt reduction, normal form techniques, integral averaging, etc. (cf. [33,34,64,111]). On the other hand, one should mention rational-valued homotopy invariants of “degree type” introduced by F.B. Fuller [57], E.N. Dancer [39] and E.N. Dancer and J.F. Toland [40, 42,41] as important tools to study the Hopf bifurcation phenomenon.

Observe that for many mathematical models of natural phenomena, very often, their closeness to the real world problems is reflected (on top of their non-linear character) in the presence of symmetries that are related to some physical or geometric regularities. For systems (3) this means: V is an (orthogonal) representation of a group Γ and f commutes with the Γ-action on V (i.e.

$$f(\alpha, \gamma x) = \gamma f(\alpha, x) \quad (\gamma \in \Gamma,\ x \in V), \tag{4}$$

in which case f is called Γ-equivariant (here Γ acts trivially on the parameter space)). In this way, we arrive at the following question: what is a link between symmetries of a system and symmetric properties of the actual dynamics? In the context relevant to our discussion, this question translates as the following *symmetric Hopf bifurcation problem*: how can one measure, predict and classify symmetric properties/minimal number of periodic solutions appearing as a result of the Hopf bifurcation?

It should be pointed out that in the symmetric setting, the characteristic roots almost always are not simple which causes significant difficulties for the application of the standard methods. To analyze the symmetric Hopf bifurcation problem, Golubitsky et al. (cf. [63, 65–67], see also [31,28,100,116,101,102,147]) suggested a method based on the Central Manifold/Lyapunov–Schmidt finite-dimensional reduction and further usage of a special singularity theory. On the other hand, if system (3) is Hamiltonian, one can use a wide spectrum of variational methods rooted in Morse theory/Lusternik–Schnirelman theory. Although very effective, these methods are not easy to use as they require a serious topological/analytical background. Also, when dealing with a concrete problem admitting a large group of symmetries, one would like to take advantage of using computer routines to handle a huge number of possible symmetry types of the bifurcating periodic solutions. From this viewpoint, it is not clear if the above methods are “open enough” to be computerized. It is our belief that the method presented in this paper is simple enough to be understood by applied mathematicians, and effective enough

(a) to be applied in a standard way to different types of symmetric dynamical systems,
(b) to provide a full topological information on the symmetric structure of the bifurcating branches of periodic solutions,
(c) to be transparent from the viewpoint of interpretation of its results, and
(d) "last but not least", to be completely computerized.

Topological degree approach

(i) *From Leray–Schauder degree to S^1-degree*. The Leray–Schauder degree theory proved itself as a powerful tool for the detection of single and multiple solutions in various types of differential equations. However, when dealing with the Hopf bifurcation phenomenon in autonomous systems through a functional analysis approach, it can only detect equilibria while it remains *blind* to non-constant periodic solutions. The reason for it can be easily explained: shifting the argument of periodic functions represents an S^1-action, which implies that finding periodic solutions to the associated operator equation constitutes an S^1-equivariant problem. By the obvious reason, the Leray–Schauder degree cannot "distinguish" between the zero-dimensional S^1-orbits ($=$ equilibria) and the one-dimensional ones ($=$ cycles), and one should look for a suitable S^1-equivariant homotopy invariant.

Speaking in a slightly more formal language, introduce the frequency β of the (unknown) periodic solution as an additional parameter and reformulate problem (3) as an operator equation in the first Sobolev space $W := H^1(S^1; V)$ as follows:

$$\mathfrak{F}(\alpha, \beta, u) = 0, \tag{5}$$

where $u \in W$, $\mathfrak{F} : \mathbb{R} \times \mathbb{R}_+ \times W \to W$ is given by

$$\mathfrak{F}(\alpha, \beta, u) := u - (L + K)^{-1}\left[Ku + \frac{1}{\beta} N_f\big(\alpha, j(u)\big)\right],$$

with $N_f(\alpha, u)(t) := f(\alpha, u(t))$, $K(u) := \frac{1}{2\pi}\int_0^{2\pi} u(s)\,\mathrm{d}s$ and $j : H^1(S^1; V) \to C(S^1, V)$ being a natural embedding into the space of continuous functions. Formula

$$(\mathrm{e}^{\mathrm{i}\tau} u)(t) := u(t + \tau) \quad (\mathrm{e}^{\mathrm{i}\tau} \in S^1,\ u \in W) \tag{6}$$

equips W with a structure of Hilbert S^1-representation and, moreover, $\mathfrak{F}$ is S^1-equivariant. Take $u \in W$, put

$$G_u := \{z \in S^1 \colon\ \mathrm{e}^{\mathrm{i}\tau} u = u\} \tag{7}$$

and call it a *symmetry* of u (commonly called the *isotropy* of u). It is easy to see that equilibria for (3) have the whole group S^1 as their symmetry, while for a non-constant periodic solution u, one has $G_u = \mathbb{Z}_l$ for some $l = 1, 2, \ldots$. Observe, by the way, that the presence of symmetry $\mathbb{Z}_l$ for a periodic solution u has a transparent geometric meaning: it clearly indicates that u is "l-folded". These simple observations suggest a natural candidate for the range of values of the "right" invariant responsible for the *existence* of non-constant periodic solutions (in particular, for the occurrence of (non-symmetric) Hopf bifurcation) – it should take its values in the free $\mathbb{Z}$-module generated by $\mathbb{Z}_l$, $l = 1, 2, \dots$, rather than in the ring $\mathbb{Z}$ (as the Leray–Schauder degree does). The corresponding construction (called S^1-degree) was suggested in [45,46] (see also Subsections 3.5 and 3.6 of the present paper for an axiomatic approach and [79,81] for a more general setting).

(ii) *From S^1-degree to twisted degree*. Let us make one more step assuming system (3) to be Γ-symmetric (cf. (4)). Put $G := \Gamma \times S^1$. Then, W is a Hilbert G-representation with the G-action given by (cf. (6))

$$\bigl((\gamma, \mathrm{e}^{\mathrm{i}\tau})u\bigr)(t) := \gamma u(t+\tau) \quad (\gamma \in \Gamma,\ \mathrm{e}^{\mathrm{i}\tau} \in S^1,\ u \in W).$$

Moreover, this time $\mathfrak{F}$ is G-equivariant. One can easily verify that the symmetries $G_u := \{g \in G\colon\ gu = u\}$ of non-constant periodic functions u are the so-called *φ-twisted l-folded* subgroups $K^{\varphi,l}$ of G given by

$$K^{\varphi,l} := \bigl\{(\gamma, z) \in K \times S^1\colon\ \varphi(\gamma) = z^l\bigr\}$$

for K being a subgroup of Γ, $\varphi : K \to S^1$ a homomorphism and $l = 1, 2, \dots$ (cf. (7)). We have now a complete parallelism with the previous situation: in the same way as the Leray–Schauder degree was not enough to establish the *existence* of non-constant periodic solutions to (3), the S^1-degree is not enough to *classify symmetries* of these solutions. Clearly, the right G-equivariant homotopy invariant should take its values in the $\mathbb{Z}$-module $A_1^t(G)$ generated by φ-twisted l-folded subgroups (more precisely, by their conjugacy classes). The *twisted G-equivariant degree* (in short, twisted degree) is a topological tool precisely needed for the above purpose.

Roughly speaking, the G-equivariant twisted degree is an object that is only slightly more complicated than the usual Leray–Schauder degree. It is a finite sequence of integers, indexed (for the convenience of the user) by the conjugacy classes (H) of φ-twisted l-folded subgroups H of G. To be more specific, given a group $G = \Gamma \times S^1$, an isometric Banach G-representation W, an open invariant subset $\Omega \subset \mathbb{R} \oplus W$ and a continuous G-equivariant map $f : (\overline{\Omega}, \partial\Omega) \to (W, W \setminus \{0\})$, one can assign to f the twisted G-equivariant degree in the following form:

$$G\text{-}\mathrm{Deg}^t(f, \Omega) = n_1(H_1) + n_2(H_2) + \cdots + n_k(H_k), \quad n_i \in \mathbb{Z}. \tag{8}$$

As we will see later on, $G\text{-}\mathrm{Deg}^t(f, \Omega)$ satisfies all the properties expected from any reasonable "degree theory", in particular, existence, homotopy invariance, excision, suspension, additivity, multiplicativity, etc. (adopted to the equivariant setting). Moreover, similarly to the Leray–Schauder degree, the twisted degree admits an axiomatic approach which allows applied mathematicians to use it without going into topological (homotopy theory, bordism theory) and analytical (equivariant transversality, normality) roots underlying its construction. Thus, it can be easily applied to equivariant settings in the same way as the Leray–Schauder degree is applied to non-symmetric situations.

(iii) *Application scheme of twisted degree*. In this survey article, we will explain how to apply the twisted G-equivariant degree to study various Hopf bifurcation problems (parameterized by $\alpha \in \mathbb{R}$) with a certain symmetry group Γ. To clarify the essence of our approach, we list below the main steps one should follow to attack problem (3).

(a) Let $(\alpha_o, 0)$ be an isolated center for (3) and let $i\beta_o$ be the corresponding purely imaginary characteristic root.

(b) Take a small 'cylinder' $\Omega \subset \mathbb{R} \times \mathbb{R}_+ \times W$ around the point $(\alpha_o, \beta_o, 0)$, construct an *auxiliary G*-invariant function $\varsigma : \overline{\Omega} \to \mathbb{R}$ (see Definition 4.3), confining solutions to (5) to the inside of Ω, i.e. ς is negative on the 'trivial' solutions $(\alpha, \beta, 0)$ and it is positive on the 'exit' set from Ω.

(c) Consider the equation

$$\mathfrak{F}_\varsigma(\alpha, \beta, u) := \big(\varsigma, \mathfrak{F}(\alpha, \beta, u)\big) = 0 \tag{9}$$

(obviously, $\mathfrak{F}_\varsigma$ decreases only one dimension and any solution to (9) is also a solution to (5)).

(d) Define the *local bifurcation invariant* $\omega(\alpha_o, \beta_o)$ containing the topological information about the symmetric nature of the bifurcation, by

$$\omega(\alpha_o, \beta_o) := G\text{-}\mathrm{Deg}^t(\mathfrak{F}_\varsigma, \Omega).$$

(e) Use the equivariant spectral properties of the linearized system at the point $(\alpha_o, \beta_o, 0)$ to extract data needed for the computation of $\omega(\alpha_o, \beta_o)$.

(f) Apply appropriate computer program (for example Maple© routines) in order to obtain *the exact value* of the local invariant $\omega(\alpha_o, \beta_o)$.

(g) Analyze $\omega(\alpha_o, \beta_o)$ in order to obtain the information describing possible branches of non-constant periodic solutions, their multiplicity and symmetric properties.

The fact that the equivariant degree approach to the symmetric Hopf bifurcation allows a computerization (based on the algebraic properties of the twisted degree) of many tedious technicalities related to algebraic nature of this problems (combined with the above mentioned axiomatic approach) constitutes one of the most significant advantages of the twisted degree method. Theoretically, this method supported by computer programming can be applied to any kind of Γ-symmetric Hopf bifurcation problem, with the group Γ being of *arbitrary size*.

(iv) *Historical roots of the twisted degree method.* Twisted equivariant degree is a part of the so-called equivariant degree, which was introduced by Ize et al. in [78] and rigorously studied in [81] for Abelian groups. The idea of the equivariant degree has emanated from different mathematical fields rooted in a variety of concepts and methods. The historical roots of the equivariant degree theory can be traced back to several mathematical fields:

(a) Borsuk–Ulam Theorems (cf. [25,9,24,38,55,81,86,99,121,127,138]; see also references in [99] and [138]);

(b) Fundamental Domains, Equivariant Retract Theory (cf. [9,81,99]; see also [3–7,82, 83,103,107,108]);

(c) Equivariant Obstruction Theory, Equivariant Bordisms, Equivariant Homotopy Groups of Spheres (cf. [37,43,87,105,130,135,139,141]; see also [12,17,18,81,99, 114,131]);

(d) Equivariant General Position Theorems (cf. [22,46,60,81,99]; see also [12,18,91,96, 115,150]);

(f) Generalized Topological Degree, Primary Degree, Topological Invariants of Equivariant Gradient Maps (cf. [39,46,57,59,62,79,81,93,99,126,132–134]);

(g) Geometric Obstruction Theory and J-Homomorphism in Multiparameter Bifurcations (cf. [2,76,77]).

In this article, we will discuss the construction of the twisted equivariant degree without entering into too much details. For more information and proofs, we refer to [19].

Overview

Let us discuss in more detail the contents of this article.

In Subsection 2.1, we include some preliminary results and the equivariant jargon frequently used later. In Subsection 2.2, we give a brief overview of some basic concepts and constructions from representation theory, explain our conventions and collect necessary information needed for the construction and usage of the equivariant degree techniques. In Subsection 2.3, we present some facts related to the notion of G-vector bundles modeled on Banach spaces and describe certain properties of (smooth) G-manifolds. In Subsection 2.4, we discuss the set of unbounded Fredholm operators of index zero $\mathfrak{F}_0(\mathbb{E},\mathbb{F})$ between two Banach spaces $\mathbb{E}$ and $\mathbb{F}$, and describe the topology on this set.

Section 3 naturally splits into four parts: (a) main ideas underlying the notion of twisted degree, (b) construction and basic properties, (c) practical computations, and (d) infinite dimensional extensions.

From the topological point of view, the (twisted) equivariant degree “measures” homotopy obstructions for an equivariant map to have equivariant extensions without zeros on a set composed of several orbit types. Therefore, in Subsection 3.1, we describe the topological ideas related to the construction of the twisted equivariant degree, i.e. the induction over orbit types, concept of fundamental domain and equivariant Kuratowski–Dugundji theorem. Since we follow the “differential viewpoint” to construct the twisted degree, in Subsection 3.2, the notions of normal and regular normal maps (i.e. equivariant replacements of “nice” representatives of homotopy classes) are introduced. In Subsection 3.3, we focus on algebraic properties of the $\mathbb{Z}$-module $A_1^t(G)$ – the range of values of the twisted degree. Namely, we discuss generators of $A_1^t(G)$ (i.e. conjugacy classes of φ-twisted l-folded subgroups of G) and, given two subgroups $L \subset H \subset G$, define purely algebraic quantities $n(L,H)$ allowing us to study an important *multiplication* structure on $A_1^t(G)$. Later on, this multiplication (“module structure”) is used in the same way as the usual multiplication in $\mathbb{Z}$ is used for the multiplicativity property of the Brouwer degree.

In Subsection 3.4, we present a construction of the twisted degree, showing that its coefficients n_i (see (8)) can be evaluated using the usual Brouwer degree of regular normal approximations restricted to fundamental domains. The properties of twisted degree providing an axiomatic approach to its usage, are listed in Subsection 3.5. The particular, nevertheless very important, case of the twisted S^1-degree is discussed in Subsection 3.6. The presented axiomatic definition of the twisted S^1-degree makes no use of the normality condition, therefore it is very close to the axiomatic definition of the Brouwer degree. What is probably more important, this definition opens a way for the practical computation of twisted degree.

To be more specific, in Subsection 3.7, based on the axiomatic definition of S^1-degree, we deduce some computational formulae for S^1-degree of mappings that naturally appear in studying the symmetric Hopf bifurcation phenomenon (cf. Problem 3.32, Theorem 3.39 and Corollary 3.40). In addition, we present the so-called Recurrence Formula (see Theorem 3.41) allowing a direct reduction of computations of the twisted degree to the ones of S^1-degree. Next, in Subsection 3.8, we introduce the concept of *basic maps* and *basic degrees*. These maps are the simplest ones having a non-trivial twisted degree and fitting our approach. Their twisted degrees, which are fully computable (based on the Recurrence Formula and the corresponding S^1-degree results), will constitute a 'library' that can be used by the computer programs in order to evaluate the twisted degree of more complicated maps. Finally, in Subsection 3.9, we establish the multiplicativity property of the twisted degree. This (functorial) property, which is similar to the multiplicativity of the Brouwer degree, plays a substantial role in developing the software needed for practical computations of twisted degree.

In Subsection 3.10, we extend in a standard way the (finite-dimensional) twisted degree to two important classes of equivariant vector fields on Banach G-representations: compact fields and condensing fields. These fields turn out to be enough to cover the applications to the Hopf bifurcation problems we are dealing with in this paper.

In Section 4, we lay down the standard steps of the twisted equivariant degree treatment for an autonomous 1-st order system of ODEs (without symmetries). As it was mentioned above, this system, in spite of having no symmetries, still leads to an S^1-equivariant fixed-point problem. In fact, we use the following strategy: in this section, we consider in detail the *occurrence* of the Hopf bifurcation in ODEs without symmetries (resp. S^1-degree approach) while, in subsequent sections, we analyze (trying to avoid any repetitions) the impact of Γ-symmetries in different classes of dynamical systems to *symmetric properties and minimal number* of bifurcating periodic solutions (resp. $\Gamma \times S^1$-equivariant twisted degree approach). In Subsection 4.1, we give a rigorous definition of the Hopf bifurcation. The corresponding setting in functional spaces is explained in Subsection 4.2. In Subsection 4.3, we introduce the local bifurcation invariant ($=$ S^1-degree of an appropriate map) and establish (in its terms) necessary and sufficient conditions for the occurrence of Hopf bifurcation. To obtain an effective formula for the local bifurcation invariant, we apply (based on the properties of S^1-degree), in Subsection 4.4, appropriate equivariant deformations to simplify the operators in question. In Subsection 4.5, by introducing the concept of the so-called crossing number, we complete the computation of the local bifurcation invariant and formulate the Hopf bifurcation result in terms of the right-hand side of the considered system only.

Section 5 is devoted to a discussion of the Hopf bifurcation problem for a systems of ODEs *with symmetries*. A functional setting for a symmetric Hopf bifurcation and the local bifurcation invariant are described in Subsection 5.1. In particular, for the purpose of estimating *precise* symmetries and a minimal number of bifurcating periodic solutions, we introduce the concept of *dominating orbit types*. The computations of the local invariant based on a reduction to the so-called product formula, are presented in Subsection 5.2 (here we essentially use the multiplicativity property of the twisted degree). The concept of an *isotypical* crossing number and the refinement of the product formula are given in Subsection 5.3. The main steps related to the usage of the twisted equivariant degree method are resumed in Subsection 5.4. In Subsection 5.5, we outline how to use the special Maple© routines, which were developed for the twisted degree computations, to compute the exact value of the local bifurcation invariant. These routines are available on the Internet through the link: http://krawcewicz.net/degree.

Section 6, based on the framework developed in the previous section, is dealing with a general symmetric Hopf bifurcation problem in functional differential equations. Following the standard steps, in Subsection 6.1, we analyze a parametrized system of symmetric delayed functional differential equations. In Subsection 6.2, we briefly discuss particularities related to the usage of the same method in symmetric neutral functional differential equations. Observe that the local bifurcation invariants can be effectively applied to study the problem of continuation of symmetric branches of non-constant periodic solutions (i.e. a global Hopf bifurcation problem) for a system of symmetric functional differential equations (see Subsection 6.3).

In Section 7, we extend the applicability of the twisted degree method to a system of functional parabolic differential equations. A general setting (as an abstract coincidence problem) in functional spaces, appropriate for studying symmetric FPDEs, is presented in Subsection 7.1. In Subsection 7.2, we discuss the Hopf bifurcation problem for a general FPDEs with symmetries. Although, basically the computations of the local bifurcation invariant follow the standard lines, several points specific to this setting are explained in Subsection 7.3.

In Section 8, we illustrate the twisted degree method by applying it to three concrete models. It should be pointed out that our choice of the examples (see the Contents) is motivated by their relative simplicity, practical meaning and the effectiveness of the analysis (via the twisted degree method) of the symmetric Hopf bifurcation phenomenon occurring in these systems. Other examples and more details can be found in [19].

In a short Appendix A, we collect the information about the groups (i.e. dihedral groups D_n and the icosahedral group A_5) and their representations, which are used in our examples presented in Section 8.

2. Auxiliary information

2.1. *Basic definitions and notations*

In this section, G stands for a compact Lie group.

DEFINITION 2.1. A *topological transformation group* is a triple (G, X, φ), where X is a Hausdorff topological space and $\varphi : G \times X \to X$ is a continuous map such that:

(i) $\varphi(g, \varphi(h, x)) = \varphi(gh, x)$ for all g, $h \in G$ and $x \in X$;

(ii) $\varphi(e, x) = x$ for all $x \in X$, where e is the identity element of G.

The map φ is called a *G-action* on X and the space X, together with a given action φ of G, is called a *G-space* (or, more precisely, *left G-action* and *left G-space*, respectively). We shall use the notation $g(x)$, $g \cdot x$ or simply gx, for $\varphi(g, x)$. For $K \subset G$ and $A \subset X$, put $K(A) := \{gx\colon\ g \in K, x \in A\}$ and for $g \in G$ we write $gA := \{gx\colon\ x \in A\}$. A set $A \subset X$

is said to be G-*invariant* if $G(A) = A$. Notice that if A is a compact set, then so is $G(A)$. Observe that on any Hausdorff topological space X one can define the *trivial action* of G by $gx = x$ for all $g \in G$ and $x \in X$.

For any $x \in X$, the closed subgroup $G_x = \{g \in G\colon\ gx = x\}$ of G is called the *isotropy group* of x and the invariant subspace $G(x) := \{gx\colon\ g \in G\}$ of X is called the *orbit* of x. Also, we say that a G-action on X is *free* if $G_x = \{e\}$ for all $x \in X$.

It is easy to see that if H is the isotropy group of $x \in X$, then $G(x)$ is homeomorphic to G/H. Denote by X/G the set of all orbits in X and consider the canonical projection $\pi : X \to X/G$ given by $\pi(x) = G(x)$. The space X/G equipped with the quotient topology induced by π is called the *orbit space* of X under the action of G.

Recall that two closed subgroups H and K of G are said to be *conjugate* in G, if $H = gKg^{-1}$ for some $g \in G$. Clearly, the conjugacy relation is an equivalence relation; denote by (H) the equivalence class of H and call (H) the *conjugacy class* of H in G. Denote by $\mathcal{O}(G)$ the set of all conjugacy classes. The set $\mathcal{O}(G)$ is partially ordered by the relation $\leqslant$ defined as follows:

$$(H) \leqslant (K) \overset{\text{Def}}{\Longleftrightarrow} \exists_{g \in G}\, gHg^{-1} \subset K. \tag{10}$$

Obviously, for a G-space X and $x \in X$, one has $G_{gx} = gG_xg^{-1}$. This gives rise to the notion of the *orbit type* of x defined as the conjugacy class (G_x).

Denote by $\mathcal{J}(X)$ the set of all orbit types occurring in X. Let us point out that according to the order relation (10), if an orbit type (G_x) is smaller than an orbit type (G_y), then the orbit $G(x)$ is “bigger” than the orbit $G(y)$. According to the famous result by Mann (cf. [112]), every action of a compact Lie group on a manifold with finitely generated homology groups has a finite number of orbit types.

The following result is known as Gleason Lemma (cf. [26]).

THEOREM 2.2. *Let X be a metric G-space such that $\mathcal{J}(X) = \{(H)\}$. Then, the orbit map $\pi : X \to X/G$ is a projection in a locally trivial fiber bundle with fiber G/H.*

Suppose that X is a finite-dimensional G-space and G is a compact Lie group. Using the definition of covering dimension (cf. [48]), the Gleason lemma and the Morita theorem (see [120]), which states that $\dim(K \times [0, 1]) = \dim K + 1$ for any metric space K, one can easily prove

PROPOSITION 2.3. *If X is a* (*metric*) *free G-space, then* $\dim(X/G) = \dim X - \dim G$.

For a G-space X and a closed subgroup H of G, we adopt the following notations:

$$\begin{aligned} X_H &:= \{x \in X\colon\ G_x = H\}, \\ X^H &:= \{x \in X\colon\ G_x \supset H\}, \\ X_{(H)} &:= \{x \in X\colon\ (G_x) = (H)\}. \end{aligned}$$

We call X^H the H-*fixed-point subset* of X.

For a closed subgroup H of G, we use $N(H) := \{g\colon\ gHg^{-1} = H\}$ to denote the *normalizer* of H in G and $W(H) := N(H)/H$ to denote the *Weyl group* of H in G. Notice that $N(H)$ is a closed subgroup of G and X^H is $N(H)$-invariant. Consequently, X^H is $W(H)$-invariant, where the $W(H)$-action on X^H is given by $gH \cdot x := gx$ for $g \in N(H)$ and $x \in X^H$. It is also clear that $W(H)$ acts freely on X_H.

Let X and Y be two G-spaces. A continuous map $f : X \to Y$ is called *G-equivariant*, or simply *equivariant*, if

$$\forall g \in G\ \forall x \in X \quad f(gx) = gf(x).$$

If the G-action on Y is trivial, an equivariant map $f : X \to Y$ is called *G-invariant*, or simply *invariant*, i.e.

$$\forall g \in G\ \forall x \in X \quad f(gx) = f(x).$$

Since, for an equivariant map $f : X \to Y$, one has $G_{f(x)} \supset G_x$, it follows that $f(X^H) \subset Y^H$ for every subgroup $H \subset G$. Consequently, the maps $f^H : X^H \to Y^H$ with $f^H := f_{|X^H}$, are well-defined and $W(H)$-equivariant for every subgroup $H \subset G$.

2.2. *Elements of representation theory*

Finite-dimensional G-representations We start with the following

DEFINITION 2.4. Let W be a finite-dimensional real (resp. complex) vector space. We say that W is a real (resp. complex) *representation* of G (in short, *G-representation*), if W is a G-space such that the *translation map* $T_g : W \to W$, defined by $T_g(v) := gv$ for $v \in W$, is an $\mathbb{R}$-linear (resp. $\mathbb{C}$-linear) operator for every $g \in G$.

It is clear that for a G-representation W, the map $T : G \to GL(W)$, given by $T(g) = T_g : W \to W$, is a continuous homomorphism. It is convenient to identify the representation W with the homomorphism $T : G \to GL(W)$. For example, a continuous homomorphism $T : G \to GL(n;\mathbb{R})$ (resp. $T : G \to GL(n;\mathbb{C})$) is called a *real* (resp. *complex*) *matrix G-representation*. For two representations W_1 and W_2, if there is an equivariant isomorphism $A : W_1 \to W_2$, we say that W_1 and W_2 are *equivalent* and write $W_1 \cong W_2$. Let W be a real (resp. complex) G-representation. An inner product (resp. Hermitian inner product) $\langle\cdot,\cdot\rangle : W \oplus W \to \mathbb{R}$ (resp. $\langle\cdot,\cdot\rangle : W \oplus W \to \mathbb{C}$) is called *$G$-invariant* if $\langle gu, gv\rangle = \langle u, v\rangle$ for all $g \in G$, u, $v \in W$. A G-representation together with a G-invariant inner product is called an *orthogonal* (resp. *unitary*) *G-representation*. It is well-known that every real (resp. complex) G-representation is equivalent to an orthogonal (resp. unitary) representation $T : G \to O(n)$ (resp. $T : G \to U(n)$).

Notice that any G-representation W is a manifold with trivial homology groups in non-zero dimensions, therefore (cf. [112]), there are only finitely many orbit types in W.

An invariant linear subspace $\widetilde{W} \subset W$ is called a *subrepresentation* of W and we say that W is an *irreducible* representation if it has no subrepresentation different from $\{0\}$ and W. Otherwise, W is called *reducible*.

The *complete reducibility theorem* states that every (finite-dimensional) G-representation W is a (not necessarily unique) direct sum of irreducible subrepresentations of W, i.e. there exist irreducible subrepresentations $\mathcal{W}^1, \dots, \mathcal{W}^m$ of V such that

$$W = \mathcal{W}^1 \oplus \mathcal{W}^2 \oplus \cdots \oplus \mathcal{W}^m.$$

Let W_1 and W_2 be two G-representations. Denote by $L^G(W_1, W_2)$ the space of all linear G-equivariant maps $A : W_1 \to W_2$, and by $GL^G(W_1, W_2)$ its subspace of all (G-equivariant) isomorphisms. Also, put $L^G(W) := L^G(W, W)$ and $GL^G(W) := GL^G(W, W)$.

Let $\mathcal{W}^1$ and $\mathcal{W}^2$ be two real or complex irreducible G-representations. Then, the Schur lemma states that every equivariant linear map $A : \mathcal{W}^1 \to \mathcal{W}^2$ is either an isomorphism or zero. Assume $\mathcal{U}$ is a *complex* irreducible G-representation. Then, it follows from Schur Lemma that every equivariant linear map $A : \mathcal{U} \to \mathcal{U}$ satisfies $A = \lambda\,\mathrm{Id}$, for some $\lambda \in \mathbb{C}$, i.e. the G-representation $\mathcal{U}$ is *absolutely irreducible* (cf. [27]). Consequently, we have that $\dim_{\mathbb{C}} L^G(\mathcal{U}^1, \mathcal{U}^2) = 1$ or 0 (where $\mathcal{U}^1$ and $\mathcal{U}^2$ are two irreducible complex G-representations). Using this fact, it can be easily proved that every complex irreducible G-representation of an Abelian compact Lie group G is one-dimensional.

In the case $\mathcal{V}$ is a *real* irreducible G-representation, the set $L^G(\mathcal{V})$ is a finite-dimensional associative division algebra over $\mathbb{R}$, so it is either $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, and we call $\mathcal{V}$ to be of *real*, *complex* or *quaternionic type*, respectively. Observe also that the type of a real irreducible G-representation is closely related to its *complexification* (cf. [27]).

REMARK 2.5 (Convention of notation). Let us explain our convention, which we use in connection to the complex and real G-representations. As long as it is possible, we use the letter V to denote a real G-representation, while the letter U is reserved for complex G-representations. In the case the type of a G-representation is not specified, we apply the letter W. Since, for a given compact Lie group G, there are only countably many irreducible G-representations (see [85,151]), we also assume that a complete list, indexed by numbers $0, 1, 2, 3, \dots$, of these irreducible representation is available, and, in the case of real G-representations, we denote them by $\mathcal{V}_0, \mathcal{V}_1, \mathcal{V}_2, \dots$ (where $\mathcal{V}_0$ always stands for the trivial irreducible G-representation), in the case of complex G-representations, by $\mathcal{U}_0, \mathcal{U}_1, \mathcal{U}_2, \dots$ (where $\mathcal{U}_0$ is the trivial complex irreducible G-representation), and in the case the type of an irreducible G-representation is not clearly specified as real or complex, we denote them by $\mathcal{W}_0, \mathcal{W}_1, \mathcal{W}_2, \dots$ (where again $\mathcal{W}_0$ is the trivial irreducible G-representation). Unspecified irreducible G-representations are denoted as follows: in the case of real representations $\mathcal{V}$, $\mathcal{V}^1$, or $\mathcal{V}^k$, in the case of complex representations $\mathcal{U}$, $\mathcal{U}^1$, or $\mathcal{U}^k$, and if the type is unknown $\mathcal{W}$, $\mathcal{W}^1$, or $\mathcal{W}^k$. We summarize our convention in Table 1.

REMARK AND NOTATION 2.6. Let Γ be a compact Lie group.

(i) Any (irreducible) complex Γ-representation can be converted in a natural way to an (irreducible) real $\Gamma \times S^1$-representation as follows. For any complex Γ-representation U and $l = 1, 2, \dots$, define a $\Gamma \times S^1$-action on U by

$$(\gamma, z)w = z^l \cdot (\gamma w), \quad (\gamma, z) \in \Gamma \times S^1,\ w \in U, \tag{11}$$

Table 1
Notational convention for real and complex G-representations

	Real	Complex	Unspecified
G-representation	$V, \mathfrak{V}$	$U, \mathfrak{U}$	$W, \mathfrak{W}$
Irreducible G-representation	$\mathcal{V}$	$\mathcal{U}$	$\mathcal{W}$
List of all irreducible G-representations	$\mathcal{V}_0, \mathcal{V}_1, \mathcal{V}_2, \dots$	$\mathcal{U}_0, \mathcal{U}_1, \mathcal{U}_2, \dots$	$\mathcal{W}_0, \mathcal{W}_1, \mathcal{W}_2, \dots$
Various irreducible G-representations	$\mathcal{V}, \mathcal{V}^1, \mathcal{V}^2, \mathcal{V}^j$	$\mathcal{U}, \mathcal{U}^1, \mathcal{U}^2, \mathcal{U}^j$	$\mathcal{W}, \mathcal{W}^1, \mathcal{W}^2, \mathcal{W}^j$

where '$\cdot$' is the complex multiplication. This real $\Gamma \times S^1$-representation is denoted by ${}^l U$. Moreover, if U is a complex irreducible Γ-representation, then ${}^l U$ is a real irreducible $\Gamma \times S^1$-representation.

(ii) Following Remark 2.5, assume that $\{\mathcal{U}_j\}$ is a *complete* list of irreducible complex Γ-representations. The (real) irreducible $\Gamma \times S^1$-representation ${}^l\mathcal{U}_j$ is called the l-th irreducible $\Gamma \times S^1$-representation *associated* with $\mathcal{U}_j$ and denoted by $\mathcal{V}_{j,l}$.

Complexification and conjugation For a real vector space V, denote by V^c the *complexification* of V given by

$$V^c := \mathbb{C} \otimes_{\mathbb{R}} V.$$

Assume further that V is a real G-representation. Then, V^c has a natural structure of a complex G-representation defined by $g(z \otimes v) = z \otimes gv$, $z \in \mathbb{C}$, $v \in V$.

For a complex vector space U, denote by $\overline{U}$ the conjugate complex vector space structure on U, i.e. the space U with the complex multiplication given by $z \cdot v := \bar{z}v$, $z \in \mathbb{C}$, $v \in U$. Notice that $GL(U) = GL(\overline{U})$, therefore, if U is a complex G-representation, then so is $\overline{U}$, which is called the G-representation *conjugate* to U. In the case of a complex matrix G-representation $T : G \to GL(n, \mathbb{C})$, the G-representation conjugate to T is given by $\overline{T} : G \to GL(n, \mathbb{C})$, where $\overline{T}(g)$ denotes the matrix obtained from $T(g)$ by replacing its entries with their conjugates.

REMARK 2.7. For a (real) vector space V, an element x of its complexification V^c can be written in a unique way as $x = 1 \otimes u + \mathrm{i} \otimes v$, where $u, v \in V$. Using this representation of elements in V^c, it is easy to show that if V has a complex structure (i.e. an isomorphism $J : V \to V$ satisfying $J^2 = -\mathrm{Id}$), then the space V^c is isomorphic (as a complex vector space) to the direct sum $V \oplus \overline{V}$, where $\overline{V}$ stands for the (complex) vector space *conjugate* to V, i.e. the space V with the complex multiplication given by $z \cdot v := \bar{z}v$. Indeed, notice that for every $x \in V^c$,

$$\begin{aligned} x &= 1 \otimes u + \mathrm{i} \otimes v \\ &= 1 \otimes \frac{u + \mathrm{i}v}{2} + \mathrm{i} \otimes \frac{v - \mathrm{i}u}{2} + 1 \otimes \frac{u - \mathrm{i}v}{2} + \mathrm{i} \otimes \frac{v + \mathrm{i}u}{2}, \end{aligned}$$

so we obtain the following direct decomposition $V^c = V_1 \oplus V_2$, where

$$V_1 = \left\{ 1 \otimes \frac{u + \mathrm{i}v}{2} + \mathrm{i} \otimes \frac{v - \mathrm{i}u}{2} \colon u, v \in V \right\},$$

$$V_2 = \left\{ 1 \otimes \frac{u - \mathrm{i}v}{2} + \mathrm{i} \otimes \frac{v + \mathrm{i}u}{2} \colon u, v \in V \right\}.$$

It is easy to verify that V_1 is $\mathbb{C}$-isomorphic to V and V_2 to $\overline{V}$.

REMARK 2.8. It is well-known (see [27]) that for a real irreducible G-representation V, the complex G-representation V^c is irreducible if and only if V is of real type. Otherwise, it follows from Remark 2.7 that if V has a natural complex structure, then V^c, as a complex G-representation, is equivalent to $V \oplus \overline{V}$. In this case V is equivalent (as a complex G-representation) to $\overline{V}$, if and only if V is of quaternionic type.

Character of representation For a finite-dimensional real (resp. complex) G-representation W, with the corresponding homomorphism $T : G \to GL(W)$, the function $\chi_W : G \to \mathbb{R}$ (resp. $\chi_W : G \to \mathbb{C}$), defined by

$$\chi_W(g) = \mathrm{Tr}\big(T(g)\big), \quad g \in G,$$

where Tr stands for the trace, is called the *character* of W. It is well-known (see, for example, [27]) that $\chi_W(ghg^{-1}) = \chi_W(h)$, for $g, h \in G$, and for two G-representations (both being either real or complex) W_1 and W_2, we have $\chi_{W_1 \oplus W_2} = \chi_{W_1} + \chi_{W_2}$. Let us point out that any G-representation is completely determined by its character.

Haar integral Recall the notion of the *Haar integral* on G, denoted by $\int_G f(g)\,\mathrm{d}\mu(g)$, where $f : G \to \mathbb{R}$ is a continuous function. There exists a unique functional $f \to \int_G f(g)\,\mathrm{d}\mu(g)$, on the space of continuous functions on G, satisfying the following properties:

(i) $\int_G (f_1(g) + f_2(g))\,\mathrm{d}\mu(g) = \int_G f_1(g)\,\mathrm{d}\mu(g) + \int_G f_2(g)\,\mathrm{d}\mu(g)$;

(ii) $\int_G cf(g)\,\mathrm{d}\mu(g) = c\int_G f(g)\,\mathrm{d}\mu(g)$, where $c \in \mathbb{R}$;

(iii) $\int_G \mathrm{d}\mu(g) = 1$;

(iv) $\int_G f(h^{-1}g)\,\mathrm{d}\mu(g) = \int_G f(gh)\,\mathrm{d}\mu(g) = \int_G f(g^{-1})\,\mathrm{d}\mu(g) = \int_G f(g)\,\mathrm{d}\mu(g)$, for all $h \in G$;

(v) $\int_G f(g)\,\mathrm{d}\mu(g) \geqslant 0$ for $f(g) \geqslant 0$, $\forall g \in G$.

Isotypical decomposition Consider a real G-representation V. The representation V can be decomposed into a direct sum

$$V = \mathcal{V}^1 \oplus \mathcal{V}^2 \oplus \cdots \oplus \mathcal{V}^m \tag{12}$$

of irreducible subrepresentations $\mathcal{V}^i$ of V, some of them may be equivalent. This direct decomposition is not "geometrically" unique and is only defined up to isomorphism. Of course, among these irreducible subrepresentations there may be distinct (non-equivalent) subrepresentations which we denote by $\mathcal{V}_{k_1}, \dots, \mathcal{V}_{k_r}$, including possibly the trivial one-dimensional representation $\mathcal{V}_0$. Let V_{k_i} be the sum of all irreducible subspaces $\mathcal{V}^o \subset V$ equivalent to $\mathcal{V}_{k_i}$. Then,

$$V = V_{k_1} \oplus V_{k_2} \oplus \cdots \oplus V_{k_n} \tag{13}$$

and the direct sum (13) is called the *isotypical decomposition* of V. In contrast to the decomposition (12), the isotypical decomposition (13) is unique. The subspaces V_{k_i} are called the *isotypical* components of V (of type $\mathcal{V}_{k_i}$ or modeled on $\mathcal{V}_{k_i}$). It is also convenient to write the isotypical decomposition (13) in the form

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_r, \quad \text{for some } r \geqslant 0, \tag{14}$$

where the isotypical components V_i are modeled on $\mathcal{V}_i$, according to the complete list of irreducible G-representations (some V_i may be trivial subspaces).

The isotypical components V_i, $i = 0, 1, 2, \dots, r$, can be also described in another way, which is more useful for infinite dimensional generalizations. Denote by $\chi_i : G \to \mathbb{R}$ the real character of the irreducible representation $\mathcal{V}_i$, $i = 0, 1, 2, \dots, r$. Define the linear map $P_i : V \to V$ by

$$P_i x = n(\mathcal{V}_i) \int_G \chi_i(g) g x \, \mathrm{d}\mu(g), \quad x \in V, \tag{15}$$

where

$$n(\mathcal{V}_i) = \begin{cases} \dim_{\mathbb{R}} \mathcal{V}_i, & \text{if } \mathcal{V}_i \text{ is of real type,} \\ \dfrac{\dim_{\mathbb{R}} \mathcal{V}_i}{2}, & \text{if } \mathcal{V}_i \text{ is of complex type,} \\ \dfrac{\dim_{\mathbb{R}} \mathcal{V}_i}{4}, & \text{if } \mathcal{V}_i \text{ is of quaternionic type.} \end{cases}$$

The number $n(\mathcal{V}_i)$ defined above, is called the *intristic dimension* of $\mathcal{V}_i$. Then (see [93] for more details):

(a) $x \in V_i \Longleftrightarrow P_i(x) = x$;
(b) $x \in V_l, l \neq i \Longrightarrow P_i(x) = 0$;
(c) $P_i \circ P_i(x) = P_i(x)$ for all $x \in V$;
(d) $P_i : V \to V$ is G-equivariant.

Consequently, every $x \in V$ can be written as

$$x = \sum_{i=0}^{r} P_i(x),$$

where $P_i(x) \in V_i$, so $\mathrm{Id} = \sum_{i=0}^{r} P_i$.

In the case U is a complex G-representation, a similar *complex isotypical decomposition* of U can be constructed, i.e.

$$U = U_0 \oplus U_1 \oplus \cdots \oplus U_s, \quad \text{for some } s \geqslant 0,$$

where the isotypical component U_j is modeled on the complex irreducible G-representation $\mathcal{U}_j$, $j = 0, 1, \dots, s$.

Decomposition of $GL^G(V)$ Let V be an orthogonal G-representation. We have the following standard algebraic facts on a decomposition of $GL^G(V)$.

PROPOSITION 2.9. (Cf. [85].) *Let (cf.* (13))

$$V = V_{k_1} \oplus \cdots \oplus V_{k_r}, \tag{16}$$

be the G-isotypical decomposition, where a component V_{k_i} is modeled on an irreducible representation $\mathcal{V}_{k_i}$. Then:

(i) $GL^G(V) = \bigoplus_{i=1}^{r} GL^G(V_{k_i})$;

(ii) *for any isotypical component V_{k_i} from* (16), *we have $GL^G(V_{k_i}) \simeq GL(m, \mathbb{F})$, where $m = \dim V_{k_i} / \dim \mathcal{V}_{k_i}$ and $\mathbb{F} \simeq GL^G(\mathcal{V}_{k_i})$, i.e. $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, depending on the type of the irreducible representation $\mathcal{V}_{k_i}$.*

Banach G-representations Recall

DEFINITION 2.10. Let W be a real (resp. complex) Banach space. We say that W is a real (resp. complex) *Banach representation* of G (in short *Banach G-representation*) if the space W is a G-space such that the translation map $T_g : W \to W$, defined by $T_g(w) = gw$ for $w \in W$, is a *bounded* $\mathbb{R}$-linear (resp. $\mathbb{C}$-linear) operator for every $g \in G$.

Clearly, every finite-dimensional G-representation is a Banach G-representation.

We say that a Banach G-representation W is *isometric* if for each $g \in G$, the translation operator $T_g : W \to W$ is an isometry, i.e. $\|T\|_g w = \|w\|$ for all $w \in W$, and we call the norm $\|\cdot\|$ a *G-invariant* norm. Using the Haar integral, it can be proved that for every Banach G-representation W, it is possible to construct a new G-invariant norm on W equivalent to the initial one.

A closed G-invariant linear subspace of W is called a Banach subrepresentation. It is also important to notice that all irreducible Banach G-representations (i.e. representations that do not contain any proper non-trivial Banach G-subrepresentations) are finite-dimensional (see [85] or [93]). Notice that for a closed subgroup $H \subset G$, the set W^H is a closed linear subspace of W.

In the case W is a real (resp. complex) Hilbert space, the inner (resp. Hermitian inner) product $\langle \cdot, \cdot \rangle$ on W is called *G-invariant* if $\langle gv, gw \rangle = \langle v, w \rangle$ for all $g \in G$, $v, w \in W$. In this case, W is called an *isometric Hilbert* (resp. *unitary Hilbert*) *G-representation*.

Consider the complete list $\{\mathcal{V}_k\colon\ k = 0, 1, 2, \dots\}$ of all irreducible G-representations (which is a countable set [27]) and let $\chi_k : G \to \mathbb{R}$ be the corresponding character of

$\mathcal{V}_k$ for $k = 0, 1, 2, \ldots$. Take a real Banach G-representation and define the linear maps $P_k : W \to W$ by

$$P_k x = n(\mathcal{V}_k) \int_G \chi_k(g) g x \, \mathrm{d}\mu(g), \quad g \in G, \ x \in W, \ k = 0, 1, 2, \ldots, \tag{17}$$

where $n(\mathcal{V}_k)$ denotes the intristic dimension of $\mathcal{V}_k$.

Then, similarly to the finite-dimensional case, we have that $P_k : W \to W$ is a *G-equivariant* (i.e. $P_k(gw) = g P_k(w)$ for $g \in G$ and $w \in W$) bounded linear projection onto the subspace

$$W_k := P_k(W). \tag{18}$$

Also (see [93]), we have the following

THEOREM 2.11. *Let W be a real isometric Banach representation of a compact Lie group G, $\mathcal{V}$ a real irreducible representation of G and χ the character of $\mathcal{V}$. Then, the linear operator $P_{\mathcal{V}} : W \to W$ defined by*

$$P_{\mathcal{V}} x = n(\mathcal{V}) \int_G \chi(g) g x \, \mathrm{d}\mu(g), \quad x \in W,$$

is a bounded G-equivariant projection on the subspace $P_{\mathcal{V}}(W)$ satisfying the following properties:

(i) *If $x \in W$ belongs to a representation space of an irreducible subrepresentation of W that is equivalent to $\mathcal{V}$, then $P_{\mathcal{V}} x = x$;*

(ii) *If $x \in W$ belongs to a representation space of an irreducible subrepresentation of W that is not equivalent to $\mathcal{V}$, then $P_{\mathcal{V}} x = 0$.*

It is an immediate consequence of Theorem 2.11 that every irreducible subrepresentation of W, which is equivalent to $\mathcal{V}_k$, is contained in W_k. The G-invariant subspace W_k is called the *isotypical component* of W corresponding to $\mathcal{V}_k$. Define the subspace

$$W_\infty := \bigoplus_k W_k \tag{19}$$

which is clearly dense in W, i.e. we have $\overline{W_\infty} = W$. Consequently, the decomposition

$$W = \overline{\bigoplus_k W_k} \tag{20}$$

is called the *isotypical decomposition* of W. Moreover, for every G-equivariant linear operator $A : W \to W$, we have that $A(W_k) \subseteq W_k$ for all $k = 0, 1, 2, \ldots$.

Using the Zorn lemma one can easily establish

PROPOSITION 2.12. *Given* (19) *and* (20), *for any finite subset $X \subset W_\infty$, the subspace* $\operatorname{span} G(X)$ *spanned by the orbits of points from X, is finite-dimensional and G-invariant.*

2.3. *G-vector bundles and G-manifolds*

Let E and B be two topological spaces, V a Banach space, and $p\colon E \to B$ a continuous map which satisfies the following conditions:

(1) For each $x \in B$, the set $p^{-1}(x)$ has a structure of a real Banach space (with topology induced from E);

(2) For each $x \in B$, there exists an open neighborhood U and a homeomorphism $\varphi_U : p^{-1}(U) \to U \times V$ such that the following diagram commutes

$$\begin{tikzcd}
p^{-1}(U) \arrow[rr, "\varphi_U"] \arrow[dr, "p"'] & & U \times V \arrow[dl, "\pi_1"', "\pi_1(u,v)=u"] \\
& U &
\end{tikzcd}$$

and, in addition,

(3) For each $x \in U$ the map $\varphi_{U,x} := (\varphi_U)_{|p^{-1}(x)} \colon p^{-1}(x) \to \{x\} \times V =: V$ is an isomorphism of Banach spaces.

Then, the triple $\xi = (p, E, B)$ is called a *Banach vector bundle* modeled on the Banach space V, E – *total space*, B – *base space*, $p^{-1}(x)$ – *fiber over* x, p – *vector bundle projection* and the pair (U, φ_U) – *local trivialization*.

A Banach vector bundle $\xi = (p, E, B)$, where E, B are two G-spaces, is called a *G-vector bundle* if

(a) p is an equivariant map;

(b) For every $g \in G$, the map $g\colon p^{-1}(x) \to p^{-1}(gx)$, defined by $g(v) = gv$, $v \in p^{-1}(x)$, is an isomorphism of Banach spaces.

A G-morphism of two G-vector bundles $\xi = (p, E, B)$ and $\xi' = (p', E', B')$ is a pair of G-equivariant maps (φ, ψ), $\varphi\colon E \to E'$ and $\psi : B \to B'$, such that the following diagram commutes

$$\begin{tikzcd}
E \arrow[r, "\varphi"] \arrow[d, "p"] & E' \arrow[d, "p'"'] \\
B \arrow[r, "\psi"] & B'
\end{tikzcd}$$

and $\varphi_x := \varphi_{|p^{-1}(x)} \colon p^{-1}(x) \to p'^{-1}(\psi(x))$ is a linear (bounded) operator. If φ and ψ are two G-equivariant homeomorphisms, we say that the G-vector bundles (p, E, B) and (p', E', B') are G-isomorphic.

If M is a G-space with a structure of a finite-dimensional smooth manifold such that the action $\varphi\colon G \times M \to M$ of G on M is a smooth map, then M is called a *G-manifold*. Given an (orthogonal) G-representation V, any open invariant subset (equipped with the action induced from V) provides an example of a G-manifold.

In the case of a G-vector bundle (p, E, B) such that E and B are G-manifolds and $p: E \to B$ is a smooth mapping admitting smooth local trivializations, we say that (p, E, B) is a *smooth G-vector bundle*. If M is a G-manifold, then it is easy to see that the *tangent vector bundle* $\tau(M)$ of M is a smooth G-vector bundle. Let M be a Riemannian G-manifold. By using the Haar integral, it is easy to construct a G-invariant Riemannian metric $\langle\cdot,\cdot\rangle : \tau(M) \times \tau(M) \to \mathbb{R}$. Suppose that M is a G-submanifold of a Riemannian G-manifold W (equipped with an invariant metric). Then, the *normal vector bundle* $\nu(M)$ of M in W is a smooth G-vector bundle (in what follows, we will use the symbol $\nu_x(M)$ to denote the normal slice at $x \in M$).

Using the standard G-vector bundle techniques one can easily establish

PROPOSITION 2.13. *Assume that N is a G-submanifold of a Riemannian G-manifold M. Then, for every G-invariant compact set $K \subset N$, there exists an invariant neighborhood U of K, called invariant tubular neighborhood of N over K, satisfying the properties*:

(a) *$U_o := U \cap N$ is an invariant neighborhood of K in N;*

(b) *there exists $\varepsilon > 0$ such that the exponential map $\exp : \nu(N, \varepsilon)_{|U_o} \to M$ is a G-diffeomorphism onto U, where $\nu(N, \varepsilon)_{|U_o}$ denotes the ε-disc bundle of the normal bundle $\nu(N)$ restricted to U_o. In particular, every point $x \in U$ can be represented by a pair $(u, v) \in \nu(N, \varepsilon)_{|U_o}$, where $u \in U_o$ and v is a vector of norm less than ε orthogonal to N at u.*

Notice that if $M = V$, where V is an orthogonal G-representation, the map $\exp : \nu(N, \varepsilon)_{|U_o} \to V$ is simply defined by $\exp(u, v) = u + v$.

We complete this subsection with the following important

THEOREM 2.14. (Cf. [84,26].) *Let M be a G-manifold and H a subgroup of G. Then*:

(i) *$M_{(H)}$ is a G-invariant submanifold of M;*

(ii) *$M_{(H)}/G$ is a manifold. If $M_{(H)}$ is connected, then $M_{(H)}/G$ is also connected;*

(iii) *If (H) is a maximal orbit type in M, then $M_{(H)}$ is closed in M;*

(iv) *If (H) is a minimal orbit type in M and M/G is connected, then $M_{(H)}/G$ is a connected, open and dense subset of M/G;*

(v) *M_H is a $W(H)$-manifold with free $W(H)$-action, and $M_H/W(H)$ is a smooth manifold,*

where the minimal and maximal orbit types are taken with respect to the partial order relation (10).

2.4. *Fredholm operators*

Let $\mathbb{E}$ and $\mathbb{F}$ be two Banach spaces. Consider on $\mathbb{E} \oplus \mathbb{F}$ the norm defined by $\|x\| + \|y\|$ for $x \in \mathbb{E}$ and $y \in \mathbb{F}$. An (unbounded in general) linear operator $L : \mathrm{Dom}(L) \to \mathbb{F}$ (where $\mathrm{Dom}(L) \subset \mathbb{E}$ is a dense subspace), is called a *Fredholm operator* if

(i) L is a *closed* operator, i.e. the graph $\mathrm{Gr}(L)$ of L given by

$$\mathrm{Gr}(L) := \bigl\{(v, Lv)\colon\ v \in \mathrm{Dom}(L)\bigr\},$$

is a closed subspace in $\mathbb{E} \oplus \mathbb{F}$,

(ii) $\mathrm{Im}(L) := \{Lv\colon\ v \in \mathrm{Dom}(L)\}$ is a closed subspace of $\mathbb{F}$,

(iii) $\dim\ker L$ and $\operatorname{codim}\operatorname{Im} L := \dim \mathbb{F}/\operatorname{Im} L$ are finite.

For a given Fredholm operator L, the number $i(L) = \dim\ker L - \dim \mathbb{F}/\operatorname{Im}(L)$ is called the *index* of L.

Let $L\colon \mathrm{Dom}(L) \subset \mathbb{E} \to \mathbb{F}$ be an unbounded Fredholm operator and denote by $\pi_1\colon \mathbb{E} \oplus \mathbb{F} \to \mathbb{E}$ (resp. $\pi_2\colon \mathbb{E} \oplus \mathbb{F} \to \mathbb{F}$) the natural projection on $\mathbb{E}$ (resp. $\mathbb{F}$). Since $\pi_1\colon \mathrm{Gr}(L) \to \mathrm{Dom}(L)$ is a bounded linear operator, which is one-to-one and onto, the subspace $\mathrm{Dom}(L)$ can be equipped with a new norm $\|\cdot\|_L$ defined by

$$\|v\|_L := \|v\|_{\mathbb{E}} + \|Lv\|_{\mathbb{F}}, \tag{21}$$

where $\|\cdot\|_{\mathbb{E}}$ (resp. $\|\cdot\|_{\mathbb{F}}$) denotes the norm in $\mathbb{E}$ (resp. in $\mathbb{F}$), i.e. the norm $\|v\|_L$ is simply the norm of $\pi_1^{-1}(v)$ in $\mathbb{E} \oplus \mathbb{F}$. Since π_1 is one-to-one, we immediately obtain that $\mathbb{E}_L := (\mathrm{Dom}(L), \|\cdot\|_L)$ is a Banach space (equipped with so-called "*graph norm*") and the operator $L\colon \mathbb{E}_L \to \mathbb{F}$ is a bounded operator. Since $\ker L$ and $\operatorname{Im} L$ are still the same, the operator $L\colon \mathbb{E}_L \to \mathbb{F}$ is a bounded Fredholm operator of the same index as the original unbounded operator $L\colon \mathrm{Dom}(L) \subset \mathbb{E} \to \mathbb{F}$.

Consider the set $\mathrm{Sub}(\mathbb{E} \oplus \mathbb{F})$ of all the closed subspaces in $\mathbb{E} \oplus \mathbb{F}$ and denote by $\mathrm{Op}(\mathbb{E}, \mathbb{F})$ the set of all (in general unbounded) densely defined closed linear operators $L\colon \mathrm{Dom}(L) \subset \mathbb{E} \to \mathbb{F}$. It is convenient to identify any such operator L with its graph $\mathrm{Gr}(L)$, thus $\mathrm{Op}(\mathbb{E}, \mathbb{F})$ can be considered as a subspace of $\mathrm{Sub}(\mathbb{E} \oplus \mathbb{F})$.

The set $\mathrm{Sub}(\mathbb{E} \oplus \mathbb{F})$ and consequently, $\mathrm{Op}(\mathbb{E}, \mathbb{F})$ can be equipped with the *Hausdorff metric*. More precisely, consider a Banach space $\mathbb{V}$ and the set $\mathfrak{B}(\mathbb{V})$ consisting of all bounded closed subsets of $\mathbb{V}$. Then, for two sets $X, Y \in \mathfrak{B}(\mathbb{V})$, define

$$D(X, Y) := \inf\bigl\{r > 0\colon\ Y \subset X + B_r(0)\bigr\}, \quad B_r(0) = \bigl\{v \in \mathbb{V}\colon\ \|v\|_{\mathbb{V}} < r\bigr\}.$$

The *Hausdorff metric* d_H on $\mathfrak{B}(\mathbb{V})$ is given by

$$d_H(X, Y) := \max\bigl\{D(X, Y), D(Y, X)\bigr\}, \quad X, Y \in \mathfrak{B}(\mathbb{V}). \tag{22}$$

Using the Hausdorff metric, define the metric on $\mathrm{Sub}(\mathbb{E} \oplus \mathbb{F})$ by

$$d(V, W) := d_H\bigl(S(V), S(W)\bigr), \quad V, W \in \mathrm{Sub}(\mathbb{E} \oplus \mathbb{F}), \tag{23}$$

where $S(V)$ (resp. $S(W)$) denotes the unit sphere in V (resp. in W).

THEOREM 2.15. *The set* $\mathfrak{F}_0(\mathbb{E}, \mathbb{F})$ (*of Fredholm operators of index zero*) *is open in* $\mathrm{Op}(\mathbb{E}, \mathbb{F})$.

Theorem 2.15 (belonging to a functional analysis folklore) provides a theoretical basis for many of the arguments justifying the functional settings used in what follows. In particular,

COROLLARY 2.16.

(a) *Let* $L \in \mathfrak{F}_0(\mathbb{E}, \mathbb{E})$ *and assume that* $0 \in \sigma(L)$, *where* $\sigma(L)$ *denotes the spectrum of* L. *Then, for all sufficiently small reals* $t > 0$, $L - t\,\mathrm{Id}$ *is a one-to-one* (*bijective onto* $\mathbb{E}$) *Fredholm operator of index zero* (*in particular,* 0 *is an eigenvalue of* L *of finite multiplicity and it is an isolated point in* $\sigma(L)$).

(b) *Let* $L \in \mathfrak{F}_0(\mathbb{E}, \mathbb{E})$ *and let* $j : \mathbb{E}_L \to \mathbb{E}$ *be a compact operator. Then, for all* $t \in \mathbb{R}$, $L - t\,\mathrm{Id} \in \mathfrak{F}_0(\mathbb{E}, \mathbb{E})$ (*in particular,* $\sigma(L)$ *is a discrete in* $\mathbb{R}$ *set composed of eigenvalues of* L, *all of them of finite multiplicity*).

Observe that Theorem 2.15 admits an equivariant version. Namely, let $\mathbb{E}$ and $\mathbb{F}$ be real isometric Banach G-representations. Denote by $\mathrm{Op}^G := \mathrm{Op}^G(\mathbb{E} \oplus \mathbb{F})$ the set of all closed G-equivariant linear operators from $\mathbb{E}$ to $\mathbb{F}$. Clearly, for $L \in \mathrm{Op}^G$, the graph $\mathrm{Gr}(L)$ is a closed invariant subspace of $\mathbb{E} \oplus \mathbb{F}$, where we assume that G acts diagonally on $\mathbb{E} \oplus \mathbb{F}$.

The space $\mathbb{E}_L$ (cf. (21)) is a Banach G-representation. It is clear that $L : \mathbb{E}_L \to \mathbb{F}$ is a continuous equivariant operator.

Equip Op^G with the metric

$$\mathrm{dist}(L_1, L_2) = d\big(\mathrm{Gr}(L_1), \mathrm{Gr}(L_2)\big), \quad L_1, L_2 \in \mathrm{Op}^G,$$

where $d(\cdot,\cdot)$ is the metric on $\mathrm{Sub}(\mathbb{E} \oplus \mathbb{F})$ defined by (23). Let $\mathfrak{F}_0^G$ be the set of all closed G-equivariant Fredholm operators of index zero from $\mathbb{E}$ to $\mathbb{F}$. Since $\mathfrak{F}_0^G = \mathfrak{F}_0 \cap \mathrm{Op}^G$, by Theorem 2.15, $\mathfrak{F}_0^G$ is an open subset of Op^G.

We complete this subsection with the following important

REMARK AND NOTATION 2.17. Consider a Fredholm operator of index zero $L : \mathrm{Dom}(L) \subset \mathbb{E} \to \mathbb{F}$. A finite-dimensional linear operator $K : \mathbb{E} \to \mathbb{F}$ is called a (finite-dimensional) *resolvent* of L if the map $L + K : \mathrm{Dom}(L) \to \mathbb{F}$ is one-to-one. Since the operator $L : \mathbb{E}_L \to \mathbb{F}$ is continuous, $L + K : \mathbb{E}_L \to \mathbb{F}$ is also continuous and one-to-one. Hence, using open mapping theorem and the fact that a compact linear perturbation of a (bounded) Fredholm operator does not change its index, we obtain that $L + K$ is surjective and $(L + K)^{-1} : \mathbb{F} \to \mathbb{E}_L$ is bounded. Consequently, by applying the natural inclusion $j : \mathbb{E}_L \hookrightarrow \mathbb{E}$, the inverse $(L + K)^{-1} : \mathbb{F} \to \mathbb{E}$ is a bounded operator. Moreover, if the natural inclusion j is compact, the inverse $(L + K)^{-1}$ is also compact.

2.5. *Bibliographical remarks*

There is a lot of excellent graduate texts and monographs devoted to Lie group theory (see, for instance, [144,29,143,44]). For the background on transformation groups and G-spaces we refer to [26,84,43]. The material related to compact Lie group representations (both finite- and infinite dimensional) can be found in [23,27,58,142,151,152].

For the concepts and constructions related to smooth equivariant topology we refer to [26,37,84] (see also [54,72] for the fiber bundle background). The proof of Theorem 2.15 can be found, for instance, in [19].

3. Twisted equivariant degree: Construction and basic properties

3.1. *Topology behind the construction: Equivariant extensions and fundamental domains*

Induction over Orbit Types As it was mentioned in the Introduction, the equivariant degree "measures" homotopy obstructions for an equivariant map to have equivariant extensions without zeros on a set composed of several orbit types. Therefore, in this subsection we briefly discuss the following problem:

$$(P_A)\quad \begin{cases} \textit{Assume that } V \textit{ is a finite-dimensional } G\textit{-representation, } Y := V \setminus \{0\}, \\ X \textit{ is a } G\textit{-space and } A \subset X \textit{ is a closed invariant subset in } X. \textit{ Let} \\ f : A \to Y \textit{ be an equivariant map. Under which conditions does there} \\ \textit{exist an equivariant extension of } f \textit{ over } X \textit{ (resp. over a } G\textit{-invariant} \\ \textit{neighborhood of } A \textit{ in } X)? \end{cases}$$

Recall the principle of constructing equivariant maps via *induction over orbit types* (see [43]). Suppose G acts on X with finitely many orbit types $\mathcal{J}(X) = \{(H_1), (H_2), \dots, (H_k)\}$, where $\mathcal{J}(X)$ is equipped with the partial order such that $(H_i) \leqslant (H_l)$ implies $i \geqslant l$ (cf. (10)). Assume that $Y^{H_l} \neq \emptyset$ for $l = 1, \dots, k$, which is a necessary condition for the existence of equivariant maps from X to Y. Define a filtration of X by closed (in X) G-invariant subsets

$$A = X_0 \subset X_1 \subset X_2 \subset \cdots \subset X_k = X,$$

where, for $l \geqslant 1$,

$$X_l := \bigl\{x \in X\colon\ (G_x) = (H_j) \text{ for some } j \leqslant l\bigr\} \cup A.$$

Suppose that $f_{l-1} : X_{l-1} \to Y$ is an equivariant map for some $l \geqslant 1$. We are interested in the existence of an equivariant extension $f_l : X_l \to Y$ of the map f_{l-1}. It is well-known (cf. [43]) that if the map $f_{l-1}^{H_l} : X_{l-1}^{H_l} \to Y^{H_l}$ admits a $W(H_l)$-equivariant extension $s : X_l^{H_l} \to Y^{H_l}$, then there exists a unique G-equivariant extension $f_l : X_l \to Y$ such that $f_{l}{}_{|X_l^{H_l}} = s$. Therefore, we arrive at the following question: *When does there exist a* $W(H_l)$*-equivariant extension* s*?* Notice that $W(H_l)$ acts freely on $X_l^{H_l} \setminus X_{l-1}^{H_l}$, thus the general problem (P_A) can be reduced to the following one:

$$(P_B)\quad \begin{cases} \textit{Let } X, A, Y \textit{ and } f \textit{ be as above and assume that } G \textit{ acts freely on } X \setminus A. \\ \textit{Find a } G\textit{-equivariant extension of } f \textit{ over } X \textit{ (resp. over a } G\textit{-invariant} \\ \textit{neighborhood of } A \textit{ in } X). \end{cases}$$

Fundamental domains The key to the extension results we are looking for is the following notion.

DEFINITION 3.1. Let a topological group Q act on a finite-dimensional metric space X. Let $D_o \subset X$ be open in its closure D. Then, D is said to be a *fundamental domain* if the following conditions are satisfied:

- (i) $Q(D) = X$;
- (ii) $g(D_o) \cap h(D_o) = \emptyset$ for $g, h \in Q$, $g \neq h$;
- (iii) $X \setminus Q(D_o) = Q(D \setminus D_o)$;
- (iv) $\dim D = \dim X/Q$, $\dim(D \setminus D_o) < \dim D$, $\dim Q(D \setminus D_o) < \dim X$, where dim stands for the covering dimension.

It turns out that

THEOREM 3.2. (Cf. [99].) *Let G be a compact Lie group and X be a finite-dimensional metric G-space on which G acts freely. Then, a fundamental domain $D \subset X$ always exists.*

IDEA OF THE PROOF. Since G is a compact Lie group acting freely,

- (a) the orbit map $p : X \to X/G$ is the fiber bundle projection (cf. Gleason lemma);
- (b) X/G is metrizable (cf., for instance, [99]);
- (c) $\dim X/G = \dim X - \dim G$ (cf. [120]).

Using (a)–(c), one can construct $T_o \subset X/G$ such that:

- (α) T_o is dense in X/G;
- (β) T_o is open in X/G;
- (γ) $\dim((X/G) \setminus T_o) < \dim X/G$;
- (δ) the fiber bundle p is trivial over T_o.

Fix a trivialization $\psi : p^{-1}(T_o) \to G \times T_o$ and put $D_o := \psi^{-1}(\{1\} \times T_o)$. Then, $D = \overline{D_o}$ is the fundamental domain. □

Equivariant Kuratowski–Dugundji theorem Let us return to problem (P_B) (recall that we assume $X \setminus A$ is a free G-subspace). By Theorem 3.2, there exists a fundamental domain $D^{(0)} \subset L^{(0)} := X \setminus A$. Let $D_o^{(0)}$ be the corresponding open subset of $D^{(0)}$ satisfying the conditions (ii)–(iv) of Definition 3.1, and let $X^{(1)} := A \cup G(D^{(0)} \setminus D_o^{(0)})$, $L^{(1)} := X^{(1)} \setminus A$. Now, by applying Theorem 3.2 to $X^{(1)} \setminus A$, we obtain $X^{(2)}$ and $L^{(2)}$, etc. Consequently, by following the same steps, we obtain a closed finite G-invariant filtration

$$X = X^{(0)} \supset X^{(1)} \supset X^{(2)} \supset \cdots \supset X^{(r)} = A.$$

PROPOSITION 3.3. (Cf. [99].) *Under the above assumptions, any G-equivariant map $A \to Y$ extends over X if for all $i \geqslant 1$, any equivariant map $X^{(i)} \to Y$ has a* (*non-equivariant*) *extension over $X^{(i)} \cup D^{(i-1)}$.*

It should be pointed out that Proposition 3.3 together with the induction over orbit types reduces the equivariant extension problem to the non-equivariant one. For example, combining Proposition 3.3 with the classical Kuratowski–Dugundji theorem (see, for example, [99]), one can easily establish the corresponding version of this result for equivariant maps.

THEOREM 3.4. (Cf. [99].) *Let V be a finite-dimensional G-representation,* $Y := V \setminus \{0\}$, *X a G-space and $A \subset X$ a closed invariant subset in X with $\mathcal{J}(X \setminus A) = \{(H_1), \dots, (H_k)\}$. Let $f : A \to Y$ be a G-equivariant map. Then*:

(i) *There always exists a G-equivariant extension of f over a G-invariant neighborhood of A in X;*

(ii) *Assume that* $\dim V^{H_l} = n_l$ *for $l = 1, \dots, k$ and* $\dim(X_{(H_l)} \setminus A)/G < n_l$ *for all $l = 1, \dots, k$. Then f extends equivariantly over X.*

Under the same assumptions on X, A and V as in Problem (P_A), we say that a G-equivariant map $\mathfrak{h} : [0, 1] \times X \to V$ (where G will always be assumed to act trivially on $[0, 1]$) is a *G-homotopy* on X between $f_0 := \mathfrak{h}(0, \cdot)$ and $f_1 := \mathfrak{h}(1, \cdot)$.

In applications one is often interested in constructing G-equivariant extensions of homotopies without zeros. We have the following immediate consequence of Theorem 3.4.

COROLLARY 3.5. *Under the same assumptions on X, A and V as in Theorem* 3.4, *if* $\dim(X_{(H_l)} \setminus A)/G < n_l - 1$ *for all $l = 1, \dots, k$, then every G-equivariant homotopy $\mathfrak{h} : [0, 1] \times A \to Y$ has an equivariant extension over* $[0, 1] \times X$.

Regular fundamental domains To make the equivariant extension techniques based on the use of fundamental domains compatible with an appropriate equivariant degree theory, a more careful analysis of a geometry of a fundamental domain is needed.

DEFINITION 3.6. Under the notations of Definition 3.1, assume there exists an open *contractible* subset $T_0 \subset X/Q$ such that the natural projection $p : X \to X/Q$ induces the homeomorphism

$$(p_{|D_o})^{-1} : T_o \to D_o. \tag{24}$$

Then, D (resp. (24)) is called a *regular fundamental domain* (resp. *lifting homeomorphism*).

THEOREM 3.7. *Let X be a G-manifold on which a compact Lie group G acts freely. Let X/G be connected. Then, there always exists a regular fundamental domain.*

The proof of Theorem 3.7 uses the same ideas as Theorem 3.2 combined with the following statement applied to X/G.

THEOREM 3.8. *Let M be a smooth connected n-dimensional manifold* (*in general noncompact*). *Take a smooth countable triangulation of M. Then, there always exists a subset T_o of M satisfying the following conditions*:

(i) *T_o is open in M;*

(ii) *T_o is dense in M;*

(iii) *T_o is contractible*;

(iv) *$M \setminus T_o$ is contained in the $n - 1$-dimensional skeleton.*

Equivariant Dugundji theorem For several reasons we will also use the following equivariant analogue of the Dugundji theorem.

PROPOSITION 3.9. (Cf. [3,4].) *Let X be a metric G-space, $A \subset X$ a closed invariant subspace, E an isometric Banach G-representation, $C \subset E$ a convex closed invariant subset and $f : A \to C$ an equivariant map. Then, there exists an equivariant extension $F : X \to C$ of f.*

In particular, given two closed disjoint and invariant subsets A and B of a metric G-space X, by applying Proposition 3.9, one can construct an invariant Urysohn function from X to $[0, 1]$ (with the trivial G-action on $[0, 1]$), which in turn allows to obtain small invariant (disjoint) neighborhoods around A and B.

3.2. *Analysis behind the construction: Regular normal approximations*

Throughout this subsection, G stands for a compact Lie group, V for an orthogonal G-representation, $\Omega \subset \mathbb{R} \oplus V$ is an open bounded G-invariant subset and $f : \mathbb{R} \oplus V \to V$ an equivariant map.

DEFINITION 3.10. Given an orthogonal G-representation V and a bounded open invariant set $\Omega \subset \mathbb{R} \oplus V$, an equivariant map $f : \mathbb{R} \oplus V \to V$ is called Ω-*admissible* if $f(x) \neq 0$ for all $x \in \partial\Omega$ (in such a case we will call (f, Ω) an *admissible pair*). Similarly, one can define an Ω-*admissible homotopy*.

The “differential viewpoint” that we follow to assign the twisted equivariant degree to an Ω-admissible map, amounts to choosing “nice” representatives of (equivariant) homotopy classes satisfying reasonable regularity/transversality condition. However, such a choice should be determined by the following two observations:

(i) Since we are dealing with group actions admitting *several* orbit types, usual transversality conditions are not reasonable. Moreover, even in the case of one orbit type, the equivariance may be in “conflict” with regularity due to the restriction requirements on the dimensions of the orbits of zeros.

(ii) Any concept of “nice” representatives should provide a way to separate the sets of zeros belonging to different orbit types (in particular, to be compatible with the principle of constructing equivariant maps via induction over orbit types), and, in addition, to be “stable” with respect to the suspension procedure (in particular, to allow a passage to infinite dimensions).

The observation (ii) leads to a concept of *normal* maps, while (i) leads to a special *regularity* requirement restricted to separate orbit types.

Normality Recall that for every (closed) subgroup $H \subset G$, the set $\Omega_{(H)}$ is a G-submanifold of Ω. In particular, one can speak about the normal G-vector bundle around $\Omega_{(H)}$.

DEFINITION 3.11. (Cf. [60,96,99].) An Ω-admissible map $f : \mathbb{R} \oplus V \to V$ is called *normal* in Ω if for every $(H) \in \mathcal{J}(\Omega)$ and every

$$x \in f^{-1}(0) \cap \Omega_H,$$

the following (H)-*normality* condition at x is satisfied: *There exists* $\delta_x > 0$ *such that for all* $w \in \nu_x(\Omega_{(H)})$ *with* $\|w\| < \delta_x$,

$$f(x + w) = f(x) + w = w$$

(here $\nu_x(\cdot)$ stands for the normal slice).

Similarly, an Ω-admissible G-homotopy $h : [0, 1] \times (\mathbb{R} \oplus V) \to V$ is called a *normal homotopy* in Ω, if for every $(H) \in \mathcal{J}(\Omega)$ and for every $(t, x) \in h^{-1}(0) \cap ([0, 1] \times \Omega_H)$, the following (H)-*normality* condition at (t, x) is satisfied: There exists $\delta_{(t,x)} > 0$ such that for all $w \in \nu_{(t,x)}([0, 1] \times \Omega_{(H)})$ with $\|w\| < \delta_{(t,x)}$,

$$h(t, x + w) = h(t, x) + w = w. \tag{25}$$

From Definition 3.11, it follows immediately (cf. observation (ii) above)

PROPOSITION 3.12. *If* $f : \mathbb{R} \oplus V \to V$ *is normal in* Ω, *then*

$$\overline{\left(f^{-1}(0) \cap \Omega_{(K)}\right)} \cap \overline{\left(f^{-1}(0) \cap \Omega_{(H)}\right)} = \emptyset \quad \textit{for all } (K), (H) \in J(\Omega), \tag{26}$$

provided $(K) \neq (H)$.

Regular normality Introduce the following

DEFINITION 3.13. (Cf. [60,96,99].) Let $f : \mathbb{R} \oplus V \to V$ be and Ω-admissible G-equivariant map. We say that f is a *regular normal* map in Ω if:

(i) f is of class C^1;
(ii) f is normal in Ω;
(iii) for every $(H) \in \mathcal{J}(f^{-1}(0) \cap \Omega)$, zero is a regular value of $f_H := f|_{\Omega_H} : \Omega_H \to V^H$.

Similarly, one can define the notion of *regular normal* homotopy.

Using induction over orbit types, Proposition 2.13 and Theorem 2.14, one can establish

PROPOSITION 3.14. (See [60,96,99].) *Each* Ω-*admissible map* $f : \mathbb{R} \oplus V \to V$ *admits an arbitrarily close regular normal approximation.*

3.3. *Algebra behind the construction: Twisted groups and Burnside modules*

Hereafter, we use the following notations: Γ stands for a compact Lie group and $G := \Gamma \times S^1$, where S^1 is the unit circle in $\mathbb{C}$.

The twisted equivariant degree we are going to construct takes its values in a certain module (denoted by $A_1^t(G)$) over the Burnside ring $A(\Gamma)$. In this subsection, we briefly discuss generators and multiplication structure of $A_1^t(G)$.

Numbers $n(L,H)$ We start our exposition of algebraic ideas underlying the construction of twisted equivariant degree with introducing certain integers intimately related to the module $A_1^t(G)$. Moreover, as we will see later on, these numbers play a substantial role in the computational formulae for the twisted degree.

Given two closed subgroups $L \subset H$ of a compact Lie group Q, define the set (cf. [74,99])

$$N(L,H) = \{g \in Q\colon\ gLg^{-1} \subset H\} \tag{27}$$

and put

$$n(L,H) := \left|\frac{N(L,H)}{N(H)}\right|, \tag{28}$$

where the symbol $|X|$ stands for the cardinality of the set X.

It turns out that

PROPOSITION 3.15. *Let $L \subset H$ be two closed subgroups of a compact Lie group such that* $\dim W(L) = \dim W(H)$. *Then, the number $n(L,H)$ is finite.*

The above statement can be easily proved using the following well-known

PROPOSITION 3.16. (Cf. [26], Corollary 5.7.) *Let $L \subset H$ be two closed subgroups of the compact Lie group Q. Then, the orbit space*

$$\frac{(Q/H)^L}{W(L)} \quad ((Q/H)^L \text{ is considered as the left } W(L)\text{-space})$$

is finite.

Observe that the number $n(L,H)$ (where $L \subset H$ are two closed subgroups of Q with $\dim W(H) = \dim W(L)$) has a very simple geometric interpretation provided by the following

LEMMA 3.17. *Let L and H be two subgroups of a compact Lie group Q such that $L \subset H$ and* $\dim W(H) = \dim W(L)$. *Then, $n(L,H)$ represents the number of different subgroups $\widetilde{H}$ in the conjugacy class (H) such that $L \subset \widetilde{H}$. In particular, if V is an orthogonal Q-representation such that (L), $(H) \in \mathcal{J}(V)$, then $V^L \cap V_{(H)}$ is a disjoint union of exactly $m = n(L,H)$ sets V_{H_j}, $j = 1,2,\dots,m$, satisfying $(H_j) = (H)$.*

NOTATION. In what follows, in the case of two orbit types (L) and (H) such that $(L) \leqslant (H)$ (with the order relation defined by (10)), we will assume that the number $n(L, H)$ corresponds to representatives L and H such that $L \subset H$. In the case the orbit types (L) and (H) are not comparable with respect to the order relation (10), we will simply put $n(L, H) = 0$.

REMARK 3.18. It should be pointed out that the numbers $n(L, H)$ are essentially of non-abelian nature: if Q is an abelian compact Lie group and $L \subset H \subset Q$ are (closed) subgroups, then $n(L, H) \equiv 1$. In a certain sense, the numbers $n(L, H)$ demonstrate a principal difference between the (twisted) equivariant degree theory for non-abelian groups from that for abelian ones.

Burnside ring $A(\Gamma)$ Recall some basic facts related to the Burnside Ring $A(\Gamma)$ of the compact Lie group Γ.

Denote by $\Phi_0(\Gamma)$ the set of conjugacy classes (H) for $H \subset \Gamma$, such that $W(H) = N(H)/H$ is finite, i.e. $\dim W(H) = 0$. Let $A(\Gamma)$ be the free abelian group generated by $(H) \in \Phi_0(\Gamma)$. There is a *multiplication* operation on $A(\Gamma)$ which induces a structure of a ring with identity on $A(\Gamma)$. In order to define the multiplication operation, observe that

$$\begin{aligned}(\Gamma/H \times \Gamma/K)_{(L)}/\,\Gamma &\cong (\Gamma/H \times \Gamma/K)_L/N(L)\\ &\subset (\Gamma/H \times \Gamma/K)^L/N(L)\\ &= (\Gamma/H^L \times \Gamma/K^L)/W(L)\\ &= \left(\frac{N(L,H)}{H} \times \frac{N(L,K)}{K}\right)\Big/ W(L).\end{aligned}$$

Since, by Proposition 3.16, the spaces Γ/H^L and Γ/K^L consist of finitely many $W(L)$-orbits and, by assumption, $W(L)$ is finite, one obtains that Γ/H^L and Γ/K^L are finite. Consequently the set $(\Gamma/H \times \Gamma/K)_{(L)}/\,\Gamma$ is finite.

The multiplication table of the generators (H) is given by the relation

$$(H) \cdot (K) = \sum_{(L) \in \Phi_0(\Gamma)} n_L(H, K)(L) \tag{29}$$

where $n_L(H, K)$ denotes the number of elements in the set $(\Gamma/H \times \Gamma/K)_{(L)}/\,\Gamma$, i.e.

$$n_L(H, K) := \left|(\Gamma/H \times \Gamma/K)_{(L)}/\,\Gamma\right|,$$

where we denote by $|X|$ the number of elements in the set X. In other words, the number $n_L(H, K)$ represents the number of the orbits of the orbit type (L) contained in the space $\Gamma/H \times \Gamma/K$.

The ring $A(\Gamma)$ is called the *Burnside Ring* of Γ. We refer to [93] for more details and proofs related to the above definition of the Burnside Ring.

The computations of the multiplication table for $A(\Gamma)$ can be effectively conducted using a simple recurrence formula:

$$n_L(H,K) = \frac{n(L,H)|W(H)|n(L,K)|W(K)| - \sum_{(\tilde{L})>(L)} n(L,\tilde{L}) n_{\tilde{L}} |W(\tilde{L})|}{|W(L)|}, \tag{30}$$

where $|X|$ denotes the number of elements in the set X.

Twisted Subgroups of $G = \Gamma \times S^1$ Here, given a compact Lie group Γ, we briefly discuss generators of the $\mathbb{Z}$-module $A_1^t(G)$, where $G := \Gamma \times S^1$. To this end, we need to classify subgroups H of G.

Consider the following diagram

$$\begin{tikzcd}
& H \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\
\Gamma & & S^1
\end{tikzcd}$$

where π_1 and π_2 are projections (homomorphisms) on Γ and S^1 respectively. Let $K := \pi_1(H)$ and consider $\ker \pi_1 = H \cap \{e\} \times S^1$, where e denotes the neutral element of Γ. If $\ker \pi_1 = \{e\} \times S^1$, then simply $H = K \times S^1$, i.e. H is a "product" subgroup. Otherwise $\ker \pi_1 = \{e\} \times \mathbb{Z}_l$, for some $l \geqslant 1$, and in this case, we are dealing with "twisted" subgroups. In the second case, we still have that $H/\ker \pi_1 \subset \Gamma \times S^1$, thus we can consider the diagram below.

$$\begin{tikzcd}
& H \arrow[d, "\mu_l"] & \\
& H/\ker\pi_1 \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\
K \arrow[rr, dotted, "\varphi"] & & S^1
\end{tikzcd}$$

Since $\pi_1 : H/\ker \pi_1 \to K$ is one-to-one and onto, we can define the homomorphism $\varphi := \pi_2 \circ \pi_1^{-1} : K \to S^1$, and consequently the subgroup $H/\ker \pi_1$ is the graph of φ, i.e.

$$H/\ker \pi_1 := \left\{(k,z) \in K \times S^1 \colon \varphi(k) = z\right\},$$

and since the subgroup H is the inverse image $\mu_l^{-1}(H/\ker \pi_1)$, we obtain:

$$H = \left\{(k,z) \in K \times S^1 \colon \varphi(k) = z^l\right\}.$$

In this case, we call the subgroup H a *twisted* (by the homomorphism $\varphi : K \to S^1$) *l-folded* subgroup which is denoted by $K^{\varphi,l}$ (for $l = 1$ we simply write K^{φ}). It is easy to see that each (closed) subgroup $K \subset \Gamma \cong \Gamma \times \{1\} \subset G$, where 1 is the unit in S^1, is (trivially) twisted by the homomorphism $\varphi : K \to S^1$ defined as $\varphi(x) \equiv 1$.

Consider a twisted l-folded subgroup $H = K^{\varphi,l}$ and the normalizer $N(H)$ of the group H. Notice that

$$\begin{aligned} N(H) = N(K^{\varphi,l}) &= \{(\gamma, w) \in \Gamma \times S^1\colon\ (\gamma, w)K^{\varphi,l}(\gamma, w)^{-1} = K^{\varphi,l}\} \\ &= \{(\gamma, w) \in \Gamma \times S^1\colon\ \gamma k \gamma^{-1} \in K,\ \varphi(\gamma k \gamma^{-1}) = \varphi(k)\ \forall k \in K\} \\ &= \{\gamma \in N(K)\colon\ \varphi(\gamma k \gamma^{-1}) = \varphi(k)\ \forall k \in K\} \times S^1 =\colon N_o \times S^1. \end{aligned}$$

Consequently,

$$\dim W(H) = \dim N_o/K + 1 \leqslant \dim W_\Gamma(K) + 1,$$

where $W_\Gamma(\cdot)$ stands for the Weyl group in Γ. In the case $(K) \in \Phi_0(\Gamma)$, i.e. $\dim W_\Gamma(K) = 0$, we have that $\dim W(H) = 1$.

REMARK 3.19. Notice that if $H = K^{\varphi,l}$ is a twisted subgroup and $(\widetilde{H}) \leqslant (H)$, then $\widetilde{H}$ is also a twisted subgroup. In particular, every subgroup $H' \in (H)$ is twisted. Consequently, it makes sense to talk about the lattice of the conjugacy classes of twisted subgroups in G. Moreover, if $\dim W_\Gamma(K) = 0$ and $L^{\psi,l'}$ is a twisted subgroup such that $(L^{\psi,l'}) \geqslant (K^{\varphi,l})$, then it is easy to see that $\dim W_\Gamma(L) = 0$.

With these preliminaries on hands, introduce the following

DEFINITION 3.20. Denote by $\Phi_1^t(G)$ the set of all conjugacy classes of the φ-twisted l-folded subgroups $H = K^{\varphi,l}$, $l = 1, 2, \ldots$, such that the Weyl group $W_\Gamma(K)$ of K in Γ is finite (i.e. $\dim W_\Gamma(K) = 0$). We call any element from $\Phi_1^t(G)$ a *twisted conjugacy class.*

In what follows, twisted conjugacy classes will perform as generators of the module $A_1^t(G)$. Clearly, the construction of twisted subgroups provides (algebraic) features allowing to classify symmetric properties of periodic solutions to dynamical systems. What is, probably, less obvious, twisted subgroups are intimately connected to *orientability* properties of the corresponding to them Weyl groups, which, in turn is extremely important for the construction of the (twisted) equivariant degree.

To be more specific, let Q be a compact Lie group and M a Q-manifold. Take $x \in M$ with the isotropy $Q_x = H$. By passing to the $W(H)$-manifold M^H, we arrive at the following question: *does the orbit $N := W(H)(x)$ admit a natural $W(H)$-invariant orientation?* Since $W(H)$ is an orientable manifold, one can choose an orientation $\mathfrak{o}_{W(H)}$ on $W(H)$ and use the action of $W(H)$ on M^H to induce an orientation $\mathfrak{o}$ on N via the diffeomorphism $\mu_{x_o}\colon W(H) \to N$ given by $\mu_{x_o}(h) := hx_o$, $h \in W(H)$, for a certain distinguished point $x_o \in N$. Clearly, in order to this construction to be natural, the orientation of N obtained in this way should not depend on a choice of the point x_o. However, this property is satisfied only if the orientation $\mathfrak{o}_{W(H)}$ is invariant with respect to *right* translations on $W(H)$. On the other hand, the constructed orientation $\mathfrak{o}$ of the orbit N is $W(H)$-invariant if and only if $\mathfrak{o}_{W(H)}$ is invariant with respect to *left* translations on $W(H)$. Consequently, we obtain that an orbit N admits a natural $W(H)$-invariant orientation if and only if $W(H)$ admits an orientation which is invariant with respect to *both* left and right translations (in this case, $W(H)$ is called *bi-orientable*).

Clearly, any reasonable equivariant degree theory allowing an "algebraic count" of orbits of periodic solutions should deal with the above mentioned *natural* $W(H)$-invariant orientation on orbits in question (or, equivalently, with bi-orientable groups $W(H)$). It is easy to see that if $W(H)$ is finite, Abelian or contains an odd number of connected components, then it is bi-orientable. On the other hand, the group O(2) presents an example of the non-bi-orientable group. Our choice of the module $A_1^t(G)$ for the range of values of the (twisted) equivariant degree is essentially based on the following

PROPOSITION 3.21. *Let Γ be a compact Lie group, $G := \Gamma \times S^1$ and $H = K^{\varphi,l}$ a twisted subgroup of G such that $(H) \in \Phi_1^t(G)$. Then*:

(i) *the connected component of $1 \in W(H)$ can be canonically identified with S^1*;

(ii) *the Weyl group $W(H)$ of H in G is bi-orientable and can be equipped with the natural orientation induced from S^1.*

$A(\Gamma)$-module $A_1^t(G)$ Take the set $\Phi_1^t(G)$ introduced in Definition 3.20 and consider the free $\mathbb{Z}$-module $A_1^t(G) := \mathbb{Z}[\Phi_1^t(G)]$. Using Proposition 3.16 and the (finiteness) argument similar to that for the Burnside ring (cf. formula (29)), one can easily establish the following

THEOREM 3.22. *There is a "multiplication" function $\cdot : A(\Gamma) \times A_1^t(G) \to A_1^t(G)$ defined on the generators $(K) \in \Phi_0(\Gamma)$ and $(H^{\varphi,l}) \in \Phi_1^t(G)$ as follows*:

$$(K) \cdot (H^{\varphi,l}) = \sum_{(L)} n_L (L^{\varphi,l}), \tag{31}$$

where the summation is taken over all subgroups L such that $W(L)$ is finite, $L = \gamma^{-1} K \gamma \cap H$ for some $\gamma \in \Gamma$ and

$$n_L = \left| \left(\frac{G}{K \times S^1} \times \frac{G}{H^{\varphi,l}} \right)_{(L^{\varphi,l})} \Big/ G \right|.$$

The $\mathbb{Z}$-module $A_1^t(G)$ equipped with the multiplication "$\cdot$" becomes an $A(\Gamma)$-module.

Below we describe the process one can follow to establish the multiplication tables for the $A(\Gamma)$-module $A_1^t(G)$, based on the same idea of a recurrence formula (30).

Let $(K) \in \Phi_0(\Gamma)$ be a generator of $A(\Gamma)$ and $(H^{\varphi,l}) \in \Phi_1^t(G)$ a generator of $A_1^t(G)$. In order to compute the coefficients n_L given in formula (31), apply the following recurrence formula

$$n_L = \Bigg[n(L,K) \big| W(K) \big| n(L^{\varphi,l}, H^{\varphi,l}) \left| \frac{W(H^{\varphi,l})}{S^1} \right|$$

$$- \sum_{(\tilde{L})>(L)} n(L^{\varphi,l}, \tilde{L}^{\varphi,l}) n_{\tilde{L}} \left| \frac{W(\tilde{L}^{\varphi,l})}{S^1} \right| \Bigg] \Bigg/ \left| \frac{W(L^{\varphi,l})}{S^1} \right| \tag{32}$$

where for a set Y, $|Y|$ denotes the number of elements in Y.

REMARK 3.23. Notice that the multiplication formulae for conjugacy classes of trivially twisted subgroups $(H) \in \Phi_1^t(G)$ (i.e. $H = H^{\varphi,1}$ with $\varphi \equiv 1$), where $(H) \in \Phi_0(\Gamma)$, coincide with the multiplication formulae for the Burnside ring $A(\Gamma)$. That means, the multiplication formulae for the Burnside ring are contained in the $A(\Gamma)$-module multiplication formulae for $A_1^t(G)$.

REMARK 3.24. Let us point out that the formula (32) allows us to implement the multiplication tables for the $A(\Gamma)$-module $A_1^t(G)$ into a computer software, which can be effectively used for the computations of the twisted degree.

3.4. *Construction*

We are now in a position to present the construction of twisted equivariant degree.

We start with the regular normal case, i.e. take an orthogonal G-representation V, open bounded G-invariant subset $\Omega \subset \mathbb{R} \oplus V$ and an Ω-admissible *regular normal* map $f : \mathbb{R} \oplus V \to V$. Fix $(H) \in \Phi_1^t(G, \Omega) := \{(H) \in \mathcal{J}(\Omega)\colon\ (H) \in \Phi_1^t(G)\}$. By Theorem 2.14(v), Ω_H is a free $W(H)$-manifold. Assuming, without loss of generality, that $\Omega_H/W(H)$ is connected, and using Theorem 3.7, one can find a regular fundamental domain D on Ω_H with the lifting homeomorphism

$$\xi := (p_{|D_o})^{-1} : T_o \to D_o$$

(cf. Definitions 3.1 and 3.6). Moreover, since f is normal, it satisfies the condition (26). Hence, one can choose D in such a way that

$$f_H^{-1}(0) \cap (D \setminus D_o) = \emptyset, \tag{33}$$

where $f_H := f|_{\Omega_H}$. Since f is regular normal, the set

$$p\bigl(f_H^{-1}(0) \cap D_o\bigr)$$

is finite, therefore, one can construct T_o providing

$$p\bigl(f_H^{-1}(0)\bigr) \subset T_o. \tag{34}$$

Since $\Omega_H/W(H)$ is a (smooth) manifold (see Theorem 2.14(v)) and T_o is open in $\Omega_H/W(H)$, T_o is a manifold as well. Since T_o is *contractible*, it is *orientable*. By assumption, $(H) \in \Phi_1^t(G)$, therefore, fix the natural orientation on $W(H)$ which is invariant with respect to both left and right translations (cf. Proposition 3.21). Also, fix an orientation on V^H inducing an orientation on $\mathbb{R} \oplus V^H$. By bi-orientability of $W(H)$, for every $x \in T_o$, the orbit $N := W(H)(\xi x)$ has a natural $W(H)$-invariant orientation. Therefore, we can assume that the orientation on T_o is chosen in such a way that for *all* $x \in T_o$, we have that the orientation on $\mathbb{R} \oplus V^H$ is obtained from the orientation on D_o followed by the orientation on N, and the lifting homeomorphism $\xi : T_o \to D_o$ preserves the orientation.

Put

$$n_H = n_H(f) := \deg(f_H \circ \xi, T_o) \tag{35}$$

(here "deg" denotes the (local) Brouwer degree with respect to zero (this degree is correctly defined according to (33) and (34))).

Furthermore, take $A_1^t(G)$ and define the twisted equivariant degree $G\text{-Deg}^t(f, \Omega)$ by:

$$G\text{-Deg}^t(f, \Omega) = \sum_{(H) \in \Phi_1^t(G, \Omega)} n_H(H), \tag{36}$$

where n_H is defined by (35).

If now $f : \mathbb{R} \oplus V \to V$ is an *arbitrary* Ω-admissible map (in general, neither normal nor regular), choose a regular normal Ω-admissible map $g : \mathbb{R} \oplus V \to V$ close to f (see Proposition 3.14) and put

$$G\text{-Deg}^t(f, \Omega) := G\text{-Deg}^t(g, \Omega). \tag{37}$$

REMARK 3.25. Some comments related to the definition of twisted degree are in order.

(i) To some extent, the bi-orientability of $W(H)$ is used to define an "index of an orbit" ("local twisted degree"), while the contractibility of T_o allows a correct definition on the whole T_o ("global twisted degree").

(ii) If we choose an orientation on D_o in such a way that ξ preserves it, then $\deg(f_H, D_o)$ is correctly defined and coincides with $\deg(f_H \circ \xi, T_o)$. In this sense, one can think of $n_H(f)$ as a "degree of f_H on a fundamental domain D".

(iii) It is possible to show that the definition of twisted degree given by (35)–(37) is independent of a choice of fundamental domains as well as regular normal approximations.

3.5. *Axiomatic approach to twisted degree*

Basic properties The twisted degree defined above satisfies all the standard properties required from any reasonable "degree theory". To see that, we need the following

DEFINITION 3.26. Let V be an orthogonal G-representation with $G = \Gamma \times S^1$ and $f : \mathbb{R} \oplus V \to V$ a regular normal map such that $f(x_o) = 0$ with $G_{x_o} = H$ and $(H) \in \Phi_1^t(G, V)$. Let

$U_{G(x_o)}$ be a G-invariant tubular neighborhood around $G(x_o)$ such that $f^{-1}(0) \cap U_{G(x_o)} = G(x_o)$ (cf. Proposition 2.13). Then, f is called a *tubular map*.

Next, fix the natural orientation on $W(H)$ which is invariant with respect to both left and right translations. Also, fix an orientation on V^H inducing an orientation on $\mathbb{R} \oplus V^H$. By bi-orientability of $W(H)$, choose the natural orientation on the orbit $W(H)(x_o)$ and assume that the orientation on the normal slice $\nu_{x_o} := \nu_{x_o}(W(H)(x_o))$ in $\mathbb{R} \oplus V^H$ is oriented in such a way that the orientation on $\mathbb{R} \oplus V^H$ is obtained from the orientation on ν_{x_o} followed by the orientation on $W(H)(x_o)$. Then, we call $n_{x_o} = \operatorname{sign} \det Df^H(x_o)|_{\nu_{x_o}}$ the *local index* of f at x_o in $U_{G(x_o)}$ (here $f^H := f|_{\Omega^H}$ and "D" stands for the differential).

PROPOSITION 3.27. *Let V be an orthogonal G-representation with $G = \Gamma \times S^1$, $\Omega \subset \mathbb{R} \oplus V$ an open bounded invariant subset and $f : \mathbb{R} \oplus V \to V$ an equivariant Ω-admissible map. Then, the twisted equivariant degree defined by* (35)–(37) *satisfies the following properties*:

(P1) (Existence) *If G-$\mathrm{Deg}^t(f, \Omega) = \sum_{(H)} n_H(H)$ is such that $n_{H_o} \neq 0$ for some $(H_o) \in \Phi_1^t(G, \Omega)$, then there exists $x \in \Omega$ with $f(x) = 0$ and $G_x \supset H_o$.*

(P2) (Additivity) *Assume that Ω_1 and Ω_2 are two G-invariant open disjoint subsets of Ω such that $f^{-1}(0) \cap \Omega \subset \Omega_1 \cup \Omega_2$. Then,*

$$G\text{-}\mathrm{Deg}^t(f, \Omega) = G\text{-}\mathrm{Deg}^t(f, \Omega_1) + G\text{-}\mathrm{Deg}^t(f, \Omega_2).$$

(P3) (Homotopy) *Suppose $h : [0,1] \times \mathbb{R} \oplus V \to V$ is an Ω-admissible G-equivariant homotopy. Then,*

$$G\text{-}\mathrm{Deg}^t(h_t, \Omega) = \text{const}$$

(*here $h_t := h(t, \cdot, \cdot)$, $t \in [0,1]$*).

(P4) (Suspension) *Suppose that W is another orthogonal G-representation and let U be an open bounded G-invariant neighborhood of* 0 *in W. Then,*

$$G\text{-}\mathrm{Deg}^t(f \times \mathrm{Id}, \Omega \times U) = G\text{-}\mathrm{Deg}^t(f, \Omega).$$

(P5) (Normalization) *Suppose f is a tubular map around $G(x_o)$ with $H := G_{x_o}$ and the local index n_{x_o} of f at x_o in a tubular neighborhood $U_{G(x_o)}$. Then,*

$$G\text{-}\mathrm{Deg}^t(f, U_{G(x_o)}) = n_{x_o}(H).$$

(P6) (Elimination) *Suppose f is normal in Ω and $\Omega_H \cap f^{-1}(0) = \emptyset$ for every $(H) \in \Phi_1^t(G, \Omega)$. Then,*

$$G\text{-}\mathrm{Deg}^t(f, \Omega) = 0.$$

(P7) (Excision) *If $f^{-1}(0) \cap \Omega \subset \Omega_0$, where $\Omega_0 \subset \Omega$ is an open invariant subset, then*

$$G\text{-}\mathrm{Deg}^t(f, \Omega) = G\text{-}\mathrm{Deg}^t(f, \Omega_0).$$

(P8) (Hopf property) *Suppose that* $\Omega_H/W(H)$ *is connected for all* $(H) \in \Phi_1^t(G,\Omega)$ *and* $\Omega_K = \emptyset$ *for all* $(K) \notin \Phi_1^t(G,\Omega)$. *Let* $f, g : \mathbb{R} \oplus V \to V$ *be two* Ω*-admissible* G*-equivariant maps such that*

$$G\text{-Deg}^t(f,\Omega) = G\text{-Deg}^t(g,\Omega).$$

Then, f *and* g *are* G*-equivariantly homotopic by an* Ω*-admissible homotopy.*

PROOF. To establish properties (P1)–(P4), (P7) in the regular normal case, one should combine Proposition 3.12 and formulae (35), (36) with appropriate properties of the (local) Brouwer degree. In the general case, take regular normal approximations sufficiently close to f (resp. h) and use the standard compactness argument.

Property (P5) (resp. (P6)) follows immediately from the regular value definition of the Brouwer degree (resp. definition of the twisted degree).

To prove property (P8), assume (without loss of generality) that f and g are regular normal. Combining Proposition 3.12 with the classical Hopf Theorem (see, for instance, [44]) applied to appropriate fundamental domains, construct "local" homotopies between f and g around the corresponding zeros and extend the obtained homotopies by G-equivariance. Next, using Theorem 3.4 (see also Corollary 3.5), extend the previous "partial homotopy" over $\Omega \times [0,1]$. The latter homotopy may have "new" zeros which can be eliminated by using Proposition 3.9. □

REMARK 3.28. Following the same lines as above, one can associate with an open Γ-invariant bounded set $U \subset W$ (W is an orthogonal Γ-representation) and a continuous Γ-equivariant map $g : (U, \partial U) \to (W, W \setminus \{0\})$ the Γ-equivariant degree $\Gamma\text{-Deg}(g,U)$ taking its values in the Burnside ring $A(\Gamma)$ and satisfying the properties similar to (P1)–(P8) with obvious modifications (see [93] and [99] for more details).

Axiomatic approach The following result provides an axiomatic approach to the twisted degree theory.

THEOREM 3.29. *There exists a unique* $A_1^t(G)$*-valued function on the set of* Ω*-admissible* G*-equivariant maps and homotopies satisfying conditions* (P1)–(P6).

PROOF. The *existence* part of Theorem 3.29 is provided by Proposition 3.27. To prove the *uniqueness* part, assume $\overline{G\text{-Deg}^t}$ is another 'degree' satisfying (P1)–(P6). Take an arbitrary admissible pair (f,Ω). By homotopy property, f can be assumed to be regular normal. By additivity (i.e. excision) and elimination properties, we can assume that $\Omega \cap f^{-1}(0)$ contains only points of the orbit types $(H) \in \Phi_1^t(G)$. Since f is regular normal, the set $\Omega \cap f^{-1}(0)$ is composed of a finite number of G-orbits. Take tubular neighborhoods isolating the above orbits (this is doable, since we have finitely many zero orbits). By additivity property, $\overline{G\text{-Deg}^t}(f,\Omega)$ is equal to the sum of degrees of restrictions of f to the tubular neighborhoods. Finally, by normalization property, the orbits in question lead to "local indices", which implies

$$\overline{G\text{-Deg}^t}(f,\Omega) = G\text{-Deg}^t(f,\Omega).$$ □

Theorem 3.29 provides the application of the twisted degree without going into topological, analytical and algebraic roots underlying its construction – this is an obvious advantage of the axiomatic approach. In addition, this theorem can be extended to equivariant maps $\mathbb{R}^n \oplus V \to V$ with G replaced by $\Gamma \times T^n$, where T^n stands for the n-dimensional torus (such a setting naturally appears in multiparameter bifurcation problems). However, Theorem 3.29 amounts to the *normality* property (see (P5)–(P6)) which (being of great theoretical importance) is easy to formulate but difficult to achieve in practice.

Fortunately, in the setting relevant to our discussion, there is an effective way "to go around the normality problem", namely:

(i) to define the twisted degree for $G = S^1$ by a list of axioms (of course, equivalent to those presented in Proposition 3.27) with the normality property not being addressed whatsoever;

(ii) for an arbitrary $G = \Gamma \times S^1$, to obtain a formula canonically reducing the computations of G-degree to those for the S^1-degree and numbers $n(L, H)$.

3.6. S^1-*degree*

Assume $\Gamma = \{e\}$, i.e. $G = \{e\} \times S^1 \cong S^1$. Recall that any Abelian compact Lie group is bi-orientable. Therefore, $A_1^t(S^1) := A_1(S^1)$ is the free $\mathbb{Z}$-module generated by $(\mathbb{Z}_k)$, $k = 1, 2, 3 \dots$.

Take an S^1-representation V, an open S^1-invariant bounded set $\Omega \subset \mathbb{R} \oplus V$, and an Ω-admissible S^1-equivariant map $f : \mathbb{R} \oplus V \to V$. Then,

$$S^1\text{-Deg}(f, \Omega) = n_{k_1}(\mathbb{Z}_{k_1}) + \cdots + n_{k_r}(\mathbb{Z}_{k_r}), \tag{38}$$

where $n_{k_i} \in \mathbb{Z}$. To provide an axiomatic approach to the degree given by (38), we need two auxiliary constructions.

Two constructions

(i) *Basic maps*. Denote by $\mathcal{V}_k$, $k = 1, 2, 3, \dots$, the (non-trivial) k-th real irreducible S^1-representation, i.e. $\mathcal{V}_k$ is the space $\mathbb{R}^2 = \mathbb{C}$ with the S^1-action given by $\gamma z := \gamma^k \cdot z$, $\gamma \in S^1$, $z \in \mathbb{C}$, and define the set

$$^k\Omega := \left\{(t, z) \in \mathbb{R} \oplus \mathcal{V}_k \colon\ |t| < 1,\ \frac{1}{2} < |z| < 2\right\} \tag{39}$$

and $b : \mathbb{R} \oplus \mathcal{V}_k \to \mathcal{V}_k$ by

$$b(t, z) := \big(1 - |z| + \mathrm{i}t\big) \cdot z, \quad (t, z) \in \mathbb{R} \oplus \mathcal{V}_k, \tag{40}$$

where "$\cdot$" denotes the complex multiplication in $\mathcal{V}_k = \mathbb{C}$. Clearly, b is S^1-equivariant and $^k\Omega$-admissible. In what follows, b is called the k-th *basic map*.

(ii) *l-folding*. For every integer $l = 1, 2, 3, \dots$, define the homomorphism $\theta_l : S^1 \to S^1$ (called *l-folding*), by $\theta_l(\gamma) = \gamma^l$, $\gamma \in S^1$, and define the induced by θ_l homomorphism $\Theta_l : A_1(S^1) \to A_1(S^1)$, by

$$\Theta_l(\mathbb{Z}_k) := (\mathbb{Z}_{kl}), \quad k = 1, 2, 3, \dots,$$

i.e. $\Theta_l(\mathbb{Z}_k) = (\theta_l^{-1}(\mathbb{Z}_k))$, where $(\mathbb{Z}_k)$ are the free generators of $A_1(S^1)$.

If $f : \mathbb{R} \oplus V \to V$ is an Ω-admissible S^1-equivariant map for a certain open bounded S^1-invariant subset $\Omega \subset \mathbb{R} \oplus V$, then for every integer $l = 1, 2, 3, \dots$, define the *associated l-folded* S^1-representation ${}^l(V)$, which is the same vector space V with the S^1-action '$\cdot$' given by

$$\gamma \cdot v := \theta_l(\gamma)v = \gamma^l v, \quad \gamma \in S^1, \ v \in V.$$

Next, the map f considered from $\mathbb{R} \oplus {}^l(V)$ to ${}^l(V)$, is S^1-equivariant as well. The set Ω considered as an S^1-subset of $\mathbb{R} \oplus {}^l(V)$ is denoted by ${}^l(\Omega)$. In what follows, we will say that the pair $(f, {}^l(\Omega))$ is the *l-folded* admissible pair *associated* with (f, Ω).

Axiomatic approach to S^1-degree

THEOREM 3.30. *There exists a unique $A_1(S^1)$-valued function S^1-*Deg *defined on Ω-admissible S^1-equivariant maps and homotopies and satisfying properties* (P1)–(P4) (*with $G = S^1$*) *as well as the following ones*:

(P5′) (Normalization) *For the* 1*-st basic map $b : \mathbb{R} \oplus \mathcal{V}_1 \to \mathcal{V}_1$,*

$$S^1\text{-Deg}(b, {}^1\Omega) = (\mathbb{Z}_1).$$

(P6′) (Elimination) *If V is a trivial S^1-representation, then*

$$S^1\text{-Deg}(f, \Omega) = 0.$$

(F) (Folding) *Let ${}^l(V)$ be the l-folded representation associated with V, and $(f, {}^l(\Omega))$ the l-folded admissible pair associated with (f, Ω). Then,*

$$S^1\text{-Deg}(f, {}^l(\Omega)) = \Theta_l\big[S^1\text{-Deg}(f, \Omega)\big].$$

PROOF. The *existence* part of Theorem 3.30 is obvious. The *uniqueness* easily follows from

LEMMA 3.31. *Under the conditions and notations of Definition* 3.26, *assume $G = S^1$ and $G_{x_o} = \mathbb{Z}_{k_o}$. Assume that S^1-*Deg *is a function provided by Theorem* 3.30. *Then,*

$$S^1\text{-Deg}(f, \Omega) = n_{x_o}(\mathbb{Z}_{k_o}),$$

where n_{x_o} is the local index of f at x_o. □

3.7. *Computational techniques for twisted degree*

Statement of the problem The goal of this subsection is to show how the axiomatic approach described in the previous two sections, being combined with some additional techniques, allows to obtain twisted degree results for an important class of equivariant maps which naturally appear in the symmetric Hopf bifurcation problems.

We start with a simple observation that every S^1-representation admits the so-called *natural complex structure*, which turns out to be a convenient setting for the discussion of Hopf bifurcation problems and a natural way of describing the S^1-action to carry out certain computations. To be more specific, let V be an (orthogonal) S^1-representation with $V^{S^1} = \{0\}$. Then, one can define on V a complex structure sensitive to the S^1-action as follows. Assume, for a moment, that $V = \mathcal{V}_k$. Then, for $z \in \mathbb{C}$, put $z = \gamma|z|$, where $\gamma = \mathrm{e}^{\mathrm{i}\theta}$ for some $\theta \in [0, 2\pi)$. The complex multiplication of $v \in \mathcal{V}_k$ by the number z is defined by

$$z \cdot v := |z|\mathrm{e}^{\frac{\mathrm{i}\theta}{k}} v. \tag{41}$$

Suppose, further, that V is (in general) reducible, and we have the following S^1-isotypical decomposition (cf. (13) and (14)):

$$V = V_{k_1} \oplus V_{k_2} \oplus \cdots \oplus V_{k_s}, \tag{42}$$

where V_{k_j} is modeled on the irreducible S^1-representation $\mathcal{V}_{k_j}$, $j = 1, 2, \ldots, s$. Since for every j, $\mathcal{V}_{k_j}$ can be equipped with the complex structure according to (41), every isotypical component from (42) also admits such a structure. In this way, we obtain on V a complex structure which we call the *natural complex structure*.

Let Γ be a compact Lie group. As we will see later on, the problem of studying Γ-symmetric Hopf bifurcations in many cases can be reduced to the following one.

PROBLEM 3.32. Let $G = \Gamma \times S^1$ and let V be an orthogonal G-representation with $V^{S^1} = \{0\}$ (here S^1 is identified with $\{e\} \times S^1$). Suppose V (considered as the S^1-representation) is equipped with the natural complex structure and put

$$\mathcal{O} := \left\{(\lambda, v) \in \mathbb{C} \oplus V\colon\ \|v\| < 2,\ \frac{1}{2} < |\lambda| < 4\right\}. \tag{43}$$

Assume G acts trivially on $\mathbb{C}$ and $\mathbb{R}$ and take a continuous map $a : S^1 \to GL^G(V)$, where $GL^G(V)$ stands for the set of all G-equivariant linear invertible maps in V. Define a G-equivariant map $f : \overline{\mathcal{O}} \to \mathbb{R} \oplus V$ by

$$f_a(\lambda, v) = \left(|\lambda|\big(\|v\| - 1\big) + \|v\| + 1, a\left(\frac{\lambda}{|\lambda|}\right)v\right), \quad (\lambda, v) \in \overline{\mathcal{O}}. \tag{44}$$

How can one compute the twisted degree $G\text{-}\mathrm{Deg}^t(f_a, \mathcal{O})$?

Our approach to attack the above problem contains the following four ingredients:

(i) *Recurrence Formula* (see Theorem 3.41) allowing a reduction of the general problem to the computation of the corresponding S^1-degree;

(ii) *Splitting Lemma* (cf. Lemma 3.37) allowing a reduction to subrepresentations;

(iii) *Homotopy Factorization* (cf. Corollaries 3.36 and 3.38) allowing a factorization of a given map through canonical representatives of the elements of $\pi_1(GL^G(k,\mathbb{C}))$ and next deformations to the so-called $\mathbb{C}$-complementing maps being natural "complex counterparts" for the k-th basic maps (cf. Definition 3.33);

(iv) *Suspension Procedure* allowing to reduce the computation of the S^1-degree of $\mathbb{C}$-complementting maps to the one of k-th basic maps (cf. Proposition 3.34).

The last three techniques come together at the end of this subsection (see Theorem 3.39 where the S^1-degree for (44) is given). Observe that the Splitting Lemma is presented in a form much more general than is needed to establish Theorem 3.39.

$\mathbb{C}$-complementing maps and suspension procedure We start with the following

DEFINITION 3.33. Let $b:\mathbb{R}\oplus\mathcal{V}_k\to\mathcal{V}_k$ (resp. $b^-:\mathbb{R}\oplus\mathcal{V}_k\to\mathcal{V}_k$) be the k-th basic map defined by (40) (resp. the map defined by $b^-(t,v)=(1-\|v\|-it)v$, $t\in\mathbb{R}$, $v\in\mathcal{V}_k$) and let ${}^k\Omega$ be defined by (39). Assume that $\mathcal{V}_k$ is equipped with the natural complex structure and $\mathcal{O}$ is given by (43) with $V:=\mathcal{V}_k$. Suppose, finally, that $f:\mathbb{C}\oplus\mathcal{V}_k\to\mathbb{R}\oplus\mathcal{V}_k$ (resp. $f^-:\mathbb{C}\oplus\mathcal{V}_k\to\mathbb{R}\oplus\mathcal{V}_k$) is defined by $f(\lambda,v)=(|\lambda|(\|v\|-1)+\|v\|+1,\lambda\cdot v)$ (resp. $f^-(\lambda,v)=(|\lambda|(\|v\|-1)+\|v\|+1,\bar{\lambda}\cdot v)$), where $\lambda\in\mathbb{C}$, $v\in\mathcal{V}_k$. Then, the pair $(f,\mathcal{O})$ (resp. $(f^-,\mathcal{O})$) is called a $\mathbb{C}$-*complementing pair* to $(b,{}^k\Omega)$ (resp. denoted by $(b^-,{}^k\Omega)$).

It is clear that $(f,\mathcal{O})$, $(f^-,\mathcal{O})$, $(b,{}^k\Omega)$ and $(b^-,{}^k\Omega)$ are admissible pairs. The following statement (basically, its proof is a careful usage of properties (P2)–(P4)) justifies the above definition.

PROPOSITION 3.34. *Let $(f,\mathcal{O})$ (resp. $(f^-,\mathcal{O})$) be a $\mathbb{C}$-complementing pair to $(b,{}^k\Omega)$ (resp. $(b^-,{}^k\Omega)$). Then f (resp. f^-) is S^1-homotopic (by an $\mathcal{O}$-admissible homotopy) to a map $\overline{f_1}$ (resp. $\overline{f_1^-}$), which is a suspension of b (resp. b^-) on an open subset containing zeros of $\overline{f_1}$ (resp. $\overline{f_1^-}$). In particular,*

$$S^1\text{-Deg}(f,\mathcal{O})=S^1\text{-Deg}(b,{}^k\Omega)=(\mathbb{Z}_k),\tag{45}$$

$$S^1\text{-Deg}(f^-,\mathcal{O})=S^1\text{-Deg}(b^-,{}^k\Omega)=-(\mathbb{Z}_k).\tag{46}$$

Homotopy factorization: Properties of $GL^G(V)$ Consider a continuous map $\varphi:S^1\to\mathbb{C}\setminus\{0\}$. Denote by $\tilde{\varphi}$ a continuous extension of φ to the space $\mathbb{C}$. Put $B:=\{z\in\mathbb{C}\colon |z|<1\}$. Clearly, the Brouwer degree $\deg(\tilde{\varphi},B)$ depends only on the map φ. Moreover, by classical Hopf theorem (see [44]), for two continuous maps $\varphi,\psi:S^1\to\mathbb{C}\setminus\{0\}$, φ is homotopic to ψ if and only if $\deg(\tilde{\varphi},B)=\deg(\tilde{\psi},B)$.

Denote by $GL(m,\mathbb{C})$ the group of all complex invertible $m\times m$-matrices. The following fact is well-known.

PROPOSITION 3.35. (*See, for instance, [93].*)

(i) *Two continuous maps* $\Phi, \Psi : S^1 \to GL(m, \mathbb{C})$, $m \geqslant 1$, *are homotopic if and only if the maps* $\varphi := \det_{\mathbb{C}} \circ \Phi$ *and* $\psi := \det_{\mathbb{C}} \circ \Psi$ *are homotopic, i.e.*

$$\deg(\tilde{\varphi}, B) = \deg(\tilde{\psi}, B).$$

(ii) *For every map* $\Phi : S^1 \to GL(m, \mathbb{C})$, *there exists* $l \in \mathbb{Z}$ *such that* Φ *is homotopic to* Φ_l *given by*

$$\Phi_l(\gamma) := \begin{bmatrix} \gamma^l & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad \gamma \in S^1. \tag{47}$$

In particular, for $\varphi_l := \det_{\mathbb{C}} \circ \Phi_l$, *one has* $\deg(\tilde{\varphi}_l, B) = l$.

Combining Proposition 3.35 with Proposition 2.9 yields

COROLLARY 3.36. *Let* V *be an orthogonal* G*-representation and* V_{k_o} *an isotypical component of* V *modeled on an irreducible* G*-representation* $\mathcal{V}_{k_o}$ *of complex type. Assume* $m = \dim V_{k_o} / \dim \mathcal{V}_{k_o}$. *Then*:

(i) $GL^G(V_{k_o}) \simeq GL(m, \mathbb{C})$;

(ii) *for each* $a \in \pi_1(GL^G(V_{k_o}))$, *there exists a representative* $\varphi_a : S^1 \to GL(m, \mathbb{C})$, *such that*

$$\varphi_a(\lambda) = \begin{bmatrix} \lambda^l & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad \lambda \in S^1,$$

for some $l \in \mathbb{Z}$. *In particular, we have an isomorphism* $\mu_{k_o} : \pi_1(GL^G(V_{k_o})) \to \mathbb{Z}$, *where* $\mu_{k_o}(a) = l$.

Splitting lemma Combining properties (P1)–(P4) of the twisted degree with standard techniques from complex analytic function theory, one can establish

LEMMA 3.37 (Splitting Lemma). *Let* V_1 *and* V_2 *be orthogonal* G*-representations and* $V = V_1 \oplus V_2$. *Assume that the* G*-isotypical decomposition of* V *contains only components modeled on irreducible* G*-representations of complex type. Suppose that* $a_j : S^1 \to GL^G(V_j)$, $j = 1, 2$, *are two continuous maps and* $a : S^1 \to GL^G(V)$ *is given by*

$$a(\lambda) = a_1(\lambda) \oplus a_2(\lambda), \quad \lambda \in S^1.$$

Assume that $\mathcal{O}$ *and* f_a *are defined by* (43) *and* (44), *respectively. Put*

$$\mathcal{O}_j := \left\{ (\lambda, v_j) \in \mathbb{C} \oplus V_j \colon \ \|v_j\| < 2, \ \frac{1}{2} < |\lambda| < 4 \right\},$$

$$f_{a_j}(\lambda, v_j) := \left(|\lambda|\big(\|v_j\| - 1\big) + \|v_j\| + 1, a_j\left(\frac{\lambda}{|\lambda|}\right)v_j \right),$$

where $j = 1, 2$ *and* $v_j \in V_j$. *Then,*

$$G\text{-Deg}^t(f_a, \mathcal{O}) = G\text{-Deg}^t(f_{a_1}, \mathcal{O}_1) + G\text{-Deg}^t(f_{a_2}, \mathcal{O}_2). \tag{48}$$

S^1-degree formulae Here, we combine the above results to compute the S^1-degree of (44). We start with the following

COROLLARY 3.38. *Let* $V := \mathcal{V}_k$ *be the* k*-th irreducible* S^1*-representation* $(k > 0)$ *equipped with the natural complex structure,* $l \in \mathbb{Z}$ *and*

$$\tilde{f}(\lambda, v) = \left(|\lambda|\big(\|v\| - 1\big) + \|v\| + 1, \left(\frac{\lambda}{|\lambda|}\right)^l v \right), \quad (\lambda, v) \in \overline{\mathcal{O}},$$

where $\mathcal{O}$ *is given by* (43). *Then,*

$$S^1\text{-Deg}(\tilde{f}, \mathcal{O}) = l(\mathbb{Z}_k).$$

PROOF. For the sake of definiteness, assume that $l > 0$ (the case $l \leqslant 0$ can be treated using similar arguments), and consider the map

$$\tilde{f} \times \mathrm{Id} : \overline{\mathcal{O}} \times \overline{B_{l-1}} \to \mathbb{R} \oplus \mathcal{V}_k \oplus \big[\underbrace{\mathcal{V}_k \oplus \cdots \oplus \mathcal{V}_k}_{l-1}\big],$$

where $B_{l-1} = \underbrace{B(\mathcal{V}_k) \times \cdots \times B(\mathcal{V}_k)}_{l-1}$ and $B(\mathcal{V}_k)$ denotes the unit ball in $\mathcal{V}_k$. Then, by suspension property,

$$S^1\text{-Deg}(\tilde{f}, \mathcal{O}) = S^1\text{-Deg}(\tilde{f} \times \mathrm{Id}, \mathcal{O} \times B_{l-1}).$$

Obviously, $\tilde{f} \times \mathrm{Id}$ is equivariantly homotopic, by an $\mathcal{O} \times B_{l-1}$-admissible homotopy, to f_a given by (44) where $v \in V = \underbrace{\mathcal{V}_k \oplus \cdots \oplus \mathcal{V}_k}_{l}$ and $a : S^1 \to GL^{S^1}(V)$ is defined by

$$a(\gamma) = \begin{bmatrix} \gamma^l & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad \gamma \in S^1.$$

Using Proposition 3.35, it is easy to see that f_a is equivariantly homotopic (by an $\mathcal{O} \times B_{l-1}$-admissible homotopy) to f_b given by

$$f_b(\lambda, v) = \left(|\lambda|\|\big(\|v\| - 1\big) + \|v\| + 1, b\left(\frac{\lambda}{|\lambda|}\right)v \right),$$

with $b : S^1 \to GL^{S^1}(V)$ defined by

$$b(\gamma) = \begin{bmatrix} \gamma & 0 & \dots & 0 \\ 0 & \gamma & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma \end{bmatrix}, \quad \gamma \in S^1.$$

Since $S^1\text{-Deg}(\tilde{f}, \mathcal{O}) = S^1\text{-Deg}(f_b, \mathcal{O} \times B_{l-1})$, by Splitting Lemma (cf. Lemma 3.37) and Proposition 3.34,

$$S^1\text{-Deg}(\tilde{f}, \mathcal{O}) = \underbrace{(\mathbb{Z}_k) + \cdots + (\mathbb{Z}_k)}_{l} = l(\mathbb{Z}_k).$$

□

By combining Proposition 3.35, Corollary 3.36, Splitting Lemma (cf. Lemma 3.37), and Corollary 3.38, we immediately obtain

THEOREM 3.39. *Let V be an orthogonal S^1-representation with $V^{S^1} = \{0\}$, admitting the isotypical decomposition* (42) *and equipped with the natural complex structure. Let $\mathcal{O}$* (*resp. f_a*) *be defined by* (43) (*resp.* (44)). *Then,*

$$S^1\text{-Deg}(f, \mathcal{O}) = \sum_{j=1}^{s} l_j (\mathbb{Z}_{k_j}), \tag{49}$$

where $l_j := \deg(\det_{\mathbb{C}} \circ a_j, B)$, $a_j(\lambda) := a(\lambda)|_{V_{k_j}} : V_{k_j} \to V_{k_j}$, for $j = 1, 2, \dots, s$ and $B := \{z \in \mathbb{C}\colon |z| < 1\}$.

As an immediate consequence of Theorem 3.39, one has

COROLLARY 3.40. *Let V and $\mathcal{O}$ be as in Theorem* 3.39. *Let $l_j \in \mathbb{Z}$, $j = 1, 2, \dots, s$, be given integers and assume that $\dim_{\mathbb{C}} V_{k_j} = m_j$. Define $f : \overline{\mathcal{O}} \to \mathbb{R} \oplus V$ by*

$$f(\lambda, v_1, \dots, v_n) = \big(|\lambda|\big(\|v\| - 1\big) + \|v\| + 1, \lambda^{l_1} v_1, \dots, \lambda^{l_s} v_s\big),$$

where $\lambda \in \mathbb{C} \setminus \{0\}$, $v_j \in V_{k_j}$. Then,

$$S^1\text{-Deg}(f, \mathcal{O}) = \sum_{j=1}^{n} m_j l_j (\mathbb{Z}_{k_j}). \tag{50}$$

Recurrence formula We complete this subsection with the so-called Recurrence Formula for the twisted G-equivariant degree, which, to some extent, can be counted as the Borsuk–Ulam theorem in the case of one free parameter. This formula allows a reduction of computations of G-equivariant twisted degree to the S^1-degree.

Let G, V, Ω and f be as in Proposition 3.27. Take $(H) \in \Phi_1^t(G)$ and put $f^H := f_{|\mathbb{R}\oplus V^H}$. Clearly (cf. Proposition 3.21(i)), f^H is S^1-equivariant and Ω^H-admissible, thus

$$S^1\text{-Deg}(f^H, \Omega^H) = n_{k_1}(\mathbb{Z}_{k_1}) + \cdots + n_{k_r}(\mathbb{Z}_{k_r}) \tag{51}$$

is correctly defined (cf. (38)). Put

$$\deg_k(f^H, \Omega^H) := n_k. \tag{52}$$

By combining Propositions 3.14, 3.15 and Lemma 3.17, one can prove

THEOREM 3.41 (Recurrence Formula). *Let* $G, V, \Omega \subset \mathbb{R} \oplus V$ *and* f *be as in Proposition* 3.27. *Assume, further, that*

$$G\text{-Deg}^t(f, \Omega) = \sum_{(H)\in\Phi_1^t(G,\Omega)} n_H \cdot (H) \tag{53}$$

is defined according to (35)–(37). *Then,*

$$n_H = \Big[\sum_k \deg_k(f^H, \Omega^H) - \sum_{(K)>(H)} n_K n(H,K)\big|W(K)/S^1\big|\Big] \Big/ \big|W(H)/S^1\big| \tag{54}$$

(*see* (52) (*resp.* (28)) *for the definition of* $\deg_k(f^H, \Omega^H)$ (*resp.* $n(H,K)$).

3.8. *General concept of basic maps*

The case of one free parameter The concept of the *basic maps*, which was introduced in Subsection 3.6 for S^1-equivariant maps (see (39) and (40)), can be generalized to G-equivariant maps, where $G = \Gamma \times S^1$. Following the same lines as in Subsection 3.6, we also define the counter part of a basic map, called a $\mathbb{C}$-*complementing* map.

DEFINITION 3.42.

(a) Under the notations explained in Remark and Notation 2.6, let $\mathcal{O} \subset \mathbb{R} \oplus \mathcal{V}_{j,l}$ be defined by

$$\mathcal{O} = \left\{(t, v)\colon \frac{1}{2} < \|v\| < 2,\ -1 < t < 1\right\}, \tag{55}$$

and $b : \overline{\mathcal{O}} \to \mathcal{V}_{j,l}$ be defined by

$$b(t, v) = \big(1 - \|v\| + \mathrm{i}t\big) \cdot v, \quad (t, v) \in \overline{\mathcal{O}}. \tag{56}$$

Then, the map b is called a *basic map* on $\mathcal{O}$, and the pair $(b, \mathcal{O})$ is called a *basic pair* for the irreducible G-representation $\mathcal{V}_{j,l}$.

(b) Let

$$\Omega = \left\{(\lambda, v) \in \mathbb{C} \oplus \mathcal{V}_{j,l}\colon\ \|v\| < 1,\ \frac{1}{2} < |\lambda| < 4\right\} \tag{57}$$

and $\mathfrak{f}\colon \overline{\Omega} \to \mathbb{R} \oplus \mathcal{V}_{j,l}$ be defined as

$$\mathfrak{f}(\lambda, v) = \bigl(|\lambda|\bigl(\|v\| - 1\bigr) + \|v\| + 1, \lambda \cdot v\bigr), \quad (\lambda, v) \in \overline{\Omega}, \tag{58}$$

where $\lambda \cdot v$ denotes the usual complex multiplication of v by λ. Then, the map $\mathfrak{f}$ is called a $\mathbb{C}$-*complementing map* on Ω, and the pair $(\mathfrak{f}, \Omega)$ is called a $\mathbb{C}$-*complementing pair* for the irreducible G-representation $\mathcal{V}_{j,l}$.

Obviously, both (56) and (58) are G-equivariant. Moreover, b is $\mathcal{O}$-admissible and $\mathfrak{f}$ is Ω-admissible. Therefore, $G\text{-Deg}^t(b, \mathcal{O})$ and $G\text{-Deg}^t(\mathfrak{f}, \Omega)$ are correctly defined. Although b and $\mathfrak{f}$ are rather different in nature (b seems to be the simplest one having a non-trivial twisted equivariant degree, while $\mathfrak{f}$ is motivated by the applications to the bifurcation theory, where it appears naturally), we have the following statement parallel to Proposition 3.34.

PROPOSITION 3.43. *Let $\mathcal{U}_j$ be a complex irreducible Γ-representation and $\mathcal{V}_{j,l} := {}^l\mathcal{U}_j$ the associated irreducible representation of $G = \Gamma \times S^1$ (cf. Remark and Notation* 2.6). *Then,*

$$G\text{-Deg}^t(b, \mathcal{O}) = G\text{-Deg}^t(\mathfrak{f}, \Omega), \tag{59}$$

where $\mathcal{O}$, b, Ω and $\mathfrak{f}$ are given by (55)–(57) *and* (58), *respectively.*

For the sequel, it is convenient to introduce

DEFINITION 3.44. Let $(b, \mathcal{O})$ be a basic pair for the irreducible G-representation $\mathcal{V}_{j,l}$. Put

$$\deg_{\mathcal{V}_{j,l}} := G\text{-Deg}^t(b, \mathcal{O}), \tag{60}$$

and call $\deg_{\mathcal{V}_{j,l}}$ the *basic degree* for $\mathcal{V}_{j,l}$.

Using the same ideas as in the proof of Corollary 3.38 in compliance with Corollary 3.36, one can easily establish a formula allowing in certain important cases a reduction to basic degrees.

PROPOSITION 3.45. *Under the notations and assumptions of Problem* 3.32 *suppose, in addition, that*:

(i) *V has a single G-isotypical component modeled on an irreducible G-representation $\mathcal{V}_{j,l}$ and $m = \dim V / \dim \mathcal{V}_{j,l}$;*

(ii) *the map a from formula* (44)*, identified with a map $S^1 \to GL(m, \mathbb{C})$* (*cf. Corollary* 3.36(ii)) *admits a representation* $\operatorname{diag}\{(\lambda^k, 1, 1, \dots, 1\}$.

Then,

$$G\text{-Deg}^t(f_a, \mathcal{O}) = k \cdot \deg_{\mathcal{V}_{j,l}}. \tag{61}$$

Observe that basic maps play an essential role in practical computations of the twisted degree. A typical degree-theoretical approach to a concrete problem is usually devised on the so-called *linearization techniques*. These techniques are based on local or global linear approximations, which can be considered as a way of transforming the original map to the one for which the degree can be easily evaluated. In the case of the equivariant degree with one free parameter, the linearization techniques are also used. However, due to the nature of the equivariant degree, it is inappropriate to expect the same type of computations for all kinds of groups. That means, in order to establish a standard method for the computations of the (twisted) equivariant degree, one needs to identify the key elements in this process. It turns out that the *basic degrees* are essential for the computations of the twisted degrees of arbitrary maps. By providing for a considered group $\Gamma \times S^1$ the tables of basic degrees, one is able to remove essential technical obstructions from the computation process of the twisted degree.

It should be pointed out that the computational formulae presented in Subsection 3.7 are enough to completely evaluate $\deg_{\mathcal{V}_{j,l}}$. Indeed, consider the representation $\mathcal{V}_{j,1}$ (the case of arbitrary $l > 1$ can be analyzed similarly). The Recurrence Formula (see Theorem 3.41) reduces the computations of $\deg_{\mathcal{V}_{j,1}}$ to $\deg_1(f^H, \Omega^H)$ and numbers $n(H, K)$, where $(H), (K) \in \Phi_1^t(G)$. Next, by Corollary 3.40, one obtains

$$\deg_1(f^H, \Omega^H) = \frac{1}{2} \dim \mathcal{V}_{j,1}^H. \tag{62}$$

Thus, to compute $\deg_{\mathcal{V}_{j,1}}$ one needs to know: (i) twisted orbit type lattice in $\mathcal{V}_{j,1}$, (ii) dimensions of H-fixed-point subspaces in $\mathcal{V}_{j,1}$ ($(H) \in \Phi_1^t(G)$), and (iii) numbers $n(H, K)$. Therefore, the computations of basic degrees can be completely computerized.

The case without free parameter Similarly to the situation of one free parameter, one could look for an appropriate concept of *basic maps* in the case without free parameters (cf. Remark 3.28). However, in the latter case, this concept turns out to be almost trivial. To be more specific, consider an irreducible orthogonal representation $\mathcal{V}_i$ of the compact Lie group Γ, and let $\mathcal{D}$ be the unit ball in $\mathcal{V}_i$. It is clear that every linear Γ-equivariant isomorphism A from $\mathcal{V}_i$ onto $\mathcal{V}_i$ is a $\mathcal{D}$-admissible Γ-equivariant map. Therefore, we can consider this type of Γ-equivariant maps as the most elementary, for which the equivariant degree $\Gamma\text{-Deg}(A, \mathcal{D})$ is defined. On the other hand, the representation $\mathcal{V}_i$ can be of one of three types: real, complex or quaternionic (cf. Proposition 2.9(ii)). It is clear that in the case $\mathcal{V}_i$ is of complex or quaternionic type, the operator A can be connected by a path in $GL^\Gamma(\mathcal{V}_i) := \{B \in L^G(\mathcal{V}_i)\colon\ B \text{ is an isomorphism}\}$ with the identity operator Id. Thus,

Γ-Deg$(A,\mathcal{D})=(\Gamma)$, which can be considered as a trivial case. In the case the representation $\mathcal{V}_i$ is of real type, the group $GL^{\Gamma}(\mathcal{V}_i)$ has two connected components, so if A does not belong to the same connected component as Id, it is homotopic to $-\,$Id, and consequently,

$$\Gamma\text{-Deg}(A,\mathcal{D})=\Gamma\text{-Deg}(-\operatorname{Id},\mathcal{D}).$$

Therefore, we call the map $-\operatorname{Id}:\mathcal{V}_i\to\mathcal{V}_i$ a *basic map* for the irreducible representation $\mathcal{V}_i$ (see [93] for more information about this type of basic maps) and put

$$\deg_{\mathcal{V}_i}:=\Gamma\text{-Deg}(-\operatorname{Id},\mathcal{D}). \tag{63}$$

It turns out that $\deg_{\mathcal{V}_i}$ can be completely evaluated using a recurrence formula similar to (53)–(54). More precisely, consider the set $\Phi_0(\Gamma,\mathcal{V}_i)$ of all the orbit types (L_k) in $\mathcal{V}_i$ such that $(L_k)\in\Phi_0(\Gamma)$. Then (see [93]),

$$\deg_{\mathcal{V}_i}=\sum_k n_{L_k}\cdot(L_k), \tag{64}$$

where

$$n_{L_{k_o}}=\frac{(-1)^{n_{k_o}}-\sum_{k=1}^{k_o-1}n(L_{k_o},L_k)\cdot n_{L_k}\cdot|W(L_k)|}{|W(L_{k_o})|} \tag{65}$$

and $n_k=\dim\mathcal{V}_i^{L_k}$.

3.9. *Multiplicativity property*

In this subsection, we connect the multiplication structures on $A(\Gamma)$ (resp. $A_1^t(G)$) to the Γ-equivariant degree (resp. twisted degree) of Γ-equivariant maps without free parameter (resp. G-equivariant maps with one free parameter). The well-known multiplicativity property of the classical Brouwer degree, which asserts that the degree of a direct product of maps is equal to the product of their degrees (in the ring $\mathbb{Z}$), can be thought of as the prototype of the theorems following below (their proofs use a regular normality argument combined with the definitions of the multiplications in $A(\Gamma)$ (resp. $A_1^t(G)$) and a careful count of orbits of zeros in the product maps).

THEOREM 3.46. *Let Γ be a compact Lie group, U and W two orthogonal Γ-representations, $\Omega_1\subset U$, $\Omega_2\subset W$ two open bounded Γ-invariant sets. Assume that $s:V\to V$ (resp. $g:W\to W$) is an Ω_1-admissible (resp. Ω_2-admissible) equivariant map. Then,*

$$\Gamma\text{-Deg}(s\times g,\Omega_1\times\Omega_2)=\Gamma\text{-Deg}(s,\Omega_1)\cdot\Gamma\text{-Deg}(g,\Omega_2), \tag{66}$$

where ‘$\cdot$’ stands for the multiplication in the Burnside ring $A(\Gamma)$ (see also Remark 3.28).

REMARK 3.47. Clearly, under the notations of Theorem 3.46, one immediately obtains the multiplicativity property for the usual Brouwer degree by putting $\Gamma = \{e\}$.

THEOREM 3.48 (Multiplicativity property). *Let* Γ *be a compact Lie group,* $G = \Gamma \times S^1$, V (*resp.* W) *an orthogonal* G*-representation* (*resp.* Γ*-representation*), $\Omega_1 \subset \mathbb{R} \oplus V$ (*resp.* $\Omega_2 \subset W$) *an invariant open bounded set and* $f_1 : \mathbb{R} \oplus V \to V$ (*resp.* $f_2 : W \to W$) *an* Ω_1*-admissible* (*resp.* Ω_2*-admissible*) *equivariant map. Then,*

$$G\text{-Deg}^t(f_1 \times f_2, \Omega_1 \times \Omega_2) = \Gamma\text{-Deg}(f_2, \Omega_2) \cdot G\text{-Deg}^t(f_1, \Omega_1),$$

where the multiplication is taken in the $A(\Gamma)$*-module* $A_1^t(G)$ (*see also Remark* 3.28).

REMARK 3.49. Obviously, in the case $\Gamma = \{e\}$, one has to take for Γ-Deg(f_2, Ω_2) the usual Brouwer degree and for "$\cdot$" the usual multiplication by an integer.

3.10. *Infinite dimensional twisted degree*

In order to apply the twisted degree theory developed above to symmetric Hopf bifurcation problems, one should extend it to reasonable classes of equivariant vector fields on *infinite dimensional* Banach G-representations. To this end, it is necessary to adopt the standard infinite dimensional techniques in the equivariant setting. Below, we briefly discuss the corresponding framework for two classes of vector fields relevant to the applications we are dealing with in this paper: (i) compact fields and (ii) condensing fields. Observe, however, that extensions to many other important classes can be easily done as well.

Throughout this subsection, W is a real isometric Banach G-representation ($G = \Gamma \times S^1$), $E := \mathbb{R} \oplus W$ is equipped with the norm $\|(\lambda, v)\| = \max\{|\lambda|, \|v\|\}$ (G acts trivially on $\mathbb{R}$) and $\pi : E \to W$ stands for the natural projection on W. Given a G-invariant subset $M \subset E$ and a G-equivariant map $F : M \to W$, define $f : M \to W$ by

$$f := \pi - F, \tag{67}$$

and call f a *G-equivariant field* on M.

Leray–Schauder twisted degree Recall the following

DEFINITION 3.50. Let $X \subset E$ be a bounded subset. Then, a continuous map $F : X \to W$ is called *compact*, if $\overline{F(X)}$ is compact in W.

THEOREM 3.51 (Equivariant Schauder approximation theorem). *Let* $X \subset E$ *be a* G*-invariant bounded subset and* $F : X \to W$ *a* G*-equivariant compact map. Then, for every* $\varepsilon > 0$, *there exists a* G*-equivariant finite-dimensional map* $F_\varepsilon : X \to W$ (*i.e. the image* $F_\varepsilon(X)$ *is contained in a finite-dimensional subrepresentation of* W) *such that*

$$\|F_\varepsilon(x) - F(x)\| < \varepsilon \quad \text{for all } x \in X. \tag{68}$$

PROOF. Take W_∞ defined by (17)–(19). By (20), there exists a finite set $N = \{v_1, \ldots, v_n\} \subset W_\infty$ such that $F(X) \subset N_\varepsilon := N + B_\varepsilon(0)$, where $B_\varepsilon(0) = \{v \in W\colon \|v\| < \varepsilon\}$. Define the functions $\nu_k : N_\varepsilon \to \mathbb{R}$, $k = 1, \ldots, n$, by $\nu_k(v) := \max\{0, \varepsilon - \|v - v_k\|\colon k = 1, 2, \ldots, n\}$, $v \in N_\varepsilon$, and put $P_\varepsilon(v) := \frac{1}{\sum_{k=1}^n \nu_k(v)} \sum_{k=1}^n \nu_k(v) v_k \in \operatorname{conv}\{v_1, \ldots, v_n\}$, $v \in N_\varepsilon$, where "conv" stands for the convex hull.

Next, define the map $\widetilde{F}_\varepsilon : X \to W$ by $\widetilde{F}_\varepsilon(x) := P_\varepsilon(F(x))$ for $x \in X$. Clearly, $\widetilde{F}_\varepsilon$ satisfies (68) and $\widetilde{F}_\varepsilon(X) \subset \operatorname{span}\{G(v_1), \ldots, G(v_n)\} =: A$. By Proposition 2.12, A is G-invariant and finite-dimensional.

To complete the proof, put $F_\varepsilon(x) := \int_G g \widetilde{F}_\varepsilon(g^{-1}x)\, \mathrm{d}\mu(g)$, $x \in X$ (here μ denotes the Haar measure on G). ☐

The construction of the Leray–Schauder equivariant degree is standard.

DEFINITION 3.52. Assume that $\Omega \subset E$ is an open bounded G-invariant subset, $F : \overline{\Omega} \to W$ is a compact equivariant map.

(a) A G-equivariant field f defined by (67) is called *compact*. Such a field f is called Ω-*admissible* if $f(x) \neq 0$ for all $x \in \partial\Omega$.

(b) Two G-equivariant Ω-admissible compact fields $f_0 = \pi - F_0$ and $f_1 = \pi - F_1$ are said to be *G-equivariantly homotopic*, if there exists a G-equivariant compact map $H : [0, 1] \times \overline{\Omega} \to W$ with $h := \pi - H$ satisfying the following conditions:
  (i) $h(0, \cdot) = f_0$, $h(1, \cdot) = f_1$;
  (ii) for each $\lambda \in [0, 1]$, $h(\lambda, \cdot)$ is Ω-admissible.

  The map h will be referred to as a *G-equivariant homotopy of compact fields* joining f_0 and f_1.

In order to define the twisted degree $G\text{-}\mathrm{Deg}^t(f, \Omega)$ of an Ω-admissible G-equivariant compact field f, take an equivariant finite-dimensional map $F_\varepsilon : \overline{\Omega} \to W$ such that $F_\varepsilon(\overline{\Omega}) \subset W_*$, $\dim W_* < \infty$ and

$$\big\| F(x) - F_\varepsilon(x) \big\| < \inf\big\{ \big\| f(y) \big\|\colon y \in \partial\Omega \big\},$$

and next, define the ε-approximation f_ε of f as $f_\varepsilon(\lambda, v) = v - F_\varepsilon(\lambda, v)$, $(\lambda, v) \in \overline{\Omega}$ (cf. Theorem 3.51). Then, put

$$G\text{-}\mathrm{Deg}^t(f, \Omega) \overset{\text{def}}{=} G\text{-}\mathrm{Deg}^t(f_\varepsilon|_{\overline{\Omega}_*}, \Omega_*) \in A_1^t(G), \tag{69}$$

where $\Omega_* := \Omega \cap (\mathbb{R} \oplus W_*)$, and call $G\text{-}\mathrm{Deg}^t(f, \Omega)$ the *Leray–Schauder twisted degree*.

By applying the standard arguments, one can verify that the definition, given by formula (69), neither depends on a choice of the equivariant approximation f_ε, nor on an invariant finite-dimensional subspace W_* (see [93] for more details).

Using Definition 3.52 and following the standard lines (cf. [81] and [93]), one can formulate and prove an analogue of Proposition 3.27 for compact fields and homotopies.

Nussbaum–Sadovskii twisted degree Denote by $\mathcal{M}$ the class of all bounded subsets of E.

DEFINITION 3.53. A function $\mu : \mathcal{M} \to [0, \infty)$ is called a *measure of noncompactness* on E if for all $A, B \in \mathcal{M}$, the following conditions are satisfied:

(a) $\mu(A) = 0 \Leftrightarrow \bar{A}$ is compact;
(b) $\mu(A) = \mu(\bar{A})$;
(c) $\mu(\mathrm{conv}(A)) = \mu(A)$;
(d) $\mu(A \cup B) = \max\{\mu(A), \mu(B)\}$;
(e) $\mu(rA) = |r| \cdot \mu(A)$, $r \in \mathbb{R}$;
(f) $\mu(A + B) \leqslant \mu(A) + \mu(B)$.

We refer to [1] and [93], where many important examples of measures of noncompactness are considered as well as their properties are studied.

DEFINITION 3.54. Let μ be a measure of noncompactness defined on $\mathcal{M}$. Take $X \subset E$ and $Y \subset W$. Consider a continuous map $F : X \to Y$ taking bounded subsets of X to bounded subsets of Y. We say that F is:

(i) a *μ-Lipschitzian map* with a constant $L \geqslant 0$, if $\mu(F(A)) \leqslant L\mu(A)$ for all bounded subsets A of X;
(ii) a *completely continuous map*, if it is μ-Lipschitzian with a constant $L = 0$;
(iii) a *condensing map*, if it is μ-Lipschitzian with a constant $L = 1$ and $\mu(F(A)) < \mu(A)$ for every bounded subset $A \subset X$ such that $\mu(A) > 0$.

REMARK 3.55.

(i) Clearly, any compact map is also completely continuous. The opposite statement is true under the assumption that the domain of the map is bounded. Also, any completely continuous map is, of course, condensing with respect to any measure of noncompactness.
(ii) Obviously, if $F : E \to W$ is a contraction map (i.e. there exists $0 \leqslant q < 1$ such that $\|F(x) - F(y)\| \leqslant q\|x - y\|$ *for all* $x, y \in E$), then F is condensing with respect to any measure of noncompactness on $\mathcal{M}$ (the applications to neutral FDEs we will be dealing with appeal exactly to this situation).

The idea of extending the equivariant Leray–Schauder degree to the class of equivariant condensing fields, is based on the same standard construction that is used in order to make such an extension in the non-equivariant case (see for example [1] and [93], Chapter 4). For the sake of completeness of our presentation, we discuss shortly some of the main steps of this construction.

DEFINITION 3.56. Let $\Omega \subset E$ be an open bounded G-invariant subset and $F : \overline{\Omega} \to W$ a condensing equivariant map with respect to some (fixed) measure of noncompactness μ on E.

(a) The map f defined by (67) is called a *condensing field*.
(b) Two G-equivariant Ω-admissible condensing fields $f_0 = \pi - F_0$ and $f_1 = \pi - F_1$ are called *G-equivariantly homotopic*, if there exists a G-equivariant condensing map $H : [0, 1] \times \overline{\Omega} \to W$ with $h := \pi - H$ satisfying the following conditions:

(i) $h(0,\cdot)=f_0$, $h(1,\cdot)=f_1$;
(ii) for each $\lambda\in[0,1]$, $h(\lambda,\cdot)$ is Ω-admissible.

The map h will be referred to as a *G-equivariant homotopy of condensing fields* joining f_0 and f_1. The map H is called a *condensing G-equivariant homotopy* between F_0 and F_1.

The following notion (see [1] for the non-equivariant setting) plays the central role in our considerations.

DEFINITION 3.57. Let $M\subset E$ be an invariant subset and $F:M\to W$ an equivariant map. A subset $Q\subset W$ is called *G-fundamental* for F, if it satisfies the following conditions:
(a) Q is non-empty, compact, convex and G-invariant;
(b) $F(M\cap Q)\subset Q$;
(c) if $x_o\in M\setminus Q$, then $x_o\notin \operatorname{conv}(F(x_o)\cup Q)$.
Observe that a G-fundamental set Q for F contains all the fixed points of F.

Similarly, one can define a notion of a G-fundamental set for a G-equivariant deformation $H:[0,1]\times M\to W$, namely, this set should be G-fundamental for each $H(\lambda,\cdot)$.

REMARK 3.58. Obviously, if $X\subset W$ is a (non-empty) compact G-invariant subset, then so is $\overline{\operatorname{conv}(X)}$. Therefore, for any G-equivariant compact map $F:M\to W$, where $M\subset E$, the set $\overline{\operatorname{conv}(F(M))}$ is G-fundamental.

It turns out (cf. [1])

LEMMA 3.59. *Let $\Omega\subset E$ be an open bounded G-invariant subset, $H:[0,1]\times\overline{\Omega}\to W$ a G-equivariant Ω-admissible condensing homotopy, $K\subset W$ an arbitrary G-invariant compact subset. Then, there exists a G-fundamental set for H containing K.*

Assume that $\Omega\subset E$ is an open bounded G-invariant subset, $\pi-F:\overline{\Omega}\to W$ a G-equivariant condensing Ω-admissible field. Let $Q\subset W$ be a G-fundamental set for F provided by Lemma 3.59. By the equivariant Dugundji theorem (see Proposition 3.9), there exists a G-equivariant extension $\widetilde{F}:E\to Q$ of the map $F|_{\overline{\Omega}\cap Q}$. Put $\overline{F}:=\widetilde{F}|_{\overline{\Omega}}$. Take the compact G-equivariant field $\pi-\overline{F}$ and put

$$G\text{-}\mathrm{Deg}^t(\pi-F,\Omega):=G\text{-}\mathrm{Deg}^t(\pi-\overline{F},\Omega). \tag{70}$$

We call $G\text{-}\mathrm{Deg}^t(\pi-F,\Omega)$ the *Nussbaum–Sadovskii equivariant degree*. Using Lemma 3.59, it is easy to show that the definition given by formula (70), neither depends on a choice of a G-fundamental set, nor on an equivariant extension $\overline{F}$ (see [1,93] for more details).

Using Definition 3.56 and following the standard lines (cf. [93] and [1]), one can formulate and prove an analogue of Theorem 3.27 for condensing maps.

3.11. *Bibliographical remarks*

Equivariant extensions via fundamental domains is one of the crucial ideas underlying our approach to the equivariant degree. In the context relevant to our discussion, a simplicial fundamental domain for an action of a cyclic group of prime order was used (probably, for the first time) in [47]. The case of an arbitrary cyclic group action on a sphere was considered by M. Krasnosel'skii [88] (see also the paper of P. Zabreiko [149], where a slightly more general case was considered). The case of an arbitrary finite group action on a subset of $\mathbb{R}^n$ was studied by Z. Balanov and S. Brodsky in [9]. In full generality, Theorems 3.2 and 3.4 were proved by Z. Balanov and A. Kushkuley in [20,21,98,99]. The concept of regular fundamental domain (as well as the proof of Theorem 3.7) was suggested in [14] (see also [19]). The detailed study of the case of linear actions of Abelian groups can be found in [80,81]. For the equivariant extension problem in a more general setting, we refer to [3,5,7].

The concept of normal/regular normal approximations (being, together with fundamental domains, another important ingredient of our approach to studying (twisted) equivariant degree) was discussed, on the one hand, by K. Gęba, W. Krawcewicz and J. Wu in [60], and on the other hand, by A. Kushkuley and Z. Balanov in [99] (cf. [96]). For other "equivariant general position" results (including those for different actions on the domain and image) closed in spirit to regular normal approximations, we refer to [99] (cf. [115,150]) and [81].

The numbers $n(L, H)$ were defined by E. Ihrig and M. Golubitsky in [74] (see also [99]). Proposition 3.15 and Lemma 3.17 were suggested in [14]. For a detailed discussion on the concept of Burnside ring (as well as for computations of several multiplication tables), we refer to [43] and [93]. Twisted subgroups of $\Gamma \times S^1$ for certain groups Γ of practical meaning, were studied in [10,11,15,17,18,14,67,65,66,91,92,145] (some additional information (including the numbers $n(L, H)$), can be found in [27,29,32,67,99]). The notion of bi-orientability was introduced by G. Peschke in [126] (see also [60]). Some important multiplication tables for $A_1^t(G)$ can be found in [10,18,92]. In fact, the $A(\Gamma)$-module structure on $A_1^t(\Gamma \times S^1)$ coincides with the appropriate restrictions of the multiplication in the Euler ring $U(\Gamma \times S^1)$ (cf. [43,59,134]), where $A(\Gamma)$ is identified with $\mathbb{Z}[\Phi(\Gamma \times S^1)] \subset U(\Gamma \times S^1)$. In the case Γ is a finite group, we have $U(\Gamma \times S^1) \simeq A(\Gamma) \oplus A_1^t(\Gamma \times S^1)$ (cf. [43]).

The concept of twisted degree was developed in [11,14,19]. Among the important "predecessors" of (twisted) S^1-equivariant degree, one should mention the rational-valued homotopy invariants introduced and studied in [57,39,40,42,41]. Actually, the approach suggested in Subsection 3.4 to construct the twisted degree, can be easily extended to a more general setting: G is an arbitrary compact Lie group (not necessarily of the form $G = \Gamma \times S^1$) and $f : \mathbb{R}^n \oplus V \to V$ is an equivariant map with $n \geqslant 1$ free parameters. In this case, the corresponding invariant (known as the *primary equivariant degree*) takes its values in the $\mathbb{Z}$-module $\mathbb{Z}[\Phi_n^+(G)]$, where $\Phi_n^+(G)$ consists of all "primary" conjugacy classes (H), i.e. $\dim W(H) = n$ and $W(H)$ is bi-orientable. It was considered by K. Gęba, W. Krawcewicz and J. Wu in [60] (following the differential topology framework), G. Peschke in [126] (using the algebraic topology ideas) and W. Krawcewicz and H. Xia in [96] (using an analytical approach). The axiomatic approach to the primary degree based on the use of (regular) fundamental domains was suggested by Z. Balanov,

W. Krawcewicz and H. Ruan in [14] (see also [19]). In turn, the primary equivariant degree is a part of the construction of equivariant degree defined by J. Ize, I. Massabó and A. Vignoli in [78] which is an equivariant homotopy class of a certain auxiliary equivariant map of G-representation spheres (a detailed exposition of this concept for *Abelian* group representations (including many non-trivial and elegant examples) can be found in [81]). In particular, this degree (as a complete equivariant homotopy invariant) takes into account the so-called “secondary” orbit types, i.e. those with $\dim W(H) < n$. Observe, however, that (i) the primary degree was defined in [60] (see also [45,46] for $G = S^1$) independently of the work [78] , and (ii) the general construction of the equivariant degree given in [78] (see also [81]), being of great theoretical importance, does not give practical hints for its computation.

The importance of S^1-basic and $\mathbb{C}$-complementing maps was indicated in [14] (see also [49,93,18]). The Splitting Lemma was proved in [14] based on ideas from [91]. The computational formulae for the S^1-degree were established in [14] (see also [46] and [93]). The Recurrence Formula (in a slightly different version) was suggested in [91] (see also [14] and [19]).

Twisted degrees of basic maps (being another important ingredient of our approach), for various irreducible representations, were considered in [10,11,18,14,91,92], where the importance of the numbers $n(L, H)$ for the computations of the basic degrees was first indicated. The multiplicativity property of the equivariant degree, which is crucial from the applications viewpoint, was studied in [17,18,91] (see also [81] for the Abelian group actions). The case without free parameter was considered in [93,141]. Finally, a systematic exposition of the equivariant degree theory for many important classes of non-Abelian group actions can be found in [19].

4. Hopf bifurcation problem for ODEs without symmetries

In this section, we start the (twisted) equivariant degree treatment of the Hopf bifurcation phenomena in (symmetric) dynamical systems. To clarify the main ingredients of our approach, we begin with the simplest situation – a parameterized by $\alpha \in \mathbb{R}$ system of ODEs $\dot{x} = f(\alpha, x)$, $x \in \mathbb{R}^N$, *without* spatial symmetries. It turns out that it is possible:

(i) to reformulate the original problem as the S^1-fixed point problem for a compact map $\mathcal{F} : \Omega \subset \mathbb{R} \oplus W \to W$, where W is an appropriate space of periodic $\mathbb{R}^N$-valued functions with the S^1-action induced by the shift of time variable;

(ii) to associate to the compact field $\mathfrak{F} := \pi - \mathcal{F}$ the S^1-degree whose nontriviality indicates the occurrence of a Hopf bifurcation (here, $\pi : \mathbb{R} \oplus W \to W$ is the natural projection).

In the next section, assuming the system to satisfy some Γ-symmetry conditions, we associate to $\mathfrak{F}$ the $\Gamma \times S^1$-equivariant twisted degree allowing to classify the bifurcating solutions according to their symmetries. The main advantage of this method rests on the fact that it can be equivalently applied (with certain cosmetic modifications) to other classes of equations (for instance, functional differential equations, neutral functional differential equations or partial functional differential equations).

4.1. *Statement of the problem*

Put $V := \mathbb{R}^N$. We begin with the following system of ODEs:

$$\begin{cases} \dot{x} = f(\alpha, x), & x \in V,\ \alpha \in \mathbb{R}, \\ x(0) = x(p), & \text{for a certain } p > 0, \end{cases} \tag{71}$$

where the function $f : \mathbb{R} \times V \to V$ is of class C^1 and satisfies the property

$$f(\alpha, 0) = 0 \quad \text{for all } \alpha \in \mathbb{R}. \tag{72}$$

The stationary point, at which we are analyzing the occurrence of small amplitude periodic solutions for the system (71), is (by the assumption (72)) the origin $0 \in V$. We say that for $\alpha = \alpha_o$, the system (71) has a *Hopf bifurcation* occurring at $(\alpha_o, 0)$ corresponding to the "limit period" $\frac{2\pi}{\beta_o}$, if there exists a family of p_s-periodic non-constant solutions $\{(\alpha_s, x_s(t))\}_{s\in\Lambda}$ (for a proper index set Λ) of (71) satisfying the conditions:

(1) The set $K := \bigcup_{s\in\Lambda}\{(\alpha_s, x_s(t)) \colon t \in \mathbb{R}\}$ contains a compact connected (infinite) set C such that $(\alpha_o, 0) \in C$,

(2) $\forall \varepsilon > 0$, $\exists\, \delta > 0$ such that

$$\forall \big(\alpha_s, x_s(t)\big) \in C \quad \sup_t \big\| x_s(t) \big\| < \delta \Rightarrow \|\alpha_o - \alpha_s\| < \varepsilon \text{ and } \left\| p_s - \frac{2\pi}{\beta_o} \right\| < \varepsilon.$$

With the system (71), associate the linearized at $x = x_o$ system

$$\dot{x} = D_x f(\alpha, x_o)x, \quad x \in V. \tag{73}$$

Taking the complexification of (73) and substituting the function $x(t) = \mathrm{e}^{\lambda t} \cdot v$ with $\lambda \in \mathbb{C}$ and $v \in V$, leads to the following *characteristic operator*

$$\Delta_{(\alpha, x_o)}(\lambda) := \lambda \,\mathrm{Id} - D_x f(\alpha, x_o) : V^c \to V^c, \tag{74}$$

where $V^c = \mathbb{C}^N$ denotes the complexification of V, $D_x f(\alpha, x_o)(z \otimes x) = z \otimes D_x f(\alpha, x_o)x$ for $z \otimes x \in V^c$, and the *characteristic equation*

$$\det\nolimits_{\mathbb{C}}\big[\Delta_{(\alpha, x_o)}(\lambda)\big] = 0, \quad \lambda \in \mathbb{C}. \tag{75}$$

A solution to (75) is called a *characteristic root* at the stationary point (α, x_o). It is obvious that the Hopf bifurcation occurs for the linearized system (73) with $\alpha = \alpha_o$ if and only if the characteristic equation (75) (for $\alpha = \alpha_o$) has a (non-zero) purely imaginary root $\lambda = \mathrm{i}\beta_o$. It will turn out that the appearance of a purely imaginary *characteristic root* can provide appropriate information to predict the occurrence of Hopf bifurcation for the non-linear system (71). In what follows, put $\Delta_\alpha(\cdot) := \Delta_{(\alpha,0)}(\cdot)$.

4.2. S^1-equivariant reformulation of the problem

Normalization of the period Introduce the unknown period p as an additional parameter to the considered system (71) of ODEs. For this purpose, substitute $u(t) := x(\frac{p}{2\pi}t)$, which leads to

$$\begin{cases} \dot{u} = \frac{p}{2\pi} f(\alpha, u), & u \in V, \\ u(0) = u(2\pi), \end{cases} \tag{76}$$

where the unknown function $u(t)$ is considered to be of the fixed 2π-period. Next, reparametrize the system (76) by introducing $\beta := \frac{2\pi}{p}$, i.e. we obtain

$$\begin{cases} \dot{u} = \frac{1}{\beta} f(\alpha, u), & u \in V, \\ u(0) = u(2\pi). \end{cases} \tag{77}$$

Setting in functional spaces and S^1-isotypical decomposition The system (77) can be reformulated as an equation in the appropriate functional spaces. Denote by $W := H^1(S^1; V)$ the first Sobolev space of V-valued functions defined on S^1, which can be identified with the corresponding space of 2π-periodic V-valued functions via the identification $S^1 \simeq \mathbb{R}/2\pi\mathbb{Z}$. Formula

$$(\zeta u)(t) := u(t + \tau) \quad (t \in \mathbb{R},\ u \in W,\ \zeta = e^{i\tau} \in S^1) \tag{78}$$

determines on W the structure of isometric Hilbert S^1-representation with respect to the standard inner product in W given by

$$\langle u, v \rangle_{H_1} := \int_0^{2\pi} \big(u(t) \bullet v(t) + u'(t) \bullet v'(t)\big)\, dt.$$

Identify the subspace of W consisting of all constant functions with V (it can be thought of as the "0-mode" of the isotypical decomposition of W). To describe other S^1-isotypical components W_l, $l = 1, 2, 3, \dots$, of the space W, observe that each component W_l is generated (as a linear space) by the functions $u(t) = \cos lt \cdot a_l + \sin lt \cdot b_l$, for $t \in [0, 2\pi]$, $a_l, b_l \in V$. Consequently, the function $u(t)$ can be viewed as

$$u(t) = e^{ilt} \cdot (x_l + iy_l), \quad x_l, y_l \in V. \tag{79}$$

Using (79), we can describe the action of $e^{i\tau} \in S^1$ on $u(t)$ as the 'complex multiplication' by $e^{il\tau}$. Moreover, each of the subspaces W_l, equipped with this S^1-action, is S^1-isomorphic to the complexification V^c of V, considered with the S^1-action defined by the l-folding, i.e. $W_l \simeq {}^lV^c$ (cf. Subsection 3.6 and Remark and Notation 2.6). Also, we can write the elements $u(t) \in W_l$ as $u(t) = \mathrm{e}^{\mathrm{i}lt} \cdot z$, $z \in \mathbb{C}^N$. Therefore, we have the following S^1-isotypical decomposition of W:

$$W = \overline{V \oplus \bigoplus_{l=1}^{\infty} W_l}. \tag{80}$$

S^1-equivariant fixed-point problem Define the following operators:

$$L : W \to L^2(S^1; V), \quad L(u) = \dot{u}, \tag{81}$$

$$j : W \to C(S^1; V), \quad j(u) = \tilde{u}, \tag{82}$$

where $u \in W$ and $C(S^1; V)$ denotes the space of 2π-periodic, V-valued continuous functions on $\mathbb{R}$ equipped with the usual sup-norm and $\tilde{u} = u$ a.e. (cf. [137]). For a function $u(t) = \mathrm{e}^{\mathrm{i}lt} \cdot z \in W_l$, we have $L(u) = \mathrm{i}lu$, i.e. $L|_{W_l} = \mathrm{i}l \, \mathrm{Id}|_{W_l}$. Put $\mathbb{R}^2_+ := \mathbb{R} \times \mathbb{R}_+$, $\mathbb{R}_+ := \{r \in \mathbb{R}\colon r > 0\}$, and let $\pi\colon \mathbb{R}^2 \oplus W \to W$ be the natural projection. Denote by N_f the Nemytsky operator associated with the function f, i.e. for every $\alpha \in \mathbb{R}$, put

$$\big(N_f(\alpha, v)\big)(t) := f\big(\alpha, v(t)\big), \quad v \in C(S^1; V). \tag{83}$$

Notice that the system (71) is equivalent to the following operator equation

$$Lu = \frac{1}{\beta} N_f\big(\alpha, j(u)\big), \quad (\alpha, \beta) \in \mathbb{R}^2_+, \ u \in W. \tag{84}$$

Equation (84) can be transformed to an S^1-equivariant fixed-point problem in $\mathbb{R}^2 \oplus W$. Define the operator $K : W \to L^2(S^1; V)$ by

$$K(u) := \frac{1}{2\pi} \int_0^{2\pi} u(t) \, \mathrm{d}t, \quad u \in W, \tag{85}$$

which is simply a projection on the subspace V of constant functions (cf. Remark and Notation 2.17). Then, it can be easily verified that the operator $L + K : W \to L^2(S^1; V)$ is an isomorphism. Put

$$\begin{aligned} \mathcal{F}(\alpha, \beta, u) &:= (L + K)^{-1} \left[\frac{1}{\beta} N_f\big(\alpha, j(u)\big) + K(u) \right], \\ \mathfrak{F}(\alpha, \beta, u) &:= u - \mathcal{F}(\alpha, \beta, u). \end{aligned} \tag{86}$$

In this way, the following equation is equivalent to (84):

$$\mathfrak{F}(\alpha, \beta, u) = 0, \tag{87}$$

where $(\alpha, \beta) \in \mathbb{R}^2_+$, $u \in W$ (cf. Subsection 3.6 and Remark and Notation 2.6).

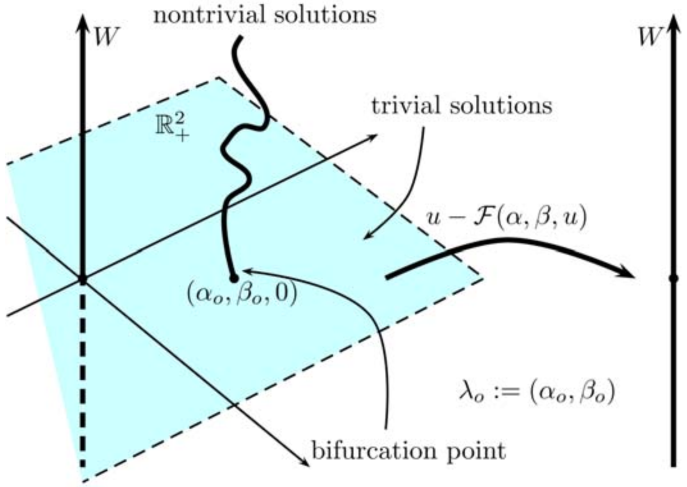


Fig. 1. Hopf bifurcation as a two-parameter fixed-point problem in functional spaces.

4.3. S^1-degree method for Hopf bifurcation problem

It is clear that every point $(\alpha, \beta, 0) \in \mathbb{R}^2_+ \times W$ is a solution to Eq. (87) (called a *trivial solution*). Denote the set $\mathbb{R}^2_+ \times \{0\} \subset \mathbb{R}^2_+ \times W$ by M. We would like to identify the so-called *bifurcation points* to (87), i.e. the trivial solutions $(\alpha_o, \beta_o, 0) \in M$ such that in every neighborhood of $(\alpha_o, \beta_o, 0)$ there is a non-trivial solution $(\alpha, \beta, u) \in \mathbb{R}^2_+ \times W$. Actually, in the case $(\alpha_o, \beta_o, 0) \in M$ is a bifurcation point, there is a "branch" in $\mathbb{R}^2_+ \times W$ of non-trivial solutions bifurcating from the trivial solution $(\alpha_o, \beta_o, 0) \in M$. The bifurcation problem for Eq. (87) is illustrated on Fig. 1.

Necessary condition for the occurrence of Hopf bifurcation By the implicit function theorem, the *necessary condition* for the point $(\alpha_o, \beta_o, 0)$ to be a bifurcation point, is that the linear operator

$$\mathrm{Id} - D_u\mathcal{F}(\alpha_o, \beta_o, 0) : W \to W$$

is not an isomorphism. In such a case, we say that $(\alpha_o, \beta_o, 0) \in M$ is a W*-singular point*. It follows directly from the definition of the map $\mathcal{F}$ that

$$a(\alpha, \beta) := \mathrm{Id} - D_u\mathcal{F}(\alpha, \beta, 0) = \mathrm{Id} - (L + K)^{-1}\left[\frac{1}{\beta} N_{D_x f}(\alpha, 0) + K\right], \tag{88}$$

where the operator $N_{D_x f}(\alpha, 0)$ is the Nemytsky operator of $D_x f(\alpha, 0)$, defined by

$$\big(N_{D_x f}(\alpha, 0) j(u)\big)(t) := D_x f(\alpha, 0) j(u)(t), \quad u \in W.$$

Clearly, $a(\alpha,\beta)(V)\subset V$ and $a(\alpha,\beta)(W_l)\subset W_l$, for all $l=1,2,3,\dots$. Put

$$a_0(\alpha,\beta):=a(\alpha,\beta)|_V,\quad a_l(\alpha,\beta):=a(\alpha,\beta)|_{W_l},\quad l=1,2,3,\dots. \tag{89}$$

It is clear that for a constant function $\bar{u}\in V\subset W$,

$$a_0(\alpha,\beta)\bar{u}=-\frac{1}{\beta}D_x f(\alpha,0)\bar{u}, \tag{90}$$

and for $e^{ilt}\cdot z\in W_l$, where $z=x+iy$, $x,y\in V$, we have (see (74) and (79))

$$a_l(\alpha,\beta)=\frac{1}{il\beta}\Delta_\alpha(il\beta). \tag{91}$$

Consequently, if $(\alpha_o,\beta_o,0)$ is a W-singular point, then at least one of the linear operators $a_l(\alpha_o,\beta_o)$, $l=0,1,2,\dots$, cannot be an isomorphism. In order to exclude the possibility for the occurrence of the so-called steady-state bifurcation (i.e. a bifurcation of constant solutions), we assume that the operator (90) is an isomorphism, i.e.

$$\det D_x f(\alpha_o,0)\neq 0. \tag{92}$$

Using (89)–(92) one can easily establish

PROPOSITION 4.1 (Necessary condition for Hopf bifurcation). *Assume the operator* $a(\alpha_o,\beta_o)\colon W\to W$ *(see* (88)*) is not an isomorphism for some* $(\alpha_o,\beta_o)\in\mathbb{R}\times\mathbb{R}_+$*, and suppose condition* (92) *is satisfied. Then, the characteristic equation* (75) *has a purely imaginary root* $il\beta_o$ *for* $\alpha=\alpha_o$ *and some* $l\in\mathbb{N}$.

DEFINITION 4.2.

(i) A point $(\alpha_o,x_o)\in\mathbb{R}\oplus\mathbb{R}^N$ is called a *center* for the system (71), if its characteristic equation (75) has a purely imaginary characteristic value $i\beta_o$ ($\beta_o>0$) for $\alpha=\alpha_o$.

(ii) A center (α_o,x_o) for the system (71) is called *isolated* if it is the only center for (71) in a certain neighborhood of (α_o,x_o) in $\mathbb{R}\oplus V$.

Sufficient conditions for the occurrence of Hopf bifurcation In what follows, it is convenient to identify $\mathbb{R}^2$ with $\mathbb{C}$, and instead of writing $(\alpha,\beta)\in\mathbb{R}^2$, we use the notation $\lambda=\alpha+i\beta$. In particular, we write $\lambda_o=\alpha_o+i\beta_o$.

Assume $(\alpha_o,0)$ is an isolated center for the system (71) with $\lambda_o=\alpha_o+i\beta_o$ (see Definition 4.2). Then, the point $(\lambda_o,0)\in\mathbb{R}^2_+\times W$ is an isolated W-singular point of $\mathfrak{F}$ (cf. (86)), i.e. for a sufficiently small $\varepsilon>0$, the linear operator $a(\lambda)=\mathrm{Id}-D_u\mathcal{F}(\lambda,0)\colon W\to W$, where $\lambda=\alpha+i\beta$ and $|\lambda-\lambda_o|<\varepsilon$, is not an isomorphism only if $\lambda=\lambda_o$. Consequently, by the implicit function theorem, there exists $r>0$ such that for all points $(\lambda,u)\in\mathbb{R}^2_+\oplus W$ with $|\lambda-\lambda_o|=\varepsilon$ and $0<\|u\|\leqslant r$, we have $u-\mathcal{F}(\lambda,u)\neq 0$. Define the set $\Omega\subset\mathbb{R}^2\oplus W$ by

$$\Omega=\bigl\{(\lambda,u)\in\mathbb{R}^2\oplus W\colon\ |\lambda-\lambda_o|<\varepsilon \text{ and } \|u\|<r\bigr\}. \tag{93}$$

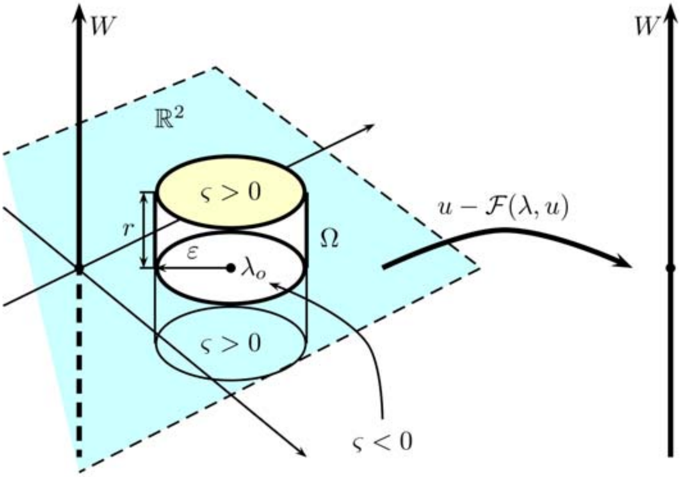


Fig. 2. Auxiliary function for Hopf bifurcation.

Put

$$\partial_0 := \overline{\Omega} \cap \left(\mathbb{R}^2 \times \{0\}\right) \quad \text{and} \quad \partial_r := \left\{(\lambda, u) \in \overline{\Omega}\colon \|u\| = r\right\}.$$

DEFINITION 4.3. An S^1-invariant continuous function $\varsigma : \overline{\Omega} \to \mathbb{R}$ is called an *auxiliary* function for $\mathfrak{F}$ in Ω, if

$$\begin{cases} \varsigma(\lambda, u) > 0 & \text{for } (\lambda, u) \in \partial_r, \\ \varsigma(\lambda, u) < 0 & \text{for } (\lambda, u) \in \partial_0 \end{cases}$$

(see Fig. 2).

The existence of an auxiliary function is provided by Proposition 3.9.

Take an auxiliary function ς and define the map $\mathfrak{F}_\varsigma : \overline{\Omega} \to \mathbb{R} \oplus W$ by

$$\mathfrak{F}_\varsigma(\lambda, u) := \left(\varsigma(\lambda, u), u - \mathcal{F}(\lambda, u)\right), \quad (\lambda, u) \in \overline{\Omega}. \tag{94}$$

Since $\mathfrak{F}_\varsigma(\lambda, u) \neq 0$ for $(\lambda, u) \in \partial\Omega$, it follows that $\mathfrak{F}_\varsigma$ is an Ω-admissible S^1-equivariant compact field, so the Leray–Schauder S^1-equivariant degree $S^1\text{-Deg}(\mathfrak{F}_\varsigma, \Omega)$ is well-defined.

THEOREM 4.4. *Given the system* (71), *assume conditions* (72) *and* (92) *are satisfied*. *Suppose*, *further*, $(\alpha_o, 0)$ *is an isolated center for* (71) (*cf. Definition* 4.2). *Take* $\mathfrak{F}$ *defined by* (86) *and construct* Ω *according to* (93). *Let* $\varsigma : \overline{\Omega} \to \mathbb{R}$ *be an auxiliary function* (*see Definition* 4.3) *and let* $\mathfrak{F}_\varsigma$ *be defined by* (94). *Assume, finally,* $S^1\text{-Deg}(\mathfrak{F}_\varsigma, \Omega) \neq 0$.

Then, there exists a branch of nontrivial solutions bifurcating from the point $(\lambda_o, 0) \in \Omega$. *More precisely, the closure of the set composed of all non-trivial solutions* $(\lambda, u) \in \Omega$ *to* (87), *i.e.*

$$\overline{\{(\lambda, u) \in \Omega\colon\ u - \mathcal{F}(\lambda, u) = 0,\ u \neq 0\}},$$

contains a compact connected subset C *such that*

$$(\lambda_o, 0) \in C \quad \text{and} \quad C \cap \partial_r \neq \emptyset.$$

The main ingredients of the proof of Theorem 4.4 are: invariant Urysohn function techniques, Zorn lemma and the following statement, which is an immediate consequence of Theorem 3 from [97], p. 170.

LEMMA 4.5. *Let* X *be a metric space,* $A, B \subset X$ *two disjoint closed sets in* X*, and* K *a compact set in* X *such that* $K \cap A \neq \emptyset \neq K \cap B$. *If the set* K *does not contain a connected component* K_o *such that* $K_o \cap A \neq \emptyset \neq K_o \cap B$, *then there exist two open disjoint sets* V_1 *and* V_2 *such that* $A \cup B \cup K \subset V_1 \cup V_2$, $A \subset V_1$ *and* $B \subset V_2$.

4.4. *Deformation of the map $\mathfrak{F}_\varsigma$: Reduction to a product map*

Theorem 4.4 reduces the study of the Hopf bifurcation for (71) to computing S^1-Deg$(\mathfrak{F}_\varsigma, \Omega)$. In this subsection, we deform $\mathfrak{F}_\varsigma$ to a field with "more computable" S^1-degree.

Complementing function and linearization Replace the set Ω (resp. the auxiliary function ς) with the set Ω_1 (resp. the so-called *complementing function* $\tilde{\varsigma} : \overline{\Omega_1} \to \mathbb{R}$) defined as

$$\Omega_1 := \left\{(\lambda, u) \in \mathbb{C} \oplus W\colon\ \frac{\varepsilon}{4} < |\lambda - \lambda_o| < \varepsilon,\ \|u\| < r\right\}. \tag{95}$$

Respectively,

$$\tilde{\varsigma} := |\lambda - \lambda_o|\big(\|u\| - r\big) + \|u\| + \varepsilon\frac{r}{2}. \tag{96}$$

For each $(\lambda, u) \in \Omega_1$, put

$$\eta(\lambda) := \lambda_o + \frac{\varepsilon(\lambda - \lambda_o)}{2|\lambda - \lambda_o|} \tag{97}$$

and define the map $\tilde{\mathfrak{F}} : \overline{\Omega_1} \to \mathbb{R} \oplus W$ by

$$\tilde{\mathfrak{F}}(\lambda, u) := \big(\tilde{\varsigma}(\lambda, u), a\big(\eta(\lambda)\big)u\big), \tag{98}$$

where $a(\cdot)$ is given by (88). Obviously, the map (98) is Ω_1-admissible S^1-equivariant and (by excision and homotopy property of twisted degree)

$$S^1\text{-Deg}(\tilde{\mathfrak{F}}, \Omega_1) = S^1\text{-Deg}(\mathfrak{F}_\varsigma, \Omega). \tag{99}$$

Finite-dimensional normalization Obviously, one may consider $a(\eta(\cdot))$ from (98) as a continuous map

$$\left\{\lambda \in \mathbb{C}\colon \frac{\varepsilon}{4} < |\lambda - \lambda_o| < \varepsilon\right\} \overset{a(\eta(\cdot))}{\longrightarrow} GL_c^{S^1}(W), \tag{100}$$

where $GL_c^{S^1}(W)$ stands for the group of S^1-equivariant linear completely continuous fields in W. Put

$$\Sigma := \left\{\lambda \in \mathbb{C}\colon |\lambda - \lambda_o| = \frac{\varepsilon}{2}\right\}. \tag{101}$$

By (97), the map (100) is completely determined by its restriction to $\Sigma \simeq S^1$. Denote by $\bar{a}(\cdot)$ the restriction of (100) to (101). Also, given $m \in \mathbb{N}$, define $W^m := V \oplus \bigoplus_{l=1}^m W_l \subset W$. Using the standard compactness argument, one can easily show

PROPOSITION 4.6. *There exists $m \in \mathbb{N}$ such that*:

(i) $W = W^m \oplus (W^m)^\perp$;

(ii) *for any $\lambda \in \Sigma$, the map $\tilde{a}(\lambda)$, defined by $\tilde{a}(\lambda) := \bar{a}(\lambda)|_{W^m} + \operatorname{Id}|_{(W^m)^\perp}$, belongs to $GL_c^{S^1}(W)$*;

(iii) *the map $\mathfrak{F}_2 : \overline{\Omega}_1 \to \mathbb{R} \oplus W$, defined by $\mathfrak{F}_2(\lambda, u) := (\tilde{\varsigma}(\lambda, u), \tilde{a}(\eta(\lambda))u)$, is the compact equivariant Ω_1-admissible field* (*cf.* (95));

(iv) S^1-Deg$(\tilde{\mathfrak{F}}, \Omega_1) = S^1$-Deg$(\mathfrak{F}_2, \Omega_1)$ (*cf.* (99)).

Isotypical deformations and finite-dimensional reduction Consider the map $\tilde{a}$ defined in Proposition 4.6(ii) more intently. Given $\lambda \in \Sigma$, put

$$\tilde{a}_0(\lambda) := \tilde{a}(\lambda)|_V; \qquad \tilde{a}_l(\lambda) := \tilde{a}(\lambda)|_{W_l}, \quad l > 0.$$

It follows immediately from (88)–(90), (92), that the map $\tilde{a}_0 : \Sigma \to GL(V) = GL(N, \mathbb{R})$ is homotopic to a constant map

$$a_0(\lambda_o) := -\frac{1}{\beta_o} D_x f(\alpha_o, 0), \tag{102}$$

where $\lambda_o = \alpha_o + \mathrm{i}\beta_o$.

Further, $\tilde{a}_l(\lambda) : W_l \to W_l$ is S^1-equivariant, therefore, $\tilde{a}_l(\lambda) \in GL^{S^1}(W_l) \simeq GL(N, \mathbb{C})$ (cf. Proposition 2.9). Thus, the map $\tilde{a}_l : \Sigma \to GL(N, \mathbb{C})$ determines an element $\mu \in$

$\pi_1(GL(N,\mathbb{C}))$. By Proposition 3.35(ii), μ contains a representative $b_l:\Sigma\to GL(N,\mathbb{C})$ of the form

$$b_l(\lambda)=\begin{bmatrix}(\lambda-\lambda_o)^{\mathfrak{k}_l} & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 1\end{bmatrix}:\mathbb{C}^N\to\mathbb{C}^N, \tag{103}$$

where $\mathfrak{k}_l$ is an integer.

Using Proposition 4.6(i), (ii), (102) and (103), define the map $\hat{a}:\Sigma\to GL_c^{S^1}(W)$ by

$$\hat{a}(\lambda):=a_0(\lambda_o)+\bigoplus_{l=1}^{m}b_l(\lambda)+\mathrm{Id}\,|_{(W^m)^\perp}, \tag{104}$$

and the compact S^1-equivariant field $\hat{\mathfrak{F}}:\overline{\Omega}_1\to W$ by $\hat{\mathfrak{F}}(\lambda,u):=(\tilde{\varsigma}(\lambda,u),\hat{a}(\eta(\lambda))u)$. Put $\Omega_o:=\Omega_1\cap(\mathbb{C}\oplus V\oplus\bigoplus_{l=1}^m W_l)$ and consider the map $\mathfrak{F}_o:=\hat{\mathfrak{F}}|_{\overline{\Omega_o}}$. By construction of the Leray–Schauder equivariant degree and the suspension property,

$$S^1\text{-Deg}(\mathfrak{F}_2,\Omega_1)=S^1\text{-Deg}(\mathfrak{F}_o,\Omega_o). \tag{105}$$

Product formula for S^1-Deg$(\mathfrak{F}_\varsigma,\Omega)$ Take $\Omega_*:=B\times\Omega_{o*}$, where $\Omega_{o*}:=\Omega_o\cap(\mathbb{C}\oplus\bigoplus_{l=1}^m W_l)$ and B is the unit ball in V, and define the map $\mathfrak{F}_{o*}:\overline{\Omega_{o*}}\to\mathbb{R}\oplus\bigoplus_{l=1}^m W_l$ by

$$\mathfrak{F}_{o*}(\lambda,v):=\left(\tilde{\varsigma}(\lambda,v),\bigoplus_{l=1}^{m}b_l\big(\eta(\lambda)\big)v_l\right)$$

(here v_l is an orthogonal projection of $v\in\bigoplus_{l=1}^m W_l$ to W_l (cf. (103))). Next, take the product map

$$\mathfrak{F}_*:=a_0(\lambda_o)\times\mathfrak{F}_{o*}:\overline{B\times\Omega_{o*}}\to V\oplus\left(\mathbb{R}\oplus\bigoplus_{l=1}^{m}W_l\right)$$

(see (102)). Using Theorem 3.48 and Remark 3.49, one obtains

$$S^1\text{-Deg}(\mathfrak{F}_o,\Omega_o)=S^1\text{-Deg}(\mathfrak{F}_*,\Omega_*)=\deg\big(a_0(\lambda_o),B\big)\cdot S^1\text{-Deg}(\mathfrak{F}_{o*},\Omega_{o*}),$$

where “deg” stands for the usual Brouwer degree. Observe, $\deg(a_0(\lambda_o),B)=\operatorname{sign}\det a_0(\lambda_o)$. Therefore, by applying Proposition 3.34, Lemma 3.37 and Theorem 3.39, we obtain

PROPOSITION 4.7. *Under the assumptions and notations of Theorem* 4.4, *one has*:

$$S^1\text{-Deg}(\mathfrak{F}_\varsigma,\Omega)=\operatorname{sign}\det a_0(\lambda_o)\cdot\sum_{l=1}^{m}\mathfrak{k}_l(\mathbb{Z}_l), \tag{106}$$

where a_0 *is given by* (90) *and the numbers* $\mathfrak{k}_l$ *are defined by*

$$\mathfrak{k}_l = \deg\bigl(\det_{\mathbb{C}}(a_l), \widetilde{D}\bigr) \tag{107}$$

for a_l *given by* (91) *and*

$$\widetilde{D} := \bigl\{\lambda \in \mathbb{C}\colon |\lambda - \lambda_o| < \varepsilon\bigr\}. \tag{108}$$

4.5. *Crossing numbers*

Under the assumption that one is able to effectively compute the numbers $\mathfrak{k}_l$ in formula (107), Proposition 4.7 together with Theorem 4.4 provide an effective way to study the Hopf bifurcation for the system (71). Fortunately, this task can be accomplished by using the information related to the S^1-equivariant spectral decomposition of $D_x f(\alpha, 0)$.

It is convenient to replace the set $\widetilde{D}$ in Proposition 4.7 with a set D defined by

$$D := \bigl\{\lambda = \alpha + \mathrm{i}\beta \in \mathbb{C}\colon \max\bigl\{|\alpha - \alpha_o|, |\beta - \beta_o|\bigr\} < \varepsilon\bigr\} \tag{109}$$

with, possibly, smaller ε. By excision and homotopy properties of Brouwer degree,

$$\mathfrak{k}_l = \deg\bigl(\det_{\mathbb{C}}(a_l), D\bigr).$$

Clearly, if the number $\mathrm{i}l\beta_o$ is not a root of the characteristic equation, then $\mathfrak{k}_l = 0$. Suppose, therefore, that $\mathrm{i}l\beta_o$ is a root of the characteristic equation (75) for $\alpha = \alpha_o$. Since $(\alpha_o, 0)$ is an isolated center, we can assume, without loss of generality, that the number ε in (109) is so small that for every α satisfying $0 < |\alpha - \alpha_o| < \varepsilon$, there is no purely imaginary root $\mathrm{i}\beta'$ of (75) with $|\beta' - l\beta_o| \leqslant \varepsilon$. Also, there exists $\delta > 0$ such that there are no characteristic roots $\tau + \mathrm{i}\beta$ of Eq. (75) for $|\alpha - \alpha_o| \leqslant \varepsilon$ belonging to the boundary of the set

$$S := \bigl\{\tau + \mathrm{i}\beta\colon 0 < \tau < \delta, |\beta - l\beta_o| < \varepsilon\bigr\}, \tag{110}$$

except for the root $\mathrm{i}l\beta$ corresponding to $\alpha = \alpha_o$ (see Fig. 3).

Using a simple Brouwer degree argument, one can show that

$$\mathfrak{k}_l = \deg(\det_{\mathbb{C}} \Delta_{\alpha_-}, S) - \deg(\det_{\mathbb{C}} \Delta_{\alpha_+}, S) \quad (\alpha_\pm := \alpha_o \pm \varepsilon). \tag{111}$$

DEFINITION 4.8. Define the following integers

$$\mathfrak{t}_l^\pm(\alpha_o, \beta_o) := \deg(\det_{\mathbb{C}} \Delta_{\alpha_\pm}, S),$$

and

$$\mathfrak{t}_l(\alpha_o, \beta_o) := \mathfrak{t}_l^-(\alpha_o, \beta_o) - \mathfrak{t}_l^+(\alpha_o, \beta_o). \tag{112}$$

The integer $\mathfrak{t}_l := \mathfrak{t}_l(\alpha_o, \beta_o)$ is called the *l-th isotypical crossing number* at (α_o, β_o).

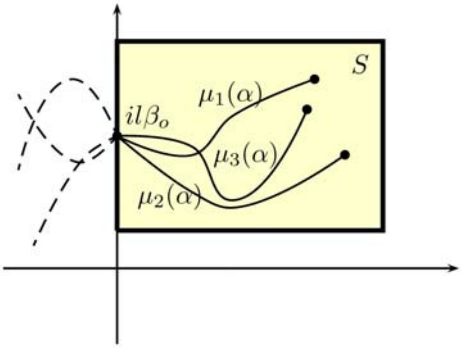

S
$\mu_1(\alpha)$
$il\beta_o$
$\mu_3(\alpha)$
$\mu_2(\alpha)$

Fig. 3. Characteristic values $\mu_k(\alpha)$ near i$l\beta_o$.

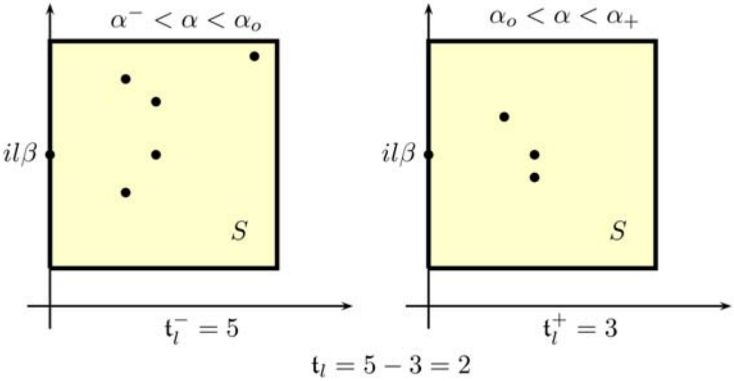

$\alpha^- < \alpha < \alpha_o$
$\alpha_o < \alpha < \alpha_+$
$il\beta$
$il\beta$
S
S
$\mathfrak{t}_l^- = 5$
$\mathfrak{t}_l^+ = 3$
$\mathfrak{t}_l = 5 - 3 = 2$

Fig. 4. The l-th isotypical crossing number.

The crossing number $\mathfrak{t}_l = \mathfrak{t}_l^- - \mathfrak{t}_l^+$ has a very simple interpretation. In the case i$l\beta_o$ is an eigenvalue of the matrix $D_x f(\alpha_o, 0)$, the number $\mathfrak{t}_l^-$ counts the number of all the eigenvalues (with multiplicity) in the set S before α crosses the value α_o, and the number $\mathfrak{t}_l^+$ counts the eigenvalues in S after α crosses α_o. The difference, which is exactly the number $\mathfrak{t}_l$, represents the net number of the eigenvalues which ‘escaped’ (if $\mathfrak{t}_l$ is positive) or ‘entered’ (if $\mathfrak{t}_l$ is negative) the set S when α was crossing α_o. This situation is illustrated on Fig. 4.

4.6. *Conclusions*

Return to the system (71). Take an isolated center $(\alpha_o, 0)$ corresponding to the purely imaginary characteristic root iβ_o, $\beta_0 > 0$. Then, Theorem 4.4 and Proposition 4.7 show that the equivariant degree S^1-Deg$(\mathfrak{F}_\varsigma, \Omega)$ is a *local Hopf bifurcation invariant* describing a possible occurrence of the bifurcation. Namely, the non-zero components of the equivariant degree provide the information about the existence of branches of non-constant periodic solutions to (71) bifurcating from the point $(\alpha_o, 0)$ (with the limit frequency β_o).

Furthermore, formulae (110)–(112), in compliance with Theorem 4.4 and Proposition 4.7, allow us to formulate an effective sufficient condition for the occurrence of the Hopf bifurcation in terms of the right-hand side of the system (71) only. Namely,

THEOREM 4.9. *Let* $V := \mathbb{R}^N$ *and let* $f : \mathbb{R} \oplus V \to V$ *and* $\alpha_o \in \mathbb{R}$ *satisfy the following conditions*:

(i) f *is of class* C^1;
(ii) $f(\alpha, 0) = 0$ *for all* $\alpha \in \mathbb{R}$;
(iii) $\det D_x f(\alpha_o, 0) \neq 0$, *where* $D_x f$ *denotes the derivative of* f *with respect to* $x \in V$;
(iv) $(\alpha_o, 0)$ *is an isolated center for the system* (71) *with* $\mathrm{i}\beta_o$ $(\beta_o > 0)$ *being the corresponding characteristic root* (*cf. Definition* 4.2 *and formulae* (74), (75));
(v) $\mathfrak{t}_1(\alpha_o, \beta_o) = \mathfrak{t}_1^-(\alpha_o, \beta_o) - \mathfrak{t}_1^+(\alpha_o, \beta_o) \neq 0$ (*see formulae* (110)–(112)).

Then, there exists a sequence (*in fact, it is a connected branch*) *of non-constant* $\frac{2\pi}{\beta_n}$-*periodic*[1] *solutions* $x_{\alpha_n}(t)$ *of the system* (71) (*with* $\alpha = \alpha_n$) *such that*

$$\begin{cases} \alpha_n \to \alpha_o, \\ x_{\alpha_n}(t) \to 0, \end{cases} \quad \text{as } n \to \infty.$$

PROOF. By conditions (i), (ii) and (iv), one can associate to (α_0, β_o) the degree $S^1\text{-Deg}(\mathfrak{F}_\varsigma, \Omega)$ (cf. formulae (93) and (94)). By Proposition 4.7 and formulae (110)–(112),

$$S^1\text{-Deg}(\mathfrak{F}_\varsigma, \Omega) = \operatorname{sign}\det a_0 \sum_{l=1}^{m} \mathfrak{k}_l(\mathbb{Z}_l) = \operatorname{sign}\det a_0 \sum_{l=1}^{m} \mathfrak{t}_l(\alpha_o, \beta_o)(\mathbb{Z}_l)$$

for some $m \geqslant 1$. By condition (v), $\mathfrak{t}_1(\alpha_o, \beta_o) \neq 0$, and the conclusion follows from Theorem 4.4 (condition (iii) excludes possible steady-state bifurcation). □

4.7. *Bibliographical remarks*

The standard approach to attack the Hopf bifurcation problem occurring in ODEs is based on the use of the Central Manifold theorem (allowing a reduction to the two-dimensional case) and further studying the properties of the Poincaré map associated with the induced system (an excellent exposition of this concept is presented in [113], see also [8,68,75]). Other techniques for studying Hopf bifurcation problems, which are based on Lyapunov–Schmidt reduction, normal forms and applications of the singularity theory can be found in [33,34,64,111]. The idea of a crossing number $\mathfrak{t}$ can be traced back to [2], where the framed bordism method was used to show that, if $\mathfrak{t}$ is odd, then a Hopf bifurcation occurs. The general result stating that, if $\mathfrak{t} \neq 0$ then a Hopf bifurcation takes place, was proved with the use of the Fuller index in [35]. One should point out that the framed bordism invariants and Fuller index can be expressed in terms of the equivariant degree (cf. [81,19]). The S^1-equivariant degree approach to the above problem described in this section was discussed in [61,93,19] (see also [46,49,81,95]). On the other hand, for Hamiltonian systems (resp. systems with first integral), important results were obtained by N.E. Dancer in [39] (resp. by E.N. Dancer and J.F. Toland in [40,42,41]).

---

[1] We do not assume here that $\frac{2\pi}{\beta_n}$ is a minimal period.

5. Hopf bifurcation problem for ODEs with symmetries

5.1. *Symmetric Hopf bifurcation and local bifurcation invariant*

Statement of the problem In this section, we extend the setting and methods developed in Section 4 to discuss the Hopf bifurcation phenomenon for the system (71) admitting a certain compact Lie group of symmetries. More precisely, suppose $V := \mathbb{R}^N$ is an orthogonal representation of a compact Lie group Γ and assume that

(A0) $f : \mathbb{R} \oplus V \to V$ is Γ-equivariant.

Also, assume that f and $\alpha_o \in \mathbb{R}$ satisfy the hypotheses (i)–(iv) from Theorem 4.9.

We are interested in describing possible symmetries of non-constant periodic solutions to the system (71) bifurcating from the origin.

$\Gamma \times S^1$-Equivariant reformulation of the problem Following the same scheme as in the non-symmetric case (see Subsection 4.2), we start with the normalization of the period (see (76) and (77)) and setting in functional spaces (see (84)). Observe that the space $W = H^1(S^1; V)$ is a natural Hilbert isometric $\Gamma \times S^1$-representation with the action given by $((\gamma, \zeta)u)(t) = \gamma u(t + \tau)$, $t \in \mathbb{R}$, $\gamma \in \Gamma$, $\zeta = e^{i\tau} \in S^1$, $u \in W$. In the same way, the space $\mathcal{C} = C(S^1; V)$ is a Banach isometric G-representation. Further, consider $\mathbb{R}^2_+$ (identified with the subspace of $\mathbb{C}$) as a trivial G-space. Then, all the operators L, j, N_f, and K (cf. (81)–(83) and (85)) turn out to be G-equivariant. Consequently, the system (71) translates as the G-equivariant fixed-point problem (87).

G-isotypical decomposition of W The S^1-isotypical decomposition (80) of W is clearly G-invariant, therefore we can use it to construct the G-isotypical decomposition of W. Consider the real (resp. complex) Γ-isotypical decomposition of V (resp. V^c):

$$V = V_0 \oplus V_1 \oplus \cdots \oplus V_r, \qquad V^c = U_0 \oplus U_1 \oplus \cdots \oplus U_s, \tag{113}$$

where the isotypical components V_i (resp. U_j) are modeled on the real (resp. complex) Γ-irreducible representations $\mathcal{V}_i$ (resp. $\mathcal{U}_j$). Notice that it is possible that there is a different number of isotypical components in the isotypical decompositions of V and V^c. The isotypical components of W_l, $l \geqslant 1$, are easily obtained (see Remark and Notation 2.6):

$$W_l = V_{0,l} \oplus V_{1,l} \oplus \cdots \oplus V_{s,l}, \tag{114}$$

where the components $V_{j,l} := {}^{l}U_j$, $j = 0, 1, \ldots, s$, are modeled on the irreducible G-representation $\mathcal{V}_{j,l}$. In this way, we obtain the following G-isotypical decomposition of W:

$$W = [V_0 \oplus V_1 \oplus \cdots \oplus V_r] \oplus \overline{\bigoplus_{l=1}^{\infty} (V_{0,l} \oplus V_{1,l} \oplus \cdots \oplus V_{s,l})}. \tag{115}$$

Dominating orbit types The following notion will be used to estimate the *minimal number* of different periodic solutions (as well as their symmetries) to the system (71) (see Theorem 5.4).

DEFINITION 5.1. An orbit type (H) in W is called *dominating,* if (H) is maximal (with respect to the usual order relation (10)) in the class of all φ-twisted *one-folded* orbit types in W (i.e. $H = K^{\varphi}$, where $\varphi : K \to S^1$ is a homomorphism, and $\dim W(K) = 0$ (see Section 3.3 for more details)).

REMARK 5.2. Let (H) be a dominating orbit type in W. Using the maximality property of (H), it is easy to see that there exists an irreducible subrepresentation $\mathcal{V} \subset W$ and a non-zero vector $u \in \mathcal{V}$ such that $G_u = H$. Consequently, the dominating orbit types in W can be easily recognized from the isotypical decomposition of W and lattices of orbit types of the corresponding to this decomposition irreducible G-representations.

REMARK 5.3. Assume there is a solution $u_o \in W$ to (77) (for $\alpha = \alpha_o$ and some $\beta > 0$), for which one has $G_{u_o} \supset H_o$. If (H_o) is a dominating orbit type in W with $H_o = K^{\varphi}$ for some $K \subset \Gamma$ and $\varphi : K \to S^1$, then, by maximality condition, $(G_{u_o}) = (K^{\varphi,l})$ with $l \geqslant 1$, and the corresponding orbit $G(u_o)$ is composed of exactly $|G/G_{u_o}|_{S^1}$ different periodic functions (where $|Y|_{S^1}$ denotes the number of S^1-orbits in Y). It is easy to check that the number of S^1-orbits in G/G_{u_o} is $|\Gamma/K|$ (where $|X|$ stands for the number of elements in X).

On the other hand, suppose that x_o is, say, a p-periodic solution to (71) canonically corresponding to the above u_o. It follows from the definition of l-folding and $\Gamma \times S^1$-action on W that x_o is also a $\frac{p}{l}$-periodic solution to (71). The pair $(x_o, \frac{p}{l})$ canonically determines an element $u_o' \in W$ being a solution to (77) (for $\alpha = \alpha_o$ and some β') satisfying the condition $G_{u_o'} = H_o$. In this way, we obtain that (71) has at least $|\Gamma/K|$ different periodic solutions with the orbit type exactly (H_o).

Sufficient conditions for symmetric Hopf bifurcation Similarly to the non-equivariant case, define the set Ω by (93). Since W is an isometric G-representation, Ω is clearly G-invariant. Combining the standard Urysohn function techniques (cf. Proposition 3.9) with Definition 4.3, construct a G-invariant auxiliary function $\varsigma : \overline{\Omega} \to \mathbb{R}$. Then, define the map $\mathfrak{F}_{\varsigma} : \overline{\Omega} \to \mathbb{R} \oplus W$ by (94), which is a G-equivariant Ω-admissible compact field. Therefore, one can assign to $\mathfrak{F}_{\varsigma}$ the twisted degree $G\text{-Deg}^{t}(\mathfrak{F}_{\varsigma}, \Omega)$.

Combining the properties of the twisted degree with Remark 5.3 and Lemma 4.5 (cf. Theorem 4.4), one can easily establish

THEOREM 5.4. *Given the system* (71), *assume conditions* (i)–(iv) *from Theorem* 4.9 *and condition* (A0) *to be satisfied. Take* $\mathfrak{F}$ *defined by* (86) *and construct* Ω *according to* (93). *Let* $\varsigma : \overline{\Omega} \to \mathbb{R}$ *be a* G*-invariant auxiliary function* (*see Definition* 4.3) *and let* $\mathfrak{F}_\varsigma$ *be defined by* (94).

(i) *Assume* $G\text{-Deg}^t(\mathfrak{F}_\varsigma, \Omega) \neq 0$, *i.e.*

$$G\text{-Deg}^t(\mathfrak{F}_\varsigma, \Omega) = \sum_{(H)} n_H(H) \quad \textit{and} \quad n_{H_o} \neq 0 \tag{116}$$

for some $(H_o) \in \Phi_1^t(G)$. *Then, there exists a branch of non-trivial solutions to* (71) *bifurcating from the point* $(\alpha_o, 0)$ (*with the limit frequency* $l\beta_o$ *for some* $l \in \mathbb{N}$). *More precisely, the closure of the set composed of all non-trivial solutions* $(\lambda, u) \in \Omega$ *to* (87), *i.e.*

$$\overline{\{(\lambda, u) \in \Omega\colon\ \mathfrak{F}(\lambda, u) = 0,\ u \neq 0\}}$$

contains a compact connected subset C *such that*

$$(\lambda_o, 0) \in C \quad \textit{and} \quad C \cap \partial_r \neq \emptyset, \quad C \subset \mathbb{R}^2_+ \times W^{H_o},$$

$(\lambda_o = \alpha_o + \mathrm{i}\beta_o)$ *which, in particular, implies that for every* $(\alpha, \beta, u) \in C$, *we have* $G_u \supset H_o$.

(ii) *If, in addition,* (H_o) *is a dominating orbit type in* W, *then there exist at least* $|G/H_o|_{S^1}$ *different branches of periodic solutions to Eq.* (71) *bifurcating from* $(\alpha_o, 0)$ (*with the limit frequency* $l\beta_o$ *for some* $l \in \mathbb{N}$). *Moreover, for each* (α, β, u) *belonging to these branches of* (*nontrivial*) *solutions one has* $(G_u) = (H_o)$ (*considered in the space* W).

REMARK 5.5.

(i) Since the order (10) on the set $\Phi_1^t(G)$ is partial, it is usually the case that there is more than one dominating orbit type in W contributing to the lower estimate of all bifurcating branches of solutions.

(ii) In addition, if there is also a coefficient $n_K \neq 0$ such that (K) is not a dominating orbit type, but $n_H = 0$ for all dominating orbit types (H) such that $(K) < (H)$, then we can also predict the existence of multiple branches by analyzing all the dominating orbit types (H) larger than (K). However, the exact orbit type of these branches (as well as the corresponding estimate) cannot be determined precisely.

Local bifurcation invariant Similarly to the non-equivariant case, the non-zero coefficient n_{H_o} in (116) indicates that in the subspace $\mathbb{R}^2_+ \oplus W^{H_o}$, there exists a branch of solutions to (87), which is “invariant” (i.e. does not disappear) under small *equivariant* perturbations of the map $\mathfrak{F}_\varsigma$. This observation justifies the following

DEFINITION 5.6. Put

$$\omega(\lambda_o) = \omega(\alpha_o, \beta_o) := G\text{-Deg}^t(\mathfrak{F}_\varsigma, \Omega) \tag{117}$$

and call it the *local equivariant topological bifurcation invariant* (or simply *local bifurcation invariant*) of the (bifurcation) point (α_o, β_o).

5.2. *Computation of local bifurcation invariant: Reduction to product formula*

Although the computations of the local bifurcation invariant (117) are, in general, parallel to the non-equivariant case (see Subsections 4.4 and 4.5), several (specifically equivariant) steps have substantial differences, which should be carefully explained.

Linearization and finite-dimensional reduction As in the non-equivariant case, we start with the linearization of the map $\mathfrak{F}_\varsigma$. Namely, using (95)–(97) and the linear map $a(\alpha, \beta)$ defined by (88), construct the map $\tilde{\mathfrak{F}} : \overline{\Omega_1} \to \mathbb{R} \oplus W$ by formula (98). Then,

$$G\text{-Deg}^t(\mathfrak{F}_\varsigma, \Omega) = G\text{-Deg}^t(\tilde{\mathfrak{F}}, \Omega_1).$$

Take Σ defined by (101) and construct W^m and $\mathfrak{F}_2 : \overline{\Omega_1} \to \mathbb{R} \oplus W$ as in Proposition 4.6. Obviously, W^m is G-invariant, so that $\mathfrak{F}_2$ is G-equivariant. Put $\Omega_o := \Omega_1 \cap (\mathbb{C} \oplus W^m)$ and $\mathfrak{F}_o := \mathfrak{F}_2|_{\overline{\Omega_o}}$. Then,

$$G\text{-Deg}^t(\mathfrak{F}_\varsigma, \Omega) = G\text{-Deg}^t(\mathfrak{F}_2, \Omega_1) = G\text{-Deg}^t(\mathfrak{F}_o, \Omega_o).$$

Product formula Put $W_{o*} := \mathbb{C} \oplus \bigoplus_{l=1}^m W_l$ and $\Omega_{o*} := \Omega_o \cap W_{o*}$. Then, without loss of generality, one can assume that $\Omega_o = B \times \Omega_{o*}$, where B is the unit ball in V, and, in addition, $\mathfrak{F}_o = a_0(\lambda_o) \times \mathfrak{F}_{o*}$, where $a_0(\lambda_o) : \overline{B} \to V$ is given by (90), $\lambda_o = \alpha_o + \mathrm{i}\beta_o$, and $\mathfrak{F}_{o*} : \overline{\Omega_{o*}} \to \mathbb{R} \oplus \bigoplus_{l=1}^m W_l$, is given by

$$\mathfrak{F}_{o*}(\lambda, u) = \left(\tilde{\varsigma}(\lambda, u), \bigoplus_{l=1}^{m} a_l\big(\eta(\lambda)\big)u \right), \tag{118}$$

where a_l is given by (91) (see also (88) and (89)).

Then, apply the multiplicativity property of the twisted G-equivariant degree (see Theorem 3.48) to obtain

$$G\text{-Deg}^t(\mathfrak{F}_\varsigma, \Omega) = \Gamma\text{-Deg}\big(a_0(\lambda_o), B\big) \cdot G\text{-Deg}^t(\mathfrak{F}_{o*}, \Omega_{o*}). \tag{119}$$

Computation of $\Gamma\text{-Deg}(a_0(\lambda_o), B)$ Consider the set $\sigma_- := \{\mu \in \sigma(a_0(\lambda_o)) : \mu < 0\}$, where $\sigma(a_0(\lambda_o))$ stands for the (real) spectrum of $a_0(\lambda_o)$ (observe, by the way, that there is a direct connection between σ_- and the set of positive eigenvalues of $D_x f(\alpha_o, 0)$). For every $\mu \in \sigma_-$, denote by $E(\mu) \subset V$ the eigenspace corresponding to μ. Since $E(\mu)$ is Γ-invariant, one can consider its Γ-isotypical decomposition

$$E(\mu) = E_0(\mu) \oplus E_1(\mu) \oplus \cdots \oplus E_r(\mu),$$

where the component $E_i(\mu)$ is modeled on the irreducible Γ-representation $\mathcal{V}_i$, $i = 0, 1, 2, \ldots, r$. Put

$$m_i(\mu) := \dim E_i(\mu)/\dim \mathcal{V}_i, \quad i = 0, 1, 2, \ldots, r. \tag{120}$$

The number $m_i(\mu)$ is called *$\mathcal{V}_i$-multiplicity* of the eigenvalue μ.

By applying the homotopy and multiplicativity properties (cf. Theorem 3.46),

$$\Gamma\text{-Deg}\big(a_0(\lambda_o), B\big) = \prod_{\mu\in\sigma_-} \prod_{i=0}^{r} (\deg_{\mathcal{V}_i})^{m_i(\mu)}. \tag{121}$$

5.3. *Computation of local bifurcation invariant: Reduction to crossing numbers and basic degrees*

In view of formulae (119) and (121), it remains to compute $G\text{-Deg}^t(\mathfrak{F}_{o*}, \Omega_{o*})$.

Isotypical representation of a_l Take the isotypical decomposition (114) and put $a_{j,l}(\lambda) := a_l(\eta(\lambda))|_{V_{j,l}} : V_{j,l} \to V_{j,l}$ ($j = 0, 1, \ldots, s$). Observe that for a fixed λ, the operator $a_{j,l}(\lambda)$ is a Γ-equivariant complex isomorphism, which can be represented (cf. Proposition 2.9) by a complex $n_j \times n_j$-matrix, where $n_j := \dim V_{j,l}/\dim \mathcal{V}_{j,l}$ (this number is the same for all $l = 1, 2, \ldots$). Put (cf. (113))

$$\Delta_{\alpha,j}(\lambda) := \Delta_\alpha(\lambda)|_{U_j} : U_j \to U_j. \tag{122}$$

Keeping in mind (91), denote by

$$\Delta^{\mathbb{C}}_{\alpha,j}(\lambda) : \mathbb{C}^{n_j} \to \mathbb{C}^{n_j}, \tag{123}$$

the $n_j \times n_j$-matrix representation of the operator $\Delta_{\alpha,j}(\lambda) : U_j \to U_j$ (cf. Proposition 2.9). Then,

$$a_{j,l}(\lambda) = \frac{1}{\mathrm{i}l\beta} \Delta^{\mathbb{C}}_{\alpha,j}(\mathrm{i}l\beta) \quad \text{with } \lambda = \alpha + \mathrm{i}\beta. \tag{124}$$

Isotypical deformations The obtained identification (124), (123) of $a_{j,l}(\cdot)$ and the fact that the linear group $GL^G(V_{j,l})$ is isomorphic to the complex linear group $GL(n_j, \mathbb{C})$, can be used to deform $a_{j,l}(\cdot)$ in a way similar to the non-equivariant case. Namely, put $\tilde{a}_{j,l}(\cdot) := a_{j,l}|_\Sigma$ (cf. (101)). Then, by Proposition 3.35, the homotopy class of the map $\tilde{a}_{j,l} : \Sigma \to GL(n_j, \mathbb{C})$ contains a representative $b_{j,l} : \Sigma \to GL(n_j, \mathbb{C})$ of the following form:

$$b_{j,l}(\lambda) = \begin{bmatrix} (\lambda - \lambda_o)^{\mathfrak{k}_{j,l}} & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} : \mathbb{C}^{n_j} \to \mathbb{C}^{n_j}, \tag{125}$$

where $\mathfrak{k}_{j,l}$ is an integer defined by

$$\mathfrak{k}_{j,l} = \deg(\det_{\mathbb{C}} \Delta^{\mathbb{C}}_{\alpha,j}, \widetilde{D}) \tag{126}$$

(cf. (103), (107) and (108)).

Using the identification (124), (123) and formula (125), define the map $\tilde{\mathfrak{F}}_{o*} : \overline{\Omega_{o*}} \to \mathbb{R} \oplus \bigoplus_{l=1}^{m} W_l$ by

$$\tilde{\mathfrak{F}}_{o*}(\lambda, v) := \left(\tilde{\varsigma}(\lambda, v), \bigoplus_{l=1}^{m} b_{j,l}\big(\eta(\lambda)\big) v_l \right) \tag{127}$$

(here v_l is an orthogonal projection of $v \in \bigoplus_{l=1}^{m} W_l$ to W_l). Combining formula (127) with Proposition 3.45 yields:

$$G\text{-Deg}^t(\mathfrak{F}_{o*}, \Omega_{o*}) = G\text{-Deg}^t(\tilde{\mathfrak{F}}_{o*}, \Omega_{o*}) = \sum_{j,l} \mathfrak{k}_{j,l} \deg_{\mathcal{V}_{j,l}}. \tag{128}$$

In this way (cf. (119), (121) and (128)), we have proved the following

PROPOSITION 5.7. *Let* $\mathfrak{F}_{\varsigma}$ *and* Ω *be as in Theorem* 5.4. *Take the isotypical decompositions* (113) *and* (114). *Then,*

$$G\text{-Deg}^t(\mathfrak{F}_{\varsigma}, \Omega) = \prod_{\mu \in \sigma_-} \prod_{i=0}^{r} (\deg_{\mathcal{V}_i})^{m_i(\mu)} \cdot \sum_{j,l} \mathfrak{k}_{j,l} \deg_{\mathcal{V}_{j,l}}, \tag{129}$$

where σ_- *denotes the* (*real*) *negative spectrum of the operator* $a_0(\lambda_o)$ (*see* (90)), $\deg_{\mathcal{V}_i}$ *is defined by* (63), $m_i(\mu)$ *stands for the* $\mathcal{V}_i$*-multiplicity of* μ (*cf.* (120)), $\mathfrak{k}_{j,l}$ *is given by* (126), $\deg_{\mathcal{V}_{j,l}}$ *is defined in* (60) ($j = 0, \ldots, s$ *and* $l = 1, \ldots, m$), *and the multiplication "$\cdot$" is taken in* $A_1^t(\Gamma \times S^1)$.

Crossing numbers To complete our computations of the local bifurcation invariant, we need to express the integers $\mathfrak{k}_{j,l}$ from (129) via the symmetric spectral properties of the operator $D_x f(\alpha, 0)$. To this end, we adopt the concept of crossing numbers to the equivariant setting.

Following the scheme used in the non-equivariant case, assume that $\mathrm{i}l\beta_o$ is a root of the characteristic equation and choose a small neighborhood S of $\mathrm{i}l\beta_o$ in the right half-plane of $\mathbb{C}$ (see (110)) such that for $|\alpha - \alpha_o| \leqslant \varepsilon$, the characteristic roots of Eq. (75) can only leave S through the 'exit' at the point $\mathrm{i}l\beta_o$ for $\alpha = \alpha_o$.

DEFINITION 5.8. Using (113) and the identification (124), (123), put

$$\mathfrak{t}^{\pm}_{j,l}(\alpha_o, \beta_o) := \deg(\det_{\mathbb{C}} \Delta^{\mathbb{C}}_{\alpha_\pm, j}, S)$$

and

$$\mathrm{t}_{j,l}(\alpha_o,\beta_o)=\mathrm{t}^{-}_{j,l}(\alpha_o,\beta_o)-\mathrm{t}^{+}_{j,l}(\alpha_o,\beta_o). \tag{130}$$

The integer (130) is called the $\mathcal{V}_{j,l}$*-isotypical crossing number* at (α_o,β_o) (cf. (112).

REMARK 5.9. The crossing number (130), if it is positive (resp. negative), indicates the algebraic count (with $\mathcal{U}_j$-multiplicities) of roots of the (isotypical) characteristic equation

$$\det\nolimits_{\mathbb{C}}\Delta_{\alpha,j}(\lambda)=0 \tag{131}$$

(i.e. the characteristic equation restricted to the isotypical component U_j of the space V^c (cf. (113)), which leave (resp. enter) the neighborhood S through the point $\mathrm{i}l\beta_o$ as α crosses α_0). For the system (71), the characteristic roots are the eigenvalues of the (complex) matrix $D_x f(\alpha,0)$. Thus, in order to determine the crossing number $\mathrm{t}_{j,l}(\alpha_o,\beta_o)$, one needs to verify what happens to the eigenvalues of $D_x f(\alpha,0)$ when α crosses α_o. For example, if there is one eigenvalue, which exits S through $\mathrm{i}l\beta_o$, and its $\mathcal{U}_j$-multiplicity is k, then $\mathrm{t}_{j,l}(\alpha_o,\beta_o)=k$.

Clearly,

$$\mathrm{t}_{j,l}(\alpha_o,\beta_o)=\mathrm{t}_{j,1}(\alpha_o,l\beta_o). \tag{132}$$

As in the non-equivariant case, a simple Brouwer degree argument leads to the relation

$$\mathfrak{t}_{j,l}=\mathrm{t}_{j,l}(\alpha_o,\beta_o). \tag{133}$$

In this way, by Proposition 5.7, we obtain the following computational formula for the local bifurcation invariant $\omega(\alpha_o,\beta_o)$.

PROPOSITION 5.10. *Under the notations of Proposition* 5.7, *we have*

$$\omega(\alpha_o,\beta_o)=\prod_{\mu\in\sigma_-}\prod_{i=0}^{r}(\deg_{\mathcal{V}_i})^{m_i(\mu)}\cdot\sum_{j,l}\mathrm{t}_{j,l}(\alpha_o,\beta_o)\deg_{\mathcal{V}_{j,l}}, \tag{134}$$

where $\mathrm{t}_{j,l}$ *is given by* (130).

5.4. *Summary of the equivariant degree method*

Theorem 5.4 together with Proposition 5.10 give an effective algorithm to classify symmetric branches of non-trivial periodic solutions to the Γ-symmetric system (77) (and, correspondingly, (71)) bifurcating from an isolated center $(\alpha_o,0)$. For the reader convenience, we sketch below the main steps one should follow.

(a) Take an isolated center $(\alpha_o, 0)$ for the system (71) with $\mathrm{i}\beta_o$ ($\beta_o > 0$) being the corresponding characteristic root (cf. Definition 4.2 and formulae (74) and (75)).

(b) Consider the operator $A := -\frac{1}{\beta_o} D_x f(\alpha_o, 0) : V \to V$, which is clearly Γ-equivariant, and put $\sigma_- := \{\mu \in \sigma(A) \colon \mu < 0\}$, where $\sigma(A)$ stands for the real spectrum of A. Notice that σ_- is composed of $-\frac{1}{\beta_o}\nu$, where ν is a positive eigenvalue of $D_x f(\alpha_o, 0)$.

(c) For every $\mu \in \sigma_-$ (or, equivalently, one can take the corresponding positive eigenvalue of $D_x f(\alpha_o, 0)$), consider the Γ-isotypical decomposition of the eigenspace $E(\mu)$:

$$E(\mu) = E_0(\mu) \oplus E_1(\mu) \oplus \cdots \oplus E_r(\mu),$$

where the component $E_i(\mu)$ is modeled on the irreducible Γ-representation $\mathcal{V}_i$, and evaluate the $\mathcal{V}_i$-multiplicity of μ, i.e. $m_i(\mu) = \dim E_i(\mu) / \dim \mathcal{V}_i$.

(d) Compute the element $\deg_0(\alpha_o, \beta_o) \in A(\Gamma)$ according to the formula:

$$\deg_0(\alpha_o, \beta_o) := \prod_{\mu \in \sigma_-} \prod_{i=0}^{r} (\deg_{\mathcal{V}_i})^{m_i(\mu)}. \tag{135}$$

(e) Find the (complex) Γ-isotypical decomposition of the complexification V^c, i.e.

$$V^c = U_0 \oplus U_1 \oplus \cdots \oplus U_s,$$

where the Γ-isotypical component U_j is modeled on the irreducible Γ-representation $\mathcal{U}_j$, and put $A_j(\alpha) := D_x f(\alpha, 0)|_{U_j}$ (here $D_x f(\alpha, 0)$ is considered as the corresponding complexification). Consider the set of all the purely imaginary eigenvalues of $A_j(\alpha_o)$ which are integer multiples of $\mathrm{i}\beta_o$:

$$\{\mathrm{i}\beta_o, \mathrm{i}k_1\beta_o, \dots, \mathrm{i}k_q\beta_o\},$$

where $k_p > 1$ are integers ($p = 1, \dots, q$).

(f) Compute the element $\deg_*(\alpha_o, \beta_o) \in A_1^t(G)$ defined as the following sum:

$$\deg_*(\alpha_o, \beta_o) = \sum_{j=0}^{s} \sum_{p=1}^{q} \mathfrak{t}_{j,k_p}(\alpha_o, \beta_o) \deg_{\mathcal{V}_{j,k_p}}, \tag{136}$$

where the numbers $\mathfrak{t}_{j,k_p}(\alpha_o, \beta_o) = \mathfrak{t}_{j,1}(\alpha_o, k_p\beta_o)$ are the $\mathcal{V}_{j,k_p}$-isotypical crossing numbers at (α_o, β_o) (cf. (132)). Obviously, all the numbers $\mathfrak{t}_{j,l}(\alpha_o, \beta_o)$ appearing in formula (134) and satisfying the condition $l \neq k_p$ ($p = 1, \dots, q$), are equal to zero.

(g) Compute the local bifurcation invariant $\omega(\alpha_o, \beta_o)$ for the system (71) according to the formula:

$$\omega(\alpha_o, \beta_o) = \deg_0(\alpha_o, \beta_o) \cdot \deg_*(\alpha_o, \beta_o) = \sum_{(H)} n_H(H), \tag{137}$$

where '$\cdot$' denotes the product in the $A(\Gamma)$-module $A_1^t(G)$.

(h) Let $\{(H_1), \dots, (H_n)\}$ be the set of all dominating orbit types in W such that the numbers n_{H_k}, $k = 1, \dots, n$, appearing in (137) are all non-zero. Then, for each $k = 1, \dots, n$, there exist at least $|\frac{G}{H_k}|_{S^1}$ different branches of non-trivial periodic solutions to (77) (with the limit frequency $l\beta_o$ for some $l \in \mathbb{N}$) bifurcating from $(\alpha_o, 0)$. Moreover, for each (α, β, u), belonging to these $|\frac{G}{H_k}|_{S^1}$ branches, one has $(G_u) = (H_k)$ (here, as usual, the symbol $|X|_{S^1}$ stands for the number of S^1-orbits in the set X). In particular, the number

$$\left|\frac{G}{H_1}\right|_{S^1} + \left|\frac{G}{H_2}\right|_{S^1} + \dots + \left|\frac{G}{H_n}\right|_{S^1}$$

gives a lower estimate for the number of branches of non-trivial solutions resulting in the (local) Hopf bifurcation in the system (77). In addition, if in formula (137) one has $n_{H_o} \neq 0$ for some *non-dominating* orbit type (H_o) in W, then Remark 5.5(ii) is applied.

5.5. *Usage of Maple© routines*

The last step of the equivariant degree method is the computation of the exact value of the local bifurcation invariant $\omega(\alpha_o, \beta_o)$ given by (134). Depending on the size of the group Γ, the number of related algebraic computations can be quite substantial and the usage of computer programming seems to be indispensable. We have created several Maple© routines, for selected groups Γ, that are able to complete these tasks, based on the properly prepared spectral data extracted from the characteristic equation (75).

Arranging equivariant spectral data We begin with the information related to the negative spectrum of the operator A (described in (b) and (c)). In order to simplify the computations of the factor (135) in formula (134), one can take advantage of the fact that $(\deg_{\mathcal{V}_i})^2 = (\Gamma)$, which is the unit in the Burnside ring $A(\Gamma)$. Therefore, it is sufficient to consider the numbers

$$\mathfrak{m}_i := \sum_{\mu \in \sigma_-} m_i(\mu), \quad \text{for } i = 0, 1, \dots, r,$$

and introduce the sequence $\{\varepsilon_0, \varepsilon_1, \dots, \varepsilon_r\}$, where

$$\varepsilon_i := \mathfrak{m}_i \bmod 2, \quad i = 0, 1, \dots, r.$$

Folding homomorphism Θ_l Define on generators $(K^{\varphi,k})$ an $A(\Gamma)$-module homomorphism $\Theta_l : A_1^t(\Gamma \times S^1) \to A_1^t(\Gamma \times S^1)$ by

$$\Theta_l\big[(K^{\varphi,k})\big] := (K^{\varphi,lk})$$

Table 2
Isotypical crossing numbers

Representation	Folding 1	Folding 2	…	Folding l
$\mathcal{U}_0$	$\mathfrak{t}_{0,1}(\alpha_o,\beta_o)$	$\mathfrak{t}_{0,2}(\alpha_o,\beta_o)$	…	$\mathfrak{t}_{0,l}(\alpha_o,\beta_o)$
$\mathcal{U}_1$	$\mathfrak{t}_{1,1}(\alpha_o,\beta_o)$	$\mathfrak{t}_{1,2}(\alpha_o,\beta_o)$	…	$\mathfrak{t}_{1,l}(\alpha_o,\beta_o)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathcal{U}_s$	$\mathfrak{t}_{s,1}(\alpha_o,\beta_o)$	$\mathfrak{t}_{s,2}(\alpha_o,\beta_o)$	…	$\mathfrak{t}_{s,l}(\alpha_o,\beta_o)$

(see Subsection 3.6). Then,

$$\deg_{\mathcal{V}_{j,l}} = \Theta_l(\deg_{\mathcal{V}_{j,1}}). \tag{138}$$

The full exact value of the local bifurcation invariant $\omega(\alpha_o,\beta_o)$ (cf. (134)) can be obtained with the usage of the Maple© routines, by applying the following formula

$$\begin{aligned}\omega(\alpha_o,\beta_o) \\ = \sum_l \Theta_l\big[\texttt{showdegree}\,[\Gamma](\varepsilon_0,\varepsilon_1,\dots,\varepsilon_r,\mathfrak{t}_{0,l}(\alpha_o,\beta_o),\mathfrak{t}_{1,l}(\alpha_o,\beta_o),\dots, \\ \mathfrak{t}_{s,l}(\alpha_o,\beta_o))\big]\end{aligned} \tag{139}$$

(see Table 2).

5.6. *Bibliographical remarks*

Symmetric Hopf bifurcation was intensively studied during the last decades. The commonly used methods were based on generic approximation, transversality theory (cf. [51–53]), reduction to normal forms and special singularity theory (see the fundamental monographs [64,67], for the recent developments we refer to [31], see also [30,32,63,66]). The case of the symmetric Hopf bifurcation from the so-called "relative equilibria" was considered in [28,100–102,116,147]). One of the advantages of the singularity theory methods is that they allow to predict stability of the bifurcating branches, while the twisted degree method is not sensitive to stability. However, it is our belief that the latter method is easier to apply (in a standard way), allows to treat "large" groups of symmetries and can be completely computerized. The $\Gamma \times S^1$-equivariant settings for an ODEs system, relevant to our discussion, was studied in [18,19,81,91,93]. The notion of a dominating orbit type (which was introduced in [10]) is commonly known in the literature as "maximal". The Maple© library for the computations of the equivariant degree, was developed by A. Biglands.

6. Symmetric Hopf bifurcation for FDEs

In this section, we apply the equivariant degree method to study the Hopf bifurcation problem in systems of symmetric functional differential equations. Since in this case, the equivariant degree setting is completely parallel to the one that was developed for ordinary differential equations, we only point out the differences that are specific for FDEs and NFDEs. In order to simplify the exposition, we restrict ourselves to Γ being a finite group. In particular, in this case the set $\Phi_1^t(G)$ coincides with $\{(H) \in \mathcal{O}(G)\colon \dim W(H) = 1\}$. We also study a continuation of a symmetric branch of non-constant periodic solutions for FDEs.

Throughout this section V stands for an orthogonal Γ-representation (Γ is finite).

6.1. *Symmetric Hopf bifurcation for FDEs with delay: General framework*

Statement of the problem Given a constant $\tau \geqslant 0$, denote by $C_{V,\tau}$ the Banach space of continuous functions from $[-\tau, 0]$ to V equipped with the usual supremum norm

$$\|\varphi\| := \sup_{-\tau \leqslant \theta \leqslant 0} \bigl|\varphi(\theta)\bigr|, \quad \varphi \in C_{V,\tau}.$$

Given a continuous function $x : \mathbb{R} \to V$ and $t \in \mathbb{R}$, define $x_t \in C_{V,\tau}$ by

$$x_t(\theta) = x(t + \theta), \quad \theta \in [-\tau, 0].$$

Clearly, the Γ-action on V induces a natural isometric Banach representation of Γ on the space $C_{V,\tau}$ with the Γ-action given by

$$(\gamma\varphi)(\theta) := \gamma\bigl(\varphi(\theta)\bigr), \quad \gamma \in \Gamma,\ \theta \in [-\tau, 0].$$

Consider the following one-parameter family of differential delayed equations

$$\dot{x} = f(\alpha, x_t), \tag{140}$$

where $x : \mathbb{R} \to V$ is a differentiable function and $f : \mathbb{R} \oplus C_{V,\tau} \to V$ satisfies the following assumptions:

(A1) f is continuously differentiable;

(A2) f is Γ-equivariant, i.e.

$$f(\alpha, \gamma\varphi) = \gamma f(\alpha, \varphi), \quad \varphi \in C_{V,\tau},\ \alpha \in \mathbb{R},\ \gamma \in \Gamma;$$

(A3) $f(\alpha, 0) = 0$ for all $\alpha \in \mathbb{R}$.

For any $x_o \in V$, we use the same symbol to denote the constant function $x_o(t) \equiv x_o$. Clearly, $(x_o)_t = x_o$ for all $t \in \mathbb{R}$. A point $(\alpha, x_o) \in \mathbb{R} \oplus V$ is said to be a *stationary point* of (140) if $f(\alpha, x_o) = 0$. In particular, by condition (A3), $(\alpha, 0)$ is a stationary point of (140) for all $\alpha \in \mathbb{R}$. Moreover, we say that a stationary point (α, x_o) is *nonsingular* if the restriction of f to the space $\mathbb{R} \oplus V \subset \mathbb{R} \oplus C_{V,\tau}$, still denoted by f, has the derivative $D_x f(\alpha, x_o) : V \to V$ (with respect to $x \in V$), which is an isomorphism. The definition of the Hopf bifurcation for the system (140) occurring at $(\alpha_o, 0)$ is exactly the same as the one given in Subsection 4.1.

In what follows, we consider a fixed stationary point $(\alpha_o, 0)$ of (140). In order to prevent the occurrence of a steady-state bifurcation at $(\alpha_o, 0)$, we assume

(A4) The stationary point $(\alpha_o, 0)$ of (140) is non-singular, i.e.

$$\det\bigl(D_x f(\alpha_o, 0)|_V\bigr) \neq 0.$$

Characteristic equation Let V^c be the complexification of the vector space V and let the Γ-isotypical decompositions of V and V^c be given by (113), where $V_0 = V^\Gamma$, V_j is modeled on the real irreducible Γ-representation $\mathcal{V}_j$, $U_0 = (V^c)^\Gamma$ and U_j is modeled on the complex irreducible Γ-representation $\mathcal{U}_j$.

Let (α, x_o) be a stationary point of (140). The linearization of (140) at (α, x_o) leads to the following *characteristic equation* for the stationary point (α, x_o)

$$\det\nolimits_{\mathbb{C}} \Delta_{(\alpha,x_o)}(\lambda) = 0, \tag{141}$$

where

$$\Delta_{(\alpha,x_o)}(\lambda) := \lambda \,\mathrm{Id} - D_x f(\alpha, x_o)(\mathrm{e}^{\lambda\cdot}\cdot)$$

is a complex linear operator from V^c to V^c, with $(\mathrm{e}^{\lambda\cdot}\cdot)(\theta, x) = \mathrm{e}^{\lambda\theta} x$ and $D_x f(\alpha, x_o)(z \otimes x) = z \otimes D_x f(\alpha, x_o)x$ for $z \otimes x \in V^c$ (cf. [145]). Put $\Delta_\alpha(\lambda) := \Delta_{(\alpha,0)}(\lambda)$.

A solution λ_o to (141) is called a *characteristic root* of (141) at the stationary point (α, x_o). It is clear that (α, x_o) is a non-singular stationary point if and only if 0 is not a characteristic root of (141) at the stationary point (α, x_o).

Recall (see Definition 4.2) that a non-singular stationary point (α, x_o) is called a *center* if (141) permits a purely imaginary root. We call (α, x_o) an *isolated center* if it is the only center in some neighborhood of (α, x_o) in $\mathbb{R} \oplus V$.

By (A2) and (A3), the operator $\Delta_\alpha(\lambda)\colon V^c \to V^c$, $\alpha \in \mathbb{R}$, $\lambda \in \mathbb{C}$, is Γ-equivariant. Consequently, for every isotypical component U_j of V^c, $j = 0, 1, \ldots, s$ (cf. (113)), $\Delta_\alpha(\lambda)(U_j) \subseteq U_j$. Put

$$\Delta_{\alpha,j}(\lambda) := \Delta_\alpha(\lambda)|_{U_j}.$$

Crossing numbers Assume the following condition to be satisfied:

(A5) The stationary point $(\alpha_o, 0)$ is an isolated center for the system (140) such that (141) permits a purely imaginary root $\lambda = \mathrm{i}\beta_o$ with $\beta_o > 0$.

Let λ be a complex root of the characteristic equation $\det_{\mathbb{C}} \Delta_{\alpha_o}(\lambda) = 0$. We use the following notations:

$$E(\lambda) := \ker \Delta_{\alpha_o}(\lambda) \subset V^c,$$

$$E_j(\lambda) := E(\lambda) \cap U_j,$$

$$m_j(\lambda) := \dim_{\mathbb{C}} E_j(\lambda) / \dim_{\mathbb{C}} \mathcal{U}_j.$$

The integer $m_j(\lambda)$ is called the *$\mathcal{U}_j$-multiplicity* of the characteristic root λ.

Define S by formula (110). Under the assumption (A5), the constants $\delta > 0$ and $\varepsilon > 0$ can be chosen so small that the following condition is satisfied:

(*) For every $\alpha \in [\alpha_o - \varepsilon, \alpha_o + \varepsilon]$, if there is a characteristic root $u + \mathrm{i}v \in \partial S$ at $(\alpha, 0)$, then $u + \mathrm{i}v = \mathrm{i}\beta_o$ and $\alpha = \alpha_o$.

Put $\alpha_\pm := \alpha_o \pm \varepsilon$. Note that $\Delta_\alpha(\lambda)$ is analytic in $\lambda \in \mathbb{C}$ and continuous in $\alpha \in [\alpha_-, \alpha_+]$ (see [71]). It follows that $\det_{\mathbb{C}} \Delta_{\alpha_o \pm \varepsilon}(\lambda) \neq 0$ for all $\lambda \in \partial S$. So, the following notation (cf. (123))

$$\mathfrak{t}^{\pm}_{j,1}(\alpha_o, \beta_o) := \deg\big(\det_{\mathbb{C}} \Delta^{\mathbb{C}}_{\alpha_\pm, j}(\cdot), S\big), \tag{142}$$

is well-defined for $0 \leqslant j \leqslant s$ (see (113)).

Following the same lines as in Subsection 5.3 (see Definition 5.8), introduce

DEFINITION 6.1. Let $\mathcal{U}_j$ be a complex Γ-irreducible representation and $\mathcal{V}_{j,l} := {}^l\mathcal{U}_j$ for $l \in \mathbb{N}$. Put

$$\mathfrak{t}_{j,1}(\alpha_o, \beta_o) := \mathfrak{t}^{-}_{j,1}(\alpha_o, \beta_o) - \mathfrak{t}^{+}_{j,1}(\alpha_o, \beta_o), \tag{143}$$

and for any integer $l > 1$,

$$\mathfrak{t}_{j,l}(\alpha_o, \beta_o) := \mathfrak{t}_{j,1}(\alpha_o, l\beta_o). \tag{144}$$

Then, $\mathfrak{t}_{j,l}(\alpha_o, \beta_o)$ is called the *$\mathcal{V}_{j,l}$-isotypical crossing number* of $(\alpha_o, 0)$ corresponding to the characteristic root $\mathrm{i}\beta_o$.

REMARK 6.2. Observe that $\mathfrak{t}_{j,l}(\alpha_o, \beta_o) = 0$ if $\mathrm{i}l\beta_o$ is not a root of (141) (cf. (142) and (143)).

In order to establish the existence of small amplitude periodic solutions bifurcating from the stationary point $(\alpha_o, 0)$, i.e. the occurrence of Hopf bifurcation at the stationary point $(\alpha_o, 0)$, and to associate with $(\alpha_o, 0)$ a *local bifurcation invariant*, we apply the standard steps for the degree-theoretical approach.

Normalization of the period Normalization of the period is obtained by making the change of variable $u(t) = x(\frac{p}{2\pi}t)$ for $t \in \mathbb{R}$. We obtain the following equation, which is equivalent to (140):

$$\dot{u}(t) = \frac{p}{2\pi} f(\alpha, u_{t,2\pi/p}), \tag{145}$$

where $u_{t,2\pi/p} \in C_{V,\tau}$ is defined by

$$u_{t,2\pi/p}(\theta) = u\left(t + \frac{2\pi}{p}\theta\right), \quad \theta \in [-\tau, 0].$$

Evidently, $u(t)$ is a 2π-periodic solution of (145) if and only if $x(t)$ is a p-periodic solution of (140). We can also introduce $\beta := \frac{2\pi}{p}$ into Eq. (145) to obtain

$$\dot{u}(t) = \frac{1}{\beta} f(\alpha, u_{t,\beta}). \tag{146}$$

$\Gamma \times S^1$-equivariant setting in functional spaces We use the standard identification $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ (with $t \leftrightarrow e^{it}$) and the operators L, j, K defined by (81), (82) and (85). Put $\mathbb{R}^2_+ := \mathbb{R} \times \mathbb{R}_+$. Take $(L+K)^{-1} : L^2(S^1; V) \to H^1(S^1; V)$ (cf. Subsection 4.2). It can be easily shown that the map $\mathcal{F} : \mathbb{R}^2_+ \times H^1(S^1; V) \to H^1(S^1; V)$ defined by

$$\mathcal{F}(\alpha, \beta, u) = (L+K)^{-1}\left[Ku + \frac{1}{\beta} N_f\big(\alpha, \beta, j(u)\big)\right], \tag{147}$$

is completely continuous, where $N_f : \mathbb{R}^2_+ \times C(S^1; V) \to L^2(S^1; V)$ is defined by

$$N_f(\alpha, \beta, v)(t) = f(\alpha, v_{t,\beta}), \tag{148}$$

and $e^{it} \in S^1$, $(\alpha, \beta, v) \in \mathbb{R}^2_+ \times C(S^1; V)$.

REMARK 6.3. Notice that, because of the presence of the delay in (140), the map N_f depends on the parameter β, which is the only difference from the setting that was discussed in Subsection 4.2.

Put $W := H^1(S^1; V)$. The space W is an isometric Hilbert $\Gamma \times S^1$-representation of the group $G = \Gamma \times S^1$ (see Subsection 5.1) and the non-linear operator $\mathcal{F}$, defined by (147), is clearly G-equivariant.

REMARK 6.4. Notice that $(\alpha, \beta, u) \in \mathbb{R}^2_+ \times W$ is a 2π-periodic solution of (146) if and only if $u = \mathcal{F}(\alpha, \beta, u)$. Consequently, the occurrence of a Hopf bifurcation at $(\alpha_o, 0)$ for Eq. (140) is equivalent to a bifurcation of 2π-periodic solutions of (146) from $(\alpha_o, \beta_o, 0)$ for some $\beta_o > 0$. On the other hand, if a bifurcation at $(\alpha_o, \beta_o, 0) \in \mathbb{R}^2_+ \times W$ takes place in (146), then we *necessarily* have that the operator $\mathrm{Id} - D_u\mathcal{F}(\alpha_o, \beta_o, 0) : W \to W$ is not an isomorphism, or equivalently, $il\beta_o$ is a purely imaginary characteristic root at $(\alpha_o, 0)$, for some $l \in \mathbb{N}$, i.e. $\det_{\mathbb{C}} \Delta_{\alpha_o}(il\beta_o) = 0$.

Sufficient condition for Hopf bifurcation Bearing in mind Remarks 6.3 and 6.4 and literally follow Subsection 5.1 (see also Subsection 4.3), one can establish

THEOREM 6.5. *Given the system* (140), *assume conditions* (A1)–(A5) *to be satisfied. Take* $\mathcal{F}$ *defined by* (147) *and construct* Ω *according to* (93). *Let* $\varsigma : \overline{\Omega} \to \mathbb{R}$ *be a* G*-invariant auxiliary function* (*see Definition* 4.3) *and let* $\mathfrak{F}_\varsigma$ *be defined by* (94). *Then, a complete analogue of the conclusion of Theorem* 5.4 *is true for the systems* (140) *and* (146).

REMARK 6.6. Following the scheme described in Subsections 5.2 and 5.3, one can establish a complete analogue of formula (134) in the setting relevant to Theorem 6.5.

6.2. *Symmetric Hopf bifurcation for neutral FDEs*

Statement of the problem We use the same notations as in Subsection 6.1.

Consider the following one-parameter family of *neutral equations*:

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[x(t) - b(\alpha, x_t)\big] = f(\alpha, x_t), \tag{149}$$

where $x : \mathbb{R} \to V$ is a continuous function,[2] and $b, f : \mathbb{R} \oplus C_{V,\tau} \to V$ satisfy the following assumptions:

(A1) b, f are continuously differentiable;

(A2) b, f are Γ-equivariant;

(A3) $b(\alpha, 0) = 0$, $f(\alpha, 0) = 0$ for all $\alpha \in \mathbb{R}$.

Also, to prevent the occurrence of the steady-state bifurcation, assume

(A4) $\det(D_x f(\alpha, 0)|_V) \neq 0$ for all $\alpha \in \mathbb{R}$.

In addition, assume that

(L) b satisfies the Lipschitz condition with respect to the second variable, i.e.

$$\exists_\kappa \ 0 \leqslant \kappa < 1, \ \text{s.t.} \ \big\|b(\alpha, \varphi) - b(\alpha, \psi)\big\| \leqslant \kappa \|\varphi - \psi\|_\infty \tag{150}$$

for all $\varphi, \psi \in C_{V,\tau}$, $\alpha \in \mathbb{R}$.

Literally following Section 6.1, one can define the notions of stationary/singular points to (149).

Characteristic equation Let (α, x_o) be a stationary point of (149). The linearization of (149) at (α, x_o) leads to the following *characteristic equation*,

$$\det_{\mathbb{C}} \Delta_{(\alpha, x_o)}(\lambda) = 0, \tag{151}$$

where

$$\Delta_{(\alpha, x_o)}(\lambda) := \lambda\big[\mathrm{Id} - D_x b(\alpha, x_o)(\mathrm{e}^{\lambda \cdot} \cdot)\big] - D_x f(\alpha, x_o)(\mathrm{e}^{\lambda \cdot} \cdot). \tag{152}$$

In what follows, assume that

(A5) The system (149) has an isolated center $(\alpha_o, 0)$ for some $\alpha_o \in \mathbb{R}$, with the corresponding purely imaginary characteristic root $\mathrm{i}\beta_o$, for $\beta_o > 0$.

We associate to the problem (149) the *local bifurcation invariant* and follow the standard steps (similar to those in Section 6.1) of the degree-theoretical treatment of the Hopf bifurcation phenomenon.

Normalization of period By making a change of variable $x(t) = u(\beta t)$ with $\beta := \frac{2\pi}{p}$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[u(t) - b(\alpha, u_{t,\beta})\big] = \frac{1}{\beta} f(\alpha, u_{t,\beta}), \tag{153}$$

---

[2] Formally speaking, we do not require in (149) that $x(t)$ is differentiable, but only $x(t) - b(\alpha, x_t)$ to be continuously differentiable.

where $u_{t,\beta} \in C_{V,\tau}$ is defined by

$$u_{t,\beta}(\theta) = u(t + \beta\theta), \quad \theta \in [-\tau, 0].$$

Evidently, $u(t)$ is a 2π-periodic solution of (153) if and only if $x(t)$ is p-periodic solution of (149).

$\Gamma \times S^1$-equivariant setting in functional spaces Put $W := H^1(S^1; V)$ and let L, j, K and N_f be defined by (81), (82), (85) and (148) respectively. Define $N_b : \mathbb{R}^2_+ \times C(S^1; V) \to L^2(S^1; V)$ by

$$N_b(\alpha, \beta, v)(t) = b(\alpha, v_{t,\beta}).$$

Take $(L + K)^{-1} : L^2(S^1; V) \to W$ and the map $\mathcal{F} : \mathbb{R}^2_+ \times W \to W$ defined by

$$\begin{aligned} \mathcal{F}(\alpha, \beta, u) = {} & (L + K)^{-1}\left[\frac{1}{\beta} N_f(\alpha, \beta, u) + K\big(u - N_b(\alpha, \beta, u)\big)\right] \\ & + N_b(\alpha, \beta, u). \end{aligned} \tag{154}$$

Observe that the map $\mathcal{F}$ is a sum of two maps, where the first map

$$(\alpha, \beta, u) \mapsto (L + K)^{-1}\left[\frac{1}{\beta} N_f(\alpha, \beta, u) + K\big(u - N_b(\alpha, \beta, u)\big)\right],$$

is completely continuous, and the second map $(\alpha, \beta, u) \mapsto N_b(\alpha, \beta, u)$ is a Banach contraction with constant κ ($0 \leqslant \kappa < 1$) (see assumptions (A1) and (L)). Therefore, $\mathcal{F}$ is a condensing map (see Definition 3.54 and Remark 3.55).

Finally, an analogue of Remark 6.4 is applied to Eq. (149).

Sufficient condition for Hopf bifurcation Following the same lines as in Subsections 4.3 and 6.1, define Ω and $\mathfrak{F}_\varsigma$ by (93) and (94) respectively. Obviously, $\mathfrak{F}_\varsigma$ is an Ω-admissible G-equivariant condensing field. Therefore, the standard Nussbaum–Sadovskii extension of the twisted degree (see Definition 3.57, Lemma 3.59 and formula (70)) can be applied to the admissible pair $(\mathfrak{F}_\varsigma, \Omega)$. In a standard way, one can easily establish

THEOREM 6.7. *Given the system* (149), *assume conditions* (A1)–(A5) *and* (L) *to be satisfied. Take $\mathcal{F}$ defined by* (154) *and construct Ω according to* (93). *Let $\varsigma : \overline{\Omega} \to \mathbb{R}$ be a G-invariant auxiliary function* (*see Definition* 4.3) *and let $\mathfrak{F}_\varsigma$ be defined by* (94). *Then, a complete analogue of the conclusion of Theorem* 5.4 *is true for the systems* (149) *and* (153). *In addition, an analogue of Remark* 6.6 *is valid.*

6.3. *Global bifurcation problems*

In this subsection, we apply the twisted degree method to a global Hopf bifurcation problem for a system of symmetric functional differential equations to analyze a continuation of symmetric branches of non-constant periodic solutions. To this end, we translate/adapt conditions (A1)–(A5) from Subsection 6.1 to the abstract setting related to symmetric global Hopf bifurcation occurring in the system (140).

To begin with, assume that the system (140) satisfies conditions (A1)–(A3). Take the Hilbert G-representation $W := H^1(S^1; V)$ and define $\mathcal{F}: \mathbb{R}^2 \oplus W \to W$ by formula (147). Put $\mathfrak{F}(\alpha, \beta, u) := u - \mathcal{F}(\alpha, \beta, u)$. Then, the conditions (A1)–(A3) translate as:

(F1) $\mathfrak{F}$ is a completely continuous G-equivariant field of class C^1 and $\mathfrak{F}(\lambda, 0) = 0$ for all $(\lambda, 0) \in \mathbb{R}^2 \oplus W$.

Put $\bar{\mathfrak{F}} := \mathfrak{F}|_{\mathbb{R}^2 \oplus W^{S^1}} : \mathbb{R}^2 \oplus W^{S^1} \to W^{S^1}$. Then, the condition (A4) translates as:

(F2) $D_w\bar{\mathfrak{F}}(\lambda, 0)$ is an isomorphism from W^{S^1} to W^{S^1} for all $\lambda \in \mathbb{R}^2$.

Condition (A5), being necessary for the occurrence of the (local) Hopf bifurcation (cf. Proposition 4.1), translates as: $D_w\mathfrak{F}(\lambda_o, 0): W \to W$ is not an isomorphism. To study the *global* behavior of branches bifurcating from $(\lambda_o, 0)$, we need the following condition to be satisfied:

(F3) The set $\Lambda := \{\lambda \in \mathbb{R}^2\colon\ D_w\mathfrak{F}(\lambda, 0): W \to W$ is not an isomorphism$\}$ is discrete in $\mathbb{R}^2$.

We are interested in a solution set to the equation

$$\mathfrak{F}(\lambda, w) = 0, \quad (\lambda, w) \in \mathbb{R}^2 \oplus W. \tag{155}$$

In the same way, as it was done in Subsection 4.3, one can define trivial/nontrivial solutions to (155) as well as its bifurcation points. Bearing in mind condition (F3), for each $\lambda_o \in \Lambda$, one can associate the local bifurcation invariant $\omega(\lambda_o)$ (cf. Section 6.1) and the *local bifurcation Theorem 6.5* can be applied.

Below we discuss the global bifurcation phenomenon for Eq. (155). Let $\mathcal{S}$ be the closure of the set of all non-trivial solutions to (155), i.e.

$$\mathcal{S} := \overline{\bigl\{(\lambda, w) \in \mathbb{R}^2 \oplus W\colon\ \mathfrak{F}(\lambda, w) = 0, w \neq 0\bigr\}}. \tag{156}$$

Notice that $(\lambda_o, 0)$ is a bifurcation point of (155) if and only if $(\lambda_o, 0) \in \mathcal{S}$. Take a connected component $\mathcal{C} \subset \mathcal{S}$. If $\mathcal{C}$ contains a bifurcation point $(\lambda_o, 0)$, then it is clearly G-invariant. Notice that, in general, $\mathcal{C}$ may be composed of several orbit types, i.e.

$$\mathcal{C} = \bigcup_{(H)} \mathcal{C}_{(H)},$$

and the global behavior of $\mathcal{C}_{(H)}$ can be different for different orbit types (H), for example some of the branches $\mathcal{C}_{(H)}$ may be bounded, while the other are unbounded.

The result following below can be considered as a *global Hopf bifurcation theorem* for the system (140).

THEOREM 6.8. *Suppose that* $\mathfrak{F}: \mathbb{R}^2 \oplus W \to W$ *satisfies the assumptions* (F1)–(F3) *and let* $\mathcal{C}$ *be a bounded connected component of* $\mathcal{S}$ (*cf.* (156)) *such that* $\mathcal{C} \cap \mathbb{R}^2 \times \{0\} \neq \emptyset$. *Then,*

$$\mathcal{C} \cap \mathbb{R}^2 \times \{0\} = \bigl\{(\lambda_1, 0), (\lambda_2, 0), \dots, (\lambda_N, 0)\bigr\},$$

and

$$\sum_{k=1}^{N} \omega(\lambda_k) = 0.$$

The proof of Theorem 6.8 can be easily obtained by combining the methods developed to prove Theorems 4.4, 5.4 and 6.5 with the standard global bifurcation techniques (see, for instance, [89,93,95,128]).

In order to present global bifurcation results for symmetric systems we will be dealing with in what follows, we need a slightly more general

THEOREM 6.9. *Suppose that* $\mathfrak{F}: \mathbb{R}^2 \oplus W \to W$ *satisfies the assumptions* (F1)–(F3) *and let* $\mathcal{C}^{(H_o)}$ *be a bounded connected component of* $\mathcal{S}^{(H_o)}$ (*cf.* (156)) *such that* $\mathcal{C}^{(H_o)} \cap \mathbb{R}^2 \times \{0\} = \{(\lambda_1, 0), (\lambda_2, 0), \dots, (\lambda_N, 0)\} \neq \emptyset$. *Suppose that*

$$\omega(\lambda_k) = \sum_{(H)} n_H^k(H).$$

Then,

$$\sum_{k=1}^{N} n_{H_o}^k = 0.$$

PROOF. Consider the map

$$\mathfrak{F}^{H_o}: \mathbb{R}^2 \oplus W^{H_o} \to W^{H_o}.$$

Put $G_o = N(H_o)$. Since the map $\mathfrak{F}^{H_o}$ is G_o-equivariant and satisfies similar to (F1)–(F3) conditions, we can define for every $\lambda_o \in \Lambda$ the local bifurcation invariant

$$\omega_o(\lambda_o) = \sum_{(K)} m_K(K) \in A_1^t(G_o)$$

associated with $\mathfrak{F}^{H_o}$. Since $\mathcal{C}^{H_o}$ is a bounded connected component of $\mathcal{S}^{H_o}$, we can apply Theorem 6.8 to obtain

$$\sum_{k=1}^{N} \omega_o(\lambda_k) = 0, \quad \omega_o(\lambda_k) = \sum_{K} m_K^k(K).$$

On the other hand, (H_o) is an orbit type in W^{H_o} and, it is easy to see that $m_{H_o}^k = n_{H_o}^k$ for all $k = 1, 2, \dots, N$. Consequently,

$$\sum_{k=1}^{N} n_{H_o}^k = 0,$$

and the statement follows. □

COROLLARY 6.10. *Suppose as in Theorem* 6.9 that $\mathfrak{F}:\mathbb{R}^2 \oplus W \to W$ *satisfies the assumptions* (F1)–(F3) *and let* $\mathcal{C}^{(H_o)}$ *be a connected component of* $\mathcal{S}^{(H_o)}$ (*cf.* (156)) *such that* $\mathcal{C}^{(H_o)} \cap \mathbb{R}^2 \times \{0\} = \{(\lambda_1, 0)\}$, *where* (H_o) *is an orbit type in* W. *If* $n^1_{H_o} \neq 0$, *then* $\mathcal{C}^{(H_o)}$ *is unbounded.*

REMARK 6.11.

(i) Observe that the conclusions of Theorems 6.8 and 6.9 and Corollary 6.10 are still true if we replace condition (F2) with the following one:

  (F2′) (a) $D_w\bar{\mathfrak{F}}(\lambda, 0)$ is an isomorphism from W^{S^1} to W^{S^1} for all $\lambda = (\alpha, \beta) \neq (\alpha_o, \beta)$, where α_o is a certain fixed number;

  (b) there is no Hopf bifurcation in (140) from the point $(\alpha_o, 0)$.

(ii) It should be pointed out that the global bifurcation results described in this subsection can be easily extended to almost all the classes of parametrized symmetric dynamical systems considered in this paper (as long as the corresponding local result is established).

6.4. *Bibliographical remarks*

The approach outlined in Section 6.1 (resp. Section 6.2) was discussed in [11] (resp. [15]) (see also [19]). The global bifurcation results given in Section 6.3 were published in [13]. Similar systems of symmetric DEs and FDEs were considered in [67,65,66,68,92,94,95, 129,140,145,146,148]. For the singularity theory treatment of the symmetric Hopf bifurcation in NFDEs, we refer to [69]. In the non-symmetric case, local and global Hopf bifurcation problems in FDEs were studied in [36,110,111,124,125,122,123].

7. Symmetric Hopf bifurcation problems for functional parabolic systems of equations

As we have seen in Sections 4–6, the twisted degree provides an effective and direct method for the classification of symmetric Hopf bifurcations in ODEs and FDEs, based on the equivariant spectral properties of the linearized equation at a stationary point. In this section, we present a setting for applications of the twisted degree to study the Hopf bifurcation problem for a continuously parametrized Γ-symmetric Fredholm coincidence equation, which is next associated to a system of non-linear functional parabolic equations with Γ-symmetries.

7.1. *Symmetric bifurcation in parameterized equivariant coincidence problems*

Delayed functional parabolic systems of our interest fall into a category of the so-called equivariant *coincidence* problems. Therefore, we start with a brief discussion of an abstract functional setting for such equivariant coincidence problems. We use the notations introduced in Subsections 2.3 and 2.4 ($G = \Gamma \times S^1$).

Functional setting for equivariant coincidence problems Let $\mathbb{E}$ and $\mathbb{F}$ be real isometric Banach G-representations, $\mathcal{P}$ a topological space equipped with the trivial G-action, and $\{L_\lambda\}_{\lambda\in\mathcal{P}} \subset \mathrm{Op}^G$ a *continuous family of equivariant Fredholm operators of index zero, parametrized by* $\mathcal{P}$, i.e. $L_\lambda \in \mathfrak{F}_0^G$ for each $\lambda \in \mathcal{P}$, and the mapping $\eta : \mathcal{P} \to \mathrm{Op}^G$ defined by $\eta(\lambda) = L_\lambda$ for $\lambda \in \mathcal{P}$, is continuous.

Define ξ by

$$\xi := \big\{(\lambda, u, y) \in \mathcal{P} \times (\mathbb{E} \oplus \mathbb{F})\colon\ u \in \mathrm{Dom}(L_\lambda),\ \ y = L_\lambda u\big\},$$

and consider $\pi : \xi \to \mathcal{P}$, $\pi(\lambda, u, y) = \lambda$ for $(\lambda, u, y) \in \xi$. It has been shown in [50] that $\pi : \xi \to \mathcal{P}$ determines a locally trivial G-vector bundle.

Let

$$\mathcal{E} := \big\{(\lambda, u) \in \mathcal{P} \times \mathbb{E}\colon\ u \in \mathbb{E}_{L_\lambda}\big\} \tag{157}$$

and let $p_1 : \xi \to \mathcal{E}$ be given by $p_1(\lambda, u, y) = (\lambda, u)$, for $(\lambda, u, y) \in \xi$. Since, for every $\lambda \in \mathcal{P}$, the projection $pr_1 : \mathrm{Gr}(L_\lambda) \to \mathbb{E}_\lambda := \mathbb{E}_{L_\lambda}$ is an equivariant isometry, the mapping $p_1 : \xi \to \mathcal{E}$ gives us the natural identification of the G-bundles ξ and $\mathcal{E}$ (we use the same symbol for a bundle and its total space).

Define the vector bundle morphism $L : \mathcal{E} \to \mathbb{F}$, where $\mathbb{F}$ is viewed as a bundle over a one-point space, by

$$L(\lambda, u) = L_\lambda u, \quad (\lambda, u) \in \mathcal{E}. \tag{158}$$

DEFINITION 7.1. Let X be a subset of $\mathcal{P}$ and let L be given by (157) and (158). An *equivariant resolvent* of L over X is a G-vector bundle morphism $K : X \times \mathbb{E} \to \mathbb{F}$ such that

(i) for every $\lambda \in X$, $K_\lambda : \mathbb{E} \to \mathbb{F}$ is a finite-dimensional linear operator;
(ii) for every $\lambda \in X$, $L_\lambda + K_\lambda : \mathbb{E}_\lambda \to \mathbb{F}$ is an isomorphism.

Denote by $\mathcal{R}^G(L, X)$ the set of all equivariant resolvents of L over X.

REMARK 7.2.

(i) In contrast to the non-equivariant case, given $\lambda_o \in \mathcal{P}$, one may have $\mathcal{R}^G(L, \{\lambda_o\}) = \emptyset$.
(ii) Also, in general it might happen that $\mathcal{R}^G(L, X) = \emptyset$, while $\mathcal{R}^G(L, \{\lambda_o\}) \neq \emptyset$ for each $\lambda_o \in X$.

In the light of Remark 7.2, the following result turns out to be useful.

LEMMA 7.3. (Cf. [90].) *Let $X \subset \mathcal{P}$ be a compact contractible set containing a point λ^* such that $\mathcal{R}^G(L, \{\lambda^*\}) \neq \emptyset$. Then, $\mathcal{R}^G(L, X) \neq \emptyset$.*

Assume the following condition to be satisfied:

(H1) There exists a compact subset $X \subset \mathcal{P}$ such that $\mathcal{R}^G(L, X) \neq \emptyset$.

Fix $K \in \mathcal{R}^G(L, X)$ and put

$$R_\lambda := (L_\lambda + K_\lambda)^{-1} \tag{159}$$

(cf. Definition 7.1(ii)).

Given a completely continuous G-equivariant map $F : \mathcal{E} \to \mathbb{F}$, consider the associated *parametrized equivariant coincidence problem* (cf. [90])

$$L_\lambda u = F(\lambda, u), \quad (\lambda, u) \in \mathcal{E}|_{X \times \mathrm{Dom}(L_\lambda)}. \tag{160}$$

Using the resolvent K, one can transform (160) into the following fixed-point problem:

$$y = \mathcal{F}(\lambda, y), \quad (\lambda, y) \in X \times \mathbb{F}, \tag{161}$$

where

$$\mathcal{F}(\lambda, y) = F(\lambda, R_\lambda y) + K_\lambda(R_\lambda y), \quad (\lambda, y) \in X \times \mathbb{F}.$$

By assumption (H1), X is compact, therefore, $\mathcal{F}$ is completely continuous.

Bifurcation invariant for the equivariant coincidence problem In what follows, $\mathbb{E}$, $\mathbb{F}$ stand for isometric Banach G-representations, $\mathcal{P} = \mathbb{R} \times \mathbb{R}_+$ and $\{L_\lambda\}_{\lambda \in \mathcal{P}}$ is a continuous family of G-equivariant Fredholm G-equivariant operators of index zero satisfying condition (H1). Fix $K \in \mathcal{R}^G(L, X)$ with R_λ, $\lambda \in \mathcal{P}$, defined by (159).

Keeping in mind the setting relevant to parametrized parabolic systems discussed in the next subsections, we will specify F from (160), assuming

(H2)
- (i) there exists a real isometric Banach G-representation $\widehat{\mathbb{E}}$ and an injective G-vector bundle morphism $J : \mathcal{E} \to \mathcal{P} \times \widehat{\mathbb{E}}$ such that $J_\lambda := J(\lambda, \cdot)$ is a compact linear operator for every $\lambda \in \mathcal{P}$;
- (ii) there exists an equivariant C^1-map $\widehat{F} : \mathcal{P} \times \widehat{E} \to \mathbb{F}$.

Put

$$F := \widehat{F} \circ J. \tag{162}$$

Obviously, F is G-equivariant and completely continuous.

Consider now the coincidence problem (160) with F defined by (162) (see also condition (H2)). In addition, assume that there exists a two-dimensional submanifold $M \subset \mathcal{P} \times \mathbb{E}^G$ (thought of as a “bifurcation surface”) satisfying the following two conditions:

(H3) M is a subset of the solution set to (160);

(H4) if $(\lambda_o, u_o) \in M$, then there exist open neighborhoods U_{λ_o} of λ_o in $\mathcal{P}$ and U_{u_o} of u_o in $\mathbb{E}^G$, and a C^1-map $\chi : U_{\lambda_o} \to \mathbb{E}^G$ such that

$$M \cap (U_{\lambda_o} \times U_{u_o}) = \mathrm{Gr}(\chi).$$

Bearing in mind condition (H3), $(\lambda, u) \in M$ is called a *trivial solution* to (160) while the other solutions to (160) are said to be *non-trivial solutions*. A point $(\lambda_o, u_o) \in M$ is called a *bifurcation point* if in each neighborhood of (λ_o, u_o), there exists a nontrivial solution to (160). In what follows, we study the existence and multiplicity of branches of non-trivial solutions bifurcating from M, and classify their symmetries.

REMARK 7.4. (i) It is clear that (λ, u) is a solution to (160) if and only if (λ, y) is a solution to the system (161), where $y = (L_\lambda + K_\lambda)u$. Moreover, the set of all trivial solutions to (161) can be represented as

$$\widetilde{M} := \big\{(\lambda, y) \in X \times \mathbb{F}\colon \big(\lambda, R_\lambda(y)\big) \in M\big\}.$$

Then, condition (H4) translates as:

(H4′) if $(\lambda_o, y_o) \in \widetilde{M}$, then there exist open neighborhoods U_{λ_o} of λ_o in $\mathcal{P}$ and U_{y_o} of y_o in $\mathbb{F}^G$ and a C^1-map $\tilde{\chi} : U_{\lambda_o} \to \mathbb{F}^G$ such that

$$\widetilde{M} \cap (U_{\lambda_o} \times U_{y_o}) = \operatorname{Gr}(\tilde{\chi}).$$

(ii) Also, $(\lambda_o, u_o) \in M$ is a bifurcation point of (160) if and only if $(\lambda_o, y_o) \in \widetilde{M}$ is a bifurcation point of (161), where $y_o = (L_{\lambda_o} + K_{\lambda_o})u_o$.

Rewrite (161) as

$$(\pi - \mathcal{F})(\lambda, y) = 0, \quad (\lambda, y) \in X \times \mathbb{F}, \tag{163}$$

where π is the projection map on $\mathbb{F}$, $\pi(\lambda, y) = y$. Clearly, $\pi - \mathcal{F}$ is a G-equivariant completely continuous field of class C^1, and

$$D_y(\pi - \mathcal{F})(\lambda, y) = \operatorname{Id} - \big(D_u F\big(\lambda, R_\lambda(y)\big)R_\lambda + K_\lambda R_\lambda\big).$$

Thus, by assumption (H2) (see also (162)), $D_y(\pi - \mathcal{F})$ is a bounded Fredholm operator of index zero. Since G acts trivially on $X \times \mathbb{F}^G$, for $(\lambda, y) \in X \times \mathbb{F}^G$, $D_y(\pi - \mathcal{F})(\lambda, y)$ is G-equivariant (in particular, it is also G-equivariant for $(\lambda, y) \in \widetilde{M}$). By the implicit function theorem, a *necessary* condition for $(\lambda_o, y_o) \in \widetilde{M}$ to be a bifurcation point, is that the derivative $D_y(\pi - \mathcal{F})(\lambda_o, y_o)$ is not an isomorphism of $\mathbb{F}$. Such a point $(\lambda_o, y_o) \in \widetilde{M}$ is called L*-singular*. An L-singular point (λ_o, y_o) is said to be *isolated*, if it is the only L-singular point in some neighborhood of (λ_o, y_o) in $\widetilde{M}$.

Finally, assume that

(H5) there exists an isolated L-singular point $(\lambda_o, y_o) \in \widetilde{M}$.

In order to associate a local bifurcation invariant $\omega(\lambda_o, u_o) \in A_1^t(G)$ to (160), take a neighborhood $\mathcal{D}_{\lambda_o}$ of (λ_o, y_o) in $\widetilde{M}$ such that (i) (λ_o, y_o) is the only L-singular point in $\mathcal{D}_{\lambda_o}$ and (ii) $\overline{\mathcal{D}}_{\lambda_o} \subset \widetilde{M} \cap (U_{\lambda_o} \times U_{y_o})$ (see (H4′)).

Choose a small number $r > 0$, define

$$U(r) := \big\{(\lambda, y) \in \mathcal{P} \times \mathbb{F}\colon \big(\lambda, \tilde{\chi}(\lambda)\big) \in \mathcal{D}_{\lambda_o},\ \big\|y - \tilde{\chi}(\lambda)\big\| < r\big\}, \tag{164}$$

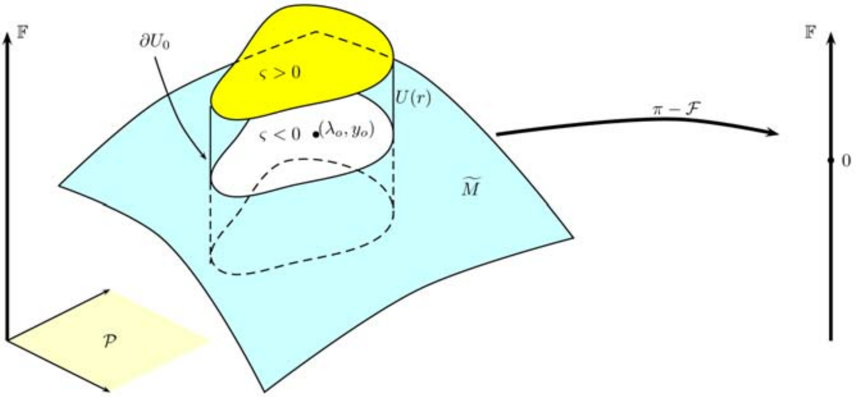


Fig. 5. Functional setting and auxiliary function for the bifurcation problem related to Eq. (163).

and introduce a G-invariant *auxiliary* function $\varsigma : \overline{U(r)} \to \mathbb{R}$ satisfying the properties

$$\begin{cases} \varsigma(\lambda, y) > 0 & \text{if } \|y - \tilde{\chi}(\lambda)\| = r, \\ \varsigma(\lambda, y) < 0 & \text{if } (\lambda, y) \in \mathcal{D}_{\lambda_o}. \end{cases} \tag{165}$$

Put

$$\partial U_0 := \big\{(\lambda, y) \in \overline{U(r)}\colon \big(\lambda, \tilde{\chi}(\lambda)\big) \in \partial \mathcal{D}_{\lambda_o}\big\} \subset \partial U(r).$$

Then, by the implicit function theorem, the above $r > 0$ can be chosen to be so small that

$$y - \mathcal{F}(\lambda, y) \neq 0, \quad \text{for } (\lambda, y) \in \partial U_0 \setminus \widetilde{M}$$

(see Fig. 5).

Define the map $\mathfrak{F}_\varsigma : \overline{U(r)} \to \mathbb{R} \oplus \mathbb{F}$ by

$$\mathfrak{F}_\varsigma(\lambda, y) := \big(\varsigma(\lambda, y), (\pi - \mathcal{F})(\lambda, y)\big), \tag{166}$$

which is clearly a $U(r)$-admissible G-equivariant completely continuous vector field. Therefore, the following *local bifurcation invariant*

$$\omega(\lambda_o, u_o) := G\text{-Deg}^t\big(\mathfrak{F}_\varsigma, U(r)\big) \in A_1^t(G) \tag{167}$$

is well-defined.

Using the standard scheme discussed in Sections 4–6, one can prove the following G-symmetric (local) bifurcation result for the equivariant coincidence problem (160).

THEOREM 7.5 (Local bifurcation theorem). *Suppose that the assumptions* (H1)–(H5) *are satisfied,* $\omega(\lambda_o, u_o)$ *is given by* (167) (*with* $\mathfrak{F}_\varsigma$ *defined by* (166), $U(r)$ *by* (164) *and* ς *satisfying* (165)). *If*

$$\omega(\lambda_o, u_o) = \sum_{(H)} n_H(H) \neq 0,$$

i.e., there is $n_{H_o} \neq 0$ *for some orbit type* (H_o), *then there exists a branch of non-trivial solutions* (λ, u) *to Eq.* (160) *bifurcating from* (λ_o, u_o) *such that* $G_u \supset H_o$.

7.2. *Hopf bifurcation for FPDEs with symmetries: Reduction to local bifurcation invariant*

Throughout this section, assume $\Gamma = \Gamma_1 \times \Gamma_2$, where Γ_1 and Γ_2 are compact Lie groups. Let $V' := \mathbb{R}^m$ (resp. $V := \mathbb{R}^n$) be an orthogonal Γ_1-representation (resp. Γ_2-representation). Assume that $\Omega \subset V'$ is an open Γ_1-invariant bounded set such that $\partial\Omega$ is C^2-smooth. Clearly, the space $L^2(\mathbb{R} \times \overline{\Omega}; V)$ is an isometric Banach Γ-representation with the Γ-action given by

$$(\gamma u)(t, x) = \gamma_2\big(u(t, \gamma_1 x)\big), \quad \gamma = (\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2.$$

Statement of the problem Consider a system of functional parabolic differential equations on $\mathbb{R} \times \overline{\Omega}$

$$\begin{cases} \frac{\partial}{\partial t} u(t, x) + P(\alpha, x)u = f(\alpha, u_t)(x), & (t, x) \in \mathbb{R} \times \Omega, \\ B(\alpha, x)u(t, x) = 0, & (t, x) \in \mathbb{R} \times \partial\Omega, \end{cases} \tag{168}$$

where $u \in L^2(\mathbb{R} \times \overline{\Omega}; V)$ satisfies appropriate differentiability requirements,[3] $u_t(\theta, x) := u(t + \theta, x)$ for $\theta \in [-\tau, 0]$ ($\tau > 0$ is a fixed constant), $\alpha \in \mathbb{R}$ is a (bifurcation) parameter, $f : \mathbb{R} \times C([-\tau, 0]; L^2(\Omega; V)) \to L^2(\Omega; V)$ is a map of class C^1, which is bounded on bounded sets, $P(\alpha, x) = [P_i(\alpha, x)]_{i=1}^n$ is a vector with components being second-order uniformly elliptic operators, i.e.

$$P_i(\alpha, x) = \nabla^T A_i(\alpha, x)\nabla + a_i(\alpha, x),$$

with $A_i(\alpha, x)$ being a continuously differentiable (with respect to α and x) $n \times n$ symmetric positive definite matrix satisfying the condition

$$\exists c_1, c_2 > 0\ \forall (\alpha, x) \in \mathbb{R} \times \overline{\Omega}\ \forall y \in V' \quad c_1\|y\| \leqslant y^T A_i(\alpha, x)y \leqslant c_2\|y\|,$$

---

[3] u is weakly differentiable with respect to t and has weak derivatives of order 2 with respect to $x \in \Omega$. More precisely, assume here that u is an element of the Sobolev space $H^{1,2}(\mathbb{R} \times \Omega; V)$ of L^2-integrable V-valued functions from $\mathbb{R} \times \Omega$ with weak L^2-integrable derivative in $\mathbb{R}$ and weak L^2-integrable derivatives of order 2 in Ω.

where ∇ stands for the gradient operator, and $a_i(\alpha, x)$ is continuous. The boundary operator $B(\alpha, x)$ is defined by either (Dirichlet conditions)

$$B(\alpha, x)u(t, x) = u(t, x)$$

or (mixed Dirichlet/Neumann conditions)

$$B(\alpha, x)u(t, x) = b(\alpha, x)u(t, x) + \frac{\partial}{\partial n}(\alpha, x)\, u(t, x),$$

where $b \in C^1(\mathbb{R} \times \partial\Omega; \mathbb{R})$, $\frac{\partial}{\partial n}(\alpha, x) = [\nu^T(x)A_i(\alpha, x)\nabla]_{i=1}^n$ ($\nu(x)$ is the outward normal vector to $\partial\Omega$ at x).

Assume that

(C1) the operators P, B and the map f are Γ-equivariant, i.e. for $\gamma \in \Gamma$,

$$\begin{aligned}
&\gamma P(\alpha, x)u = P(\alpha, x)\gamma u, && x \in \Omega,\ \theta \in [-\tau, 0],\\
&\gamma f\big(\alpha, v(\theta, x)\big) = f\big(\alpha, \gamma v(\theta, x)\big), && x \in \Omega,\ \theta \in [-\tau, 0],\\
&\gamma B(\alpha, x)u = B(\alpha, x)\gamma u, && x \in \partial\Omega.
\end{aligned}$$

We use the standard identification $S^1 \simeq \mathbb{R}/2\pi\mathbb{Z}$ and introduce the following notation

$$\mathcal{H}^{1,2}_{B(\alpha)} = \big\{\varphi \in H^{1,2}(S^1 \times \Omega; V):\ B(\alpha, x)\varphi = 0\big\}, \tag{169}$$

where $H^{k,\ell}(S^1 \times \Omega; V)$ stands for the Sobolev space of V-valued functions with weak (L^2-integrable) derivatives of order k in S^1 and of order l in Ω. Put

$$\mathbb{E} = \mathbb{F} = L^2(S^1 \times \Omega; V), \quad \mathcal{P} = \mathbb{R} \times \mathbb{R}_+, \quad \widehat{\mathbb{E}} = C\big(S^1; L^2(\Omega; V)\big), \tag{170}$$

where $\widehat{\mathbb{E}}$ is equipped with the usual supremum norm.

We are looking for p-periodic solutions to the system (168) for an unknown period $p > 0$.

Normalization of the period As usual, we replace the original problem with the equivalent one, for which we are looking for 2π-periodic solutions. Namely, let $\beta := \frac{2\pi}{p}$ and

$$v(t, x) = u\left(\frac{1}{\beta}t, x\right).$$

Then, the original problem is reduced to finding non-trivial solutions (α, β, v) for the system

$$\begin{cases}
\frac{\partial}{\partial t}v(t, x) + \frac{1}{\beta}P(\alpha, x)v = \frac{1}{\beta}f(\alpha, v_{t,\beta})(x), & (t, x) \in \mathbb{R} \times \Omega,\\
B(\alpha, x)v(t, x) = 0, & (t, x) \in \mathbb{R} \times \partial\Omega,\\
v(t, x) = v(t + 2\pi, x), & (t, x) \in \mathbb{R} \times \Omega,
\end{cases} \tag{171}$$

where

$$v_{t,\beta}(\theta,x) := v(t+\beta\theta,x) \quad \text{for } (\theta,x)\in[-\tau,0]\times\Omega.$$

$\Gamma\times S^1$*-setting in functional spaces* Below we reformulate the system (171) as an equivariant parametrized coincidence problem based on the general discussion in Subsection 7.1.

For every $\lambda := (\alpha,\beta)\in\mathcal{P}$, define the subspace

$$\operatorname{Dom}(L_\lambda) := \bigl\{u\in\mathbb{E}\colon\ u\in\mathcal{H}^{1,2}_{B(\alpha)}\bigr\},$$

and the operator

$$L_\lambda : \operatorname{Dom}(L_\lambda)\subset\mathbb{E}\to\mathbb{E}$$

(cf. (170)) by

$$L_\lambda v(t,x) := \frac{\partial}{\partial t}v(t,x) + \frac{1}{\beta}P(\alpha,x)v$$

(cf. (169)–(171)).

Notice that $\mathbb{E}$, $H^{1,2}(S^1\times\Omega;V)$ and $\widehat{\mathbb{E}}$ are isometric Banach G-representations, where S^1 acts in a standard way by shifting the time argument t. It is also clear (see [106]) that each (unbounded) linear operator L_λ, for $\lambda\in\mathcal{P}$, is a closed G-equivariant Fredholm operator of index zero, and the orthogonal projection on the (finite-dimensional) kernel of L_λ is a G-equivariant resolvent K of L_λ. Therefore, $R^G(L,\{\lambda\})\neq\emptyset$ for any $\lambda\in\mathcal{P}$, and, by Lemma 7.3, the condition (H1) is satisfied for every compact subset $X\subset\mathcal{P}$.

On the other hand, since

$$\frac{1}{\beta}f(\alpha,v_{t,\beta})\in L^2(\Omega;V) \quad \text{for } v_{t,\beta}\in C\bigl([-\tau,0];L^2(\Omega;V)\bigr),$$

we have the continuous map $N_f:\mathcal{P}\times\widehat{\mathbb{E}}\to L^2(\Omega,V)$ defined by

$$N_f(\alpha,\beta,v)(t) := \frac{1}{\beta}f(\alpha,v_{t,\beta}).$$

Define $\widehat{F}:\mathcal{P}\times\widehat{\mathbb{E}}\to\mathbb{F}$ by

$$\widehat{F}(\lambda,v)(t,x) := i\circ N_f(\alpha,\beta,v)(t)(x) = \frac{1}{\beta}f(\alpha,v_{t,\beta})(x), \quad \lambda=(\alpha,\beta),$$

where i denotes the natural embedding $\widehat{\mathbb{E}}\hookrightarrow\mathbb{F}$. The continuous differentiability of f implies that $\widehat{F}$ is continuously differentiable. Since the following composition of the embeddings

$$H^{1,2}(S^1\times\Omega;V)\hookrightarrow H^{2/3,0}(S^1\times\Omega;V)\hookrightarrow C\bigl(S^1;L^2(\Omega;V)\bigr)=\widehat{\mathbb{E}}$$

is compact (cf. [106]), we have the following embedding

$$J:\mathcal{E}\longrightarrow\mathcal{P}\times\widehat{\mathbb{E}}$$

where $J_\lambda:\mathbb{E}_\lambda\to\widehat{\mathbb{E}}$ is a compact operator for all $\lambda\in\mathcal{P}$. Therefore, $\widehat{F}$ and J satisfy the condition (H2) from Subsection 7.1. In particular, $F:\mathcal{E}\to\mathbb{F}$ defined by $F=\widehat{F}\circ J$ is a G-equivariant completely continuous map of class C^1.

As a consequence, finding a periodic solution $v\in H^{1,2}(S^1\times\Omega;V)$ for the system (171) is equivalent to solving the following parametrized coincidence problem (cf. (160))

$$L_\lambda v=F(\lambda,v),\qquad \lambda\in X,\tag{172}$$

where X is a given compact subset of $\mathcal{P}$.

Γ-symmetric steady-state solutions to (168) In order to describe a manifold of trivial solutions to (171) (cf. conditions (H3) and (H4) from Subsection 7.1), at which we expect the occurrence of a Hopf bifurcation, observe that the functions $u(t,x)$ from $H^{1,2}(S^1\times\Omega;V)$, constant with respect to the first variable, can be identified with functions from $H^2(\Omega;V)$, where $H^2(\Omega;V)$ denotes the Sobolev space of V-valued functions with weak (L^2-integrable) derivatives of order 2 in Ω. Clearly, for a function $u(t,x)$, constant with respect to t, we have $u_t(\theta,x)=u(\theta,x)$ for all $t\in\mathbb{R}$.

With these preliminaries on hands, introduce the following

DEFINITION 7.6. Let (α_o,u_o) be a solution to (168) satisfying the following conditions:

(i) $u_o\in H^2(\Omega,V)$;

(ii) $\gamma u_o=u_0$ for all $\gamma\in\Gamma$;

(iii) $\begin{cases}P(\alpha_o,x)u_o=f(\alpha_o,u_o)(x) & \text{in }\Omega,\\ B(\alpha_o,x)u_o=0 & \text{on }\partial\Omega.\end{cases}$

Then, (α_o,u_o) is called a Γ-*symmetric steady-state* solution to (168).

Introduce the following spaces

$$\begin{aligned}
\mathfrak{X}_{\alpha_o}&:=\big\{\omega\in H^2(\Omega;V):\ B(\alpha_o,x)\omega=0\big\},\\
\mathfrak{X}^c_{\alpha_o}&:=\big\{\omega\in H^2(\Omega;V^c):\ B(\alpha_o,x)\omega=0\big\},\\
\mathfrak{C}&:=C\big([-\tau,0];L^2(\Omega;V)\big),\\
\mathfrak{C}^c&:=C\big([-\tau,0];L^2(\Omega;V^c)\big),
\end{aligned}$$

where V^c stands for the complexification of the Γ-representation V. Identify $L^2(\Omega;V)$ with the subspace of $\mathfrak{C}$ consisting of constant $L^2(\Omega;V)$-valued functions, and $L^2(\Omega;V^c)$ with the subspace of $\mathfrak{C}^c$ consisting of constant $L^2(\Omega;V^c)$-valued functions. Denote by $\bar{f}(\alpha,\cdot)$ the restriction of $f(\alpha,\cdot)$ to $L^2(\Omega;V)$ and define

$$\mathcal{L}_{\alpha_o}:=P(\alpha_o,x)-\mathfrak{d}f_{\alpha_o}:\mathfrak{X}_{\alpha_o}\subset L^2(\Omega;V)\to L^2(\Omega;V),\tag{173}$$

where $\mathfrak{d}f_{\alpha_o}(\varphi):=D_u\bar{f}(\alpha_o,u_o)\varphi$ for $\varphi\in\mathfrak{X}_{\alpha_o}$.

In what follows, we apply the same symbols to denote the complexified operators $P(\alpha, x)$, $\mathfrak{d} f_\alpha$, and $B(\alpha_o, x)$, i.e. the operators:

$$\begin{aligned}
&P(\alpha, x) : \mathfrak{X}^c_\alpha \subseteq L^2(\Omega; V^c) \to L^2(\Omega; V^c),\\
&\mathfrak{d} f_\alpha : \mathfrak{C}^c \to L^2(\Omega; V^c),\\
&B(\alpha_o, x) : H^2(\Omega; V^c) \to L^2(\partial\Omega, V^c).
\end{aligned}$$

Suppose that (α_o, u_o) is a Γ-symmetric steady-state solution to (168) with $\alpha = \alpha_o$. We say that (α_o, u_o) is *non-singular* if $0 \notin \sigma(\mathcal{L}_{\alpha_o})$, where $\sigma(\mathcal{L}(\alpha_o))$ denotes the spectrum of $\mathcal{L}_{\alpha_o}$. Assume that

(C2) (α_o, u_o) is a nonsingular Γ-symmetric steady-state solution.

Then, by the implicit function theorem, there exists a continuously differentiable function $u(\alpha)$ for $\alpha \in (\alpha_o - \eta, \alpha_o + \eta)$ (for a sufficiently small $\eta > 0$) such that $(\alpha, u(\alpha))$ is a Γ-symmetric steady-state solution to (168) for each α (the condition $\gamma u(\alpha) = u(\alpha)$ for all $\gamma \in \Gamma$ (cf. Definition 7.6(ii)) is provided by the Γ-equivariance of $\mathcal{L}_\alpha$). In what follows, assume that $\{(\alpha, u(\alpha)) \colon \alpha \in (\alpha_o - \eta, \alpha_o + \eta)\}$ is a fixed family of steady-state Γ-symmetric solutions near (α_o, u_o). Since $(\alpha, \beta, u(\alpha))$, $\alpha \in (\alpha_o - \eta, \alpha_o + \eta)$, $\beta \in \mathbb{R}_+$, is clearly a solution to (171) belonging to $\mathcal{P} \times \mathbb{E}^G$, consider it as a *trivial solution*. Moreover, define the map $\chi \colon (\alpha_o - \eta, \alpha_o + \eta) \times \mathbb{R}_+ \to \mathbb{E}^G$ by $\chi(\alpha, \beta) = (\alpha, \beta, u(\alpha))$. Then, the set of (nonsingular) Γ-symmetric steady-state solutions to (168) (cf. condition (C2)), gives rise to a manifold $M \subset \mathcal{P} \times \mathbb{E}^G$, $M := \{(\alpha, \beta, u(\alpha)) \colon \alpha \in (\alpha_o - \eta, \alpha_o + \eta),\ \beta \in \mathbb{R}_+\}$ (defined locally) satisfying the conditions (H3) and (H4).

Characteristic equation Let $(\alpha, u(\alpha))$ be a non-singular Γ-symmetric steady-state solution to (168). The linearization of (168) at $(\alpha, u(\alpha))$ leads to the following *characteristic equation* (near $(\alpha_o, u(\alpha_o))$)

$$\Delta_{\alpha; u(\alpha)}(\lambda) w := \lambda w + P(\alpha, x) w - \mathfrak{d} f_\alpha(e^{\lambda \cdot} w) = 0, \quad \lambda \in \mathbb{C}, \tag{174}$$

where the *characteristic operator* $\Delta_{\alpha; u(\alpha)} : \mathfrak{X}^c_\alpha \to L^2(\Omega; V^c)$ is defined (by (174)) using the standard complexifications of $P(\alpha, x)$ and $\mathfrak{d} f_\alpha(\varphi)$.

Observe that $\mathfrak{X}^c_\alpha$ equipped with the H^2-norm is a complex Hilbert space such that the embedding $\mathfrak{X}^c_\alpha \hookrightarrow L^2(\Omega; V^c)$ is compact, $P(\alpha, x)$ is an elliptic self-adjoint operator (thus, it is a bounded Fredholm operator of index zero from $\mathfrak{X}^c_\alpha$ to $L^2(\Omega; V^c)$), and $\mathfrak{d} f_\alpha(e^{\lambda \cdot} \cdot)$ is a bounded linear operator for all $\lambda \in \mathbb{C}$, therefore $\Delta_{\alpha; u(\alpha)}(\lambda) : \mathfrak{X}^c_\alpha \to L^2(\Omega; V^c)$, where $\mathfrak{X}^c_\alpha$ is equipped with the H^2-norm, is a bounded Fredholm operator of index zero.[4] Consequently, $\Delta_{\alpha; u(\alpha)}(\lambda)$ is a closed (unbounded) Fredholm operator of index zero from $L^2(\Omega; V^c)$ to itself.

A number $\lambda \in \mathbb{C}$ is called a *characteristic root* of the system (168) at a Γ-symmetric steady-state solution $(\alpha, u(\alpha))$ if $\ker \Delta_{\alpha; u(\alpha)}(\lambda) \neq \{0\}$. It is clear that a Γ-symmetric steady-state solution $(\alpha, u(\alpha))$ is non-singular if and only if 0 is not a characteristic root of

---

[4] $\Delta_{\alpha; u(\alpha)}(\lambda)$ is a sum of the compact operator $\lambda \,\mathrm{Id} - \mathfrak{d} f_\alpha(e^{\lambda \cdot} \cdot) : \mathfrak{X}^c_\alpha \to L^2(\Omega; V^c)$ and the Fredholm operator $P(\alpha, x)$ of index zero.

(168) at $(\alpha, u(\alpha))$. We say that a non-singular Γ-symmetric steady-state solution (α_o, u_o), $u_o = u(\alpha_o)$, is a *center* if it has a purely imaginary characteristic root $\mathrm{i}\beta_o$, $\beta_o > 0$, i.e. $\ker \Delta_{\alpha_o;u_o}(\mathrm{i}\beta_o) \neq \{0\}$. A center (α_o, u_o) is called *isolated* if it is the only center in some neighborhood of (α_o, u_o) in $\mathbb{R} \oplus L^2(\Omega; V)$.

For the purpose of studying the local Hopf bifurcation problem for (168), assume:

(C3) There exists a Γ-symmetric steady-state solution $(\alpha_o, u_o) \in \mathbb{R} \oplus L^2(\Omega; V)$ to (168), which is an isolated center such that $\mathrm{i}\beta_o$ (for $\beta_o > 0$) is a characteristic root of (168) for $\alpha = \alpha_o$.

Thus, condition (H5) from Subsection 7.1 is satisfied. Also, notice that (C3) is the necessary condition for the occurrence of the Hopf bifurcation at (α_o, u_o), while (C2) excludes the appearance of the "steady-state" bifurcation.

Apply the "eigenspace" reduction to describe the characteristic roots of the system (168) at the Γ-symmetric steady-state solution $(\alpha, u(\alpha))$. Denote by $\sigma_\alpha \subset \mathbb{R}$ the spectrum of the self-adjoint operator $P(\alpha, x) : \mathfrak{X}^c_\alpha \subseteq L^2(\Omega; V^c) \to L^2(\Omega; V^c)$. Since $P(\alpha, x)$ is a uniformly elliptic differential operator, the spectrum σ_α is discrete and all the eigenvalues $\mu^\alpha_{\mathfrak{k}}$ are real of finite multiplicity and such that

$$\mu^\alpha_o < \mu^\alpha_1 < \ldots < \mu^\alpha_{\mathfrak{k}} < \cdots .$$

Using the fact that for every $r > 0$, the number ir is not in the spectrum σ_α of $P(\alpha, x)$, the (auxiliary) operator $S : L^2(\Omega; V^c) \to L^2(\Omega; V^c)$ defined by

$$Sw = \mathrm{i}rw, \quad w \in L^2(\Omega; V^c),$$

is a Γ-equivariant resolvent (cf. Remark and Notation 2.17 and Definition 7.1) of $P(\alpha, x)$, i.e. the (bounded) inverse $\widetilde{R}_{\alpha,r} := [P(\alpha, x) + S]^{-1} : L^2(\Omega; V^c) \to L^2(\Omega; V^c)$ exists for all $\alpha \in \mathbb{R}$ and is Γ-equivariant. On the other hand, since $P(\alpha, x) + S$, as a (bounded) operator from the space $\mathfrak{X}^c_\alpha$ (equipped with the Sobolev H^2-norm) is also invertible and the embedding $\mathfrak{X}^c_\alpha \hookrightarrow L^2(\Omega; V^c)$ is compact, we obtain that the inverse $\widetilde{R}_{\alpha,r} : L^2(\Omega; V^c) \to L^2(\Omega; V^c)$ (i.e. $\widetilde{R}_{\alpha,r}$ is considered here as the inverse of the *unbounded* operator $P(\alpha, x) + S$ from $L^2(\Omega; V^c)$ into itself) is a compact Γ-equivariant operator. Eq. (174) can be rewritten as

$$\tilde{\Delta}^r_{\alpha;u(\alpha)}(\lambda)w := w - \mathfrak{d} f_\alpha\big(\mathrm{e}^{\lambda\cdot}\widetilde{R}_{\alpha,r}(w)\big) + (\lambda - \mathrm{i}r)\widetilde{R}_{\alpha,r}(w) = 0. \tag{175}$$

It is clear that $\lambda \in \mathbb{C}$ is a characteristic root of the system (168) at the Γ-symmetric steady-state solution $(\alpha, u(\alpha))$ if and only if $\ker \tilde{\Delta}^r_{\alpha;u(\alpha)}(\lambda) \neq \{0\}$. Also, $\tilde{\Delta}^r_{\alpha;u(\alpha)}(\lambda)$ is an analytic function in λ (cf. [145]), hence all the characteristic roots λ are isolated. Moreover, $\tilde{\Delta}^r_{\alpha;u(\alpha)}(\lambda)$ is a Γ-equivariant compact field, thus it is a bounded Γ-equivariant Fredholm operator of index zero.

Denote by $E^\alpha_{\mathfrak{k}} \subset L^2(\Omega; V^c)$ the eigenspace of $P(\alpha, x)$ corresponding to $\mu^\alpha_{\mathfrak{k}} \in \sigma_\alpha$, and let $\mathfrak{p}^\alpha_{\mathfrak{k}} : L^2(\Omega; V^c) \to E^\alpha_{\mathfrak{k}}$ be the orthogonal projection. Consequently, for every $w \in$

$L^2(\Omega; V^c)$, we have $w = \sum_{\mathfrak{k}=0}^{\infty} \mathfrak{p}^{\alpha}_{\mathfrak{k}}(w)$. By substituting $w = \sum_{\mathfrak{k}=0}^{\infty} \mathfrak{p}^{\alpha}_{\mathfrak{k}}(w)$ into (175), we obtain

$$\sum_{\mathfrak{k}=0}^{\infty}\left[\mathfrak{p}^{\alpha}_{\mathfrak{k}}(w) - \frac{1}{\mu^{\alpha}_{\mathfrak{k}} + \mathrm{i}r}\mathfrak{d} f_{\alpha}\big(\mathrm{e}^{\lambda\cdot}\mathfrak{p}^{\alpha}_{\mathfrak{k}}(w)\big) + \frac{\lambda - \mathrm{i}r}{\mu^{\alpha}_{\mathfrak{k}} + \mathrm{i}r}\mathfrak{p}^{\alpha}_{\mathfrak{k}}(w)\right] = 0. \tag{176}$$

Denote by $F^{\alpha}_{\mathfrak{k}}$ the subspace of $\mathfrak{C}$ spanned by functions of the type $t \to \varphi(t)w$, where $\varphi \in C([-\tau, 0]; \mathbb{C})$ and $w \in E^{\alpha}_{\mathfrak{k}}$. We need the following hypothesis (cf. [90,119])

(C4) $\mathfrak{d} f_{\alpha}(F^{\alpha}_{\mathfrak{k}}) \subset E^{\alpha}_{\mathfrak{k}}$ for all Γ-symmetric steady-state solutions $(\alpha, u(\alpha))$ and $\mathfrak{k} = 0, 1, 2, \ldots$.

REMARK 7.7. The condition (C4) is required mainly to simplify the computation of the characteristic roots through a reduction to isotypical components of $L^2(\Omega, V^c)$ (see also [117,118]). One can check that the reaction–diffusion systems with delay of the type considered in [34–36] satisfy (C4). In the case of a parabolic system of Γ-symmetric PDEs (without delay), or the reaction–diffusion logistic equation with delay, (C4) is automatically satisfied.

Under the assumption (C4), Eq. (176) can be reduced to the following sequence of equations

$$\mathfrak{p}^{\alpha}_{\mathfrak{k}}(w) - \frac{1}{\mu^{\alpha}_{\mathfrak{k}} + \mathrm{i}r}\mathfrak{d} f_{\alpha}\big(\mathrm{e}^{\lambda\cdot}\mathfrak{p}^{\alpha}_{\mathfrak{k}}(w)\big) + \frac{\lambda - \mathrm{i}r}{\mu^{\alpha}_{\mathfrak{k}} + \mathrm{i}r}\mathfrak{p}^{\alpha}_{\mathfrak{k}}(w) = 0, \quad \mathfrak{k} = 0, 1, \ldots. \tag{177}$$

Equation (177) can be written in the equivalent form as

$$(\mu^{\alpha}_{\mathfrak{k}} + \lambda)\mathfrak{p}^{\alpha}_{\mathfrak{k}}(w) + \mathfrak{d} f_{\alpha}\big(\mathrm{e}^{\lambda\cdot}\rho^{\alpha}_{\mathfrak{k}}(w)\big) = 0, \quad \mathfrak{k} = 0, 1, \ldots. \tag{178}$$

Under the assumptions (C1)–(C4), for any compact subset $X \subset \mathcal{P}$, the system (168) leads to a parametrized equivariant coincidence problem of the type (160) satisfying the conditions (H1)–(H5) (cf. (172)). Hence, given an isolated center (α_o, u_o) with the corresponding characteristic root $i\beta_o$ (cf. condition (C3)), and following the scheme outlined in Subsection 7.1 (cf. (164)–(167)), one can associate to (α_o, β_o, u_o) the local bifurcation invariant $\omega(\lambda_o, u_o) = \omega(\alpha_o, \beta_o, u_o) \in A^t_1(\Gamma \times S^1)$.

7.3. *Computation of local bifurcation invariant*

In order to obtain an effective computational formula for $\omega(\alpha_o, \beta_o, u_o)$, we need to discuss the so-called *negative spectrum* and *crossing numbers*.

Negative spectrum For the sake of simplicity, assume that Γ is a finite group and, moreover, Γ_1 acts trivially on V'. Assume that V and V^c admit the isotypical decompositions (113) with $V_0 = V^{\Gamma}$ and $U_0 = (V^c)^{\Gamma}$. Moreover, these decompositions are required to correspond to a complete list of the irreducible Γ-representations $\mathcal{V}_0, \mathcal{V}_1, \ldots, \mathcal{V}_r$ (resp.

$\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_s$). As in the previous sections, we do not exclude the possibility that some of these Γ-isotypical components are trivial. Clearly, (113) induces the Γ-isotypical decompositions

$$L^2(\Omega;\, V) = \bigoplus_{i=0}^{r} \mathfrak{V}_i, \qquad L^2(\Omega;\, V^c) = \bigoplus_{j=0}^{s} \mathfrak{U}_j, \tag{179}$$

where $\mathfrak{V}_i$ (resp. $\mathfrak{U}_j$) denotes the Γ-isotypical component of $L^2(\Omega;\, V)$ (resp. $L^2(\Omega;\, V^c)$) modeled on $\mathcal{V}_i$ (resp. $\mathcal{U}_j$).

Consider the operator $P(\alpha_o, x) : \mathfrak{X}_{\alpha_o} \subseteq L^2(\Omega;\, V) \to L^2(\Omega;\, V)$ and let K be the orthogonal projection on its kernel. Then, K is a Γ-equivariant resolvent of $P(\alpha_o, x)$. Put $\widetilde{R}_{\alpha_o} := [P(\alpha_o, x) + K]^{-1}$ and define

$$\mathcal{A} := \mathrm{Id} - \frac{1}{\beta_o} \widetilde{R}_{\alpha_o} \circ D_u \bar{f}(\alpha_o, u_o) - \widetilde{R}_{\alpha_o} K : L^2(\Omega;\, V) \to L^2(\Omega;\, V), \tag{180}$$

which is clearly a compact field. Put $\mathcal{A}^i := \mathcal{A}_{|\mathfrak{V}_i} : \mathfrak{V}_i \to \mathfrak{V}_i$ and let $\sigma_-(\mathcal{A})$ denote the set of all negative eigenvalues of the operator $\mathcal{A}$. Since $\mathcal{A}$ is a compact field, it is a Fredholm operator of index 0, the set $\sigma_-(\mathcal{A})$ is finite and all the eigenvalues in $\sigma_-(\mathcal{A})$ are of finite multiplicity. Thus, for $\mu \in \sigma_-(\mathcal{A})$, define

$$\begin{aligned} E(\mu) &:= \bigcup_{k=1}^{\infty} \ker[\mathcal{A} - \mu\, \mathrm{Id}]^k, \\ E_i(\mu) &:= \bigcup_{k=1}^{\infty} \ker[\mathcal{A}^i - \mu\, \mathrm{Id}_{|\mathfrak{V}_i}]^k, \\ m_i(\mu) &:= \dim E_i(\mu) / \dim \mathcal{V}_i, \end{aligned} \tag{181}$$

where $E_i(\mu)$ is referred to as a $\mathcal{V}_i$*-isotypical* eigenspace of the operator $\mathcal{A}$ and the integer $m_i(\mu)$ is called the $\mathcal{V}_i$*-multiplicity* of μ.

Crossing numbers Put $\tilde{\Delta}^r_{\alpha;u(\alpha),j}(\lambda) := \tilde{\Delta}^r_{\alpha;u(\alpha)}(\lambda)|_{\mathfrak{U}_j}$ (cf. (175) and (179)). For a characteristic root λ of the system (168) at the Γ-symmetric steady-state solution (α_o, u_o), we use the following notations:

$$\begin{aligned} E_j(\lambda) &:= \bigcup_{k=1}^{\infty} \ker\big[\tilde{\Delta}^r_{\alpha;u(\alpha),j}(\lambda)\big]^k, \\ m_j(\lambda) &:= \dim E_j(\lambda) / \dim \mathcal{U}_j, \end{aligned} \tag{182}$$

where the integer $m_j(\lambda)$ is called the $\mathcal{U}_j$*-multiplicity* of the characteristic root λ. Notice that, since $\tilde{\Delta}^r_{\alpha;u(\alpha),j}(\lambda)$ is a Fredholm operator of index 0, $m_j(\lambda) < \infty$ for all characteristic roots λ.

Let $(\alpha_o, u_o) \in \mathbb{R} \oplus L^2(\Omega; V)$ be an isolated center with $\mathrm{i}\beta_o$ $(\beta_o > 0)$ being the corresponding characteristic root as described in condition (C3). Define the set

$$\mathcal{S} = \left\{\tau + \mathrm{i}\beta\colon\ 0 < \tau < \delta, |\beta - \beta_o| < \varepsilon\right\} \subset \mathbb{C},$$

where $\delta > 0$ and $\varepsilon > 0$ are so small numbers that for all $\tau + \mathrm{i}\beta \in \partial\mathcal{S}$ and $\alpha \in [\alpha_o - \varepsilon, \alpha_o + \varepsilon]$, $\ker \Delta_{\alpha;u(\alpha)}(\tau + \mathrm{i}\beta) \neq \{0\}$ implies $\alpha = \alpha_o$ and $\tau + \mathrm{i}\beta = \mathrm{i}\beta_o$. Put $\alpha_\pm := \alpha_o \pm \varepsilon$ and denote by $\mathfrak{s}_\pm$ the set of all characteristic roots $\lambda \in \mathcal{S}$ for $\alpha = \alpha_\pm$, i.e.

$$\mathfrak{s}_\pm := \left\{\lambda \in \mathcal{S}\colon\ \ker \Delta_{\alpha_\pm;u(\alpha_\pm)}(\lambda) \neq \{0\}\right\}.$$

Since $\ker \Delta_{\alpha_\pm;u(\alpha_\pm)}(\lambda) = \ker \tilde{\Delta}^r_{\alpha_\pm;u(\alpha_\pm)}(\lambda)$ and $\tilde{\Delta}^r_{\alpha_\pm;u(\alpha_\pm)}(\lambda)$ is an analytic function in λ, the sets $\mathfrak{s}_\pm$ are finite. Then, for $j = 0, 1, 2, \ldots, s$ (corresponding to the complex Γ-irreducible representations $\mathcal{U}_j$), put

$$\mathfrak{t}_j^\pm(\alpha_o, \beta_o, u_o) := \sum_{\lambda \in \mathfrak{s}_\pm} m_j(\lambda) \tag{183}$$

(cf. (182)).

DEFINITION 7.8. The *$\mathcal{V}_{j,l}$-isotypical crossing numbers* of (α_o, β_o, u_o) are defined as follows:

$$\mathfrak{t}_{j,1}(\alpha_o, \beta_o, u_o) := \mathfrak{t}_j^-(\alpha_o, \beta_o, u_o) - \mathfrak{t}_j^+(\alpha_o, \beta_o, u_o), \tag{184}$$

where $\mathfrak{t}_j^\pm(\alpha_o, \beta_o, u_o)$ are given by (183), and in the case $\mathrm{i}l\beta_o$ is also a characteristic root of (168) at (α_o, u_o) for some integer $l > 1$ (cf. Definition 4.8),

$$\mathfrak{t}_{j,l}(\alpha_o, \beta_o, u_o) := \mathfrak{t}_{j,1}(\alpha_o, l\beta_o, u_o).$$

Local bifurcation result By applying the standard steps, one can establish the following formula for the local bifurcation invariant:

$$\begin{aligned}&\omega(\alpha_o, \beta_o, u_o)\\ &\quad := \left(\prod_{\mu \in \sigma_-(\mathcal{A})} \prod_{i=0}^{r} (\deg_{\mathcal{V}_i})^{m_i(\mu)}\right) \cdot \sum_{j,l} \mathfrak{t}_{j,l}(\alpha_o, \beta_o, u_o) \deg_{\mathcal{V}_{j,l}}.\end{aligned} \tag{185}$$

Combining the concept of the dominating orbit types with Theorem 7.5 (see also Remark 5.2) and using the same argument as in Sections 4 and 5 (cf. Theorems 4.9 and 5.4), one can prove the following

THEOREM 7.9. *Suppose that the system* (168) *satisfies the assumption* (C1) *and* (C4), *and suppose that* (α_o, u_o) *is a* Γ*-symmetric steady-state solution to* (168) (*cf. Definition* 7.6) *satisfying* (C2)–(C3), $\omega(\alpha_o, \beta_o, u_o)$ *is given by* (167) (*with* $\lambda_o = (\alpha_o, \beta_o)$, $\mathfrak{F}_\varsigma$ *defined by* (166), $U(r)$ *by* (164) *and* ς *satisfying* (165)). *Assume* (*cf.* (185)) $\omega(\alpha_o, \beta_o, u_o) \neq 0$, *i.e.*

$$\omega(\alpha_o, \beta_o, u_o) = \sum_{(H)} n_H(H) \quad \text{and} \quad n_{H_o} \neq 0 \tag{186}$$

for some $(H_o) \in \Phi_1^t(G)$.

(i) *Then, there exists a branch of non-trivial solutions to* (168) *with symmetries at least* H_o (*considered in the space* $\mathbb{F}$ (*cf.* (170))) *bifurcating from the point* (α_o, u_o) (*with the limit frequency* $l\beta_o$ *for some* $l \in \mathbb{N}$).

(ii) *If, in addition,* (H_o) *is a dominating orbit type in* $\mathbb{F}$, *then there exist at least* $|G/H_o|_{S^1}$ *different branches of periodic solutions to Eq.* (168) *bifurcating from* (α_o, u_o). *Moreover, for each* (α, β, u) *belonging to these branches of* (*nontrivial*) *solutions, one has* $(G_u) = (H_o)$ (*considered in the space* $\mathbb{F}$).

7.4. *Bibliographical remarks*

The general functional framework for abstract coincidence problem for unbounded parametrized Fredholm operators was studied in [90]. The equivariant situation was considered in [16]. For additional references, see [34–36,117,118].

8. Applications

In this section we consider three concrete models of practical meaning, which can be studied using the methods developed in the previous sections.

Throughout this section, V stands for an orthogonal representation of a group Γ which, for the sake of simplicity, is supposed to be one of the following *finite* groups: the dihedral groups D_8, D_{12} and the icosahedral group (i.e. the alternating group of five elements) A_5. The notations used in this section are explained in Appendix A. For all other details, we refer to [19].

8.1. Γ-*symmetric FDEs describing configurations of identical oscillators*

Model Consider a ring of 8 identical cells (for example, being chemical oscillators) coupled symmetrically by diffusion between the adjacent cells. Denote by $x^j(t)$ the concentration of the chemical substance in the j-th cell. The coupling is taking place between adjacent cells connected by the edges of a regular 8-gon. More precisely, the coupling strength between cells is represented by a function $h : \mathbb{R} \to \mathbb{R}$, in general non-linear. Since the coupling strength between the adjacent j-th and i-th cells may be non-linear as well, and depending on the concentrations x^j and x^i, we assume that it is of the form

$$h\big(x^j(t)\big)\big(g(\alpha, x_t^j) - g(\alpha, x_t^i)\big).$$

This term is supported by the ordinary law of diffusion, which simply means that the chemical substance moves from a region of greater concentration to a region of less concentration, at rate proportional to the gradient of the concentration.

The following system of delayed differential equations represents the simplest mathematical model for the above process:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -\alpha x(t) + \alpha H\big(x(t)\big) \cdot C\big(G\big(x(t-1)\big)\big), \tag{187}$$

where

$$x = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^8 \end{bmatrix}, \quad H(x) = \begin{bmatrix} h(x^1) \\ h(x^2) \\ \vdots \\ h(x^8) \end{bmatrix}, \quad G(x) = \begin{bmatrix} g(x^1) \\ g(x^2) \\ \vdots \\ g(x^8) \end{bmatrix},$$

the product '$\cdot$' is defined on the vectors by component-wise multiplication, and

(G1) The functions $h, g : \mathbb{R} \to \mathbb{R}$ are continuously differentiable, $h(t) \neq 0$ for all $t \in \mathbb{R}$, $g(0) = 0$, $g'(0) > 0$ and C is a symmetric 8×8-matrix of the type

$$C = \begin{bmatrix} c & d & 0 & 0 & 0 & 0 & 0 & d \\ d & c & d & 0 & 0 & 0 & 0 & 0 \\ 0 & d & c & d & 0 & 0 & 0 & 0 \\ 0 & 0 & d & c & d & 0 & 0 & 0 \\ 0 & 0 & 0 & d & c & d & 0 & 0 \\ 0 & 0 & 0 & 0 & d & c & d & 0 \\ 0 & 0 & 0 & 0 & 0 & d & c & d \\ d & 0 & 0 & 0 & 0 & 0 & d & c \end{bmatrix} \tag{188}$$

which clearly commutes with an orthogonal D_8-representation.

Therefore, conditions (A1)–(A3) from Subsection 6.1 are satisfied for the system (187).

Characteristic values To verify the remaining conditions (A4)–(A5), consider the linearization of the system (187) at $(\alpha, 0)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -\alpha x(t) + \alpha h(0)g'(0)C\big(x(t-1)\big),$$

and put

$$\eta := h(0)g'(0). \tag{189}$$

Thus, the condition (A4) for the system (187) amounts to

$$\prod_{i=0}^{4}[-\alpha + \alpha\eta\xi_i] \neq 0, \tag{190}$$

where $\xi_0, \xi_1, \xi_2, \xi_3, \xi_4$ are the eigenvalues of the matrix C given by

$$\xi_i := c + 2d\cos\frac{i2\pi}{8}, \quad i = 0, 1, 2, 3, 4, \tag{191}$$

i.e.

$$\begin{aligned} &\xi_0 := c + 2d, && \xi_1 := c + d\sqrt{2}, \\ &\xi_2 := c, && \xi_3 := c - d\sqrt{2}, \\ &\xi_4 := c - 2d, \end{aligned}$$

and the eigenspace $E(\xi_i)$, for $i = 0, 1, 2, 3$ corresponding to the eigenvalue ξ_i is equivalent to the irreducible D_8-representation $\mathcal{V}_i$, while $E(\xi_4)$ is equivalent to $\mathcal{V}_5$ (cf. Appendix A and [19]). Moreover,

$$\Delta_\alpha(\lambda) = (\lambda + \alpha)\,\mathrm{Id} - \alpha\eta e^{-\lambda} C.$$

Therefore, a number $\lambda \in \mathbb{C}$ is a characteristic root of (141) at the stationary point $(\alpha_o, 0)$ if and only if

$$\det\nolimits_{\mathbb{C}} \Delta_\alpha(\lambda) = \prod_{i=1}^{k} [\lambda + \alpha - \alpha\eta\xi_i e^{-\lambda}] = 0. \tag{192}$$

To find a root $\lambda \in \mathbb{C}$ of the system (192), consider the following equation

$$\lambda + \alpha - \alpha\eta\xi_o e^{-\lambda} = 0, \tag{193}$$

where ξ_o is an eigenvalue of C. Obviously, $\xi_o \neq 0$ (otherwise $\lambda = -\alpha \in \mathbb{R}$ cannot be purely imaginary). A similar reason forces

$$\alpha \neq 0. \tag{194}$$

By direct computation, $\lambda = i\beta$ is a purely imaginary root of (193) if

$$\begin{cases} \cos\beta = \dfrac{1}{\eta\xi_o}, \\ \sin\beta = -\dfrac{1}{\alpha\eta\xi_o}\beta. \end{cases} \tag{195}$$

For the sake of simplicity, in our computations of the local bifurcation invariants, we take concrete values of the numbers c, d and η, involved in (195). For example, put

$$c = 2, \quad d = 3, \quad \eta = 6.$$

Then, clearly $|\frac{1}{\xi_i\eta}| < 1$ for $i = 0, 1, 2, 3, 4$, there exists $\beta_i \in (0, \pi]$ such that $\cos\beta_i = \frac{1}{\eta\xi_i}$, and, in addition, it is possible to find a unique $\alpha_i = -\beta_i\cot\beta_i$ (which clearly satisfies conditions (A4) and (A5) from Subsection 6.1). The approximate values of the pairs of solutions (α_o, β_o) to (195) are:

$$\begin{aligned}
(\alpha_0, \beta_0) &= (-0.03229787433, 1.549961486),\\
(\alpha_1, \beta_1) &= (-0.04123910887, 1.544095051),\\
(\alpha_2, \beta_2) &= (-0.1243798137, 1.487366240),\\
(\alpha_3, \beta_3) &= (0.1226042867, 1.645182058),\\
(\alpha_4, \beta_4) &= (0.06724485806, 1.612475059).
\end{aligned}$$

Crossing numbers In order to determine the value of the crossing number associated with a purely imaginary characteristic root $\lambda_o = \mathrm{i}\beta_o$, compute (by implicit differentiation) $\frac{\mathrm{d}}{\mathrm{d}\alpha}u(\alpha)$. By following the same lines as in [19], we obtain that

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}u|_{\alpha=\alpha_o} = \frac{\beta_o^2}{\alpha_o((\alpha_o+1)^2+\beta_o^2)},$$

thus

$$\operatorname{sign}\frac{\mathrm{d}}{\mathrm{d}\alpha}u|_{\alpha=\alpha_o} = \operatorname{sign}\alpha_o.$$

Consider an eigenvalue $\xi_k \in \sigma(C)$. Then, there is exists a purely imaginary characteristic root $\mathrm{i}\beta_k$, $\beta_k > 0$, of the characteristic equation (192) for $\alpha = \alpha_k$, where

$$\cos\beta_k = -\frac{1}{\eta\xi_k}, \qquad \alpha_k = \frac{\beta_k}{\eta\xi_k \sin\beta_k}$$

and consequently,

$$\begin{aligned}
&\text{if } \alpha_k > 0, \qquad \text{then } \mathfrak{t}_{j,1}(\alpha_k, \beta_k) = -m_j(\mathrm{i}\beta_k) = -1;\\
&\text{if } \alpha_k < 0, \qquad \text{then } \mathfrak{t}_{j,1}(\alpha_k, \beta_k) = m_j(\mathrm{i}\beta_k) = 1.
\end{aligned}$$

Notice that in our case, for the chosen values of c, d and η, all the crossing numbers $\mathfrak{t}_{j,l}(\alpha_k, \beta_k), l > 1$, are zero.

Negative spectrum σ_- We have $A := a_0(\alpha_k, \beta_k) = -\alpha_k \operatorname{Id} + \alpha_k h(0)g'(0)C = -\alpha_k \operatorname{Id} + \alpha_k \eta C$ (cf. (89), (90) and Subsection 5.2), so

$$\sigma(A) = \{\mu\colon\ \mu = -\alpha_k + \alpha_k\eta\xi,\ \xi \in \sigma(C)\}.$$

By direct computations, for the bifurcation points (α_k, β_k) we have the following sets σ_-:

$$\begin{aligned}
&(\alpha_0, \beta_0): && \sigma_- = \{-\alpha_0 + \alpha_0\eta\xi_0, -\alpha_0 + \alpha_0\eta\xi_1, -\alpha_0 + \alpha_0\eta\xi_2\},\\
&(\alpha_1, \beta_1): && \sigma_- = \{-\alpha_1 + \alpha_1\eta\xi_0, -\alpha_1 + \alpha_1\eta\xi_1, -\alpha_1 + \alpha_1\eta\xi_2\},\\
&(\alpha_2, \beta_2): && \sigma_- = \{-\alpha_2 + \alpha_2\eta\xi_0, -\alpha_2 + \alpha_2\eta\xi_1, -\alpha_2 + \alpha_2\eta\xi_2\},\\
&(\alpha_3, \beta_3): && \sigma_- = \{-\alpha_3 + \alpha_3\eta\xi_3, -\alpha_3 + \alpha_3\eta\xi_4\},\\
&(\alpha_4, \beta_4): && \sigma_- = \{-\alpha_4 + \alpha_4\eta\xi_3, -\alpha_4 + \alpha_4\eta\xi_4\}.
\end{aligned}$$

Maple© input data Consequently (for the chosen values of c, d and η), we obtain the following input data for the computation of the exact value of the local bifurcation invariants

(α_o, β_o)	$\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6$	$\mathfrak{t}_0, \mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3, \mathfrak{t}_4, \mathfrak{t}_5, \mathfrak{t}_6$
(α_0, β_0)	1,1,1,0,0,0,0	1,0,0,0,0,0,0
(α_1, β_1)	1,1,1,0,0,0,0	0,1,0,0,0,0,0
(α_2, β_2)	1,1,1,0,0,0,0	0,0,1,0,0,0,0
(α_3, β_3)	0,0,0,1,0,1,0	0,0,0,−1,0,0,0
(α_4, β_4)	0,0,0,1,0,1,0	0,0,0,0,0,−1,0

Table of results Finally, based on the fact that (D_8^d), (D_8), $(\widetilde{D}_4^d)$, (D_4^d), $(\mathbb{Z}_8^{t_1})$, $(\mathbb{Z}_8^{t_2})$, $(\mathbb{Z}_8^{t_3})$, $(\widetilde{D}_2^d)$ and (D_2^d) are the dominating orbit types in W, we are able to establish the following table of the local bifurcation invariants for each the bifurcation point (α_i, β_i), $i = 0, 1, 2, 3, 4$:

(α_i, β_i)	$\omega(\alpha_i, \beta_i)$	# of branches
(α_0, β_0)	$-(D_8) + (\widetilde{D}_2) + (D_2) - (\widetilde{D}_1) - (D_1) - (\mathbb{Z}_2) + (\mathbb{Z}_1)$	1
(α_1, β_1)	$-(\mathbb{Z}_8^{t_1}) + (\widetilde{D}_2^d) + (D_2^d) - (\widetilde{D}_1^z) - (\widetilde{D}_1) - (D_1^z) - (D_1) - (\mathbb{Z}_2^-) + 2(\mathbb{Z}_1)$	10
(α_2, β_2)	$-(\widetilde{D}_4^d) - (D_4^d) - (\mathbb{Z}_8^{t_2}) + (\widetilde{D}_2^z) + (\widetilde{D}_2) + (D_2^z) + (D_2) + (\mathbb{Z}_4^d)$ $- (\widetilde{D}_1^z) - (\widetilde{D}_1) - (D_1^z) - (D_1) - 2(\mathbb{Z}_2) + 2(\mathbb{Z}_1)$	6
(α_3, β_3)	$(\mathbb{Z}_8^{t_3}) - (\widetilde{D}_2^d) - (D_2^d) + (\widetilde{D}_1^z) + (\widetilde{D}_1) + (D_1^z) + (D_1) + (\mathbb{Z}_2^-) - 2(\mathbb{Z}_1)$	10
(α_4, β_4)	$-(D_8^d) + (D_4) + (\widetilde{D}_1^z) - (D_1)$	1

8.2. *Hopf bifurcation in symmetric configuration of transmission lines*

In this subsection, we consider two simple types of symmetric configurations for the lossless transmission line models. For all details we refer to [19].

Configuration 1: Internal coupling Consider a cube of symmetrically coupled lossless transmission line networks between two recipients and two power stations. Assume all coupled networks are identical, each of which is a uniformly distributed lossless transmission line with the inductance L_s and parallel capacitance C_s per unit length. To derive

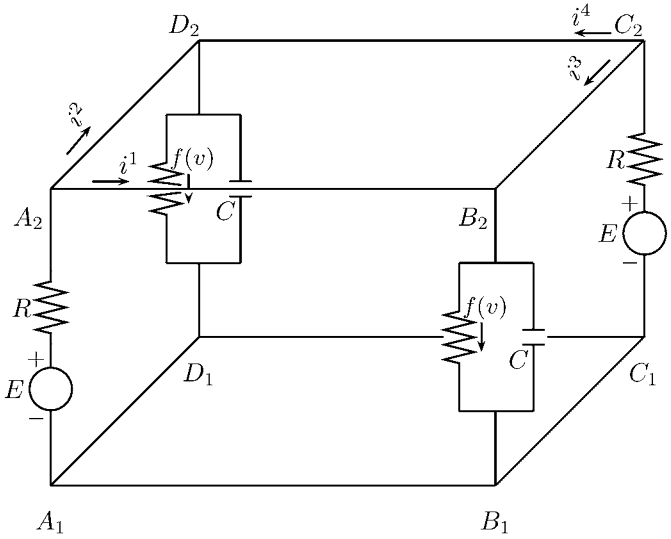


Fig. 6. Symmetric model of transmission lines: internal coupling.

the network equations, place the x-axis in the direction of the line with two ends of the normalized line at $x = 0$ and $x = 1$ (see Fig. 6).[5]

Denote by $i^j(x,t)$ the current flowing in the j-th line at time t and distance x down the line and $v^j(x,t)$ the voltage across the line at t and x, for $j = 1,2,3,4$. It is well-known that (see, for instance, [109]) the functions $i^j := i^j(x,t)$ and $v^j := v^j(x,t)$ obey the *telegrapher's equation*

$$\begin{cases} \dfrac{\partial v^j}{\partial x} = -L_s \dfrac{\partial i^j}{\partial t}, \\ C_s \dfrac{\partial v^j}{\partial t} = -\dfrac{\partial i^j}{\partial x}, \end{cases} \tag{196}$$

with the boundary conditions determined by Kirchoff law

$$\begin{cases} E = v_0^1 + (i_0^1 + i_0^2)R, \\ i_1^1 + i_1^3 = f(v_1^1) + C\dfrac{\mathrm{d}v_1^1}{\mathrm{d}t}, \\ E = v_0^3 + (i_0^3 + i_0^4)R, \\ i_1^2 + i_1^4 = f(v_1^2) + C\dfrac{\mathrm{d}v_1^2}{\mathrm{d}t}, \\ v_0^1 = v_0^2, \quad v_0^3 = v_0^4, \\ v_1^1 = v_1^3, \quad v_1^2 = v_1^4, \end{cases} \tag{197}$$

---

[5] This example of internal coupling can be easily generalized to a coupling of N recipients and N power stations with $N > 2$.

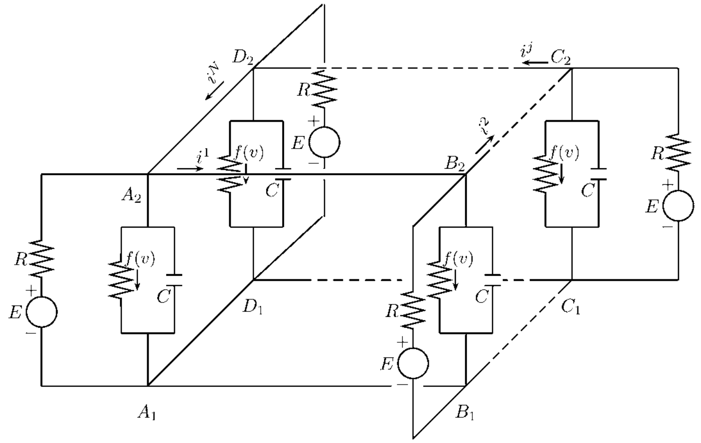


Fig. 7. Symmetric model of transmission lines: external coupling.

where $i^j_\delta = i^j_\delta(t) := i^j(\delta, t)$, $v^j_\delta = v^j_\delta(t) := v^j(\delta, t)$ for $\delta \in \{0, 1\}$, E is the constant direct current voltage and $f(v^j_1)$ is the current through the non-linear resistor in the direction shown in Fig. 6. Assume that

(E1) the boundary value problem (196)–(197) admits a unique solution $(v^j_*, i^j_*) := (v^j_*(x,t), i^j_*(x,t))$, for $j = 1, 2, 3, 4$ such that $\frac{\partial i^j_*}{\partial x} = \frac{\partial v^j_*}{\partial x} = 0$ (the so-called *equilibrium point*).

By applying standard transformations (cf. [19]) we obtain from (196) the following system of neutral functional differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[x(t) - Qx(t-r)\big] = -\frac{2}{bC}x(t) - \frac{2}{bC}Qx(t-r) - \frac{1}{C}G\big(x(t)\big) + \frac{1}{C}QG\big(x(t-r)\big), \tag{198}$$

where

$$r = \frac{2}{a}, \quad x(t) = \begin{bmatrix} v^1_1(t) \\ v^2_1(t) \end{bmatrix}, \quad G\big(x(t)\big) = \begin{bmatrix} g(v^1_1(t)) \\ g(v^2_1(t)) \end{bmatrix},$$
$$Q = \begin{bmatrix} \frac{b}{2R+b} & -\frac{2R}{2R+b} \\ -\frac{2R}{2R+b} & \frac{b}{2R+b} \end{bmatrix}.$$

Configuration 2: External coupling A second example of symmetric coupling was considered in [145], where N recipients are mutually coupled via lossless transmission line network which are interconnected by a common resistor R_o between neighboring recipients, and extensively connected with N power stations (see Fig. 7).

Denote by $i^j(x,t)$ the current flowing in the j-th line at time t and distance x down the line and $v^j(x,t)$ the voltage across the line at t and x, for $j = 1, \dots, N$. The same telegrapher's equation (196) holds for $i^j(x,t)$ and $v^j(x,t)$. However, the boundary conditions need to be modified for this external coupling. For $j = 1, \dots, N$,

$$\begin{cases} E = v_0^j + i_0^j R, \\ i_1^j = f(v_1^j) + C\dfrac{\mathrm{d}v_1^j}{\mathrm{d}t} - \big(I^{j-1}(t) - I^j(t)\big), \\ v_1^j - v_1^{j+1} = I^j(t)R_o, \end{cases} \tag{199}$$

where $I^0(t) := I^N(t)$, $v^{N+1} := v^1$, I^j's are the so-called coupling terms (see [145]).

For mathematical simplicity, assume that (cf. (E1))

(E2) the boundary value problem (196) and (199) admits a unique equilibrium point $(v_*^j, i_*^j) := (v_*^j(x,t), i_*^j(x,t))$, for $j = 1, \dots, N$.

By applying similar standard transformations, the system (196) with the boundary conditions (199) can be translated to

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\big[x(t) - \alpha x(t-r)\big] = -\frac{1}{bC}Px(t) - \frac{1}{bC}\alpha Px(t-r) - \frac{1}{C}G\big(x(t)\big) \\ + \frac{1}{C}\alpha G\big(x(t-r)\big), \end{aligned} \tag{200}$$

where

$$r = \frac{2}{a}, \quad x(t) = \begin{bmatrix} v_1^1(t) \\ \vdots \\ v_1^N(t) \end{bmatrix}, \quad G\big(x(t)\big) = \begin{bmatrix} g(v_1^1(t)) \\ \vdots \\ g(v_1^N(t)) \end{bmatrix},$$

$$\alpha = -\frac{R-b}{R+b}, \quad P = \begin{bmatrix} 1 + \frac{2b}{R_o} & -\frac{b}{R_o} & 0 & \dots & 0 & -\frac{b}{R_o} \\ -\frac{b}{R_o} & 1 + \frac{2b}{R_o} & -\frac{b}{R_o} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{b}{R_o} & 0 & 0 & \dots & -\frac{b}{R_o} & 1 + \frac{2b}{R_o} \end{bmatrix}.$$

Notice that (200) is a D_N-symmetric system in the following sense: consider D_N acting on $V := \mathbb{R}^N$ by permuting the coordinates of vectors

$$x = \begin{bmatrix} v^1 \\ \vdots \\ v^N \end{bmatrix} \in V,$$

then the system (200) is symmetric with respect to the D_N-action on V.

Symmetric configuration of transmission lines Being motivated by the above two generic models of symmetric couplings, we consider a general symmetric system of functional differential equations and discuss several crucial elements related to computations of its associated bifurcation invariant, which are the prerequisite for the usage of our Maple© package.

Combine the two coupling models in the following symmetric neutral functional differential equation

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\bigl[x(t)-\alpha Qx(t-r)\bigr]\\&\quad=-P_1x(t)-\alpha QP_2x(t-r)-aG\bigl(x(t)\bigr)+a\alpha QG\bigl(x(t-r)\bigr)\end{aligned}\tag{201}$$

where a and r are positive constants, $\alpha\in\mathbb{R}$ is the bifurcation parameter, and

$$x(t)=\begin{bmatrix}x^1(t)\\x^2(t)\\\vdots\\x^n(t)\end{bmatrix}\in\mathbb{R}^n,\qquad G\bigl(x(t)\bigr)=\begin{bmatrix}g(x^1(t))\\g(x^2(t))\\\vdots\\g(x^n(t))\end{bmatrix}\in\mathbb{R}^n.$$

For definitiveness, assume $n=12$, $r=1$, and

(H1) $g:\mathbb{R}\to\mathbb{R}$ is continuously differentiable, $g(0)=g'(0)=0$;

(H2) $\Gamma=D_{12}$ acts on $\mathbb{R}^{12}$ by permuting the coordinates of vectors $x\in V:=\mathbb{R}^{12}$, meaning geometrically it permutes the vertices of the 12-gon;

(H3)

(i) Q, P_1, P_2 are 12×12-matrices of the type:

$$C=\begin{bmatrix}c&d&0&0&0&0&0&0&0&0&0&d\\d&c&d&0&0&0&0&0&0&0&0&0\\0&d&c&d&0&0&0&0&0&0&0&0\\0&0&d&c&d&0&0&0&0&0&0&0\\0&0&0&d&c&d&0&0&0&0&0&0\\0&0&0&0&d&c&d&0&0&0&0&0\\0&0&0&0&0&d&c&d&0&0&0&0\\0&0&0&0&0&0&0&0&d&c&d&0\\0&0&0&0&0&0&0&d&c&d&0&0\\0&0&0&0&0&0&0&0&d&c&d&0\\0&0&0&0&0&0&0&0&0&d&c&d\\d&0&0&0&0&0&0&0&0&0&d&c\end{bmatrix}.\tag{202}$$

(ii) for the matrix Q, we choose $c=-\frac{1}{3}$, $d=\frac{1}{4}$; for P_1, $c=\frac{1}{2}$, $d=\frac{1}{3}$; and for P_2, $c=25$, $d=10$;

(H4) we will consider α satisfying $|\alpha|\cdot\|Q\|=\frac{5}{6}|\alpha|<1$.

REMARK 8.1.

(i) It is clear that the system (201) is symmetric with respect to the D_{12}-action on V and $(\alpha,0)$ is a stationary point for all α.

(ii) Q, P_1, P_2 are pairwisely commuting matrices. They can be diagonalized simultaneously. In other words, Q, P_1, P_2 share the same eigenspaces with respect to a certain choice of a basis of V. We use the symbols ξ, ζ and η to denote the eigenvalues of Q, P_1, and P_2 (respectively) corresponding to the same eigenvector $v \in V$.

(iii) One can verify that in the case $\zeta\eta > 0$, ζ and η satisfy $\sqrt{\zeta\eta} \neq \frac{2k+1}{2}\pi$ for any $k \in \mathbb{Z}$.

In this way, we are dealing here with a D_{12}-symmetric system of functional differential equations, and we are interested in studying a Hopf bifurcation phenomenon.

Let us organize all the eigenvalues of the matrices Q, P_1 and P_2 according to their D_{12}-isotypical type (see Appendix A):

k	ξ_k	ζ_k	η_k	Eigenspace
0	$\frac{1}{6}$	$\frac{7}{6}$	45	$\mathcal{V}_0$
1	$-\frac{1}{3}+\frac{1}{4}\sqrt{3}$	$\frac{1}{2}+\frac{1}{3}\sqrt{3}$	$25+10\sqrt{3}$	$\mathcal{V}_1$
2	$-\frac{1}{12}$	$\frac{5}{6}$	35	$\mathcal{V}_2$
3	$-\frac{1}{3}$	$\frac{1}{2}$	25	$\mathcal{V}_3$
4	$-\frac{7}{12}$	$\frac{1}{6}$	15	$\mathcal{V}_4$
5	$-\frac{1}{3}-\frac{1}{4}\sqrt{3}$	$\frac{1}{2}-\frac{1}{3}\sqrt{3}$	$25-10\sqrt{3}$	$\mathcal{V}_5$
6	$-\frac{5}{6}$	$-\frac{1}{6}$	5	$\mathcal{V}_7$

Characteristic equation and isolated centers By linearizing the system (201) at $x = 0$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[x(t) - \alpha Q x(t-1)\big] = -P_1 x(t) - \alpha Q P_2 x(t-1),$$

and substituting $x = \mathrm{e}^{\lambda t} v$ for $\lambda \in \mathbb{C}$, $0 \neq v \in V$, leads to

$$\lambda \mathrm{e}^{\lambda t} v - \alpha Q \lambda \mathrm{e}^{\lambda(t-1)} v = -P_1 \mathrm{e}^{\lambda t} v - \alpha Q P_2 \mathrm{e}^{\lambda(t-1)} v$$

i.e.

$$[\lambda \,\mathrm{Id} - \alpha Q \lambda \mathrm{e}^{-\lambda} + P_1 + \alpha Q P_2 \mathrm{e}^{-\lambda}] v = 0.$$

Therefore, we have the following characteristic equation for the system (201)

$$\det_{\mathbb{C}} \Delta_{(\alpha,0)}(\lambda) = 0, \tag{203}$$

where

$$\Delta_{(\alpha,0)}(\lambda) := (\lambda \,\mathrm{Id} - \alpha Q \lambda \mathrm{e}^{-\lambda}) + P_1 + \alpha Q P_2 \mathrm{e}^{-\lambda}.$$

Table 3
Some of the first bifurcation points (α_o, β_o), their isotypical types and the corresponding negative spectrum σ_-

k	$\mathcal{V}_k$	α_o	β_o	σ_-
0	$\mathcal{V}_0$	0.3134847613	2.044048292	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$
1	$\mathcal{V}_1$	0.5409015727	2.014360183	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$
2	$\mathcal{V}_2$	−0.7179480745	1.924503245	μ_0, μ_1
3	$\mathcal{V}_3$	−0.2206845553	1.774570762	μ_0
4	$\mathcal{V}_4$	−0.1796805216	1.572004370	μ_0
5	$\mathcal{V}_5$	−0.2247337034	1.340355130	μ_0
6	$\mathcal{V}_7$	−0.2821763604	1.197438321	μ_0, μ_1
0	$\mathcal{V}_0$	−0.6602221152	4.841664958	μ_0, μ_1
1	$\mathcal{V}_1$	−1.163028327	4.818957370	μ_0, μ_1
2	$\mathcal{V}_2$	1.638785602†	4.751099765	excluded
3	$\mathcal{V}_3$	0.5502165702	4.636440934	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$
4	$\mathcal{V}_4$	0.4889847314	4.460684627	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$
5	$\mathcal{V}_5$	0.6255618890	4.194073856	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$
6	$\mathcal{V}_7$	0.7498744480	3.996406449	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$
0	$\mathcal{V}_0$	1.039854303	7.829632575	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$
1	$\mathcal{V}_1$	1.837568325†	7.808625800	excluded
2	$\mathcal{V}_2$	−2.607182120†	7.743452649	excluded
3	$\mathcal{V}_3$	−0.8769191325	7.623492141	μ_0, μ_1
4	$\mathcal{V}_4$	−0.7600522375	7.417224874	μ_0, μ_1
5	$\mathcal{V}_5$	−0.8856961981	7.097079227	μ_0, μ_1
6	$\mathcal{V}_7$	−0.9713506495	6.886959465	μ_0, μ_1

†These values of α_o have to be excluded because they are not in the required interval $(-\frac{6}{5}, \frac{6}{5})$.

Next, we need to find possible values of α such that (203) has a purely imaginary root $\mathrm{i}\beta$ for some $\beta > 0$, so we could detect potential bifurcation points $(\alpha, \beta, 0) \in (-\frac{6}{5}, \frac{6}{5}) \times \mathbb{R}_+ \times W$.

By Remark 8.1(ii), when restricted to the same eigenspace of Q, P_1 and P_2, the characteristic equation (203) reduces to the following algebraic equation

$$(\lambda + \zeta)\mathrm{e}^{\lambda} - \alpha\xi(\lambda - \eta) = 0. \tag{204}$$

By replacing in (204) λ with $\mathrm{i}\beta$ and separating the real and imaginary parts, we obtain

$$\tan(\beta) = \frac{\beta(\zeta + \eta)}{\beta^2 - \zeta\eta}, \qquad \alpha = \frac{\zeta \sin\beta + \beta\cos\beta}{\beta\xi}. \tag{205}$$

The system (205) has infinitely many solutions, which can be easily estimated. In Table 3, we present some of the first values of the bifurcation points (α_o, β_o) arranged according to the isotypical type $\mathcal{V}_k$ of the eigenspace $E(\xi_k)$.

Negative spectrum σ_- Consider a bifurcation point (α_o, β_o). The spectrum $\sigma(A(\alpha_o, \beta_o))$ of the linear operator $A(\alpha_o, \beta_o) := -\frac{1}{\beta}_o D_x f(\alpha_o, 0)|_V$ is given by

$$\sigma\big(A(\alpha_o, \beta_o)\big) = \left\{\mu_k = \frac{1}{\beta}_o (\zeta_k + \alpha_o \xi_k \eta_k)\colon\ k = 0, 1, 2, 3, 4, 5, 6\right\}.$$

Therefore, the negative spectrum σ_- of $A(\alpha_o, \beta_o)$ can be obtained easily by inspection (see Table 3). The $\mathcal{V}_i$-multiplicity $m_i(\mu)$ for each $\mu \in \sigma_-$ is 1.

Crossing numbers $\mathfrak{t}_{k,1}$ For a given bifurcation point (α_o, β_o), in order to evaluate the crossing number $\mathfrak{t}_{k,1}(\alpha_o, \beta_o)$, we substitute $\lambda = u + \mathrm{i}v$ in (204), separate the real and imaginary parts, and apply implicit differentiation with respect to α at α_o, $u = 0$, $v = \beta_o$. If $\beta_o > 1$, then one can show (see [19]) that

$$\operatorname{sign}\left(\frac{\mathrm{d}u}{\mathrm{d}\alpha}(\alpha_o)\right) = \operatorname{sign}(\alpha_o).$$

Consequently, if (α_o, β_o) satisfies (205) with $\xi = \xi_k$, $\zeta = \zeta_k$ and $\eta = \eta_k$, for some $k = 0, 1, 2, 3, 4, 5, 6$, then

$$\mathfrak{t}_{k,1}(\alpha_o, \beta_o) = -\operatorname{sign}(\alpha_o).$$

Table of results In Table 4 we list the exact values of the local bifurcation invariants $\omega(\alpha_o, \beta_o)$ for all bifurcation points indicated in Table 3. The dominating orbit types in W are (D_{12}), (D_{12}^d), $(\mathbb{Z}_{12}^{t_1})$, $(\mathbb{Z}_{12}^{t_2})$, $(\mathbb{Z}_{12}^{t_3})$, $(\mathbb{Z}_{12}^{t_4})$, $(\mathbb{Z}_{12}^{t_5})$, $(\widetilde{D}_6^d)$, and (D_6^d).

8.3. *Global continuation of bifurcating branches*

Here, we apply the general results presented Subsection 6.3 to study the global behavior of branches of periodic solutions to the system (187) (with the dihedral symmetries D_8) occurring as a result of the (local) Hopf bifurcation.

By applying the same setting as in Subsection 6.1, the problem of finding periodic solutions to (187) can be reformulated as an equation of type (155), where $\mathfrak{F}$ is given by

$$\mathfrak{F}(\alpha, \beta, u) = u - \mathcal{F}(\alpha, \beta, u), \tag{206}$$

and

$$\mathcal{F}(\alpha, \beta, u)(t) = (L + K)^{-1}\left[-\frac{\alpha}{\beta}\big(u(t) - H\big(u(t)\big)\big)\cdot CG\big(u(t - \beta)\big)\big) + Ku\right]. \tag{207}$$

Throughout this subsection, we assume that the condition (G1) of Subsection 8.1 is satisfied with $c = -3$, $d = 1$, and $\eta := h(0)g'(0) = 2$.

Table 4
Classification of the D_{12}-symmetric Hopf bifurcation in (201) for selected bifurcation points

k	(α_o, β_o)	σ_-	$\omega(\alpha_0, \beta_o)$	#
0	(0.3134847613,2.044048292)	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$	$-(D_{12}) + (D_6) + 2(D_4) + (\widetilde{D}_3)$ $- (D_3) - (\widetilde{D}_2) - (D_2) - (\mathbb{Z}_4)$ $- (\widetilde{D}_1) + (D_1) + (\mathbb{Z}_2)$	1
1	(0.5409015727,2.014360183)	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$	$-(\mathbb{Z}_{12}^{t_1}) + (\mathbb{Z}_6^{t_1}) - (\widetilde{D}_2^d) + (D_2^d)$	6
2	(−0.7179480745,1.924503245)	μ_0, μ_1	$-(\mathbb{Z}_{12}^{t_2}) - (D_4^d) - (D_4^{\hat{d}}) + (\mathbb{Z}_4^d) + (\widetilde{D}_1^z)$ $+ (D_1^z) + (\widetilde{D}_1) + (D_1) - 2(\mathbb{Z}_1)$	4
3	(−0.2206845553,1.774570762)	μ_0	$-(\widetilde{D}_6^d) - (D_6^d) - (\mathbb{Z}_{12}^{t_3}) + (\mathbb{Z}_6^d)$	6
4	(−0.1796805216,1.572004370)	μ_0	$-(\mathbb{Z}_{12}^{t_4}) - (D_4^z) - (D_4) + (\mathbb{Z}_4)$	4
5	(−0.2247337034,1.340355130)	μ_0	$-(\mathbb{Z}_{12}^{t_5}) - (\widetilde{D}_2^d) - (D_2^d) + (\mathbb{Z}_2^-)$	4
6	(−0.2821763604,1.197438321)	μ_0, μ_1	$-(D_{12}^d) + (\widetilde{D}_1^z) + (D_1) - (\mathbb{Z}_1)$	1
0	(−0.6602221152,4.841664958)	μ_0, μ_1	$-(D_{12}) + (\widetilde{D}_1) + (D_1) - (\mathbb{Z}_1)$	1
1	(−1.163028327,4.818957370)	μ_0, μ_1	$-(\mathbb{Z}_{12}^{t_1}) - (\widetilde{D}_2^d) - (D_2^d) + (\widetilde{D}_1^z)$ $+ (D_1^z) + (\widetilde{D}_1) + (D_1) + (\mathbb{Z}_2^-) - 2(\mathbb{Z}_1)$	6
3	(0.5502165702,4.636440934)	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$	$-(\widetilde{D}_6^d) + (D_6^d) - (\mathbb{Z}_{12}^{t_3}) + (\widetilde{D}_3^z)$ $- (D_3^z) + (\widetilde{D}_3) - (D_3) + (\mathbb{Z}_6^d)$ $- (\widetilde{D}_1^z) + (D_1^z) - (\widetilde{D}_1) + (D_1)$	6
4	(0.4889847314,4.460684627)	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$	$-(\mathbb{Z}_{12}^{t_4}) + (D_4^z) + (D_4) + (\mathbb{Z}_6^{t_2})$ $- (\widetilde{D}_2^z) - (\widetilde{D}_2) - (\mathbb{Z}_4) + (\mathbb{Z}_2)$	4
5	(0.6255618890,4.194073856)	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$	$-(\mathbb{Z}_{12}^{t_5}) + (\mathbb{Z}_6^{t_1}) - (\widetilde{D}_2^d) + (D_2^d)$	4
6	(0.7498744480,3.996406449)	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$	$(D_{12}^d) - (D_6) - 2(D_4^d) - (\widetilde{D}_3^z)$ $+ (D_3) + (\widetilde{D}_2^z) + (D_2) + (\mathbb{Z}_4^d)$ $+ (\widetilde{D}_1^z) - (D_1) - (\mathbb{Z}_2)$	1
0	(1.039854303,7.829632575)	$\mu_2, \mu_3, \mu_4, \mu_5, \mu_7$	$-(D_{12}) + (D_6) + 2(D_4) + (\widetilde{D}_3)$ $- (D_3) - (\widetilde{D}_2) - (D_2) - (\mathbb{Z}_4)$ $- (\widetilde{D}_1) + (D_1) + (\mathbb{Z}_2)$	1
3	(−0.8769191325,7.623492141)	μ_0, μ_1	$-(\widetilde{D}_6^d) - (D_6^d) - (\mathbb{Z}_{12}^{t_3}) + (\mathbb{Z}_6^d) + (\widetilde{D}_1^z)$ $+ (D_1^z) + (\widetilde{D}_1) + (D_1) - 2(\mathbb{Z}_1)$	6
4	(−0.7600522375,7.417224874)	μ_0, μ_1	$-(\mathbb{Z}_{12}^{t_4}) - (D_4^z) - (D_4) + (\mathbb{Z}_4) + (\widetilde{D}_1^z)$ $+ (D_1^z) + (\widetilde{D}_1) + (D_1) - 2(\mathbb{Z}_1)$	4
5	(−0.8856961981,7.097079227)	μ_0, μ_1	$-(\mathbb{Z}_{12}^{t_5}) - (\widetilde{D}_2^d) - (D_2^d) + (\widetilde{D}_1^z)$ $+ (D_1^z) + (\widetilde{D}_1) + (D_1) + (\mathbb{Z}_2^-) - 2(\mathbb{Z}_1)$	4
6	(−0.9713506495,6.886959465)	μ_0, μ_1	$-(D_{12}^z) + (\widetilde{D}_1^z) + (D_1^z) - (\mathbb{Z}_1)$	1

Global Hopf bifurcation in a system with dihedral symmetries D_8 The system (187) is symmetric with respect to the dihedral group $\Gamma = D_8$ acting on $V = \mathbb{R}^8$ by permuting the coordinates of vectors. In addition, $\sigma(C)$ is composed of negative numbers: $\xi_0 := -1$, $\xi_1 := -3 + \sqrt{2}$, $\xi_2 := -3$, $\xi_3 := -3 - \sqrt{2}$, $\xi_4 := -5$ and clearly $\eta\xi_o < -1$ for all $\xi_o \in \sigma(C)$. Therefore, for every $\xi_o \in \sigma(C)$, the system of Eqs. (195) has a unique solution (α_o, β_o) such that $\beta_o \in (\frac{\pi}{2}, \pi)$. We are interested in the point (α_4, β_4) (for $\xi_o = \xi_4$), which can be easily approximated, i.e. $(\alpha_4, \beta_4) \simeq (0.1679381755, 1.670963748)$. By applying the same steps as in Subsection 8.1, one can verify that the set σ_- for the point (α_4, β_4) is empty. Since the eigenspace $E(\xi_4)$ is equivalent to $\mathcal{V}_5$, we obtain that the $\mathcal{V}_5$-isotypical crossing number $\mathfrak{t}_{5,1}(\alpha_4, \beta_4) = -1$, which allows us to compute the exact value of the local bifurcation invariant

$$\omega(\alpha_4, \beta_4) = -(D_8^d).$$

Let us point out that (D_8^d) is a dominating orbit type in W. Therefore, by the local bifurcation theorem (cf. Theorem 5.4), there exists a branch $\mathcal{C} := \mathcal{C}_{(D_8^{d,l})}$ of non-constant periodic solutions of the orbit type $(D_8^{d,l})$ (for some $l \geqslant 1$), bifurcating from (α_4, β_4). Notice that $D_8^d \subset D_8^{d,l}$.

Assume in addition, the following conditions to be satisfied

(B1) $tg(t)/h(t) > 0$ for all $t \neq 0$; $\lim_{t\to\infty} tg(t)/h(t) = \infty$.

(B2) There exist constants $A, B > 0$ and $\delta, \gamma > 0$ with $1 > \delta + \gamma$ such that

$$\left|h(t)\right| \leqslant A + B|t|^\delta, \qquad \left|g(t)\right| \leqslant A + B|t|^\gamma.$$

Let us point out, that under the above assumptions one has (cf. Remark 6.11):

(a) if (α, x_o) is a stationary point for the system (187) and $\alpha \neq 0$, then $x_o = 0$;

(b) if a Hopf bifurcation occurs at (α_o, x_o), then $\alpha_o \neq 0$.

Suppose that (α, β, x) belongs to $\mathcal{C}$. Then, the D_8^d-symmetry properties of $x(t)$ can be translated as follows: $x(t) = (x^0(t), x^1(t), \dots, x^7(t))^T$ is a $\frac{2\pi}{\beta}$-periodic solution such that

$$x^k(t) = x^{k+1}\left(t - \frac{\pi}{\beta}\right) \quad (\text{mod } 8), \tag{208}$$

and

$$x^k(t) = x^{n-k-1}\left(t - \frac{\pi}{\beta}\right) \quad (\text{mod } 8). \tag{209}$$

Combining (208), (209) with condition (B1) and applying the same argument as in [94], one can easily show that the periods $p = \frac{2\pi}{\beta}$ of solutions $(\alpha, \beta, x) \in \mathcal{C}$ satisfy the inequality $2 < p < 4$. This fact immediately implies $\mathcal{C} \cap \mathbb{R}^2 \times \{0\} = (\alpha_4, \beta_4, 0)$.

Consequently, $\mathcal{C} \subset \mathbb{R} \times (\pi/2, \pi) \times W$. Further, using the assumption (B2), one can easily show that there exists a constant $M > 0$ such that for every periodic solution $x(t)$ to (187) $\sup\{\|x(t)\|: t \in \mathbb{R}\} \leqslant M$ (see [19]). Thus,

$$\mathcal{C} \subset \mathbb{R} \times (\pi/2, \pi) \times \left\{x \in W: \|x\| \leqslant M\right\}.$$

Finally, the system (187) has no non-constant periodic solution for $\alpha = 0$, from which it follows

$$\mathcal{C} \subset (0, \infty) \times (\pi/2, \pi) \times \bigl\{x \in W\colon \|x\| \leqslant M\bigr\}.$$

However, by Corollary 6.10, the connected component $\mathcal{C}$ is unbounded, therefore

$$[\alpha_{n/2}, \infty) \subset \bigl\{\alpha\colon (\alpha, \beta, x) \in \mathcal{C}\bigr\}.$$

In this way we obtained the following

THEOREM 8.2. *Consider the system* (187) *and suppose the conditions* (G1) *from Subsection* 8.1 (*with* $c = -3$, $d = 1$ *and* $\eta = 2$), (B1) *and* (B2) *are satisfied. Then, the branch* $\mathcal{C} := \mathcal{C}_{D_8^{d,l}}$ *of periodic solutions bifurcating from* $(\alpha_4, \beta_4, 0)$ *is unbounded and satisfies*

$$[\alpha_4, \infty) \subset \bigl\{\alpha\colon (\alpha, \beta, x) \in \mathcal{C}\bigr\}.$$

8.4. *Symmetric system of Hutchinson model in population dynamics*

Hutchinson model of an n species ecosystem The standard model for the dynamics of a simple (single) population (in terms of its density) is the Verhulst equation (cf. [73,70])

$$\dot{v} = \alpha v\left(1 - \frac{v}{K}\right), \tag{210}$$

which is based on the idea that the population grows exponentially at low densities and saturates towards the carrying capacity K (of resources) at high densities. The Hutchinson model (of a single species) is obtained from (210) by taking into account a delayed response to the remaining resources,

$$\dot{v}(t) = \alpha v(t)\left(1 - \frac{v(t-\tau)}{K}\right), \tag{211}$$

where $\tau > 0$ is a presumed delay constant and α refers to the intrinsic growth rate.

Consider an ecosystem composed of n species interacting with each other (according to a certain symmetry) by competing (or cooperating) over shared resources such as food and habitats, while maintaining a self-inhibiting nature (meaning self-limiting in respond to rare resources and self-reproducing to abundant resources). A mathematical treatment for such a community model was developed by Levins in [104], where one attaches a loop diagram in order to carry out a loop analysis for this community type situation.

In such a system (which is illustrated on Fig. 8), the coefficient a_{jj} describes the self-inhibiting nature of the j-th species, and $a_{ij} < 0$ (resp. $a_{ij} > 0$) is the competing (resp. cooperating) coefficient between species i and j. Also, observe that $a_{ij} = a_{ji}$.

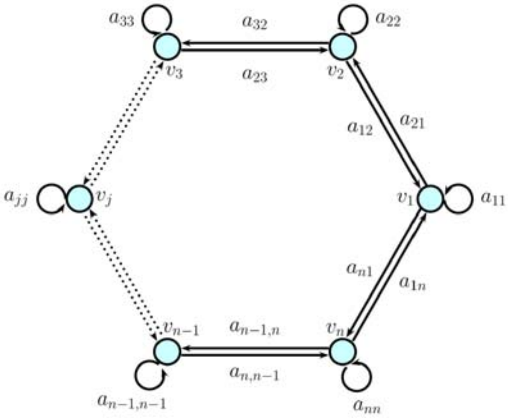


Fig. 8. System with dihedral symmetries.

Introduce

$$C = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{212}$$

and call it the *community matrix*. This community ecosystem can be described by the following equations,

$$\dot{v}(t) = \alpha C v(t) \cdot \left(1 - \frac{v(t-\tau)}{K}\right), \tag{213}$$

where '$\cdot$' stands for the component-wise "vector-multiplication" $u \cdot v = [u_1 v_1, \dots, u_n v_n]^T$ for $u = [u_1, \dots, u_n]^T$ and $v = [v_1, \dots, v_n]^T$ (cf. Subsection 8.1).

By applying the standard substitution, (213) is transformed to the equivalent system

$$\dot{u}(t) = -\alpha C u(t-\tau) \cdot \big[1 + u(t)\big], \tag{214}$$

where $u(t) = v(t)/K - 1$ is, in fact, a population saturation index with respect to the available resources. In order to study the system (214) in a heterogeneous environment, we add to (214) a spatial diffusion term, which leads to the following reaction–diffusion equations

$$\frac{\partial}{\partial t} u(x,t) = d \frac{\partial^2}{\partial x^2} u(x,t) - \alpha C u(x, t-1)\big[1 + u(x,t)\big], \tag{215}$$

where $d > 0$ is a spatial diffusion coefficient.

Symmetric system of the Hutchinson model Consider a symmetric system of n species Hutchinson model of the form (215) (for $t > 0$ and $x \in (0, \pi)$)

$$\begin{cases} \dfrac{\partial}{\partial t} u(x,t) = \dfrac{\partial^2}{\partial x^2} u(x,t) - \alpha C u(x, t-1) \cdot \big[1 + u(x,t)\big], \\ \dfrac{\partial}{\partial x} u(x,t) = 0, \quad x = 0, \pi, \end{cases} \tag{216}$$

where $u : [0, \pi] \times \mathbb{R} \to \mathbb{R}^n$ is a population saturation index (cf. (214)), $d > 0$ is a spatial diffusion coefficient and $\alpha \neq 0$ is the intrinsic growth rate (cf. (211)), which is considered as a bifurcation parameter, and C is a (symmetric) community matrix describing the interaction among the species. For simplicity, in what follows, we assume that $d = 1$.

Assume that

(A1) The geometrical configuration described by the system (216) has the icosahedral symmetry group A_5. The group A_5 permutes twenty vertices of the related dodecahedron, which means it acts on $\mathbb{R}^{20}$ by permuting the coordinates of the vectors $x \in \mathbb{R}^{20}$. The (symmetric) matrix C commuting with this A_5-action is of the type:

$$C = \begin{bmatrix}
1 & 2 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\
0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\
0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & 0 & 2 & 1
\end{bmatrix}. \tag{217}$$

Under the assumption (A1), the space $V := \mathbb{R}^{20}$ becomes an orthogonal A_5-representation and the system (216) is symmetric with respect to the action of the group $\Gamma = A_5$.

Characteristic equation and isolated centers At the A_5-symmetric steady-state solution $(\alpha, 0)$ the system (216) has the linearization

$$\begin{cases} \frac{\partial}{\partial t}u(x,t) = \frac{\partial^2}{\partial x^2}u(x,t) - \alpha C u(x,t-1), \\ \frac{\partial}{\partial x}u(x,t) = 0, \quad x = 0, \pi. \end{cases} \tag{218}$$

Since the matrix C is symmetric, it is completely diagonalizable with respect to a basis composed of its eigenvectors. Consider the spectrum $\sigma(C)$ of the matrix C,

$$\sigma(C) = \{\xi_0 = 7, \xi_1 = 1, \xi_2 = -3, \xi_3 = 3, \xi_4 = 1 + 2\sqrt{5}, \xi_5 = 1 - 2\sqrt{5}\}$$

and denote by $\mathfrak{E}(\xi_k) \subset V$ the eigenspace of ξ_k. One can verify that

$$\begin{aligned} &\mathfrak{E}(\xi_0) = \mathcal{V}_0, \quad \mathfrak{E}(\xi_1) = \mathcal{V}_1, \quad \mathfrak{E}(\xi_2) = \mathcal{V}_1, \\ &\mathfrak{E}(\xi_3) = \mathcal{V}_2, \quad \mathfrak{E}(\xi_4) = \mathcal{V}_3, \quad \mathfrak{E}(\xi_5) = \mathcal{V}_4. \end{aligned}$$

Then,

$$L^2\big([0,\pi]; V\big) = \bigoplus_{k=0}^{5} L^2\big([0,\pi]; \mathfrak{E}(\xi_k)\big), \tag{219}$$

and $w \in L^2([0,\pi]; V)$ can be represented as $w(x) = \sum_k w_k(x)$, where $w_k \in L^2([0,\pi];$ $\mathfrak{E}(\xi_k))$. In a similar way, one also has

$$L^2\big([0,\pi]; V^c\big) = \bigoplus_{k=0}^{5} L^2\big([0,\pi]; \mathfrak{E}^c(\xi_k)\big), \tag{220}$$

where $\mathfrak{E}^c(\xi_k)$ denotes the complexification of the eigenspace $\mathfrak{E}(\xi_k)$.

Notice that $(\alpha, 0)$ is a Γ-symmetric steady-state solution to (216) for all (non-zero) α. Thus, we can take the set $(\alpha, \beta, 0)$, $\alpha \neq 0$, for the manifold $M \subset \mathcal{P} \times \mathbb{E}^G$ described in Subsection 7.1 (see also Subsection 7.2). Moreover, $(\alpha_o, 0)$ is non-singular if $0 \notin \sigma(\mathcal{L}_{\alpha_o})$, where $\mathcal{L}_{\alpha_o} := \frac{\partial^2}{\partial x^2} - \alpha_o C : H_0^2([0,\pi]; V) \to L^2([0,\pi); V]$ with $H_0^2([0,\pi]; V)$ being the subspace of $H^2([0,\pi]; V)$ consisting of functions u satisfying $u(0) = u(\pi) = 0$. One can easily verify that if

$$-\frac{\alpha_o \xi_k}{d} \neq m^2 \quad \text{for all } k = 0, 2, \dots, 5 \text{ and } m = 0, 1, \dots,$$

then $(\alpha_o, 0)$ is a non-singular Γ-symmetric steady-state solution, i.e. $(\alpha_o, 0)$ satisfies the condition (C2) from Subsection 7.2.

A number $\lambda \in \mathbb{C}$ is a characteristic root of the system (216) at a Γ-symmetric steady-state solution $(\alpha, 0) \in \mathbb{R} \oplus V$ if there exists a non-zero function $v \in L^2([0,\pi]; V^c)$ such that

$$\Delta_\alpha(\lambda)v(x) := \lambda v(x) - \frac{\partial^2}{\partial x^2} v(x) + \alpha e^{-\lambda} C v(x) = 0, \tag{221}$$

where we put $\Delta_\alpha := \Delta_{\alpha;0}$ (cf. (174)).

By using the decomposition (220), v can be written as $v(x) = \sum_{k=0}^{5} v_k(x)$, for $v_k(x) \in E(\xi_k)$. Consequently, (221) yields

$$\Delta_\alpha(\lambda)v(x) = \sum_k \left(\lambda v_k(x) - \frac{\partial^2}{\partial x^2} v_k(x) + \alpha e^{-\lambda} \xi_k v_k(x) \right) = 0. \tag{222}$$

Next, by using the point spectrum $\{\zeta_m := m^2\}_{m=0}^{\infty}$ of the (scalar-valued) Laplace operator $L := -\frac{\partial^2}{\partial x^2}$ and the corresponding eigenspaces $E(\zeta_m)$, we can write $v_k(x) = \sum_m v_{k,m}(x)$, for $v_{k,m} \in E(\zeta_m)$, thus

$$\Delta_\alpha(\lambda)v(x) = \sum_{k,m} \big(\lambda v_{k,m}(x) + m^2 v_{k,m}(x) + \alpha e^{-\lambda} \xi_k v_{k,m}(x) \big) = 0. \tag{223}$$

Therefore, $\lambda \in \mathbb{C}$ is a characteristic root of (216) at the Γ-symmetric steady-state solution $(\alpha, 0)$, if

$$\lambda + m^2 + \alpha \xi_k e^{-\lambda} = 0, \quad \text{for } k = 1, \ldots, 5 \text{ and } m = 0, 1, \ldots. \tag{224}$$

Computations for the local bifurcation $\Gamma \times S^1$*-invariant* In order to find the values α_o for which the condition (C3) from Subsection 7.2 holds, we need to find purely imaginary roots $\lambda = i\beta$ ($\beta > 0$) of (224). Assume that $(\alpha, 0)$ is a non-singular Γ-symmetric steady-state solution to (216) (in particular, $\alpha \neq 0$).

- **Computation for purely imaginary roots** $\lambda = i\beta$ ($\beta > 0$)

By substituting $\lambda = i\beta$ into (224),

$$\begin{cases} m^2 + \alpha \xi_k \cos\beta = 0, \\ \beta - \alpha \xi_k \sin\beta = 0, \end{cases} \quad \text{for } k = 0, \ldots, 5. \tag{225}$$

In the case $m = 0$,

$$\begin{cases} \beta := \beta_{\nu,0,k} = \dfrac{\pi}{2} + \nu\pi, \\ \alpha := \alpha_{\nu,0,k} = (-1)^\nu \dfrac{\beta}{\xi_k}, \end{cases}$$

for $k = 0, \ldots, 5$ and $\nu = 0, 1, \ldots$. Consequently,

$$\operatorname{sign} \alpha_{\nu,0,k} = (-1)^\nu \operatorname{sign} \xi_k. \tag{226}$$

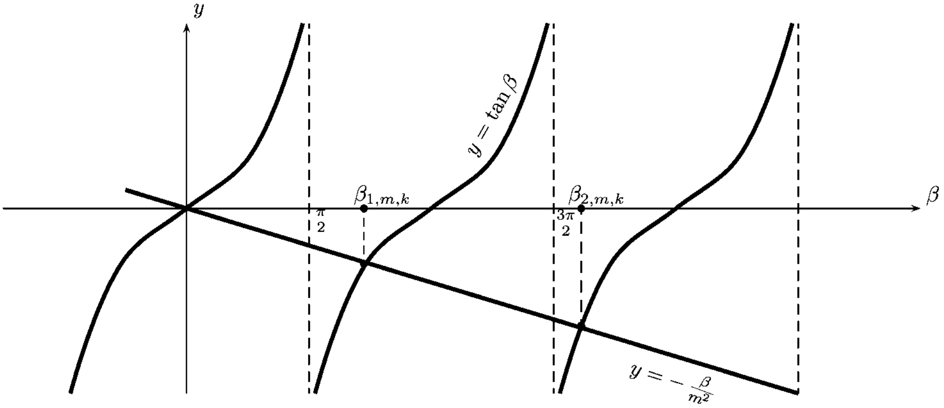


Fig. 9. Purely imaginary roots of the characteristic equation.

In the case $m \neq 0$ (thus $\cos\beta \neq 0$ by the first equation in (225)),

$$\tan\beta = -\frac{\beta}{m^2}, \tag{227}$$

$$\alpha = -\frac{m^2}{\xi_k \cos\beta}, \tag{228}$$

Eq. (227) has infinitely many positive solutions, which are denoted by $\{\beta_{\nu,m,k}\}_{\nu=1}^{\infty}$ (see Fig. 9). The corresponding solution α of (228) is denoted by $\alpha_{\nu,m,k}$. In Table 5 we list approximate values of the roots $(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$ for $m = 0, 1, 2, 3$ and $\nu = 0, 1, 2, 3, 4$.

Also, notice that $\operatorname{sign}\cos\beta_{\nu,m,k} = (-1)^{\nu}$, thus by (228) for $m > 0$,

$$\operatorname{sign}\alpha_{\nu,m,k} = (-1)^{\nu+1}\operatorname{sign}\xi_k. \tag{229}$$

• **Computation for** $\operatorname{sign}\frac{\mathrm{d}}{\mathrm{d}\alpha}w(\alpha)|_{\alpha=\alpha_{\nu,m,k}}$

Put $\alpha_o := \alpha_{\nu,m,k}$ and $\beta_o := \beta_{\nu,m,k}$. In order to determine the value of the crossing number $t_{j,1}(\alpha_o, \beta_o, 0)$, compute $\frac{\mathrm{d}}{\mathrm{d}\alpha}w(\alpha)|_{\alpha=\alpha_o}$ by implicit differentiation.

By substituting $\lambda = w + \mathrm{i}v$ into (224), then, differentiating with respect to α, and substituting $\alpha = \alpha_o$, $w = 0$ and $v = \beta_o$, we obtain

$$\operatorname{sign}\frac{\mathrm{d}w}{\mathrm{d}\alpha}|_{\alpha=\alpha_o=\alpha_{\nu,m,k}} = \begin{cases} (-1)^{\nu}\operatorname{sign}\xi_k, & \text{if } m = 0,\\ (-1)^{\nu+1}\operatorname{sign}\xi_k, & \text{if } m = 1, 2, \ldots . \end{cases} \tag{230}$$

Therefore, for $m = 0$

$$\mathfrak{t}_{j,1}(\alpha_o, \beta_o) = \begin{cases} (-1)^{\nu+1}\operatorname{sign}\xi_k m_j(\mathrm{i}\beta_o), & j = j_{\beta_o},\\ 0, & j \neq j_{\beta_o}, \end{cases} \tag{231}$$

Table 5
List of approximate values of several bifurcation points (α_o, β_o) for (216)

(ν, m)	$(\alpha_{\nu,m,0}, \beta_{\nu,m,0})$	$(\alpha_{\nu,m,1}, \beta_{\nu,m,1})$	$(\alpha_{\nu,m,2}, \beta_{\nu,m,2})$	$(\alpha_{\nu,m,3}, \beta_{\nu,m,3})$	$(\alpha_{\nu,m,4}, \beta_{\nu,m,4})$	$(\alpha_{\nu,m,5}, \beta_{\nu,m,5})$
(0,0)	(0.224, 1.57)	(1.57, 1.57)	(−0.523, 1.57)	(0.523, 1.57)	(0.286, 1.57)	(−0.451, 1.57)
(1,0)	(−0.673, 4.71)	(−4.71, 4.71)	(1.57, 4.71)	(−1.57, 4.71)	(−0.857, 4.71)	(1.35, 4.71)
(2,0)	(1.12, 7.85)	(7.85, 7.85)	(−2.62, 7.85)	(2.62, 7.85)	(1.43, 7.85)	(−2.25, 7.85)
(3,0)	(−1.57, 11.0)	(−11.0, 11.0)	(3.67, 11.0)	(−3.67, 11.0)	(−2.00, 11.0)	(3.16, 11.0)
(4,0)	(2.01, 14.1)	(14.1, 14.1)	(−4.70, 14.1)	(4.70, 14.1)	(2.57, 14.1)	(−4.05, 14.1)
(1,1)	(0.323, 2.03)	(2.26, 2.03)	(−0.752, 2.03)	(0.752, 2.03)	(0.411, 2.03)	(−0.649, 2.03)
(2,1)	(−0.730, 4.91)	(−5.10, 4.91)	(1.70, 4.91)	(−1.70, 4.91)	(−0.928, 4.91)	(1.46, 4.91)
(3,1)	(1.13, 7.98)	(7.94, 7.98)	(−2.64, 7.98)	(2.64, 7.98)	(1.45, 7.98)	(−2.28, 7.98)
(4,1)	(−1.38, 11.1)	(−9.62, 11.1)	(3.20, 11.1)	(−3.20, 11.1)	(−1.75, 11.1)	(2.76, 11.1)
(1,2)	(0.679, 2.57)	(4.76, 2.57)	(−1.58, 2.57)	(1.58, 2.57)	(0.866, 2.57)	(−1.37, 2.57)
(2,2)	(−0.960, 5.35)	(−6.72, 5.35)	(2.24, 5.35)	(−2.24, 5.35)	(−1.22, 5.35)	(1.93, 5.35)
(3,2)	(1.32, 8.30)	(9.28, 8.30)	(−3.09, 8.30)	(3.09, 8.30)	(1.69, 8.30)	(−2.66, 8.30)
(4,2)	(−1.90, 11.3)	(−13.3, 11.3)	(4.43, 11.3)	(−4.43, 11.3)	(−2.42, 11.3)	(3.82, 11.3)
(1,3)	(1.35, 2.84)	(9.42, 2.84)	(−3.14, 2.84)	(3.14, 2.84)	(1.71, 2.84)	(−2.70, 2.84)
(2,3)	(−1.52, 5.72)	(−10.6, 5.72)	(3.55, 5.72)	(−3.55, 5.72)	(−1.93, 5.72)	(3.04, 5.72)
(3,3)	(1.79, 8.66)	(12.5, 8.66)	(−4.16, 8.66)	(4.16, 8.66)	(2.28, 8.66)	(−3.59, 8.66)
(4,3)	(−1.99, 11.7)	(−13.9, 11.7)	(4.63, 11.7)	(−4.63, 11.7)	(−2.53, 11.7)	(3.99, 11.7)

and for $m \neq 0$

$$\mathfrak{t}_{j,1}(\alpha_o, \beta_o) = \begin{cases} (-1)^{\nu} \operatorname{sign} \xi_k \, m_j(\mathrm{i}\beta_o), & j = j_{\beta_o}, \\ 0, & j \neq j_{\beta_o}. \end{cases} \tag{232}$$

On the other hand, we have

$$m_j(\mathrm{i}\beta_{\nu,m,k}) = m_j(\xi_k) = \begin{cases} 1 & \text{for } j = k,\ k \neq 1, \\ 2 & \text{for } j = k,\ k = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Negative spectrum σ_- For a given bifurcation point $(\alpha_o, \beta_o, 0)$, we need to determine the negative spectrum σ_- of the operator $\mathcal{A}$ (given by (180) with $u_o = 0$). One can easily verify that the spectrum $\sigma(\mathcal{A})$ of the operator $\mathcal{A}$ is given by

$$\sigma(\mathcal{A}) = \left\{ \zeta_{m,k} := \frac{m^2 + \alpha_o \xi_k}{m^2 + 1} \colon m = 0, 1, 2, 3, \dots; k = 0, 1, \dots, 5 \right\}$$

where the isotypical multiplicity of $\zeta_{m,k}$ is the same as that of ξ_k. Consequently, we can define

$$m_i := \sum_{\mu \in \sigma_-} m_i(\mu), \quad i = 0, 1, 2, 3, 4, 5.$$

For the considered bifurcation points, the values $(m_0, m_1, m_2, m_3, m_4, m_5)$ are listed in Table 6.

Table 6
Numbers $(m_0, m_1, m_2, m_3, m_4)$ for the bifurcation points $(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$

(ν, m, k)	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
$(0, 0, k)$	[0, 1, 0, 0, 1]	[0, 3, 0, 0, 3]	[2, 1, 2, 2, 0]	[0, 2, 0, 0, 2]	[0, 1, 0, 0, 1]	[2, 1, 2, 2, 0]
$(1, 0, k)$	[3, 1, 2, 2, 0]	[6, 3, 4, 6, 0]	[0, 3, 0, 0, 3]	[4, 2, 3, 3, 0]	[3, 1, 2, 3, 0]	[0, 3, 0, 0, 3]
$(2, 0, k)$	[0, 2, 0, 0, 2]	[0, 5, 0, 0, 6]	[5, 2, 3, 4, 0]	[0, 3, 0, 0, 4]	[0, 3, 0, 0, 3]	[4, 2, 3, 4, 0]
$(3, 0, k)$	[4, 2, 3, 3, 0]	[8, 4, 6, 8, 0]	[0, 4, 0, 0, 4]	[6, 2, 4, 5, 0]	[4, 2, 3, 4, 0]	[0, 4, 0, 0, 4]
$(4, 0, k)$	[0, 3, 0, 0, 3]	[0, 7, 0, 0, 8]	[6, 3, 4, 6, 0]	[0, 4, 0, 0, 5]	[0, 3, 0, 0, 3]	[6, 3, 4, 5, 0]
$(1, 1, k)$	[0, 1, 0, 0, 2]	[0, 3, 0, 0, 3]	[3, 1, 2, 3, 0]	[0, 2, 0, 0, 2]	[0, 2, 0, 0, 2]	[3, 1, 2, 2, 0]
$(2, 1, k)$	[3, 1, 2, 2, 0]	[6, 3, 4, 6, 0]	[0, 3, 0, 0, 3]	[4, 2, 3, 4, 0]	[3, 1, 2, 3, 0]	[0, 3, 0, 0, 3]
$(3, 1, k)$	[0, 2, 0, 0, 2]	[0, 5, 0, 0, 6]	[5, 2, 3, 4, 0]	[0, 3, 0, 0, 4]	[0, 3, 0, 0, 3]	[4, 2, 3, 4, 0]
$(4, 1, k)$	[4, 2, 3, 3, 0]	[8, 4, 6, 8, 0]	[0, 4, 0, 0, 4]	[5, 2, 4, 5, 0]	[4, 2, 3, 4, 0]	[0, 3, 0, 0, 4]
$(1, 2, k)$	[0, 2, 0, 0, 2]	[0, 4, 0, 0, 5]	[4, 2, 3, 3, 0]	[0, 3, 0, 0, 3]	[0, 2, 0, 0, 2]	[4, 2, 3, 3, 0]
$(2, 2, k)$	[3, 1, 2, 3, 0]	[7, 3, 5, 7, 0]	[0, 3, 0, 0, 3]	[4, 2, 3, 4, 0]	[3, 2, 2, 3, 0]	[0, 3, 0, 0, 3]
$(3, 2, k)$	[0, 2, 0, 0, 3]	[0, 6, 0, 0, 6]	[5, 0, 4, 5, 0]	[0, 4, 0, 0, 4]	[0, 3, 0, 0, 3]	[5, 2, 3, 4, 0]
$(4, 2, k)$	[4, 2, 3, 4, 0]	[8, 4, 7, 8, 0]	[0, 4, 0, 0, 4]	[6, 3, 4, 5, 0]	[5, 2, 3, 4, 0]	[0, 4, 0, 0, 4]
$(1, 3, k)$	[0, 3, 0, 0, 3]	[0, 6, 0, 0, 6]	[5, 2, 4, 5, 0]	[0, 4, 0, 0, 4]	[0, 3, 0, 0, 3]	[5, 2, 3, 4, 0]
$(2, 3, 0)$	[4, 2, 3, 3, 0]	[8, 4, 6, 8, 0]	[0, 4, 0, 0, 4]	[5, 2, 4, 5, 0]	[4, 2, 3, 4, 0]	[0, 4, 0, 0, 4]
$(3, 3, 0)$	[0, 3, 0, 0, 3]	[0, 7, 0, 0, 7]	[6, 3, 4, 5, 0]	[0, 4, 0, 0, 4]	[0, 3, 0, 0, 3]	[6, 2, 4, 5, 0]
$(4, 3, 0)$	[4, 2, 3, 4, 0]	[8, 4, 7, 8, 0]	[0, 4, 0, 0, 5]	[6, 3, 4, 6, 0]	[5, 2, 3, 4, 0]	[0, 4, 0, 0, 4]

Consequently, in order to compute the value of the local bifurcation invariant $\omega(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$ we need to use the following Maple© procedure

$$\text{showdegree[A5]}(\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \mathfrak{t}_0, \mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3, \mathfrak{t}_4, \mathfrak{t}_5)$$

where $\varepsilon_i = m_i \bmod 2$, and (since only ξ_2 and ξ_5 are negative)

$$\mathfrak{t}_j := \mathfrak{t}_j(\alpha_{\nu,m,k}, \beta_{\nu,m,k}) = \begin{cases} (-1)^\nu & \text{if } j=k,\ k=0,3,4, \\ (-1)^{\nu+1} & \text{if } j=k,\ k=2,5, \\ (-1)^\nu 2 & \text{if } j=k,\ k=1. \end{cases}$$

Preparation of input data for the Maple© routines Based on the above computations we are now in a position to prepare the input data for the Maple© routines in order to evaluate (for selected bifurcation points $(\alpha_{\nu,m,k}, \beta_{\nu,m,k}, 0)$) the exact value of the bifurcation invariant

$$\begin{aligned} &\omega(\alpha_{\nu,m,k}, \beta_{\nu,m,k}) \\ &\quad = \texttt{showdegree[A5]}(\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \mathfrak{t}_0, \mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3, \mathfrak{t}_4, \mathfrak{t}_5), \end{aligned}$$

which is presented in Table 7.

Table of results In Table 8, we present the classification of the A_5-symmetric Hopf bifurcation in (216) for selected bifurcation points $(\alpha_{\nu,m,k}, \beta_{\nu,m,k}, 0)$.

Table 7
Input data $(\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \mathfrak{t}_0, \mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3, \mathfrak{t}_4, \mathfrak{t}_5)$ for selected bifurcation points $(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$

$(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$	(ν, m, k)	$\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$	$\mathfrak{t}_0, \mathfrak{t}_1, \mathfrak{t}_2, \mathfrak{t}_3, \mathfrak{t}_4$
(0.224, 1.57)	(0, 0, 0)	0, 1, 0, 0, 1	−1, 0, 0, 0, 0
(0.286, 1.57)	(0, 0, 4)	0, 1, 0, 0, 1	0, 0, 0, −1, 0
(−0.673, 4.71)	(1, 0, 0)	1, 1, 0, 0, 0	−1, 0, 0, 0, 0
(1.57, 4.71)	(1, 0, 2)	0, 0, 1, 1, 0	0, 1, 0, 0, 0
(1.35, 4.71)	(1, 0, 5)	0, 1, 0, 0, 1	0, 0, 0, 0, 1
(7.85, 7.85)	(2, 0, 1)	0, 1, 0, 0, 0	0, −1, 0, 0, 0
(1.43, 7.85)	(2, 0, 4)	0, 1, 0, 0, 1	0, 0, 0, −1, 0
(−11.0, 11.0)	(3, 0, 1)	0, 0, 0, 0, 0	0, 1, 0, 0, 0
(−2.00, 11.0)	(3, 0, 4)	0, 0, 1, 0, 0	0, 0, 0, 1, 0
(2.01, 14.1)	(4, 0, 0)	0, 1, 0, 0, 1	−1, 0, 0, 0, 0
(4.70, 14.1)	(4, 0, 3)	0, 0, 0, 0, 1	0, 0, −1, 0, 0
(0.323, 2.03)	(1, 1, 0)	0, 1, 0, 0, 0	0, 0, −1, 0, 0
(−0.649, 2.03)	(1, 1, 5)	1, 1, 0, 0, 0	0, 0, 0, 0, 1
(1.70, 4.91)	(2, 1, 2)	0, 1, 0, 0, 1	0, 1, 0, 0, 0
(−0.928, 4.91)	(2, 1, 4)	1, 1, 0, 1, 0	0, 0, 0, −1, 0
(7.94, 7.98)	(3, 1, 1)	0, 1, 0, 0, 0	0, 1, 0, 0, 0
(1.45, 7.98)	(3, 1, 4)	0, 1, 0, 0, 1	0, 0, 0, 1, 0
(−3.20, 11.1)	(4, 1, 3)	0, 0, 0, 0, 0	0, 0, −1, 0, 0
(2.76, 11.1)	(4, 1, 5)	0, 1, 0, 0, 0	0, 0, 0, 0, 1
(4.76, 2.57)	(1, 2, 1)	0, 0, 1, 1, 0	0, 1, 0, 0, 0
(−1.37, 2.57)	(1, 2, 5)	0, 0, 1, 1, 0	0, 0, 0, 0, 1
(2.24, 5.35)	(2, 2, 2)	0, 1, 0, 0, 1	0, 1, 0, 0, 0
(1.93, 5.35)	(2, 2, 5)	0, 1, 0, 0, 1	0, 0, 0, 0, 1
(−3.09, 8.30)	(3, 2, 2)	1, 0, 0, 1, 0	0, −1, 0, 0, 0
(1.69, 8.30)	(3, 2, 4)	0, 1, 0, 0, 1	0, 0, 0, 1, 0
(−2.42, 11.3)	(4, 2, 4)	1, 0, 1, 0, 0	0, 0, 0, −1, 0
(3.82, 11.3)	(4, 2, 5)	0, 0, 0, 0, 0	0, 0, 0, 0, 1
(−3.14, 2.84)	(1, 3, 2)	1, 0, 0, 1, 0	0, −1, 0, 0, 0
(−2.70, 2.84)	(1, 3, 5)	1, 0, 1, 0, 0	0, 0, 0, 0, −1
(3.55, 5.72)	(2, 3, 2)	0, 0, 0, 0, 0	0, 1, 0, 0, 0
(−1.93, 5.72)	(2, 3, 4)	0, 0, 1, 0, 0	0, 0, 0, −1, 0
(4.16, 8.66)	(3, 3, 3)	0, 0, 0, 0, 0	0, 0, 1, 0, 0
(−3.59, 8.66)	(3, 3, 5)	0, 0, 0, 1, 0	0, 0, 0, 0, −1
(−4.63, 11.7)	(4, 3, 3)	0, 1, 0, 0, 0	0, 0, −1, 0, 0
(−2.53, 11.7)	(4, 3, 4)	1, 0, 1, 0, 0	0, 0, 0, −1, 0

8.5. *Bibliographical remarks*

The model presented in Subsection 8.1 was rigorously studied in [11] . Symmetric configuration of transmission lines presented in Subsection 8.2, was analyzed in details in [15] The global bifurcation results given in Subsection 8.3 were published in [13]. For more details we refer to [19]. Similar systems of symmetric DEs and FDEs were also considered in [67,65,66,68,92,94,95,129,140,145,146,148]. In the non-symmetric case, local and global Hopf bifurcation problems in FDEs were studied in [36,110,111,124,125,122,123].

The results presented in Subsection 8.4 were taken from [16]. For population ecology background, which was used in this section, we refer to [73,136,56,70]. An equation, sim-

Table 8
Classification of the A_5-symmetric Hopf bifurcation in (216) for selected bifurcation points $(\alpha_{\nu,m,k}, \beta_{\nu,m,k}, 0)$

$(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$	(ν, m, k)	$\omega(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$	# branches
(0.224, 1.57)	$(0, 0, 0)$	$-(A_5) + 2(A_4) + 2(D_3) + (\mathbb{Z}_5) - 2(\mathbb{Z}_3) - 2(\mathbb{Z}_2) + (\mathbb{Z}_1)$	1
(0.286, 1.57)	$(0, 0, 4)$	$-(D_5^z) + (D_3^z) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5) + (V_4^-) + (\mathbb{Z}_3^t) - (\mathbb{Z}_1)$	48
(−0.673, 4.71)	$(1, 0, 0)$	$(A_5) - 2(A_4) - 2(D_3) + 3(\mathbb{Z}_3) + 3(\mathbb{Z}_2) - 2(\mathbb{Z}_1)$	1
(1.57, 4.71)	$(1, 0, 2)$	$(A_4) - (D_3^z) - (D_3) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5^{t_2}) + (V_4^-) + (\mathbb{Z}_3^t) + (\mathbb{Z}_3) + (\mathbb{Z}_2^-) + (\mathbb{Z}_2) - 2(\mathbb{Z}_1)$	55
(1.35, 4.71)	$(1, 0, 5)$	$(D_5^z) - (D_3^z) - (\mathbb{Z}_5^{t_2}) - (\mathbb{Z}_5) - (V_4^-) - (\mathbb{Z}_3^t) + (\mathbb{Z}_1)$	48
(7.85, 7.85)	$(2, 0, 1)$	$(A_4) + (D_3^z) + (D_3) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5^{t_2}) + (V_4^-) - (\mathbb{Z}_3^t) - 3(\mathbb{Z}_3) - 3(\mathbb{Z}_2^-) - 3(\mathbb{Z}_2) + 4(\mathbb{Z}_1)$	55
(1.43, 7.85)	$(2, 0, 4)$	$-(D_5^z) + (D_3^z) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5) + (V_4^-) + (\mathbb{Z}_3^t) - (\mathbb{Z}_1)$	48
(−11.0, 11.0)	$(3, 0, 1)$	$(A_4) + (D_3^z) + (D_3) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5^{t_2}) + (V_4^-) + (\mathbb{Z}_3^t) - (\mathbb{Z}_3) - (\mathbb{Z}_2^-) - (\mathbb{Z}_2)$	55
(−2.00, 11.0)	$(3, 0, 4)$	$-(D_5^z) - (D_3^z) - (\mathbb{Z}_5^{t_1}) + (V_4^-) - (\mathbb{Z}_3^t) - (\mathbb{Z}_2) + (\mathbb{Z}_1)$	48
(2.01, 14.1)	$(4, 0, 0)$	$-(A_5) + 2(A_4) + 2(D_3) + (\mathbb{Z}_5) - 2(\mathbb{Z}_3) - 2(\mathbb{Z}_2) + (\mathbb{Z}_1)$	1
(4.70, 14.1)	$(4, 0, 3)$	$-(A_4^{t_1}) - (A_4^{t_2}) - (D_5) - (D_3) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5^{t_2}) + (\mathbb{Z}_5) - (V_4^-) + 4(\mathbb{Z}_3^t) + (\mathbb{Z}_3) + 2(\mathbb{Z}_2^-) + 5(\mathbb{Z}_2) - 5(\mathbb{Z}_1)$	50
(0.323, 2.03)	$(1, 1, 0)$	$(A_4^{t_1}) + (A_4^{t_2}) - (D_5) + (D_3) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5^{t_2}) + (V_4^-) - 4(\mathbb{Z}_3^t) - (\mathbb{Z}_3) - 2(\mathbb{Z}_2^-) - 3(\mathbb{Z}_2) + 4(\mathbb{Z}_1)$	50
(−0.649, 2.03)	$(1, 1, 5)$	$-(D_5^z) + (D_3^z) - (\mathbb{Z}_5^{t_2}) + (V_4^-) - (\mathbb{Z}_3^t) - (\mathbb{Z}_3) - 2(\mathbb{Z}_2^-) - (\mathbb{Z}_2) + 2(\mathbb{Z}_1)$	48
(1.70, 4.91)	$(2, 1, 2)$	$-(A_4) - (D_3^z) - (D_3) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5^{t_2}) - (V_4^-) - (\mathbb{Z}_3^t) + (\mathbb{Z}_3) + (\mathbb{Z}_2^-) + (\mathbb{Z}_2)$	55
(−0.928, 4.91)	$(2, 1, 4)$	$(D_5^z) - (D_3^z) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5) - (V_4^-) - (\mathbb{Z}_3^t) + (\mathbb{Z}_1)$	48
(7.94, 7.98)	$(3, 1, 1)$	$-(A_4) - (D_3^z) - (D_3) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5^{t_2}) - (V_4^-) + (\mathbb{Z}_3^t) + 3(\mathbb{Z}_3) + 3(\mathbb{Z}_2^-) + 3(\mathbb{Z}_2) - 4(\mathbb{Z}_1)$	55
(1.45, 7.98)	$(3, 1, 4)$	$(D_5^z) - (D_3^z) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5) - (V_4^-) - (\mathbb{Z}_3^t) + (\mathbb{Z}_1)$	48
(−3.20, 11.1)	$(4, 1, 3)$	$-(A_4^{t_1}) - (A_4^{t_2}) - (D_5) - (D_3) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5^{t_2}) - (V_4^-) + 2(\mathbb{Z}_2)$	50
(2.76, 11.1)	$(4, 1, 5)$	$(D_5^z) - (D_3^z) + (\mathbb{Z}_5^{t_2}) - (V_4^-) + (\mathbb{Z}_3^t) + (\mathbb{Z}_3) + 2(\mathbb{Z}_2^-) + (\mathbb{Z}_2) - 2(\mathbb{Z}_1)$	48
(4.76, 2.57)	$(1, 2, 1)$	$(A_4) - (D_3^z) - (D_3) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5^{t_2}) + (V_4^-) + (\mathbb{Z}_3^t) + (\mathbb{Z}_3) + (\mathbb{Z}_2^-) + (\mathbb{Z}_2) - 2(\mathbb{Z}_1)$	55
(−1.37, 2.57)	$(1, 2, 5)$	$-(D_5^z) - (D_3^z) + (\mathbb{Z}_5^{t_2}) + (\mathbb{Z}_5) + (V_4^-) + (\mathbb{Z}_3^t) + (\mathbb{Z}_3) + 2(\mathbb{Z}_2^-) - 2(\mathbb{Z}_1)$	48
(2.24, 5.35)	$(2, 2, 2)$	$-(A_4) - (D_3^z) - (D_3) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5^{t_2}) - (V_4^-) - (\mathbb{Z}_3^t) + (\mathbb{Z}_3) + (\mathbb{Z}_2^-) + (\mathbb{Z}_2)$	55
(1.93, 5.35)	$(2, 2, 5)$	$(D_5^z) - (D_3^z) - (\mathbb{Z}_5^{t_2}) - (\mathbb{Z}_5) - (V_4^-) - (\mathbb{Z}_3^t) + (\mathbb{Z}_1)$	48
(−3.09, 8.30)	$(3, 2, 2)$	$(A_4) + (D_3^z) + (D_3)_-(\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5^{t_2}) + (V_4^-) - (\mathbb{Z}_3^t) - 3(\mathbb{Z}_3) - 3(\mathbb{Z}_2^-) - 3(\mathbb{Z}_2) + 4(\mathbb{Z}_1)$	55
(1.69, 8.30)	$(3, 2, 4)$	$(D_5^z) - (D_3^z) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5) - (V_4^-) - (\mathbb{Z}_3^t) + (\mathbb{Z}_1)$	48
(−2.42, 11.3)	$(4, 2, 4)$	$-(D_5^z - (D_3^z) - (\mathbb{Z}_5^{t_1}) + (V_4^-) - (\mathbb{Z}_3^t) - (\mathbb{Z}_2) + (\mathbb{Z}_1)$	48
(3.82, 11.3)	$(4, 2, 5)$	$(D_5^z) + (D_3^z) + (\mathbb{Z}_5^{t_2}) + (V_4^-) + (\mathbb{Z}_3^t) - 2(\mathbb{Z}_2^-)$	48

(*continued on next page*)

Table 8
(*continued*)

$(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$	(ν, m, k)	$\omega(\alpha_{\nu,m,k}, \beta_{\nu,m,k})$	# branches
(−3.14, 2.84)	$(1, 3, 2)$	$(A_4) + (D_3^z) + (D_3) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5^{t_2}) + (V_4^-) - (\mathbb{Z}_3^t)$ $- 3(\mathbb{Z}_3) - 3(\mathbb{Z}_2^-) - 3(\mathbb{Z}_2) + 4(\mathbb{Z}_1)$	55
(−2.70, 2.84)	$(1, 3, 5)$	$-(D_5^z) - (D_3^z) - (\mathbb{Z}_5^{t_2}) + (V_4^-) - (\mathbb{Z}_3^t) - (\mathbb{Z}_2) + (\mathbb{Z}_1)$	48
(3.55, 5.72)	$(2, 3, 2)$	$(A_4) + (D_3^z) + (D_3) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5^{t_2}) + (V_4^-) + (\mathbb{Z}_3^t)$ $- (\mathbb{Z}_3) - (\mathbb{Z}_2^-) - (\mathbb{Z}_2)$	55
(−1.93, 5.72)	$(2, 3, 4)$	$(D_5^z) + (D_3^z) + (\mathbb{Z}_5^{t_1}) - (V_4^-) + (\mathbb{Z}_3^t) + (\mathbb{Z}_2) - (\mathbb{Z}_1)$	48
(4.16, 8.66)	$(3, 3, 3)$	$(A_4^{t_1}) + (A_4^{t_2}) + (D_5) + (D_3) + (\mathbb{Z}_5^{t_1}) + (\mathbb{Z}_5^{t_2}) + (V_4^-) - 2(\mathbb{Z}_2)$	50
(−3.59, 8.66)	$(3, 3, 5)$	$-(D_5^z) - (D_3^z) + (\mathbb{Z}_5^{t_2}) + (\mathbb{Z}_5) - (V_4^-) + (\mathbb{Z}_3^t) + (\mathbb{Z}_3)$ $+ 4(\mathbb{Z}_2^-) + (\mathbb{Z}_2) - 3(\mathbb{Z}_1)$	48
(−4.63, 11.7)	$(4, 3, 3)$	$(A_4^{t_1}) + (A_4^{t_2}) - (D_5) + (D_3) - (\mathbb{Z}_5^{t_1}) - (\mathbb{Z}_5^{t_2})$ $+ (V_4^-) - 4(\mathbb{Z}_3^t) - (\mathbb{Z}_3) - 2(\mathbb{Z}_2^-) - 3(\mathbb{Z}_2) + 4(\mathbb{Z}_1)$	50
(−2.53, 11.7)	$(4, 3, 4)$	$-(D_5^z) - (D_3^z) - (\mathbb{Z}_5^{t_1}) + (V_4^-) - (\mathbb{Z}_3^t) - (\mathbb{Z}_2) + (\mathbb{Z}_1)$	48

ilar to (215) (for a single species in one dimensional space), was also studied in [145] for the stability property of the periodic solutions bifurcating from 0.

Appendix A

Dihedral group D_N

Represent the dihedral group D_N of order $2N$ as the group of rotations $1, \xi, \xi^2, \ldots, \xi^{N-1}$ of the complex plane (where ξ is the multiplication by $e^{\frac{2\pi i}{N}}$) plus the reflections κ, $\kappa\xi$, $\kappa\xi^2, \ldots, \kappa\xi^{N-1}$ with κ being the operator of complex conjugation described by the matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

Let us introduce some special subgroups of D_N and of $D_N \times S^1$. For a positive integer k with $k|N$, and $\gamma := e^{\frac{2\pi i}{k}}$, we have:

(i) the subgroups $D_k = \{1, \gamma, \gamma^2, \ldots, \gamma^{k-1}, \kappa, \kappa\gamma, \ldots, \kappa\gamma^{k-1}\}$ and their isomorphic copies

$$D_{k,j} = \{1, \gamma, \gamma^2, \ldots, \gamma^{k-1}, \kappa\xi^j, \kappa\xi^j\gamma, \ldots, \kappa\xi^j\gamma^{k-1}\} \subset D_N,$$

$j = 0, 1, \ldots, \frac{N}{k} - 1$, which are all conjugate if $\frac{N}{k}$ is odd, but split into two conjugacy classes (D_k) and $(\widetilde{D}_k)$, where $\widetilde{D}_k := D_{k,1}$, if $\frac{N}{k}$ is even;

(ii) the twisted subgroups of $\mathbb{Z}_k = \{1, \gamma, \gamma^2, \ldots, \gamma^{k-1}\}$, which are the subgroups

$$\mathbb{Z}_k^{t_r} := \{(1, 1), (\gamma, \gamma^r), (\gamma^2, \gamma^{2r}), \ldots, (\gamma^{k-1}, \gamma^{(k-1)r})\},$$

with $r \in \{0, 1, \dots, k-1\}$. Obviously, for $0 < r < \frac{k}{2}$, $\kappa \mathbb{Z}_k^{t_r} \kappa = \mathbb{Z}_k^{t_{k-r}}$, i.e. $\mathbb{Z}_k^{t_r}$ and $\mathbb{Z}_k^{t_{k-r}}$ are conjugate;

(iii) the twisted subgroups

$$D_k^z = \big\{(1,1), (\gamma,1), \dots, (\gamma^{k-1},1), (\kappa,-1), (\kappa\gamma,-1), \dots, (\kappa\gamma^{k-1},-1)\big\},$$

of D_k (and similarly for the $D_{k,j}$ and, in particular, for $\widetilde{D}_k$);

(iv) another set of twisted subgroups in the case $k = 2m$:

First,

$$D_{2m}^d = \big\{(1,1), (\gamma,-1), (\gamma^2,1), \dots, (\gamma^{k-1},-1), (\kappa,1), (\kappa\gamma,-1), \dots, \\ (\kappa\gamma^{k-1},-1)\big\}$$

and similarly

$$\widetilde{D}_{2m}^d = \big\{(1,1), (\gamma,-1), (\gamma^2,1), \dots, (\gamma^{k-1},-1), (\kappa\xi,1), (\kappa\xi\gamma,-1), \\ \dots, (\kappa\xi\gamma^{k-1},-1)\big\}.$$

If $\frac{N}{2m} =: l$ is even, then

$$\begin{aligned} D_{2m}^{\hat{d}} &:= \big\{(1,1), (\gamma,-1), (\gamma^2,1), \dots, (\gamma^{k-1},-1), (\kappa,-1), (\kappa\gamma,1), \\ &\qquad \dots, (\kappa\gamma^{k-1},1)\big\} \\ &= \mathrm{e}^{-\frac{2\pi \mathrm{i}}{4m}} D_{2m}^d \mathrm{e}^{\frac{2\pi \mathrm{i}}{4m}} = \xi^{-\frac{l}{2}} D_{2m}^d \xi^{\frac{l}{2}}, \end{aligned}$$

i.e. $D_{2m}^{\hat{d}}$ and D_{2m}^d are conjugate, but not in the case of $l := \frac{N}{2m}$ being odd. Observe that for $\frac{N}{2m} \in \{3, 5, 7, \dots\}$,

$$\xi^{-\frac{1}{2}(\frac{N}{2m}+1)} D_{2m}^{\hat{d}} \xi^{\frac{1}{2}(\frac{N}{2m}+1)} = \widetilde{D}_{2m}^d, \quad \text{i.e.} \quad D_{2m}^{\hat{d}} \sim \widetilde{D}_{2m}^d;$$

(v) the subgroup, for $k = 2m$,

$$\mathbb{Z}_{2m}^d := \mathbb{Z}_{2m}^{t_m} = \big\{(1,1), (\gamma,-1), (\gamma^2,1), \dots, (\gamma^{2m-1},-1)\big\}.$$

In the case $m = 1$ we will write $\mathbb{Z}_2^-$ instead of $\mathbb{Z}_2^d$.

Irreducible representations of dihedral groups

Let us describe the list of all real irreducible representations of D_n.

(a0) Clearly, there is a one-dimensional trivial representation $\mathcal{V}_0$. In this case,

$$\deg_{\mathcal{V}_0} = -(D_n).$$

(a1) For every integer number $1 \leqslant j < n/2$, there is an orthogonal representation $\mathcal{V}_j$ of D_n on $\mathbb{C}$ given by

$$\gamma z := \gamma^j \cdot z, \quad \text{for } \gamma \in \mathbb{Z}_n \text{ and } z \in \mathbb{C};$$
$$\kappa z := \bar{z},$$

where $\gamma^j \cdot z$ denotes the usual complex multiplication. Put

$$h := \gcd(j, n) \quad \text{and} \quad m := n/h. \tag{233}$$

Put $j_n := [(n+1)/2]$.

(a2) There is a representation $\mathcal{V}_{j_n}$ given by the homomorphism $c : D_n \to \mathbb{Z}_2$, such that $\ker c = \mathbb{Z}_n$.

(a3) For n even, there is an irreducible representation $\mathcal{V}_{j_n+1}$ given by $d : D_n \to \mathbb{Z}_2$ such that $\ker d = D_{n/2}$.

(a4) Also, for n even, there is an irreducible representation $\mathcal{V}_{j_n+2}$ given by $\hat{d} : D_n \to \mathbb{Z}_2$, where $\ker \hat{d} = \widetilde{D}_{n/2}$.

Icosahedral group A_5

Consider the icosahedral group $\mathbb{I}$ consisting of symmetries (preserving orientation) of a regular icosahedron. The group $\mathbb{I}$ is isomorphic to the alternating group of five elements A_5 and it has 60 elements. Let us list (nontrivial) representatives of the conjugacy classes of the subgroups in A_5:

$$\mathbb{Z}_2 = \{(1), (12)(34)\},$$
$$\mathbb{Z}_3 = \{(1), (123), (132)\},$$
$$V_4 = \{(1), (12)(34), (13)(24), (23)(14)\},$$
$$\mathbb{Z}_5 = \{(1), (12345), (13524), (14253), (15432)\},$$
$$D_3 = \{(1), (123), (132), (12)(45), (13)(45), (23)(45)\},$$
$$A_4 = \{(1), (12)(34), (13)(24), (14)(23), (123), (132), (124), (142), (134), (143), (234), (243)\},$$
$$D_5 = \{(1), (12345), (13524), (14253), (15432), (12)(35), (13)(45), (14)(23), (15)(24), (25)(34)\}.$$

We have the following list, up to conjugacy, of the twisted (one-folded) subgroups H^{φ} of $A_5 \times S^1$, where H is a subgroup of A_5, $\varphi : H \to S^1$ a group homomorphism, and $H^{\varphi} = \{(h,z) \in H \times S^1\colon\ \varphi(h) = z\}$:

$$\begin{aligned}
\mathbb{Z}_2^- &= \bigl\{\bigl((1),1\bigr), \bigl((12)(34),-1\bigr)\bigr\},\\
V_4^- &= \bigl\{\bigl((1),1\bigr), \bigl((12)(34),-1\bigr), \bigl((13)(24),-1\bigr), \bigl((23)(14),1\bigr)\bigr\},\\
\mathbb{Z}_5^{t_1} &= \bigl\{\bigl((1),1\bigr), \bigl((12345),\xi\bigr), \bigl((13524),\xi^2\bigr), \bigl((14253),\xi^3\bigr), \bigl((15432),\xi^4\bigr)\bigr\},\\
\mathbb{Z}_5^{t_2} &= \bigl\{\bigl((1),1\bigr), \bigl((12345),\xi^2\bigr), \bigl((13524),\xi^4\bigr), \bigl((14253),\xi\bigr), \bigl((15432),\xi^3\bigr)\bigr\},\\
\mathbb{Z}_3^{t} &= \bigl\{\bigl((1),1\bigr), \bigl((123),\gamma\bigr), \bigl((132),\gamma^2\bigr)\bigr\},\\
D_3^{z} &= \bigl\{\bigl((1),1\bigr), \bigl((123),1\bigr), \bigl((132),1\bigr), \bigl((12)(45),-1\bigr), \bigl((13)(45),-1\bigr),\\
&\qquad \bigl((23)(45),-1\bigr)\bigr\},\\
A_4^{t_1} &= \bigl\{\bigl((1),1\bigr), \bigl((12)(34),1\bigr), \bigl((13)(24),1\bigr), \bigl((14)(23),1\bigr), \bigl((123),\gamma\bigr),\\
&\qquad \bigl((132),\gamma^2\bigr), \bigl((124),\gamma^2\bigr), \bigl((142),\gamma\bigr), \bigl((134),\gamma\bigr), \bigl((143),\gamma^2\bigr),\\
&\qquad \bigl((234),\gamma^2\bigr), \bigl((243),\gamma\bigr)\bigr\},\\
A_4^{t_2} &= \bigl\{\bigl((1),1\bigr), \bigl((12)(34),1\bigr), \bigl((13)(24),1\bigr), \bigl((14)(23),1\bigr), \bigl((123),\gamma^2\bigr),\\
&\qquad \bigl((132),\gamma\bigr), \bigl((124),\gamma\bigr), \bigl((142),\gamma^2\bigr), \bigl((134),\gamma^2\bigr), \bigl((143),\gamma\bigr), \bigl((234),\gamma\bigr),\\
&\qquad \bigl((243),\gamma^2\bigr)\bigr\},\\
D_5^{z} &= \bigl\{\bigl((1),1\bigr), \bigl((12345),1\bigr), \bigl((13524),1\bigr), \bigl((14253),1\bigr), \bigl((15432),1\bigr),\\
&\qquad \bigl((12)(35),-1\bigr), \bigl((13)(45),-1\bigr), \bigl((14)(23),-1\bigr), \bigl((15)(24),-1\bigr),\\
&\qquad \bigl((25)(34),-1\bigr)\bigr\},
\end{aligned}$$

where $\xi = \mathrm{e}^{\frac{2\pi}{5}\mathrm{i}}$, $\gamma = \mathrm{e}^{\frac{2\pi\mathrm{i}}{3}}$.

Irreducible representations of A_5

There are exactly 5 irreducible representations of A_5: $\mathcal{V}_0$ – the trivial representation, $\mathcal{V}_1$ – the natural 4-dimensional representation of A_5, $\mathcal{V}_2$ – the 5-dimensional representation of A_5, and two 3-dimensional representations $\mathcal{V}_3$ and $\mathcal{V}_4$. We distinguish these two representations by their characters χ_3 and χ_4, i.e. we assume that

$$\chi_3\bigl((12345)\bigr) = \frac{1+\sqrt{5}}{2}, \qquad \chi_4\bigl((12345)\bigr) = \frac{1-\sqrt{5}}{2}.$$

Notice that all the irreducible Γ-representations considered in this appendix are of real type. Therefore, the corresponding to $\mathcal{V}_j$ irreducible $\Gamma \times S^1$-representation $\mathcal{V}_{j,l}$ can be easily constructed from the formula $\mathcal{V}_{j,l} = {}^{l}\mathcal{V}_j^c$.

Acknowledgment

The authors would like to thank H. Ruan, S. Rybicki and H. Steinlein for inspiring discussions related to the topics of this paper.

References

[1] R.R. Akhmerov, M.I. Kamenskiĭ, A.S. Potapov, A.E. Rodkina and B.N. Sadovskiĭ, *Measures of Noncompactness and Condensing Operators*, Birkhäuser, Basel (1992).

[2] J.C. Alexander and J.A. Yorke, *Global bifurcations of periodic orbits*, Amer. J. Math. **100** (1978), 263–292.

[3] S.A. Antonyan, *Retracts in categories of G-spaces*, Izv. Akad. Nauk Armyan. SSR Ser. Mat. **15** (1980), 365–378, 417 (in Russian); English translation in Soviet J. Contemp. Math. Anal. **15** (1980), 30–43.

[4] S.A. Antonyan, *Retracts in the category of G-spaces*, Bull. Acad. Polon. Sci. Sér. Sci. Math. **28** (1980), 613–618 (1981) (in Russian).

[5] S.A. Antonyan, *Equivariant generalization of Dugundji's theorem*, Mat. Zametki **38** (1985), 608–616, 636 (in Russian); English translation in Math. Notes **38** (1985), 844–848.

[6] S. Antonian, *An equivariant theory of retracts*, Aspects of Topology, London Math. Soc. Lecture Note Ser., Vol. 93, Cambridge University Press, Cambridge (1985), 251–269.

[7] S.A. Antonyan, *A characterization of equivariant absolute extensors and the equivariant Dugundji theorem*, Houston J. Math. **31** (2005), 451–462 (electronic).

[8] V.I. Arnol'd, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Grundlehren Math. Wiss., Vol. 250, Springer (1983).

[9] Z. Balanov and S. Brodsky, *Krasnosel'skiĭ's comparison principle and continuation of equivariant mappings*, Functional Analysis, no. 23, Ul'yanovsk. Gos. Ped. Inst., Ul'yanovsk (1984), 18–31 (in Russian).

[10] Z. Balanov, M. Farzamirad and W. Krawcewicz, *Symmetric systems of van der Pol equations*, Topol. Methods Nonlinear Anal. **27** (2006), 29–90.

[11] Z. Balanov, M. Farzamirad, W. Krawcewicz and H. Ruan, *Applied equivariant degree, part II: Symmetric Hopf bifurcation for functional differential equations*, Discrete Contin. Dyn. Syst. Ser. A **16** (2006).

[12] Z. Balanov and W. Krawcewicz, *Remarks on the equivariant degree theory*, Topol. Methods Nonlinear Anal. **13** (1999), 91–103.

[13] Z. Balanov, W. Krawcewicz and H. Ruan, *Applied equivariant degree, part III: Global symmetric Hopf bifurcation for functional differential equations*, Proc. Latv. Acad. Sci. Sect. B Nat. Exact Appl. Sci. **59** (2005), 234–240.

[14] Z. Balanov, W. Krawcewicz and H. Ruan, *Applied equivariant degree, part I: An axiomatic approach to primary degree*, Discrete Contin. Dyn. Syst. Ser. A **15** (2006), 983–1016.

[15] Z. Balanov, W. Krawcewicz and H. Ruan, *Hopf bifurcation in a symmetric configuration of lossless transmission lines*, Nonlinear Anal. Real World Applications **8** (2007), 1144–1170.

[16] Z. Balanov, W. Krawcewicz and H. Ruan, *G.E. Hutchinson's delay logistic system with symmetries and spatial diffusion*, Nonlinear Anal. Real World Applications **9** (2008), 154–182.

[17] Z. Balanov, W. Krawcewicz and H. Steinlein, *Reduced $SO(3) \times S^1$-equivariant degree with applications to symmetric bifurcation problems*, Nonlinear Anal. **47** (2001), 1617–1628.

[18] Z. Balanov, W. Krawcewicz and H. Steinlein, *$SO(3) \times S^1$-equivariant degree with applications to symmetric bifurcation problems: The case of one free parameter*, Topol. Methods Nonlinear Anal. **20** (2002), 335–374.

[19] Z. Balanov, W. Krawcewicz, H. Steinlein, *Applied Equivariant Degree*, AIMS Series on Differential Equations & Dynamical Systems, American Institute of Mathematical Sciences (December 2006).

[20] Z. Balanov and A. Kushkuley, *Comparison principle for equivariant maps*, Abstracts Amer. Math. Soc. **17** (1988), 98.

[21] Z. Balanov, A. Kushkuley and P. Zabreiko, *Geometric methods in degree theory for equivariant maps*, Preprint No. 137/1990, Ruhr University, Bochum (1990).

[22] T. Bartsch, *Topological Methods for Variational Problems with Symmetries*, Lecture Notes in Math., Vol. 1560, Springer-Verlag, Berlin (1993).

[23] A.O. Barut and R. Rączka, *Theory of Group Representations and Applications*, World Scientific Publishing Co., Singapore (1986).

[24] Yu.G. Borisovich and T.N. Fomenko, *Homological methods in the theory of periodic and equivariant maps*, Global Analysis and Mathematical Physics, Voronezh University Press, Voronezh (1987), 3–25 (in Russian); English translation in Global Analysis – Studies and Applications, III, Lecture Notes in Math., Vol. 1334, Springer-Verlag, Berlin (1988), 21–41.

[25] K. Borsuk, *Über die Zerlegung einer euklidischen n-dimensionalen Vollkugel in n Mengen*, Verhandlungen des Internationalen Mathematiker Kongresses Zürich 1932, II. Band: Sektions-Vorträge, Orel Füssli, Zürich (1932), 192–193.

[26] G.E. Bredon, *Introduction to Compact Transformation Groups*, Academic Press, New York–London (1972).

[27] T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*, Springer-Verlag, New York–Berlin (1985).

[28] D. Chan and I. Melbourne, *A geometric characterization of resonance in Hopf bifurcation from relative equilibria*, Physica D **234** (2007), 98–104.

[29] C. Chevalley, *Theory of Lie Groups. I*, Princeton University Press, Princeton (1946).

[30] P. Chossat, Y. Demay and G. Iooss, *Interaction de modes azimutaux dans le problème de Couette–Taylor*, Arch. Ration. Mech. Anal. **99** (1987), 213–248.

[31] P. Chossat and R. Lauterbach, *Methods in Equivariant Bifurcations and Dynamical Systems*, Advanced Series in Nonlinear Dynamics, Vol. 15, World Scientific Publishing Co., Inc., River Edge, NJ (2000).

[32] P. Chossat, R. Lauterbach and I. Melbourne, *Steady-state bifurcation with* $\mathrm{O}(3)$*-symmetry*, Arch. Ration. Mech. Anal. **113** (1990), 313–376.

[33] S.-N. Chow and J.K. Hale, *Methods of Bifurcation Theory*, Springer-Verlag, New York (1982).

[34] S.-N. Chow and J. Mallet-Paret, *Integral averaging and bifurcation*, J. Differential Equations **26** (1977), 112–159.

[35] S.-N. Chow and J. Mallet-Paret, *The Fuller index and global Hopf bifurcation*, J. Differential Equations **29** (1978), 66–85.

[36] S.-N. Chow, J. Mallet-Paret and J.A. Yorke, *Global Hopf bifurcation from a multiple eigenvalue*, Nonlinear Anal. **2** (1978), 753–763.

[37] P.E. Conner and E.E. Floyd, *Differentiable Periodic Maps*, Springer-Verlag, Berlin (1964).

[38] E.N. Dancer, *Symmetries, degree, homotopy indices and asymptotically homogeneous problems*, Nonlinear Anal. **6** (1982), 667–686.

[39] E.N. Dancer, *A new degree for* S^1*-invariant gradient mappings and applications*, Ann. Inst. H. Poincaré Anal. Non Linéaire **2** (1985), 329–370.

[40] E.N. Dancer and J.F. Toland, *Real transformations with polynomial invariants*, J. Geom. Phys. **19** (1996), 99–122.

[41] E.N. Dancer and J.F. Toland, *The index change and global bifurcation for flows with first integrals*, Proc. London Math. Soc. **66** (1993), 539–567.

[42] E.N. Dancer and J.F. Toland, *Equilibrium states in the degree theory of periodic orbits with a first integral*, Proc. London Math. Soc. **63** (1991), 569–594.

[43] T. tom Dieck, *Transformation Groups*, Walter de Gruyter & Co., Berlin (1987).

[44] B.A. Dubrovin, A.T. Fomenko and S.P. Novikov, *Modern Geometry – Methods and Applications. Part II. The Geometry and Topology of Manifolds*, Springer-Verlag, New York (1985).

[45] G. Dylawerski, *An* S^1*-degree and* S^1*-maps between representation spheres*, Algebraic Topology and Transformation Groups, T. tom Dieck, ed., Lecture Notes in Math., Vol. 1361, Springer-Verlag, Berlin (1988), 14–28.

[46] G. Dylawerski, K. Gęba, J. Jodel and W. Marzantowicz, *An* S^1*-equivariant degree and the Fuller index*, Ann. Polon. Math. **52** (1991), 243–280.

[47] S. Eilenberg, *On a theorem of P.A. Smith concerning fixed points for periodic transformations*, Duke Math. J. **6** (1940), 428–437.

[48] R. Engelking, *General Topology*, PWN – Polish Scientific Publishers, Warsaw (1977).

[49] L.H. Erbe, W. Krawcewicz, K. Gęba and J. Wu, *S^1-degree and global Hopf bifurcation theory of functional differential equations*, J. Differential Equations **98** (1992), 277–298.

[50] L.H. Erbe, W. Krawcewicz and J. Wu, *Leray–Schauder degree for semilinear Fredholm maps and periodic boundary value problems of neutral equations*, Nonlinear Anal. **15** (1990), 747–764.

[51] B. Fiedler, *Global Hopf bifurcation in porous catalysts*, Equadiff 82 (Würzburg, 1982), Lecture Notes in Math., Vol. 1017, Springer-Verlag, Berlin (1983), 177–184.

[52] B. Fiedler, *An index for global Hopf bifurcation in parabolic systems*, J. Reine Angew. Math. **359** (1985), 1–36.

[53] B. Fiedler, *Global Bifurcation of Periodic Solutions with Symmetry*, Lecture Notes in Math., Vol. 1309, Springer-Verlag, Berlin (1988).

[54] A.T. Fomenko and D.B. Fuks, *A Course in Homotopic Topology*, Nauka, Moscow (1989) (in Russian).

[55] T.N. Fomenko, *Algebraic properties of some cohomological invariants of equivariant mappings*, Mat. Zametki **50** (1991), 108–117, 160 (in Russian); English translation in Math. Notes **50** (1991), 731–737 (1992).

[56] H.I. Freedman, *Deterministic Mathematical Models in Population Ecology*, Marcel Dekker, New York (1980).

[57] F.B. Fuller, *An index of fixed point type for periodic orbits*, Amer. J. Math. **89** (1967), 133–148.

[58] W. Fulton and J. Harris, *Representation Theory: A First Course*, Graduate Texts in Math., Vol. 129, Springer-Verlag, New York (1991).

[59] K. Gęba, *Degree for gradient equivariant maps and equivariant Conley index*, Topological Nonlinear Analysis, II (Frascati, 1995), Progr. Nonlinear Differential Equations Appl., Vol. 27, Birkhäuser Boston, Boston (1997), 247–272.

[60] K. Gęba, W. Krawcewicz and J.H. Wu, *An equivariant degree with applications to symmetric bifurcation problems. I. Construction of the degree*, Proc. London. Math. Soc. (3) **69** (1994), 377–398.

[61] K. Gęba and W. Marzantowicz, *Global bifurcation of periodic solutions*, Topol. Methods Nonlinear Anal. **1** (1993), 67–93.

[62] K. Gęba, I. Massabò and A. Vignoli, *Generalized topological degree and bifurcation*, Nonlinear Functional Analysis and Its Applications (Maratea, 1985), NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., Vol. 173, Reidel, Dordrecht (1986), 55–73.

[63] M. Golubitsky and W.F. Langford, *Pattern formation and bistability in flow between counterrotating cylinders*, Phys. D **32** (1988), 362–392.

[64] M. Golubitsky and D.G. Schaeffer, *Singularities and Groups in Bifurcation Theory, Vol. I*, Applied Mathematical Sciences, Vol. 51. Springer-Verlag, New York (1985).

[65] M. Golubitsky and I.N. Stewart, *Hopf bifurcation in the presence of symmetry*, Arch. Ration. Mech. Anal. **87** (1984/85), 107–165.

[66] M. Golubitsky and I.N. Stewart, *Hopf bifurcation with dihedral group symmetry: coupled nonlinear oscillators*, Multiparameter Bifurcation Theory (Arcata, CA, 1985), Contemp. Math., Vol. 56, Amer. Math. Soc., Providence (1986), 131–137.

[67] M. Golubitsky, I.N. Stewart and D.G. Schaeffer, *Singularities and Groups in Bifurcation Theory, Vol. II*, Applied Mathematical Sciences, Vol. 69, Springer-Verlag, New York (1988).

[68] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Applied Mathematical Sciences, Vol. 42, Springer-Verlag, New York (1983).

[69] S. Guo and J.S.W. Lamb, *Equivariant Hopf bifurcation for neutral functional differential equations*, Proc. Amer. Math. Soc., in press.

[70] K.P. Hadeler and G. Bocharov, *Where to put delays in population models, in particular in the neutral case*, Can. Appl. Math. Q. **11** (2003), 159–173.

[71] J.K. Hale, *Theory of Functional Differential Equations*, second edition, Applied Mathematical Sciences, Vol. 3, Springer-Verlag, New York (1977).

[72] D. Husemoller, *Fibre Bundles*, McGraw-Hill Book Co., New York (1966).

[73] G.E. Hutchinson, *An Introduction to Population Ecology*, Yale University Press, New Haven (1978).

[74] E. Ihrig and M. Golubitsky, *Pattern selection with* $\mathrm{O}(3)$ *symmetry*, Phys. D **13** (1984), 1–33.

[75] G. Iooss, *Bifurcation of Maps and Applications*, North-Holland Mathematics Studies, Vol. 36, North-Holland Publishing Co., Amsterdam (1979).

[76] J. Ize, *Bifurcation theory for Fredholm operators*, Mem. Amer. Math. Soc. **7** (174) (1976).

[77] J. Ize, *Topological bifurcation*, Topological Nonlinear Analysis, Progr. Nonlinear Differential Equations Appl., Vol. 15, Birkhäuser Boston, Boston (1995), 341–463.

[78] J. Ize, I. Massabò and A. Vignoli, *Degree theory for equivariant maps. I*, Trans. Amer. Math. Soc. **315** (1989), 433–510.

[79] J. Ize, I. Massabò and A. Vignoli, *Degree theory for equivariant maps, the general* S^1*-action*, Mem. Amer. Math. Soc. **100** (1992), no. 481.

[80] J. Ize and A. Vignoli, *Equivariant degree for Abelian actions. II. Index computations*, Topol. Methods Nonlinear Anal. **7** (1996), 369–430.

[81] J. Ize and A. Vignoli, *Equivariant Degree Theory*, de Gruyter Series in Nonlinear Analysis and Applications, Vol. 8, Walter de Gruyter & Co., Berlin (2003).

[82] J. Jaworowski, *Extensions of* G*-maps and Euclidean* G*-retracts*, Math. Z. **146** (1976), 143–148.

[83] J. Jaworowski, *An equivariant extension theorem and* G*-retracts with a finite structure*, Manuscripta Math. **35** (1981), 323–329.

[84] K. Kawakubo, *The Theory of Transformation Groups*, The Clarendon Press, Oxford University Press, New York (1991).

[85] A.A. Kirillov, *Elements of the Theory of Representations*, Grundlehren Math. Wiss., Band 220, Springer-Verlag, Berlin (1976).

[86] K. Komiya, *Fixed point indices of equivariant maps and Möbius inversion*, Invent. Math. **91** (1988), 129–135.

[87] C. Kosniowski, *Equivariant cohomology and stable cohomotopy*, Math. Ann. **210** (1974), 83–104.

[88] M.A. Krasnosel'skiĭ, *On computation of the rotation of a vector field on the* n*-dimensional sphere*, Dokl. Akad. Nauk SSSR (N.S.) **101** (1955), 401–404 (in Russian).

[89] M.A. Krasnosel'skiĭ and P.P. Zabreĭko, *Geometrical Methods of Nonlinear Analysis*, Grundlehren Math. Wiss., Band 263, Springer-Verlag, Berlin (1984).

[90] W. Krawcewicz, T. Spanily and J. Wu, *Hopf bifurcation for parametrized equivariant coincidence problems and parabolic equations with delays*, Funkcial. Ekvac. **37** (1994), 415–446.

[91] W. Krawcewicz, P. Vivi and J. Wu, *Computational formulae of an equivariant degree with applications to symmetric bifurcations*, Nonlinear Stud. **4** (1997), 89–119.

[92] W. Krawcewicz, P. Vivi and J. Wu, *Hopf bifurcations of functional differential equations with dihedral symmetries*, J. Differential Equations **146** (1998), 157–184.

[93] W. Krawcewicz and J. Wu, *Theory of Degrees with Applications to Bifurcations and Differential Equations*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley & Sons, New York (1997).

[94] W. Krawcewicz and J. Wu, *Theory and applications of Hopf bifurcations in symmetric functional differential equations*, Nonlinear Anal. **35** (1999), 845–870.

[95] W. Krawcewicz, J. Wu and H. Xia, *Global Hopf bifurcation theory for condensing fields and neutral equations with applications to lossless transmission problems*, Canad. Appl. Math. Q. **1** (1993), 167–220.

[96] W. Krawcewicz and H. Xia, *An analytic definition of the equivariant degree*, Izv. Vyssh. Uchebn. Zaved. Mat. **6** (1996), 37–53 (in Russian); English translation in Russian Math. (Iz. VUZ) **40** (1996), 34–49.

[97] K. Kuratowski, *Topology, Vol. II*, Academic Press, New York; PWN – Polish Scientific Publishers, Warsaw (1968).

[98] A. Kushkuley and Z. Balanov, *A comparison principle and extension of equivariant maps*, Manuscripta Math. **83** (1994), 239–264.

[99] A. Kushkuley and Z. Balanov, *Geometric Methods in Degree Theory for Equivariant Maps*, Lecture Notes in Math., Vol. 1632, Springer-Verlag, Berlin (1996).

[100] J.S.W. Lamb and I. Melbourne, *Normal form theory for relative equilibria and relative periodic solutions*, Trans. Amer. Math. Soc. **359** (2007), 4537–4556.

[101] J.S.W. Lamb, I. Melbourne and C. Wulff, *Hopf bifurcation from relative periodic solutions; secondary bifurcation from meandering spirals*, J. Difference Equ. Appl. **12** (2006), 1127–1145.

[102] J.S.W. Lamb, I. Melbourne and C. Wulff, *Bifurcation from periodic solutions with spatiotemporal symmetry, including resonances and mode interactions*, J. Differential Equations **191** (2003), 377–407.

[103] R. Lashof, *The equivariant extension theorem*, Proc. Amer. Math. Soc. **83** (1981), 138–140.

[104] R. Levins, *Evolution in communities near equilibrium, Ecology and Evolution of Communities*, Belknap Press of Harvard University Press, Cambridge, MA (1975), 16–50.

[105] L.G. Lewis Jr., J.P. May, M. Steinberger and J.E. McClure, *Equivariant Stable Homotopy Theory*, Lecture Notes in Math., Vol. 1213, Springer-Verlag, Berlin (1986).

[106] J.L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications, Vol. 1*, Dunod, Paris (1968).

[107] M. Madirimov, *On Jaworowski's theorem on the extension of equivariant mappings*, Uspekhi Mat. Nauk **39** (3) (1984) (237), 237–238 (in Russian); English translation in Russian Math. Surveys **39** (3) (1984), 209–210.

[108] M. Madirimov, *Dimension and Retracts in the Theory of Topological Transformation Groups*, Fan, Tashkent (1986) (in Russian).

[109] P.C. Magnusson, G.C. Alexander and V.K. Tripathi, *Transmission Lines and Wave Propagation*, third edition, CRC Press, Boca Raton, FL (1992).

[110] J. Mallet-Paret, *Generic periodic solutions of functional differential equations*, J. Differential Equations **25** (1977), 163–183.

[111] J. Mallet-Paret and J.A. Yorke, *Snakes: oriented families of periodic orbits, their sources, sinks, and continuation*, J. Differential Equations **43** (1982), 419–450.

[112] L.N. Mann, *Finite orbit structure on locally compact manifolds*, Michigan Math. J. **9** (1962), 87–92.

[113] J.E. Marsden and M. McCracken, *The Hopf Bifurcation and Its Applications*, Applied Mathematical Sciences, Vol. 19, Springer-Verlag, New York (1976).

[114] W. Marzantowicz and C. Prieto, *Computation of the equivariant* 1*-stem*, Nonlinear Anal. **63** (2005), 513–524.

[115] C. McCrory, *Stratified general position*, Algebraic and Geometric Topology (Santa Barbara, CA, 1977), Lecture Notes in Math., Vol. 664, Springer-Verlag, Berlin (1978), 142–146.

[116] I. Melbourne, *Symmetry and symmetry breaking in dynamical systems*, Encycl. Math. Phys., Vol. 5, J.-P. Francoise, G.L. Naber and S.T. Tsou, eds., Elsevier, Oxford (2006), pp. 184–190.

[117] M.C. Memory, *Bifurcation and asymptotic behavior of solutions of a delay-differential equation with diffusion*, SIAM J. Math. Anal. **20** (1989), 533–546.

[118] M.C. Memory, *Stable and unstable manifolds for partial functional-differential equations*, Nonlinear Anal. **16** (1991), 131–142.

[119] M.C. Memory, *Invariant manifolds for partial functional-differential equations*, Mathematical Population Dynamics (New Brunswick, NJ, 1989), Lecture Notes in Pure and Appl. Math., Vol. 131, Marcel Dekker, New York (1991), 223–232.

[120] K. Morita, *On the dimension of normal spaces. I*, Japan J. Math. **20** (1950), 5–36.

[121] L. Nirenberg, *Comments on nonlinear problems*, Mathematische (Catania), Vol. 36 (1981), 109–119.

[122] R.D. Nussbaum, *A global bifurcation theorem with applications to functional differential equations*, J. Funct. Anal. **19** (1975), 319–338.

[123] R.D. Nussbaum, *Global bifurcation of periodic solutions of some autonomous functional differential equations*, J. Math. Anal. Appl. **55** (1976), 699–725.

[124] R.D. Nussbaum, *A Hopf global bifurcation theorem for retarded functional differential equations*, Trans. Amer. Math. Soc. **238** (1978), 139–164.

[125] J.C.F. de Oliveira, *Hopf bifurcation for functional differential equations*, Nonlinear Anal. **4** (1980), 217–229.

[126] G. Peschke, *Degree of certain equivariant maps into a representation sphere*, Topology Appl. **59** (1994), 137–156.

[127] P. Rabier, *Topological degree and theorem of Borsuk for general covariant mappings with applications*, Nonlin. Anal. **16** (1991), 399–420.

[128] P.H. Rabinowitz, *Some global results for nonlinear eigenvalue problems*, J. Funct. Anal. **7** (1971), 487–513.

[129] R.H. Rand and P.J. Holmes, *Bifurcation of periodic motions in two weakly coupled van der Pol oscillators*, Internat. J. Non-Linear Mech. **15** (1980), 387–399.

[130] R.L. Rubinsztein, *On the equivariant homotopy of spheres*, Dissertationes Math. (Rozprawy Mat.) **134** (1976), 48 pp.

[131] Y. Rudyak, *On Thom Spectra, Orientability, and Cobordism*, Springer Monographs in Mathematics, Springer-Verlag, Berlin (1998).

[132] S. Rybicki, *A degree for S^1-equivariant orthogonal maps and its applications to bifurcation theory*, Nonlinear Anal. **23** (1994), 83–102.

[133] S. Rybicki, *Applications of degree for S^1-equivariant gradient maps to variational nonlinear problems with S^1-symmetries*, Topol. Methods Nonlinear Anal. **9** (1997), 383–417.

[134] S. Rybicki, *Degree for equivariant gradient maps*, Milan J. Math. **73** (2005), 103–144.

[135] G.B. Segal, *Equivariant stable homotopy theory*, Actes du Congrès International des Mathématiciens (Nice, 1970), Tome 2, Gauthier-Villars, Paris (1971), 59–63.

[136] J.M. Smith, *Models in Ecology*, Cambridge University Press, Cambridge (1974).

[137] S.L. Sobolev, *Some Applications of Functional Analysis in Mathematical Physics*, Translations of Mathematical Monographs, Vol. 90, American Mathematical Society, Providence, RI (1991).

[138] H. Steinlein, *Borsuk's antipodal theorem and its generalizations and applications: a survey*, Méth. topologiques en analyse nonlinéaries, A. Granas, ed., Sémin. Math. Sup., Vol. 95, Montréal (1985), 166-235.

[139] R.E. Stong, *Notes on Cobordism Theory*, Princeton University Press, Princeton; University of Tokyo Press, Tokyo (1968).

[140] D.W. Storti and R.H. Rand, *Dynamics of two strongly coupled relaxation oscillators*, SIAM J. Appl. Math. **46** (1986), 56–67.

[141] H. Ulrich, *Fixed Point Theory of Parametrized Equivariant Maps*, Lecture Notes in Math., Vol. 1343, Springer-Verlag, Berlin (1988).

[142] N.Ja. Vilenkin and A.U. Klimyk, *Representation of Lie Groups and Special Functions. Recent Advances*, Kluwer Academic Publishers Group, Dordrecht (1995).

[143] È.B. Vinberg, *Compact Lie Groups*, Izdat. Moskov. Univ., Moscow (1967) (in Russian).

[144] F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Graduate Texts in Mathematics, Vol. 94, Springer-Verlag, New York–Berlin (1983).

[145] J. Wu, *Theory and Applications of Partial Functional-Differential Equations*, Applied Mathematical Sciences, Vol. 119, Springer-Verlag, New York (1996).

[146] J. Wu, *Symmetric functional-differential equations and neural networks with memory*, Trans. Amer. Math. Soc. **350** (1998), 4799–4838.

[147] C. Wulff, J.S.W. Lamb and I. Melbourne, *Bifurcation from relative periodic solutions*, Ergodic Theory Dynam. Systems **21** (2001), 605–635.

[148] H. Xia, *Equivariant degree and global Hopf bifurcation for NFDEs with symmetry*, Ph.D. Thesis, University of Alberta, Edmonton (1994).

[149] P.P. Zabreĭko, *On the theory of periodic vector fields*, Vestnik Yaroslav. Univ. Vyp. **2** (1973), 24–30 (in Russian).

[150] E.C. Zeeman, *Seminar on Combinatorial Topology*, Institut des Hautes Etudes Scientifiques, Paris (1963).

[151] D.P. Zhelobenko, *Compact Lie Groups and Their Representations*, Nauka, Moscow (1970) (in Russian).

[152] D.P. Zhelobenko and A.I. Shtern, *Representations of Lie Groups*, Nauka, Moscow (1983) (in Russian).

This page intentionally left blank

CHAPTER 2

Nonautonomous Differential Systems in Two Dimensions

Roberta Fabbri and Russell Johnson

Dipartimento di Sistemi e Informatica, Università di Firenze, Italy
E-mails: *{johnson, fabbri}@dsi.unifi.it*

Luca Zampogni

Dipartimento di Matematica e Informatica, Università di Perugia, Italy
E-mail: *zampoglu@dipmat.unipg.it*; *zampogni@math.unifi.it*

Contents

1. Introduction . . . 135
2. Preliminaries . . . 142
3. The projective flow . . . 152
4. Algebro-geometric Sturm–Liouville coefficients . . . 180
    4.1. Basic results on the Sturm–Liouville operator . . . 184
    4.2. The Sturm–Liouville equation as a nonautonomous differential system . . . 193
    4.3. The projective flow in the Sturm–Liouville setting . . . 197
    4.4. The inversion problem . . . 219
5. Genericity of exponential dichotomy . . . 250
    5.1. Proof of Theorem 5.1 (Density result) . . . 255
References . . . 262

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 4
Edited by F. Battelli and M. Fečkan

© 2008 Elsevier B.V. All rights reserved

This page intentionally left blank

1. Introduction

In the last four decades, the field of nonautonomous dynamical systems has developed from a loose collection of results and techniques into a coherent area with a recognizable and characteristic set of methods and problems. Among the more classical areas of pure and applied mathematics which it has influenced are spectral theory, control theory, bifurcation theory, functional differential equations, monotone systems, and others as well.

There is a certain amount of conceptual overlap between the fields of nonautonomous dynamical systems and stochastic dynamical systems. In the latter area, one places emphasis on the use of dynamical techniques to study stochastic differential equations. This point of view has been applied systematically and with particular success by L. Arnold and his school. It is natural to think of the two fields – nonautonomous dynamics and stochastic dynamics – as occupying complementary parts of the theory of differential equations with time-varying coefficients. It may not be unreasonable to use a term such as "random dynamics" to cover their union.

Be that as it may, the following considerations allow one to distinguish the two fields.

First of all, a nonautonomous dynamical system is frequently defined in terms of the solutions of a differential equation whose coefficients vary with time, but in a nonstochastic way. The limiting case is that of equations with periodic time dependence; however, one is also interested in equations whose coefficients satisfy Bohr almost periodicity, Birkhoff recurrence, or still weaker recurrence properties. One can express the same idea in a different way by saying that the time dependence of the coefficients exhibits randomness which is "between" periodicity and stochasticity.

A second distinction occurs on the methodological level. Ergodic techniques, for example the Oseledets theory, have an important place in the study of stochastic dynamical systems. This is true in the study of the nonautonomous ones, as well. However, in the latter area, techniques of topological dynamics are of equal importance. There is a reason for this. Namely, a Bebutov-type construction can often be applied to the right-hand side of a nonautonomous differential equation, giving rise to a compact hull. The solutions of the original equation together with those of the equations in the hull then form a continuous dynamical system. Of course one has a vast array of results available to study this system; they are associated with the names of Bronstein, Ellis, Furstenberg, Sell, and many other researchers. These results, however, cannot usually be applied to a stochastic dynamical system since, for such a system, the equivalent of the "hull" has only a measurable structure (but see [69]).

We will explain the Bebutov construction later on. Before doing so, we make some simple observations concerning *autonomous* differential equations, with the purpose of putting the succeeding discussion into perspective.

It is well known that the solutions of an autonomous differential equation

$$x' = f(x), \quad x \in \mathbb{R}^n \tag{1.1}$$

define a local flow on $\mathbb{R}^n$ if f is Lipschitz continuous. They define a global flow if f is complete; that is, if for each $x_0 \in \mathbb{R}^n$, the solution $x(t)$ of (1.1) satisfying $x(0) = x_0$ is defined on the whole real line $-\infty < t < \infty$. This simple observation is very useful if one wishes to study the asymptotic behavior of the solutions of Eq. (1.1). In particular, a coherent and articulated answer to the question "where do the solutions go?" can be given using the concepts of α- and ω-limit sets, attractors and repellers, basins of attraction/repulsion, etc.

The situation is quite different if f is time dependent: $f = f(t, x)$, $(t \in \mathbb{R}, x \in \mathbb{R}^n)$. It is of course true that, if f is T-periodic for some $T > 0$, then one can introduce the period map defined by solving

$$x' = f(t, x) \tag{1.2}$$

for initial conditions $x(0) \in \mathbb{R}^n$ and continuing them for time T. One thus obtains a discrete flow or discrete dynamical system, at least if all solutions of (1.2) exist for all $t \in \mathbb{R}$. The study of such a system is in many (but not all) ways analogous to that of the continuous dynamical system defined by the solutions of the autonomous equation (1.1).

If f is aperiodic, however, the period map is not available. The fact is that such objects as ω-limit sets do not in general provide much information about the asymptotic behavior of the solutions of (1.2). One can adjoin the equation $t' = 1$ to (1.2), thereby obtaining an autonomous equation on $\mathbb{R} \times \mathbb{R}^n$ which defines a dynamical system on that space. However, all α- and ω-limit sets are empty. Moreover, the t-recurrence properties of f are not reflected in the extended vector field. The addition of the equation $t' = 1$ to (1.2) therefore is unsatisfying if one wishes to use dynamical methods to study Eq. (1.2).

Fortunately, it turns out that very often a simple and direct construction – first studied systematically by Bebutov ([12], also [143]) – permits one to use the full panoply of methods of dynamical systems theory to study the solutions of a nonautonomous differential equation (1.2). See, e.g., [33,94,143]. Let us illustrate the Bebutov construction in the case of a linear system

$$x' = A(t)x, \quad x \in \mathbb{R}^n, t \in \mathbb{R} \tag{1.3}$$

where A is a bounded, uniformly continuous function with values in the space $\mathbb{M}_n$ of $n \times n$ real matrices. Introduce the space $\mathcal{C} = C(\mathbb{R}, \mathbb{M}_n) = \{C : \mathbb{R} \to \mathbb{M}_n \mid C$ is continuous and bounded$\}$, equipped with the topology of uniform convergence on compact subsets of $\mathbb{R}$. For each $t \in \mathbb{R}$, let $\tau_t : \mathcal{C} \to \mathcal{C}$ be the translation:

$$\tau_t(C)(\cdot) = C(t + \cdot) \quad (C \in \mathcal{C}).$$

Then $\{\tau_t \mid t \in \mathbb{R}\}$ defines a flow on $\mathcal{C}$. That is $\tau_0 = \mathrm{id}y$; $\tau_t \circ \tau_s = \tau_{t+s}$ for all $t, \in \mathbb{R}$; and $\tau : \mathcal{C} \times \mathbb{R} \to \mathcal{C}$: $\tau(C, t) = \tau_t(C)$ is continuous. We call $\{\tau_t\}$ the *Bebutov flow* on $\mathcal{C}$.

Now view $A(\cdot)$ as an element of $\mathcal{C}$. Let $\Omega = \mathrm{cls}\{\tau_t(A) \mid t \in \mathbb{R}\}$. Then Ω is compact since A is uniformly continuous. Let us outline a proof of this fact. Let $\{t_n\} \subset \mathbb{R}$ be a sequence. We first show that the sequence $\{\tau_{t_n}(A)\}$ contains a subsequence which converges in $\mathcal{C}$. For, let $k \geqslant 1$, let $I_k = [-k, k] \subset \mathbb{R}$. By the Ascoli–Arzelá theorem, there is a subsequence $t_{n_1}, \ldots, t_{n_j}, \ldots$ of $\{t_n\}$ such that $\{t_{n_j}(A)\}$ converges uniformly on I_k. We label this sequence t_j^k $(j = 1, 2, \ldots)$, then set $s_k = t_k^k$ $(k = 1, 2, \ldots)$. Clearly $\{\tau_{s_k}(A)\}$ converges in $\mathcal{C}$. If $\{\omega_n\} \subset \Omega$ is a sequence, we approximate $\omega_n \in \mathcal{C}$ by $\tau_{t_n}(A)$ for an appropriate $t_n \in \mathbb{R}$, then extract a convergent subsequence of $\{\omega_n\}$ by applying the Ascoli–Arzelá theorem and a Cantor diagonal argument to the sequence $\{\tau_{t_n}(A)\}$.

Note further that Ω admits a dense orbit. Also Ω is invariant with respect to the Bebutov flow. If $\omega \in \Omega$, then it makes sense to speak of the differential equation $x' = \omega(t)x$. A notational device is convenient. Define $\tilde{A} : \Omega \to \mathbb{M}_n$: $\tilde{A}(\omega) = \omega(0)$. Then $\tilde{A}(\tau_t(\omega)) = \omega(t)$ for all $\omega \in \Omega$, $t \in \mathbb{R}$. Moreover if we write $\omega_0 = A \in \Omega$, then $\tilde{A}(\tau_t(\omega_0)) = A(t)$, and so the differential equation $x' = \tilde{A}(\tau_t(\omega_0))x$ coincides with (1.3).

Let us abuse notation and write $A : \Omega \to \mathbb{M}_n$: $A(\omega) = \omega(0)$. We consider the family of equations

$$x' = A\big(\tau_t(\omega)\big)x \quad (x \in \mathbb{R}^n, \omega \in \Omega). \tag{1.3_ω}$$

This family is indexed by the compact, Bebutov-invariant set $\Omega \subset \mathcal{C}$, and contains the initial equation (1.3) = (1.3_{ω_0}).

Let $\omega \in \Omega$, and let $\Phi_\omega(t)$ be the fundamental matrix solution of (1.3_ω). The collection $\{\Phi_\omega(t) \mid \omega \in \Omega\}$ defines a map $\Phi : \Omega \times \mathbb{R} \to \mathrm{GL}(n, \mathbb{R}) : (\omega, t) \mapsto \Phi_\omega(t)$. The map Φ is a *cocycle*; i.e., it satisfies the relation

$$\Phi_{\omega\cdot t}(s)\Phi_\omega(t) = \Phi_\omega(t+s) \quad (\omega \in \Omega; t, s \in \mathbb{R}).$$

The cocycle relation follows from uniqueness of solutions, which holds for Eqs. (1.3_ω). Now, using the cocycle relation, one verifies the following assertion.

For each $t \in \mathbb{R}$ define $\tilde{\tau} : \Omega \times \mathbb{R}^n \to \Omega \times \mathbb{R}^n$ via the formula

$$\tilde{\tau}_t(\omega, x) = \big(\tau_t(\omega), \Phi_\omega(t)x\big);$$

then $(\Omega \times \mathbb{R}^n, \{\tilde{\tau}_t\})$ is a flow. Note that the first coordinate on the right does not depend on x, and that the second coordinate on the right is linear in x.

The flow $\{\tilde{\tau}_t\}$ is a nonautonomous analogue of the one-parameter group $\{\hat{\tau}_t = \mathrm{e}^{At}\}$ generated by the linear autonomous equation $x' = Ax$. The systematic study of the flow $\{\tilde{\tau}_t\}$ on $\Omega \times \mathbb{R}^n$ has led to new insights concerning the properties of solutions of linear nonautonomous ODEs. Experience has shown that, although some problems involving linear systems (1.3) can be treated without introducing the hull Ω and the flow $\{\tilde{\tau}_t\}$, others seem to require it for effective study.

Next, consider a nonlinear nonautonomous equation (1.2). Suppose that, for each compact set $K \subset \mathbb{R}^n$, f is uniformly continuous and uniformly bounded on $\mathbb{R} \times K$. For simplicity, suppose also that f is uniformly Lipschitz on each set $\mathbb{R} \times K$, in the sense that there is a constant M_K such that

$$\big\| f(t,x) - f(t,y) \big\| \leqslant M_K \|x - y\|$$

for all $t \in \mathbb{R}, x, y \in K$.

In this case, one can carry out a Bebutov-type construction to obtain a compact, translation-invariant set $\Omega = \Omega_f$ of functions $\omega(t, x)$, each having the properties satisfied by f which are listed above. The flow is defined by $\tau_t(\omega)(\cdot, x) = \omega(\cdot + t, x)$. Furthermore, if $\omega \in \Omega, x_0 \in \mathbb{R}$, and $x(t)$ is the solution of the equation $x' = \omega(t, x)$ which satisfies $x(0) = x_0$, then the locally defined mappings given by $\tilde{\tau}_t(\omega, x_0) = (\tau_t(\omega), x(t))$ define a local flow on $\Omega \times \mathbb{R}^n$. This flow is global if all solutions of all equations $x' = \omega(t, x)$ exist on $(-\infty, \infty)$. In this case, one obtains a dynamical system on $\Omega \times \mathbb{R}^n$, the behavior of whose orbits reflects the behavior of the solutions of the various equations $x' = \omega(t, x)$, and also reflects the t-recurrence properties of f. These last properties are encoded in the translation flow $\{\tau_t\}$ on Ω.

One can now systematically apply the methods of topological dynamics, ergodic theory etc. to study the solutions of the family of equations $x' = \omega(t, x)$ $(\omega \in \Omega)$. This procedure not infrequently leads to insights concerning the original equation which might not otherwise have come to light. Of course the remark made above for linear systems applies to the nonlinear ones, as well.

Constructions of Bebutov-type can also be carried out for wide classes of parabolic or hyperbolic partial differential equations and for functional differential equations with time-varying coefficients. As a consequence, one can effectively use “dynamical” methods to study such equations. See, e.g., [4,119,124,125] for examples illustrating this approach. We will not give more details here because our attention in this paper will be focused on linear ordinary differential equations, usually in dimension 2.

We turn to an outline of the contents of this paper. Consider the linear, nonautonomous, two-dimensional differential equation

$$x' = A(t)x, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ t \in \mathbb{R}. \tag{1.4}$$

We are interested in functions $A(\cdot)$ which depend aperiodically on time; on the other hand we will impose certain recurrence properties on A. In fact, we will always assume that A is stationary and ergodic in a sense to be defined later (Section 2). Often we will assume that A is a Birkhoff recurrent function of t, or even that A is Bohr almost periodic. Even under such restrictive hypotheses, we will see that the solutions of (1.4) exhibit remarkable properties which have no analogues in the case of periodic systems.

In Section 3, we will study the minimal subsets of the projective flow defined by Eq. (1.4). We will always assume that $A(\cdot)$ is Birkhoff recurrent and usually that it is Bohr almost periodic. A characterization of the minimal subsets of the projective flow constitutes one approach to generalizing to recurrent systems (1.4) the standard Floquet theory for periodic linear systems.

In more detail, let $\mathcal{C} = C(\mathbb{R}, \mathbb{M}_2)$ with the compact-open topology, and let $\Omega \subset \mathcal{C}$ be the result of the Bebutov construction applied to $A(\cdot)$. By carrying out the various steps discussed previously, we obtain a family of differential systems

$$x' = A\big(\tau_t(\omega)\big)x; \tag{1.4$_\omega$}$$

the original equation (1.4) is in this family.

Let $\mathbb{P}$ be the set of lines through the origin in $\mathbb{R}^2$. Thus $\mathbb{P}$ is the one-dimensional real projective space and is topologically a circle. The family (1.4$_\omega$) induces a flow on $\Sigma =$

$\Omega \times \mathbb{P}$, in the following way. Let $\Phi_\omega(t)$ be the fundamental matrix solution of (1.4_ω). For each $t \in \mathbb{R}$, set

$$\hat{\tau}_t(\omega, l) = \big(\tau_t(\omega), \Phi_\omega(t)l\big) \quad (\omega \in \Omega, l \in \mathbb{P}).$$

It is easy to verify that $(\Sigma, \{\hat{\tau}_t\})$ is indeed a flow. We set the goal of classifying those compact, nonempty, $\{\hat{\tau}_t\}$ – invariant subsets $M \subset \Sigma$ which are minimal in the sense that, for each $m \in M$, the orbit $\{\hat{\tau}_t(m) \mid t \in \mathbb{R}\}$ is dense in M.

Such minimal sets are quite easy to understand if A is periodic. In the case when A is Bohr almost periodic, however, a minimal subset $M \subset \Sigma$ may have properties quite unlike those occurring when A is periodic. In the recurrent case, there are even more possibilities for M. To simplify the discussion, we will assume, with few exceptions, that A is almost periodic.

One striking phenomenon which may be present in the projective flow of an almost periodic family (1.4_ω) is that of almost automorphy in the sense of Bochner and Veech [20,150,151]. We will see that various types of almost automorphic, non-almost periodic minimal sets can be realized as invariant subsets of the projective flow defined by an almost periodic family (1.4_ω). Moreover, we will see that these almost automorphic minimal sets illustrate in a concrete way such phenomena as Marcus–Moore disconjugacy, unbounded mean motion, and strange nonchaotic attractors. Examples illustrating these phenomena are of “M-V type”, where the letters indicate the names of Millionscikov and Vinograd. They constructed the first examples of almost periodic families (1.4_ω) containing equations which are not almost reducible [108,152]. It later turned out that these and related examples give rise to almost automorphic subsets of Σ as mentioned above.

We note further that the projective flow $(\Sigma, \{\hat{\tau}_t\})$ of an almost periodic family (1.4_ω) may contain an almost automorphic minimal set which is not of M-V type. For instance, there are families (1.4_ω) for which the Favard separation property is violated and for which each Eq. (1.4_ω) admits a bounded solution. In these circumstances, it may happen that Σ contains an almost automorphic, non-almost periodic minimal set M; for reasons which we will discuss, M cannot be of M-V type. Examples illustrating this phenomenon were constructed in [158] and, later and independently, in [73]. The phenomenon itself was nicely clarified by Ortega and Tarallo [126].

There are other classes of almost periodic families (1.4_ω) whose projective flows admit minimal sets with no analogue in the periodic case. For example, Σ may admit a unique “invariant two-sheet”. Examples of this sort were predicted in [83] and were written down in [88,123].

Furthermore, there are situations in which the projective flow Σ is itself minimal but is quite unlike those arising when Σ is minimal and A is periodic (these last are all of Kronecker type). We will consider several classes of examples. One consists of families (1.4_ω) for which Σ is minimal, uniquely ergodic, and is a proximal extension of Ω ([61], also [87]). Another class is illustrated by a recent example of Bjerklov [13,14], for which Σ is minimal and the family (1.4_ω) is weakly hyperbolic. Such examples exhibit the property of “Li–Yorke chaos”. See [17] for more information on this topic.

We close this introductory discussion of the contents of Section 3 by noting that the measure-theoretic (ergodic) structure of the projective flow of Eqs. (1.4_ω) is more completely understood than is the topologico-dynamical structure. See [122,121] for a discussion and classification of the ergodic measures of the projective flow when $n = 2$. See also [9] and references therein for a complete theory of the ergodic measures of the projective flow for n-dimensional linear systems.

In Section 4, we discuss some recent developments in the spectral theory and inverse spectral theory of the classical Sturm–Liouville operator. The material reviewed here is found in the Ph.D. thesis of one of us [157] and in [96,95]. The methods of nonautonomous differential equations are systematically exploited both in the formulation of and in the proofs of the results we consider. Other techniques, in particular certain constructions from the classical theory of algebraic curves, are also found to be very useful.

The theory we present is motivated in part by the progress in the inverse spectral theory of the one-dimensional Schrödinger operator which took place in the 1970s. In fact, it was discovered that there are basic and nonobvious relations between the inverse spectral theory of the Schrödinger operator, the theory of algebraic curves, and the solutions of a certain hierarchy of nonlinear partial differential equations. One of the equations of this hierarchy is the K-dV (Korteweg–de Vries) equation

$$\frac{\partial u}{\partial t} = 3u\frac{\partial u}{\partial x} - \frac{1}{2}\frac{\partial^3 u}{\partial x^3}$$

of shallow water-wave theory.

It turns out that the inverse spectral theory of the Sturm–Liouville operator also can be studied using the theory of algebraic curves, in particular certain facts concerning generalized Jacobians. To work out the relation between these a priori separate fields, it is convenient to first develop certain aspects of the *direct* spectral theory of the Sturm–Liouville operator. To this end, one uses facts from the theory of nonautonomous differential equations. Among the “nonautonomous” facts which are found useful are the following:

- The equivalence between the resolvent of a Sturm–Liouville operator and the existence of an exponential dichotomy for the associated nonautonomous linear system.
- The relation between the spectrum of a Sturm–Liouville operator and the rotation number of the associated nonautonomous linear system.
- The Kotani theory for Sturm–Liouville type operators.

We will discuss these matters. We will in addition give a brief discussion of the Gilbert–Pearson theory [60] in the context of the Sturm–Liouville operator. This theory has been found to be of basic importance in the study of the one-dimensional Schrödinger operator and we think it is important to indicate how the Gilbert–Pearson theory can be extended to the more general Sturm–Liouville operator.

Our inverse spectral theory takes these and other facts as a starting point. Making use of the Weyl m-functions, one then can make a connection between a Sturm–Liouville operator having certain spectral properties, and a generalized Jacobian. One is able to give a concrete description of the operators having these spectral properties. We will discuss all these matters in Section 4.

In analogy to the relation between the Schrödinger operator and the K-dV equation, there is a close connection between the special Sturm–Liouville operator defined by the equation

$$\left(-\frac{\mathrm{d}^2}{\mathrm{d}x^2}+1\right)\varphi = Ey(x)\varphi$$

and the Camassa–Holm equation

$$4\frac{\partial u}{\partial t} = \frac{\partial^3 u}{\partial x^2 \partial t} + 2u\frac{\partial^3 u}{\partial x^3} + 4\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x^2} - 24u\frac{\partial u}{\partial x}.$$

Using our methods, it is possible to work out an ample family of quasi periodic solutions of the Camassa–Holm equation and of other equations in the Camassa–Holm hierarchy. These facts are discussed in [157,156]; for earlier work see [2,3,58]. For reasons of space we will not take up this theme in the present paper.

In the final Section 5 of this paper, we consider the following basic question. Let $(\Omega,\mathbb{R})$ be a fixed minimal flow. Let $\mathrm{sl}(2,\mathbb{R})$ be the set of all 2×2 real matrices with trace zero. Let $\mathcal{C}_0 = C(\Omega, \mathrm{sl}(2,\mathbb{R}))$ be the set of continuous functions $A:\Omega\to \mathrm{sl}(2,\mathbb{R})$ with the uniform topology induced by some fixed norm on $\mathrm{sl}(2,\mathbb{R})$. Each $A\in\mathcal{C}_0$ defines a family of two-dimensional differential systems of type (1.4_ω):

$$x' = A\big(\tau_t(\omega)\big)x, \quad x\in\mathbb{R}^2$$

where $\operatorname{tr} A(\cdot) = 0$. We ask "how many" matrix functions $A(\cdot)$ give rise to a family (1.4_ω) which exhibits an exponential dichotomy. In particular, we ask if the set of $\mathrm{sl}(2,\mathbb{R})$-valued functions $\{A\}$ for which the family (1.4_ω) has an exponential dichotomy is open and dense in $\mathcal{C}_0$.

Related questions have been considered by Millionscikov [109], Novikov [118], Cong [116,117] and many other authors. We will not give an encyclopedia of the results which have been obtained. Rather, we will consider in some detail a particular result concerning quasi-periodic minimal flows, in the C^0-category.

It is known that, for certain minimal flows $(\Omega,\mathbb{R})$ with compact metric phase space Ω, the set $\mathcal{C}_{\mathrm{ED}} = \{A\in\mathcal{C}_0 \mid \text{the family } (1.4_\omega) \text{ admits an exponential dichotomy}\}$ is *not* dense in $\mathcal{C}_0$. In fact, this is true if Ω is the circle and the flow $\{\tau_t\}$ is defined by a rigid rotation; this case occurs exactly when (1.4_ω) is a family of periodic equations with fixed period and the coefficient matrices are related one to another by translations. It also occurs whenever the image of the so-called Schwarzmann homomorphism is isomorphic to $\{0\}$ or $\mathbb{Z}$. This condition in ensured if the first integer Čech cohomology group $H^1(\Omega,\mathbb{Z})$ is, modulo torsion, equal to $\{0\}$ or $\mathbb{Z}$.

However, if Ω is the phase space of a nonperiodic, almost periodic minimal flow, then the image of the Schwarzmann homeomorphism is dense in $\mathbb{R}$ (it coincides with the so-called frequency module $\mathcal{M}$ of $(\Omega,\mathbb{R})$). In this case, the above criterion breaks down. If $(\Omega,\mathbb{R})$ is a "limit-periodic" flow, then it can be shown that $\mathcal{C}_{\mathrm{ED}}\subset\mathcal{C}_0$ is open and dense. This is not a completely straightforward matter, but arguments of Moser [110] can be adapted to do the job.

In the case when $(\Omega, \mathbb{R})$ is quasi-periodic – that is, when Ω is a k-dimensional torus and the flow is determined by a Kronecker winding – we consider the question as to whether $\mathcal{C}_{\mathrm{ED}}$ is dense in $\mathcal{C}_0$ (we note that $\mathcal{C}_{\mathrm{ED}}$ is open in $\mathcal{C}_0$ by standard perturbation results for exponential dichotomies; see, e.g., [29]). We will discuss a result to the effect that, if the Kronecker winding is defined by a Liouville frequency vector, then $\mathcal{C}_{\mathrm{ED}}$ is indeed dense in $\mathcal{C}_0$. This result is proved using methods worked out in [77,44,50].

This result can be viewed as included in a line of research which aims at determining the nature of the set of elements $A \in \mathcal{C}_0$ for which the family (1.4_ω) has a positive Lyapunov exponent. See, e.g., [117,100,15], and for definitive information [19]. We note, however, that positivity of the Lyapunov exponent is much weaker than the dichotomy property.

On the other hand, the result we discuss should be viewed as in some sense orthogonal to the impressive body of theorems worked out in, e.g., [98,39,102,135], and papers by other authors. These theorems concern the case when A is an analytic (or at least highly regular) function on the torus $\Omega = \mathbb{T}^k$, and the frequency vector satisfies a Diophantine condition. Several of these results consist of statements to the following effect: in a suitable parameterized family of analytic quasi-periodic differential systems with Diophantine frequency vector, there is a measure-theoretically large set of parameter values for which the system is reducible. That is, there is an analytic quasi-periodic change of variables $x = B(\omega \cdot t)y$ such that $x' = A(\omega \cdot t)x$ is transformed into $y' = A_0 y$ where A_0 is a constant matrix. Interestingly enough, there is often a topologically large set of parameter values for which reducibility does *not* hold. Instead, an almost reducibility property holds; moreover the projective flow corresponding to such an almost reducible system seems to have interesting properties.

Our results concerning density of $\mathcal{C}_{\mathrm{ED}} \subset \mathcal{C}_0$ would appear to break down when one passes from the C^0-category to the C^1-category, at least when $k \geqslant 3$. (Our methods may produce a density result in the C^r-category if r is less than 1.) The question of C^1-density of the dichotomy property (for any kind of frequency vector, Liouville or Diophantine or “other”) has not been clarified, and only nondefinitive indications concerning the real facts of the case are at this point available [23,26,30]. We mention that, very recently, A. Avila, J. Bochi and D. Damanik have proved the density result for all frequencies in the C^0 category. Their methods seem specific to this category, though they can also deal with certain non-almost periodic base flows. We thank Prof. David Damanik for this information.

We finish the introduction by listing some notation which will be used in this paper.

If $k \geqslant 1$, let $\mathbb{T}^k = \mathbb{R}^k/\mathbb{Z}^k$ denote the standard k-torus. Thus $\mathbb{T}^1$ is the circle, which can be identified with $\{\mathrm{e}^{2\pi \mathrm{i}\psi} \mid 0 \leqslant \psi < 1\}$.

We use $|\cdot|$ to indicate all norms which arise, and $\langle , \rangle$ to indicate all inner products.

2. Preliminaries

We review basic concepts and fix notation which will be used in the succeeding sections.

Let us first recall the definition of a (continuous, real) *flow*. Let Ω be a metric space. For each $t \in \mathbb{R}$, let $\tau_t : \Omega \to \Omega$ be a homeomorphism. Suppose that the following conditions are satisfied:

- $\tau_0(\omega) = \omega$, for all $\omega \in \Omega$;

- $\tau_{t+s}(\omega) = \tau_t \circ \tau_s(\omega)$ for all $t, s \in \mathbb{R}$, $\omega \in \Omega$;
- the map $\tau : \Omega \times \mathbb{R} \to \Omega : (\omega, t) \mapsto \tau_t(\omega)$ is continuous.

We call the pair $(\Omega, \{\tau_t\})$ a *flow* or *dynamical system* on Ω. We will also use the notation $(\Omega, \mathbb{R})$ to denote a flow, or Ω resp. $\{\tau_t\}$ if there is no doubt concerning the identity of $\{\tau_t\}$ resp. Ω. We will often write $\omega \cdot t$ in place of $\tau_t(\omega)$.

Here is some standard terminology concerning flows. If $\omega \in \Omega$, then the *orbit* through ω is $\{\omega \cdot t \mid t \in \mathbb{R}\}$. If $\Omega_* \subset \Omega$, then Ω_* is *invariant* if for every $\omega \in \Omega_*$, the orbit through ω is contained in Ω_*. Let $(\tilde{\Omega}, \{\tilde{\tau}_t\})$ be another flow. A continuous map $\pi : \tilde{\Omega} \to \Omega$ is a *flow homomorphism* if $\pi \circ \tilde{\tau}_t = \tau_t \circ \pi$ for all $t \in \mathbb{R}$. If π is also homeomorphism of $\tilde{\Omega}$ onto Ω, then π is a *flow isomorphism*.

Some flows of interest to us arise via the Bebutov construction, which we described in the Introduction. We repeat it briefly. Let $n \geqslant 1$, and let $\mathcal{C} = \{C\colon \mathbb{R} \to \mathbb{M}_n \mid C \text{is continuous and bounded}\}$. Endow $\mathcal{C}$ with the compact-open topology. For each $t \in \mathbb{R}$ and $C \in \mathcal{C}$, set $\tau_t(C)(\cdot) = C(t + \cdot)$; then τ_t is the t-translation. Then $(\mathcal{C}, \{\tau_t\})$ is a flow. If $A \in \mathcal{C}$ is uniformly continuous, define $\Omega = \operatorname{cls}\{\tau_t(A) \mid t \in \mathbb{R}\}$. Then Ω is compact and is invariant with respect to the Bebutov flow $\{\tau_t\}$. We call Ω the *hull* of A. As we pointed out in the Introduction, the map $\tilde{A} : \Omega \to \mathbb{M}_n\colon \tilde{A}(\omega) = \omega(0)$ is continuous and has the property that, if $\omega_0 = A$, then $\tilde{A}(\omega_0 \cdot t) = A(t)$ for all $t \in \mathbb{R}$. In this case it is convenient to abuse notation and write A instead of $\tilde{A}$.

Certain specific types of flow will be important in what follows; we review the relevant definitions. Let Ω be a *compact* metric space, and let $(\Omega, \{\tau_t\})$ be a flow.

First of all, the flow $(\Omega, \{\tau_t\})$ is called *minimal* or *Birkhoff recurrent* if for each $\omega \in \Omega$, the orbit $\{\omega \cdot t \mid t \in \mathbb{R}\}$ is dense in Ω. If Ω is the hull of a uniformly continuous function $A : \mathbb{R} \to \mathbb{M}_n$, then Ω is minimal if and only if A satisfies the Birkhoff recurrence condition, namely if to each compact set $K \subset \mathbb{R}$ and each $\varepsilon > 0$ there corresponds a relatively dense set $\{t_k \mid k \in \mathbb{Z}\} \subset \mathbb{R}$ such that, if $k \in \mathbb{Z}$, then $|A(t + t_k) - A(t)| \leqslant \varepsilon$ for all $t \in K$.

A special class of minimal flows is determined by those which are *Bohr almost periodic* or simply *almost periodic*. By definition, a flow $(\Omega, \mathbb{R})$ satisfies this condition if it is minimal and if there is a metric d on Ω (which is compatible with the topology on Ω) such that, if $\omega_1, \omega_2 \in \Omega$, then $d(\omega_1 \cdot t, \omega_2 \cdot t) = d(\omega_1, \omega_2)$ for all $t \in \mathbb{R}$. If Ω is the hull of a uniformly continuous function $A \in \mathcal{C}$, then $(\Omega, \mathbb{R})$ is Bohr almost periodic if and only if for each $\varepsilon > 0$ there is a relatively dense set $\{t_k \mid k \in \mathbb{Z}\}$ such that, if $k \in \mathbb{Z}$, then $|A(t + t_k) - A(t)| \leqslant \varepsilon$ for all $t \in \mathbb{R}$. One sometimes calls $\{t_k\}$ a set of almost periods of A. See, e.g., [5,21,54].

It turns out that, if $(\Omega, \mathbb{R})$ is minimal and Bohr almost periodic, then Ω may be given the structure of a compact Abelian topological group with dense subgroup $\mathbb{R}$ in such a way that the group structure and the flow structure are compatible. In fact, if $\omega_0 \in \Omega$ and $t_1, t_2 \in \mathbb{R}$, define $(\omega_0 \cdot t_1) * (\omega_0 \cdot t_2) = \omega_0 \cdot (t_1 + t_2)$. It is an exercise to show that $*$ extends to a group operation on Ω with the stated properties. Moreover ω_0 is the identity with respect to this group structure.

To each minimal almost periodic flow $(\Omega, \mathbb{R})$ there corresponds the so-called frequency module $\mathcal{M} = \mathcal{M}_\Omega$. This object is constructed in the following way. View Ω as a compact Abelian topological group with dense subgroup $\mathbb{R}$. Let $\hat{\Omega}$ be the group of continuous homomorphisms (*characters*) $\chi : \Omega \to \mathbb{T}^1$. Since $\mathbb{R} \subset \Omega$ is a dense subgroup, a character χ is determined by its restriction $\chi_{\mathbb{R}}$ to $\mathbb{R}$, and this restriction has the form $\chi_{\mathbb{R}}(t) = e^{2\pi i\lambda t}$ for some $\lambda \in \mathbb{R}$. The smallest $\mathbb{Z}$-submodule of $\mathbb{R}$ containing all the numbers λ arising in the this way is the *frequency module* of Ω. Due to the fact that Ω is a compact metric Abelian group, the frequency module is a countable subgroup of $\mathbb{R}$.

We mention the following examples. Let $\Omega = \mathbb{T}^k$ be the k-torus. Let $\gamma_1, \dots \gamma_k$ be real numbers which are independent over the rational field $\mathbb{Q}$. If $\omega \in \Omega$, write $\omega = (\exp 2\pi \mathrm{i}\psi_1, \dots, \exp 2\pi \mathrm{i}\psi_k)$, then define for each $t \in \mathbb{R}$

$$\tau_t(\omega) = \bigl(\exp 2\pi \mathrm{i}(\psi_1 + \gamma_1 t), \dots, \exp 2\pi \mathrm{i}(\psi_k + \gamma_k t)\bigr).$$

Below we will sometimes write $\psi = (\psi_1, \dots, \psi_k)$ and also write $\tau_t(\psi) = \psi + \gamma t = (\psi_1 + \gamma_1 t, \dots, \psi_k + \gamma_k t)$. One says that $(\Omega, \{\tau_t\})$ is a *quasi-periodic* flow or a *Kronecker* flow. One also speaks of a Kronecker winding. It is minimal and almost periodic. The numbers $\gamma_1, \dots, \gamma_g$ are called the *frequencies* of the flow.

Second, set $T_0 = 1, T_1 = n_1, \dots, T_k = n_k T_{k-1}, \dots$ where $n_1 > 1, \dots, n_k > 1, \dots$ are integers. Consider the subgroup $\mathcal{M}$ of $\mathbb{R}$ generated by $T_0^{-1}, \dots, T_k^{-1}, \dots$

$$\mathcal{M} = \left\{ \frac{l_0}{T_0} + \frac{l_1}{T_1} + \cdots + \frac{l_r}{T_r} \;\middle|\; r \geqslant 1; l_0, \dots, l_r \in \mathbb{Z} \right\}.$$

Then $\mathcal{M}$ is the frequency module of an almost periodic minimal flow. Such a flow is called *limit periodic*. If Ω is the hull of an almost periodic function $A : \mathbb{R} \to \mathbb{M}_n$, then $(\Omega, \mathbb{R})$ is limit periodic if and only if A is the uniform limit of a sequence of periodic functions.

Still another type of flow is distinguished by the property of *almost automorphy*. Let $(\Omega_0, \mathbb{R})$ be an almost periodic minimal flow. Let $(\Omega, \mathbb{R})$ be another minimal flow, and let $\pi : \Omega \to \Omega_0$ be a flow homomorphism. Note that π is necessarily surjective since Ω_0 is minimal. Suppose that there is a point $\omega_0 \in \Omega_0$ such that the inverse image $\pi^{-1}(\omega_0)$ reduces to a single point in Ω. Then $(\Omega, \mathbb{R})$ is called an almost automorphic extension of $(\Omega_0, \mathbb{R})$, or an *almost automorphic* flow. If Ω is the hull of an uniformly continuous bounded function $A : \mathbb{R} \to \mathbb{M}_n$, then $(\Omega, \mathbb{R})$ is almost automorphic if (but not only if) the function A satisfies the Bochner condition. Namely, if $\{t_k\} \subset \mathbb{R}$ is a sequence such that $|t_k| \to \infty$ and $A(t_k + \cdot) \to A(\cdot)$ in $\mathcal{C}$, then $A(-t_k + \cdot) \to A(\cdot)$ in $\mathcal{C}$. If the Bochner condition is satisfied, then there is an almost periodic minimal flow $(\Omega_0, \mathbb{R})$ together with a flow homomorphism $\pi : \Omega \to \Omega_0$ such that $\pi^{-1}\pi(A) = \{A\}$ [151]. A function $A(\cdot)$ which satisfies the Bochner condition is said to be an *almost automorphic function*.

Almost automorphic minimal sets were studied by Veech [150,151]. See [144,145,154] for a review of their properties and a discussion of their presence in various applied situations.

Next we review some basic definitions from ergodic theory. Let Ω be a compact metric space and let $(\Omega, \{\tau_t\})$ be a flow.

A regular Borel probability measure ν on Ω is called $\{\tau_t\}$-*invariant*, or simply *invariant* if there is no doubt concerning the identity of $\{\tau_t\}$, if for each Borel subset $B \subset \Omega$ and each $t \in \mathbb{R}$, there holds $\nu(\tau_t(B)) = \nu(B)$. An invariant measure ν is called *ergodic* if it is indecomposable in the sense that, whenever $B \subset \Omega$ is a Borel set such that $\mu(\tau_t(B) \triangle B) = 0$ for all $t \in \mathbb{R}$ ($\triangle =$ symmetric difference of sets), then either $\nu(B) = 0$ or $\nu(B) = 1$.

One can show that, if Ω is a compact metric space and $(\Omega, \mathbb{R})$ is a flow, then there exists an ergodic measure ν on Ω. In fact, one can first use a classical construction of Krylov and

Bogoliubov to determine an *invariant* measure ν on Ω; see, e.g., [55,113]. We will make use of this construction below. Second, one notes that the set of invariant measures forms a weak-$*$ compact, convex set of the vector space of all Borel regular signed measures on Ω. By the Krein–Mil'man Theorem this set admits an extreme point. It is easy to see that such an extreme point is ergodic.

If Ω is a compact metric space and ν is a Borel regular measure on Ω, then the *topological support* $\operatorname{Supp}\nu$ of ν is by definition the complement in Ω of the largest open set $U \subset \Omega$ which has zero ν-measure. This concept is well defined and $\operatorname{Supp}\nu$ is compact. If ν is $\{\tau_t\}$-invariant, then $\operatorname{Supp}\nu$ is invariant. Observe that if $U \subset \operatorname{Supp}\nu$ is relatively open then $\nu(U) > 0$. If $(\Omega, \mathbb{R})$ is minimal then $\operatorname{Supp}\nu = \Omega$ for every invariant measure ν on Ω.

It may happen that a flow $(\Omega, \mathbb{R})$ admits a unique invariant measure ν. This is the case, for example, if $(\Omega, \mathbb{R})$ is minimal and almost periodic; in fact, the normalized Haar measure ν on Ω is the unique invariant measure. If $(\Omega, \mathbb{R})$ admits a unique invariant measure ν, then ν is ergodic. If $(\Omega, \mathbb{R})$ is minimal and admits a unique invariant (ergodic) measure ν, then $(\Omega, \mathbb{R})$ is said to be *strictly ergodic*.

We state a version of the Birkhoff Ergodic Theorem which is adequate for our purposes. For a formulation and proof of the general Birkhoff theorem see [113]. We include a refinement of the usual Birkhoff theorem in point (ii) of Theorem 2.1; for a proof of this statement see, e.g., [55].

THEOREM 2.1. *Let Ω be a compact metric space with flow $\{\tau_t\}$, and let ν be an ergodic measure on Ω.*

(i) *Let $f \in L^1(\Omega, \nu)$, then for ν-a.a. $\omega \in \Omega$ the time averages $\lim_{t\to\infty} \frac{1}{t}\int_0^t f(\omega\cdot s)\,\mathrm{d}s$ and $\lim_{t\to-\infty} \frac{1}{t}\int_0^t f(\omega\cdot s)\,\mathrm{d}s$ exist and equal $\int_\Omega f\,\mathrm{d}\nu$.*

(ii) *If Ω admits a unique ergodic measure ν and if f is continuous, then*

$$\lim_{|t|\to\infty} \frac{1}{t}\int_0^t f(\omega\cdot s)\,\mathrm{d}s = \int_\Omega f\,\mathrm{d}\nu$$

for all $\omega \in \Omega$, and the limit is uniform in Ω.

Next we state some facts concerning linear nonautonomous differential systems. We adopt the following point of view. Let Ω be a compact metric space, and let $(\Omega, \mathbb{R})$ be a flow. Let $A : \Omega \to \mathbb{M}_n$ be a continuous function. Consider the family of differential systems

$$x' = A(\omega\cdot t)x, \quad x \in \mathbb{R}^n,\ \omega \in \Omega. \tag{2.1$_\omega$}$$

We have already seen that, via the Bebutov construction, a linear differential equation $x' = A(t)x$ with uniformly continuous coefficient matrix $A(\cdot)$ gives rise to a family of this type. We will study the solutions of Eqs. (2.1$_\omega$) using the methods of topological dynamics and ergodic theory.

Let $\omega \in \Omega$, and let $\Phi_\omega(t)$ be the fundamental matrix solution of (2.1$_\omega$). The collection $\{\Phi_\omega(t) \mid \omega \in \Omega\}$ defines a map $\Phi : \Omega \times \mathbb{R} \to \mathrm{GL}(n, \mathbb{R}) : (\omega, t) \mapsto \Phi_\omega(t)$. We noted in Section 1 that Φ is a cocycle and that the maps $\tilde{\tau}_t : \Omega \times \mathbb{R}^n \to \Omega \times \mathbb{R}^n : \tilde{\tau}_t(\omega, x) = (\tau_t(\omega), \Phi_\omega(t)x)$ define a flow on $\Omega \times \mathbb{R}^n (t \in \mathbb{R})$.

Let $\mathbb{P}^{n-1}$ be the $(n-1)$ dimensional space of lines through the origin in $\mathbb{R}^n$. For each $t \in \mathbb{R}$, define $\hat{\tau} : \Omega \times \mathbb{P}^{n-1} \to \Omega \times \mathbb{P}^{n-1}$: $\hat{\tau}_t(\omega, l) = (\tau_t(\omega), \Phi_\omega(t)l)$. Here $l \in \mathbb{P}^{n-1}$ is acted on in the natural way by $\Phi_\omega(t) \in GL(n, \mathbb{R})$. Using the cocycle identity again, one can show that $(\Omega \times \mathbb{P}^{n-1}, \{\hat{\tau}_t\})$ is a flow. This is the *projective flow* defined by Eqs. (2.1_ω). If one wishes to study the "angular behavior" of the solutions of Eqs. (2.1_ω), the projective flow is one natural object to study. We will write

$$\Sigma = \Omega \times \mathbb{P}^{n-1}, \quad \pi : \Sigma \to \Omega : (\omega, l) \mapsto \omega$$

for $\omega \in \Omega, l \in \mathbb{P}^{n-1}$. The projection π is a flow homomorphism.

Let us now briefly discuss some key concepts concerning the family (2.1_ω). These concepts will later be specialized to the case of interest in this paper, namely that when $n = 2$.

Fix $\omega \in \Omega$. A number $\beta \in \mathbb{R}$ is called a *Lyapunov exponent* of Eq. (2.1_ω) if there is a nonzero vector $x_0 \in \mathbb{R}^n$ such that

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \ln\left|\Phi_\omega(t)x_0\right| = \beta.$$

It is well-known that each fixed Eq. (2.1_ω) admits finitely many Lyapunov exponents $\beta_1 < \cdots < \beta_k$ where $1 \leqslant k \leqslant n$. Moreover, there is a filtration of $\mathbb{R}^n$ consisting of subspaces $\{0\} = V_0 \subset V_1 \subset \cdots \subset V_k = \mathbb{R}^n$ such that

$$V_r = \{0\} \cup \left\{0 \neq x_0 \in \mathbb{R}^n \;\middle|\; \overline{\lim_{t\to\infty}} \frac{1}{t} \ln\left|\Phi_\omega(t)x_0\right| \leqslant \beta_r\right\} \quad (1 \leqslant r \leqslant k).$$

See, e.g., [8]. Difficulties are caused by the fact that neither the Lyapunov exponents nor the filtration need to be continuous with respect to variation of ω.

The Oseledets theory furnishes information concerning this point and others. We do not give the most general version of the Oseledets theorem; the one we do give is, however, more than sufficient for our purposes. Let ν be an ergodic measure on Ω.

THEOREM 2.2 (Oseledets). *One can find a ν-measurable, invariant subset $\Omega_* \subset \Omega$ with $\nu(\Omega_*) = 1$, and real numbers $\beta_1 < \cdots < \beta_k$ with $1 \leqslant k \leqslant n$, with the following properties.*

(i) *If $\omega \in \Omega_*$ then for each $r = 1, 2, \ldots, k$ there is a nonzero vector $x_0 \in \mathbb{R}^n$ such that the limits $\lim_{t\to\infty} \frac{1}{t} \ln|\Phi_\omega(t)x_0|$ and $\lim_{t\to-\infty} \frac{1}{t} \ln|\Phi_\omega(t)x_0|$ both exist and equal β_r.*

(ii) *For each $\omega \in \Omega_*$ and each $r = 1, 2, \ldots, k$, the set*

$$W_r(\omega) = \{0\} \cup \left\{0 \neq x_0 \in \mathbb{R}^n \;\middle|\; \lim_{|t|\to\infty} \frac{1}{t} \ln\left|\Phi_\omega(t)x_0\right| = \beta_r\right\}$$

is a vector subspace of $\mathbb{R}^n$, of dimension $d_r \in \{1, 2, \ldots, n\}$. One has $d_1 + \cdots + d_k = n$. The mapping $\omega \mapsto W_r(\omega)$ is measurable as a function from Ω_* to the Grassmann manifold $\mathrm{Gr}(n, d_r)$ of d_r-dimensional subspaces of $\mathbb{R}^n$.

(iii) *For each* $r = 1, 2, \dots, k$, *the "measurable subbundle"*

$$W_r = \bigcup \bigl\{(\omega, x_0) \mid x_0 \in W_r(\omega), \omega \in \Omega_*\bigr\}$$

is invariant in the sense that $\Phi_\omega(t) W_r(\omega) = W_r(\omega \cdot t)$ *for all* $\omega \in \Omega_*$, $t \in \mathbb{R}$.

(iv) *For each* $\omega \in \Omega_*$, *there holds* $\beta_1 + \cdots + \beta_k = \lim_{|t|\to\infty} \frac{1}{t} \ln \det \Phi_\omega(t)$.

Point (i) defines the "Oseledets spectrum" of the family (2.1_ω). Point (iv) states that ν-a.a. systems (2.1_ω) are both forward and backward Lyapunov regular. For proofs of the Oseledets theorem see [129,8,94]. When $n = 2$ it is possible to give a quite simple proof of a strengthened version of the Oseledets Theorem. We will discuss this point later on.

Next we introduce the concept of *exponential dichotomy* [29,137]. Again let Ω be a compact metric space, and let $(\Omega, \mathbb{R})$ be a real flow. Let $A : \Omega \to \mathbb{M}_n$ be a continuous map, and consider the corresponding family of Eqs. (2.1_ω).

DEFINITIONS 2.3.

(a) The family $\{(2.1_\omega) \mid \omega \in \Omega\}$ is said to have an exponential dichotomy (ED) over Ω if there are positive constants η, δ together with a continuous, projection-valued function $P : \Omega \to \mathbb{M}_n$ (thus $P(\omega)^2 = P(\omega)$ for all $\omega \in \Omega$) such that the following estimates hold:
  (i) $|\Phi_\omega(t) P(\omega) \Phi_\omega(s)^{-1}| \leqslant \eta e^{-\delta(t-s)}$, $t \geqslant s$,
  (ii) $|\Phi_\omega(t)(I - P(\omega)) \Phi_\omega(s)^{-1}| \leqslant \eta e^{\delta(t-s)}$, $t \leqslant s$.

(b) The *dynamical* or *Sacker–Sell spectrum* of the family (2.1_ω) is $\{\lambda \in \mathbb{R} \mid$ the translated equations $x' = [-\lambda I + A(\omega \cdot t)]x$ do *not* admit an exponential dichotomy over $\Omega\}$.

If Eqs. (2.1_ω) admit an exponential dichotomy over Ω, then it is easy to see that $\Omega \times \mathbb{R}^n$ splits into a Whitney sum of the vector subbundles $W_1 = \{(\omega, x) \in \Omega \times \mathbb{R}^n \mid x \in \operatorname{Im} P(\omega)\}$ and $W_2 = \{(\omega, x) \in \Omega \times \mathbb{R}^n \mid x \in \operatorname{Ker} P(\omega)\}$. Of course it can happen that $P(\omega) = 0$ or $P(\omega) = I$. If Ω is connected then the dimension of $\operatorname{Im} P(\omega)$ does not depend on Ω.

The following theorem is due to Sacker and Sell [137,138]; see also Selgrade [142].

THEOREM 2.4. *Let* Ω *be a compact connected metric space, and let* $(\Omega, \mathbb{R})$ *be a flow. Let* $A : \Omega \to \mathbb{M}_n$ *be a continuous map.*

(a) *The dynamical spectrum of the family* (2.1_ω) *is a finite union* $[a_1, b_1] \cup \cdots \cup [a_k, b_k]$ *of compact subintervals of* $\mathbb{R}$, *where* $1 \leqslant k \leqslant n$. *Some or all of these intervals may degenerate to points.*

(b) *For each* $r = 1, 2, \dots, k$ *set*

$$W_r = \left\{ (\omega, x) \in \Omega \times \mathbb{R}^n \;\middle|\; x_0 = 0 \text{ or } \Bigl(\limsup_{t\to\infty}, \liminf_{t\to\infty}\Bigr) \frac{1}{t} \ln\bigl|\Phi_\omega(t) x_0\bigr| \in [a_r, b_r] \right.$$
$$\left. \text{and } \Bigl(\limsup_{t\to-\infty}, \liminf_{t\to-\infty}\Bigr) \frac{1}{t} \ln\bigl|\Phi_\omega(t) x_0\bigr| \in [a_r, b_r] \right\}.$$

Then W_r is a topological vector subbundle of $\Omega \times \mathbb{R}^n$. One has

$$\Omega \times \mathbb{R}^n = W_1 \oplus W_2 \oplus \cdots \oplus W_k \quad (\textit{Whitney sum}).$$

There are various relations between the "Oseledets bundles" and the "Sacker–Sell bundles", and between the Oseledets spectrum and the dynamical spectrum. See [94] for a discussion of this issue.

Let us now consider the case $n = 2$. As we will see, there is no loss of generality for present purposes in assuming that the trace $\operatorname{tr} A(\omega) = 0$ for all $\omega \in \Omega$. Thus we assume that

$$A : \Omega \to \mathrm{sl}(2, \mathbb{R})$$

is a continuous function. Let us introduce the basis $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ of $\mathrm{sl}(2, \mathbb{R})$. We expand $A(\cdot)$ with respect to this basis:

$$A = \begin{pmatrix} a & -b + c \\ b + c & -a \end{pmatrix}$$

where $a, b, c : \Omega \to \mathbb{R}$ are continuous functions.

Consider the family of equations

$$x' = A(\omega \cdot t)x, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ \omega \in \Omega. \tag{2.2$_\omega$}$$

Let us express these equations in polar coordinates $r^2 = x_1^2 + x_2^2$, $\theta = \arctan \frac{x_2}{x_1}$:

$$\frac{r'}{r} = a(\omega \cdot t) \cos 2\theta + c(\omega \cdot t) \sin 2\theta,$$

$$\theta' = b(\omega \cdot t) - a(\omega \cdot t) \sin 2\theta + c(\omega \cdot t) \cos 2\theta.$$

If we think of θ as a π-periodic angular variable on the real projective space $\mathbb{P}$, then the θ-equation gives a concrete way of studying the projective flow defined by Eqs. (2.1$_\omega$). In fact, let $l_0 \in \mathbb{P}$, $\omega_0 \in \Omega$, and suppose that l_0 is parameterized by $\theta_0 \in \mathbb{R}$. Let $\theta(t)$ satisfy the θ-equation with $\theta(0) = \theta_0$ and $\omega = \omega_0$. Then $\hat{\tau}_t(\omega_0, l_0) = \hat{\tau}_t(\omega_0, \theta_0) = (\tau_t(\omega), \theta(t))$. It is understood that $\theta(t)$ is identified with the line through the origin in $\mathbb{R}^2$ which it parametrizes.

Next let us write

$$f : \Sigma \to \mathbb{R}\colon\ f(\omega, \theta) = a(\omega) \cos 2\theta + c(\omega) \sin 2\theta,$$

$$g : \Sigma \to \mathbb{R}\colon\ g(\omega, \theta) = b(\omega) - a(\omega) \sin 2\theta + c(\omega) \cos 2\theta.$$

Then Eqs. (2.2$_\omega$) are expressed in polar coordinates r, θ as follows:

$$\frac{r'}{r} = f\bigl(\hat{\tau}_t(\omega, \theta)\bigr), \tag{2.3$_\omega$}$$

$$\theta' = g\big(\hat{\tau}_t(\omega, \theta)\big). \tag{2.4_ω}$$

Let $0 \neq x_0 \in \mathbb{R}^2$ have polar angle θ_0, and let $\theta(t)$ be the solution of (2.4_ω) satisfying $\theta(0) = \theta_0$. Let $r(t) = |x(t)|$, where $x(t)$ satisfies (2.2_ω) with $x(0) = x_0$. Then

$$\frac{1}{t} \ln r(t) = \frac{1}{t} \ln r(0) + \frac{1}{t} \int_0^t f\big(\hat{\tau}_s(\omega, \theta_0)\big)\, \mathrm{d}s. \tag{2.5}$$

Thus there is a close connection between the Lyapunov exponents of Eq. (2.2_ω) and the time averages of f. Using the Birkhoff Theorem, we can thus relate the Lyapunov exponents of Eq. (2.2_ω) to the averages of f with respect to the $\{\tau_t\}$-ergodic measures μ on Σ.

Following [78], we indicate how an analysis of the ergodic averages of f permits one to prove the Oseledets Theorem 2.2 and the Sacker–Sell Theorem 2.4 in the case $n = 2$. We sketch the necessary arguments in the case when $(\Omega, \mathbb{R})$ admits a unique invariant measure ν (which is then ergodic). As mentioned earlier, this condition is valid if Ω is minimal and almost periodic. As we go along, we will illustrate the Krylov–Bogoliubov construction, together with some simple but useful techniques.

Note first that there exists an ergodic measure μ on Σ. The image of μ under the projection $\pi : \Sigma \to \Omega$ is an invariant measure, hence equals ν. We distinguish two possibilities:

(a) $\int_\Sigma f\, \mathrm{d}\mu = 0$,

(b) $\int_\Sigma f\, \mathrm{d}\mu \neq 0$.

Here f is the function occurring in (2.3_ω).

Suppose that (a) holds. We will prove that all Lyapunov exponents of all Eqs. (2.2_ω) are zero.

PROPOSITION 2.5. *Suppose that there is an ergodic measure μ on Σ such that $\int_\Sigma f\, \mathrm{d}\mu = 0$. Let $\varepsilon > 0$ be given. Then there exists $T > 0$ such that, if $\omega \in \Omega$ and if $x_0 \in \mathbb{R}^2$ satisfies $|x_0| = 1$, then*

$$-\varepsilon \leqslant \frac{1}{t} \ln\big|\Phi_\omega(t) x_0\big| \leqslant \varepsilon \quad \textit{whenever } |t| > T.$$

PROOF. We make systematic use of Eq. (2.5), which relates the asymptotic behavior of the quantities $\frac{1}{t} \ln |\Phi_\omega(t) x_0|$ to the time averages of f. Suppose for contradiction that there are sequences $(\omega_k) \subset \Omega$, $(x_k) \subset \mathbb{R}^2$, and $(t_k) \subset \mathbb{R}$ such that $|x_k| = 1$, $|t_k| \to \infty$, and

$$\left| \frac{1}{t_k} \int_0^{t_k} f\big(\hat{\tau}_s(\omega_k, l_k)\big)\, \mathrm{d}s \right| \geqslant \varepsilon.$$

Here $l_k \in \mathbb{P}$ is the line through the origin in $\mathbb{R}^2$ containing x_k. Passing to subsequences if necessary, we can assume that $\lim_{k\to\infty} \frac{1}{t_k} \int_0^{t_k} f(\hat{\tau}_s(\omega_k, l_k))\, \mathrm{d}s$ exists. Call the limit $\Lambda(f)$. Clearly $\Lambda(f) \neq 0$.

Let $C(\Sigma) = \{h : \Sigma \to \mathbb{R} \mid h \text{is continuous}\}$. Let $\mathcal{H}$ be a countable dense $\mathbb{Q}$-vector subspace of $C(\Sigma)$. Using a Cantor diagonal argument, one can determine a subsequence (t_m) of (t_k) such that for each $h \in \mathcal{H}$ the limit

$$\lim_{m\to\infty} \frac{1}{t_m} \int_0^{t_m} h\bigl(\hat{\tau}_s(\omega_m, l_m)\bigr)\, \mathrm{d}s$$

exists. Call the limit $\Lambda(h)$ ($h \in \mathcal{H}$). It is easy to see that Λ defines a $\mathbb{Q}$-linear, bounded functional on $\mathcal{H}$, and that $|\Lambda| = 1$. This functional extends uniquely to a $\mathbb{R}$-linear functional – which we also call Λ – on $C(\Sigma)$. One sees that $|\Lambda| \leqslant 1$, and in fact $|\Lambda| = 1$ because, if we let h be the function on Σ whose only value is 1, then $\Lambda(h) = 1$. Moreover, Λ is invariant in the sense that, if $h \in C(\Sigma)$, $t \in \mathbb{R}$, and $(\hat{\tau}_t h)(\sigma) =: h(\hat{\tau}_{-t}(\sigma))$ for each $\sigma \in \Sigma$, then

$$\Lambda(\hat{\tau}_t h) = \Lambda(h).$$

By the Riesz representation theorem, there is a unique Borel regular probability measure μ_Λ on Σ such that $\Lambda(h) = \int_\Sigma h\, \mathrm{d}\mu_\Lambda$ for all $h \in C(\Sigma)$. One checks that μ_Λ is invariant. Clearly $\int_\Sigma f\, \mathrm{d}\mu_\Lambda \neq 0$.

Now let I be the weak-$*$ compact, convex subset of the dual space $C^*(\Sigma)$ which consists of the set of $\{\hat{\tau}_t\}$-invariant measures on Σ (or rather, the linear functionals which correspond to them under the Riesz theorem). Applying the Choquet theory [131] to the set I, and taking account of the fact that ergodic measures on Σ correspond to extreme points of I, one sees that there exists an ergodic measure μ_1 on Σ such that $\int_\Sigma f\, \mathrm{d}\mu_1 \neq 0$. Write $\beta = \int_\Sigma f\, \mathrm{d}\mu_1$.

Next, use the Birkhoff ergodic theorem to find a μ_1-measurable, invariant set $\Sigma_1 \subset \Sigma$ with $\mu_1(\Sigma_1) = 1$, such that, if $\sigma \in \Sigma_1$, then

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f\bigl(\hat{\tau}_s(\sigma)\bigr)\, \mathrm{d}s = \lim_{t\to-\infty} \frac{1}{t} \int_0^t f\bigl(\hat{\tau}_s(\sigma)\bigr)\, \mathrm{d}s = \beta.$$

Let $\Omega_1 = \pi(\Sigma_1)$; then Ω_1 is ν-measurable, invariant and $\nu(\Omega_1) = 1$. In a similar way, we can find a μ-measurable, invariant set $\Sigma_* \subset \Sigma$ with $\mu(\Sigma_*) = 1$, such that, if $\sigma \in \Sigma_*$, then

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f\bigl(\hat{\tau}_s(\sigma)\bigr)\, \mathrm{d}s = \lim_{t\to-\infty} \frac{1}{t} \int_0^t f\bigl(\hat{\tau}_s(\sigma)\bigr)\, \mathrm{d}s = 0.$$

Let $\Omega_* = \pi(\Sigma_*)$; then Ω_* is also ν-measurable with $\nu(\Omega_*) = 1$.

Choose $\omega \in \Omega_1 \cap \Omega_*$, then pick points $(\omega, l_1) \in \Sigma_1$ and $(\omega, l_*) \in \Sigma_*$. Clearly $l_1 \neq l_*$. Let $x_1 \in l_1$ and $x_* \in l_*$ have norm 1; then x_1 and x_* are linearly independent. Let $x_1(t) = \Phi_\omega(t)x_1$, $x_*(t) = \Phi_\omega(t)x_*$. Let $\Psi(t)$ be the matrix solution of Eq. (2.2_ω) whose columns are $x_1(t)$ and $x_*(t)$. By Liouville's formula $\det\Psi(t) = \det\Psi(0) = const.$ for all $t \in \mathbb{R}$. However, using relation (2.5), we see that $\lim_{t\to\infty} \det\Psi(t) = 0$ if $\beta < 0$ and that $\lim_{t\to-\infty} \det\Psi(t) = 0$ if $\beta > 0$. This is a contradiction. The proof of Proposition 2.5 is complete. □

Proposition 2.5 states that, if there is an ergodic measure μ on Σ for which $\int_\Sigma f\,\mathrm{d}\mu = 0$, then $\lim_{|t|\to\infty} \frac{1}{t}\ln|\Phi_\omega(t)x| = 0$, uniformly in $\omega \in \Omega$ and $x \in \mathbb{R}^2$ with $|x| = 1$. Thus all Lyapunov exponents of all Eqs. (2.2_ω) are zero, and this holds in a uniform way. Now let us describe what happens in the case when there is an ergodic measure μ_1 on Σ for which $\int_\Sigma f\,\mathrm{d}\mu_1 = \beta \neq 0$, i.e., the case (b).

First of all, using the Birkhoff ergodic theorem, we can find a μ_1-measurable, invariant set $\Sigma_1 \subset \Sigma$ with $\mu_1(\Sigma_1) = 1$ such that $\lim_{|t|\to\infty} \frac{1}{t}\int_0^t f(\hat{\tau}_s(\sigma))\,\mathrm{d}s = \beta$ for all $\sigma \in \Sigma_1$. Let $\Omega_1 = \pi(\Sigma_1)$, so that Ω_1 is ν-measurable, invariant, and $\nu(\Omega_1) = 1$. We claim that, for each $\omega_1 \in \Omega_1$, the fiber $\pi^{-1}(\omega_1) \cap \Sigma_1$ contains exactly one point. For, suppose that $(\omega_1, l_1) \in \Sigma_1$ and $(\omega_1, l_*) \in \Sigma_1$, where $l \neq l_*$. Then, using Liouville's formula and relation (2.5) as was done in the last part of the proof of Proposition 2.5, we argue to a contradiction. We thus conclude that Σ_1 is a ν-measurable invariant section in Σ when Σ is viewed as a fiber bundle over Ω.

Further reasoning of this elementary sort can be used to prove the following facts; see [78]. First, there is exactly one other ergodic measure μ_2 on Σ. Second, $\int_\Sigma f\,\mathrm{d}\mu_2 = -\beta$. Third, there is a ν-measurable, invariant section $\Sigma_2 \subset \Sigma$ such that $\mu_2(\Sigma_2) = 1$. The measurable invariant sections define measurable subbundles $W_{1,2} \subset \Omega \times \mathbb{R}^2$ of fiber dimension one.

There are now just two possibilities.

(i) The invariant measurable sections Σ_1, Σ_2 extend to continuous invariant sections of Σ. In this case, W_1 and W_2 are clearly topological subbundles of $\Omega \times \mathbb{R}^2$ of fiber dimension one. One can show that $\Omega \times \mathbb{R}^2 = W_1 \oplus W_2$ (Whitney sum). It is now easily verified that the dynamical spectrum of the family (2.4_ω) coincides with the two point set $\{-\beta, \beta\}$. See [78] for details.

(ii) The invariant sections Σ_1, Σ_2 do not extend to continuous sections of Σ. In this case, one can show that the dynamical spectrum of the family (2.4_ω) coincides with the interval $[-\beta, \beta]$. On the other hand, the Oseledets spectrum is the two point set $\{-\beta, \beta\}$. The Oseledets bundles are (ν-a.e.) the bundles W_1 and W_2. See again [78] for details.

Examples of type (ii) are particularly interesting and invariably give rise to projective flows with interesting structure. As noted in the Introduction, Millionščikov [108] and Vinograd [152] constructed examples of families (2.4_ω) for which $(\Omega, \mathbb{R})$ is an almost periodic minimal set and for which the conditions in (ii) turned out to be true. They were motivated by the search for almost periodic linear differential systems which are not almost reducible, and which are therefore not regular in the sense of Lyapunov. In what follows, we will refer to the number β as the Lyapunov exponent of the family (2.1_ω). The above discussion shows that, if $\beta = 0$, then $\lim_{|t|\to\infty} \frac{1}{t}\ln|\Phi_\omega(t)x_0| = 0$, uniformly with respect to $\omega \in \Omega$ and $x_0 \in \mathbb{R}^2$ satisfying $|x_0| = 1$.

Let us finish this section by defining and briefly discussing the rotation number for Eqs. (2.2_ω). This quantity was defined, and certain of its properties were worked out, in [89]. See also [81] for a further discussion of properties of the rotation number and for a generalization of this concept. It is related to the so-called integrated density of states [147]. A discrete version of the rotation number – the fiber rotation number – was worked out in [68].

The rotation number is most easily defined and analyzed using Eqs. (2.4_ω) for the polar angle θ. Let Ω be a compact metric space, let $(\Omega, \mathbb{R})$ be a flow, and let ν be an ergodic measure on Ω. Consider the function g in (2.4_ω): if $\omega \in \Omega$, $\theta_0 \in \mathbb{R}$, and if $\theta(t)$ is the solution of Eq. (2.4_ω) with $\theta(0) = \theta_0$, then

$$\theta(t) = \theta_0 + \int_0^t g\big(\hat{\tau}_s(\omega_0)\big)\,\mathrm{d}s. \tag{2.6}$$

We state the

DEFINITION 2.6. The *rotation number* α_ν of Eqs. (2.2_ω) with respect to the ergodic measure ν on Σ is

$$\alpha_\nu = \lim_{t\to\infty} \frac{\theta(t)}{t}.$$

Of course one has to prove that this definition makes sense, i.e. that the limit exists in an appropriate sense. It turns out that there is a ν-measurable, invariant set $\Omega_\nu \subset \Omega$ with $\nu(\Omega_\nu) = 1$ such that, if $\omega \in \Omega_\nu$, $\theta_0 \in \mathbb{P} = \mathbb{R}/\pi\mathbb{Z}$ and $\theta(t)$ is the solution of (2.4_ω) with $\theta(0) = \theta_0$, then $\lim_{|t|\to\infty} \frac{\theta(t)}{t}$ exists and does not depend on the choice of $(\omega, \theta_0) \in \Omega_\nu \times \mathbb{P}$. In fact, for each ergodic measure μ on Σ which is a lift of ν in the sense that $\pi(\mu) = \nu$, one has

$$\lim_{|t|\to\infty} \frac{\theta(t)}{t} = \int_\Sigma g\,\mathrm{d}\mu. \tag{2.7}$$

In particular, the right-hand side of (2.7) does not depend on the choice of the ergodic lift μ of ν. The proofs of these statements use (2.6), the Birkhoff ergodic theorem, and certain special arguments. See [89] for details.

It turns out that α has good continuity properties. We discuss one which turns out to be significant. A proof of the following statement can be based on an argument of Krylov–Bogoliubov type; see [89] for arguments of the appropriate type. There are various other continuity results concerning the rotation number. Some of these will be stated later when they are needed.

PROPOSITION 2.7. *Let Ω be a compact metric space, let $(\Omega, \{\tau_t\})$ be a flow, and let ν_k be a sequence of measures which are ergodic with respect to $\{\tau_t\}$. Suppose that ν is a $\{\tau_t\}$-ergodic measure on Ω, and suppose that $\nu_k \to \nu$ in the weak-$*$ sense. Then $\alpha_{\nu_k} \to \alpha_\nu$.*

3. The projective flow

Let us first recall some basic facts concerning a linear periodic system

$$x' = A(t)x, \quad x \in \mathbb{R}^n \tag{3.1}$$

where $A(t) = A(t + 2\pi)$ is a continuous function with values in $\mathbb{M}_n$. Let $\Phi(t)$ be the fundamental matrix solution of (3.1). The asymptotic behavior of the solution of (3.1) is determined by the Floquet matrix $\Phi(2\pi)$ and its iterates. For example, knowledge of the real parts of the logarithms of the eigenvalues and of the generalized eigenspaces of $\Phi(2\pi)$ permit one to determine which solutions of (3.1) exhibit exponential growth or decay, and with which rates.

We wish to understand the directional variation of the solutions of (3.1). Though the concept of "directional variation" does not have immediate significance, it can be rendered meaningful using the imaginary parts of the logarithms of the eigenvalues of $\Phi(2\pi)$ together with its generalized eigenspaces. Alternatively, one can use the Floquet "theory" to determine an invertible, 4π-periodic matrix function $L = L(t)$ such that the change of variables $x = L(t)y$ transforms (3.1) into

$$y' = Ry$$

for a constant real matrix R. Using the generalized eigenspaces and the imaginary parts of the eigenvalues of R, one can develop a theory concerning the rates with which solutions of (3.1) change direction with time, and related matters.

Suppose now that $A(\cdot)$ is nonperiodic. We seek information concerning the asymptotic behavior of the solutions of (3.1). We have seen in Section 2 that the concepts of Lyapunov exponent and exponential dichotomy allow one to develop profound theories concerning the exponential growth of the solutions of (3.1). However it is not clear how one might develop a reasonable theory of the directional variation of those solutions. In this regard, one might for example ask for an analogue of the Floquet theory in the context of nonperiodic systems (3.1). However, it is well known that one cannot hope in general to reduce (3.1) to a system with constant coefficients by mean of a bounded and boundedly invertible change of variables $x = L(t)y$. The search for an analogue of the period matrix $\Phi(2\pi)$ is even more problematic. In fact, it seems difficult to elaborate this concept in a useful way even in the case when $A(\cdot)$ is almost periodic and there exist "almost periods" with a definite meaning.

We therefore try to formulate a theory of the directional variation of the solutions of (3.1) in the nonperiodic case from another point of view. Suppose that a Bebutov-type construction has been carried out, with the following results: a compact metric space Ω, a flow $\{\tau_t\}$ on Ω, a continuous function $A : \Omega \to \mathbb{M}_n$, and a family of equations

$$x' = A\big(\tau_t(\omega)\big)x \tag{3.1_ω}$$

which contains (3.1). Let $\mathbb{P}$ be the $(n-1)$-dimensional projective space of lines through the origin in $\mathbb{R}^n$. Let $\Sigma = \Omega \times \mathbb{P}$. We define a flow $\{\hat{\tau}_t\}$ on Σ as was done in Section 2. Namely, let $\Phi_\omega(t)$ be the fundamental matrix solution of (3.1_ω). Let $\omega \in \Omega$, $l \in \mathbb{P}$ and set

$$\hat{\tau}_t(\omega, l) = \big(\tau_t(\omega), \Phi_\omega(t)l\big)$$

where $\Phi_\omega(t)$ acts linearly on l for each $t \in \mathbb{R}$.

We propose to study the directional variation of the solutions of Eqs. (3.1_ω) by collecting information about the compact invariant subsets of Σ. In particular, we would like to classify the minimal subsets of $(\Sigma, \mathbb{R})$.

When $n \geqslant 3$, the study of the projective flow $(\Sigma, \mathbb{R})$ is still in its infancy, though there is a substantial body of particular results. We remark that there is no reason to restrict attention to the projective flow. It is of interest to substitute $\mathbb{P}$ with other compact homogeneous spaces H of the general linear group or of a Lie subgroup thereof, and study the flow on $\Omega \times H$ induced by Eqs. (3.1_ω). See [46–49,81,90–93,120].

Let us now consider the case when $n = 2$. Then $\mathbb{P}$ is the projective circle, and Σ is a trivial circle bundle over Ω with projection $\pi : \Sigma \to \Omega$: $(\omega, l) \mapsto \omega$. Even in this relatively simple case, we will find diverse phenomena which illustrate the complexity of the directional variation of the solutions of linear differential systems with nonperiodic coefficients.

It is easy to classify the possible projective flows $(\Sigma, \mathbb{R})$ when $n = 2$ and when Ω is the Bebutov hull of a periodic function $A(\cdot)$. This classification can serve as a guide to the study of $(\Sigma, \mathbb{R})$ in more general cases. So return to Eq. (3.1) for a moment, and assume that $A(t + 2\pi) = A(t)$ for all $t \in \mathbb{R}$. We view A as an element of $\mathcal{C} = \{C : \mathbb{R} \to \mathbb{M}_2 \mid C \text{ is continuous and bounded}\}$ with the compact-open topology. We carry out the Bebutov construction beginning with A and obtain a compact hull Ω which is homeomorphic to a circle. In fact, we can think of $\omega \in \Omega$ as an angular coordinate on this circle, so that a generic element of Ω has the form A_ω where $A_\omega(\cdot) = A(\omega + \cdot)$. The space $\Sigma = \Omega \times \mathbb{P}$ is a two-torus.

Let us consider the following model equations:

(α) $x' = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix} x$, $a > 0$;

(β) $x' = \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix} x$, $b \in \mathbb{R}$;

(γ) $x' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x$

where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. We view the coefficients matrices as constant functions defined on Ω. Let us examine the corresponding projective flows.

Case (α) There are exactly two minimal subsets $M_1, M_2 \subset \Sigma$; these are flow isomorphic to Ω via the projection $\pi : \Sigma \to \Omega$. In fact, if $l_1 \subset \mathbb{R}^2$ denotes the x_1-axis and $l_2 \subset \mathbb{R}^2$ denotes the x_2-axis, then $M_i = \{(\omega, l_i) \mid \omega \in \Omega\}$, $i = 1, 2$. The orbits in Σ which lie off $M_1 \cup M_2$ exhibit an hyperbolic structure.

Case (β) If $\frac{b}{2\pi}$ is rational, then the flow on Σ laminates in an obvious sense into uncountably many minimal subsets. If $b = 0$ these are defined by conditions of the type $l = constant$. If $\frac{b}{2\pi} = \frac{p}{q}$ where $q \geqslant 1$ and the fraction is expressed in lowest terms, then each minimal set defines a closed curve in Σ which winds q times around Ω as one traverses the circle $\mathbb{P}$. On the other hand, if $\frac{b}{2\pi}$ is irrational, then Σ itself is a minimal set, and the flow on Σ is a Kronecker winding. The orbits in Σ exhibit an elliptic structure.

Case (γ) Let $l_1 \in \mathbb{R}^2$ be the x_1-axis, and let $M = \{(\omega, l_1) \mid \omega \in \Omega\}$. Then $M \subset \Sigma$ is the unique minimal subset of Σ. The orbits of Σ which lie off M exhibit a parabolic structure.

It is rather easy to show that, when $A(\cdot)$ is 2π-periodic, then there is a flow isomorphism between the projective flow defined by Eqs. (3.1_ω) and one of the flows considered in the cases (α), (β) and (γ). To see this, note first that the projective flow $(\Sigma, \mathbb{R})$ does not change if a 2π-periodic, continuous function of the form $\delta(t)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is added to $A(t)$, so choosing $\delta(t) = -\frac{1}{2} \operatorname{tr} A(t)$, we can arrange that the trace $\operatorname{tr} A(t)$ is zero of all $t \in \mathbb{R}$.

Assuming this, the Floquet theory assures the existence of a 4π-periodic change of variables $x = L(t)y$ which brings (3.1) to the form

$$y' = Ry$$

where R has one of the forms (α), (β), (γ) (this uses the fact that $L(\cdot)$ can be chosen to have values in $\mathrm{SL}(2,\mathbb{R})$). Now, the projective flow is insensitive to the 2π – respectively 4π – periodicity of the function L. We take advantage of this fact as follows: viewing ω as an angular coordinate on the circle Ω, we define $h : \Sigma \to \Sigma$: $h(\omega, l) = (\omega, L(\omega)l)$. One can check that h defines a flow isomorphism between the flow $(\Sigma, \mathbb{R})$ defined by the appropriate model equation, and the $(\Sigma, \mathbb{R})$ defined by the family (3.1_ω).

Let us now turn to a discussion of the projective flow in the nonperiodic case. For the rest of Section 3, we will assume that $(\Sigma, \mathbb{R})$ is a minimal flow. Frequently we will assume in addition that it is Bohr almost periodic. Let $A : \Omega \to \mathbb{M}_2$ be a continuous function. The family of Eqs. (3.1_ω):

$$x' = A(\omega \cdot t)x$$

defines a flow $(\Sigma, \mathbb{R})$.

One might ask for a classification up to flow isomorphism of all possible projective flows, for all such continuous maps A. This seems to be a completely hopeless endeavor. We set a more modest goal, and ask for a description of the possible minimal subsets of Σ. Progress along these lines has been made. We will describe this progress, and will encounter along the way a rich collection of interesting minimal sets which supplies examples illustrating various phenomena which often are not related in an a priori way to properties of two-dimensional linear differential systems.

Before beginning, let us observe that the projective flow $(\Sigma, \mathbb{R})$ is unchanged if the diagonal matrix function $-\frac{1}{2}\operatorname{tr} A(\cdot)\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ is added to $A(\cdot)$. Hence for present purposes, there is no loss of generality in assuming that the trace $\operatorname{tr} A(\cdot) = 0$, that is, that A takes values is the Lie algebra $\mathrm{sl}(2,\mathbb{R})$. For the rest of Section 3, we will assume that A is a continuous, $\mathrm{sl}(2,\mathbb{R})$-valued function on Ω.

We begin the discussion with a lemma which will be applied several times in what follows. We give a proof which uses some basic techniques in the theory of nonautonomous differential systems.

LEMMA 3.1. *Let $(\Omega, \mathbb{R})$ be a minimal flow, and let $f : \Omega \to \mathbb{R}$ be a continuous function. Suppose that there is an invariant measure ν_0 on Ω such that $\int_\Omega f\,\mathrm{d}\nu_0 = 0$. Then one of the following conditions holds.*

(i) *There is a continuous function $F : \Omega \to \mathbb{R}$ such that $F(\omega \cdot t) - F(\omega) = \int_0^t f(\omega \cdot s)\,\mathrm{d}s$ $(\omega \in \Omega, t \in \mathbb{R})$.*

(ii) *There is a residual subset $\Omega_* \subset \Omega$ such that, if $\omega_* \in \Omega_*$, then*

$$\limsup_{t\to\infty} \int_0^t f(\omega_* \cdot s)\,\mathrm{d}s = \infty, \qquad \liminf_{t\to\infty} \int_0^t f(\omega_* \cdot s)\,\mathrm{d}s = -\infty,$$

$$\limsup_{t\to-\infty}\int_0^t f(\omega_*\cdot s)\,\mathrm{d}s=\infty,\qquad \liminf_{t\to-\infty}\int_0^t f(\omega_*\cdot s)\,\mathrm{d}s=-\infty.$$

Condition (i) *holds if and only if there is a point* $\omega_0\in\Omega$ *such that* $\int_0^t f(\omega_0\cdot s)\,\mathrm{d}s$ *is bounded. Furthermore, if* (i) *holds, then* $\int_\Omega f\,d\nu=0$ *for every invariant measure* ν *on* Ω.

PROOF. We define a flow $\{\tilde\tau_t\mid t\in\mathbb{R}\}$ on $\Omega\times\mathbb{R}$ as follows: $\tilde\tau_t(\omega,x)=(\tau_t(\omega),x+\int_0^t f(\omega\cdot s)\,\mathrm{d}s)$; it is easy to check that $\{\tilde\tau_t\}$ is indeed a flow. Let $\tilde\pi:\Omega\times\mathbb{R}\to\Omega$ be the projection; it is a flow homomorphism.

Let us now first assume that there is a point $\omega_0\in\Omega$ such that $\int_0^t f(\omega_0\cdot s)\,\mathrm{d}s$ is bounded. The orbit $\{\tilde\tau_t(\omega_0,0)\mid t\in\mathbb{R}\}\subset\Omega\times\mathbb{R}$ is bounded, hence its closure D is compact and $\{\tilde\tau_t\}$-invariant. Let $M\subset D$ be a minimal subset. We claim that M defines a section of $\Omega\times\mathbb{R}$; that is, there exists a continuous function $\varphi:\Omega\to\mathbb{R}$ such that $M=\{(\omega,\varphi(\omega))\mid \omega\in\Omega\}$.

To see this, suppose that there exist a point $\omega\in\Omega$ and points $x_1,x_2\in\mathbb{R}$ such that $(\omega,x_i)\in M$ $(i=1,2)$. Let $\delta=|x_2-x_1|$, and let $T_\delta:\Omega\times\mathbb{R}\to\Omega\times\mathbb{R}$ be the map $T_\delta(\omega,x)=(\omega,x+\delta)$. Then $T_\delta(M)=M$ by invariance and minimality of M. Suppose that $\delta\neq0$. Then $T_{n\delta}(M)=M$ $(n=0,\pm1,\pm2,\ldots)$. But then M is not compact. This contradiction proves that $\delta=0$. It follows that there is a function $\varphi:\Omega\to\mathbb{R}$ such that $M=\{(\omega,\varphi(\omega))\mid\omega\in\Omega\}$. The continuity of φ follows from the compactness of M.

Let us now define $F(\omega)=\varphi(\omega)$ $(\omega\in\Omega)$. Then for each $\omega\in\Omega$ and $t\in\mathbb{R}$, we have $F(\omega\cdot t)-F(\omega)=\int_0^t f(\omega\cdot s)\,\mathrm{d}s$. That is, F satisfies the condition (i).

It is clear that, if (i) holds, then $\int_0^t f(\omega\cdot s)\,\mathrm{d}s$ is bounded for every $\omega\in\Omega$. So we have verified the equivalence of (i) with the boundedness of the integrals $\int_0^t f(\omega\cdot s)\,\mathrm{d}s$.

Next, suppose that condition (i) does not hold. We will prove that the oscillation conditions (ii) are valid. To do so, we will adapt to the present (simple) situation techniques which are used by Sacker and Sell to prove a basic result in the theory of exponential dichotomies ([137]; see also Selgrade [142] and Mañé [105]).

To facilitate comparison with [137], we introduce the family of one-dimensional linear systems

$$x'=f(\omega\cdot t)x,\quad x\in\mathbb{R}\tag{3.2$_\omega$}$$

Let $\varphi(\omega,t)=\exp\int_0^t f(\omega\cdot s)\,\mathrm{d}s$ be the solution of (3.2$_\omega$) satisfying $\varphi(\omega,0)=1$. Equations (3.2$_\omega$) define a flow $\{\tilde\tau_t\mid t\in\mathbb{R}\}$ on $\Omega\times\mathbb{R}$, as follows: $\tilde\tau_t(\omega,x)=(\tau_t(\omega),\varphi(\omega,t)x)$. Here $\{\tilde\tau_t\}$ has a different meaning as compared to the preceding lines.

Let us assume that there is a constant c such that $\int_0^t f(\omega\cdot s)\,\mathrm{d}s\leqslant c$ for all $\omega\in\Omega$ and $t\geqslant0$; we will presently see that this assumption leads to a contradiction. One checks that if our assumption holds, then $\int_0^t f(\omega\cdot s)\,\mathrm{d}s\geqslant -c$ for all $\omega\in\Omega, t\leqslant0$. Since condition (i) does not hold, we have that $\int_0^t f(\omega\cdot s)\,\mathrm{d}s$ is unbounded above as $t\to-\infty$, for each $\omega\in\Omega$. This means that, for each $\omega\in\Omega$, the only solution $x(t)$ of (3.2$_\omega$) which is bounded on all of $\mathbb{R}$ is the zero solution $x(t)\equiv0$.

We continue to assume that $\int_0^t f(\omega\cdot s)\,\mathrm{d}s\leqslant c$ for all $\omega\in\Omega, t\geqslant0$. We claim that $\varphi(\omega,t)$ decays to zero in a uniform exponential manner as $t\to\infty$. To see this, we first show that, for each $\omega \in \Omega$, $\varphi(\omega, t) \to 0$ as $t \to \infty$. For, if were not the case, there would exist $\delta > 0$, $\omega \in \Omega$, and a sequence $t_k \to \infty$ such that $\varphi(\omega, t_k) \geqslant \delta$ $(k = 1, 2, \ldots)$. Passing to a subsequence if necessary, we can assume that $\tau_{t_k}(\omega) \to \bar{\omega}$ and $\varphi(\omega, t_k) \to \bar{x}$ where $\bar{x} \neq 0$. Consider the omega-limit set with respect to the flow $\{\tilde{\tau}_t\}$ of the point $(\omega, 1) \in \Omega \times \mathbb{R}$. This set is compact and invariant, and contains the point $(\bar{\omega}, \bar{x})$. It follows that $\varphi(\bar{\omega}, t)\bar{x}$ is bounded for all $t \in \mathbb{R}$; in fact, $\varphi(\bar{\omega}, t)\bar{x} \leqslant \frac{c}{\delta}$ $(t \in \mathbb{R})$. However, the only bounded solution of $(2_{\bar{\omega}})$ is the zero solution, hence $\bar{x} = 0$. This contradiction shows that $\varphi(\omega, t) \to 0$ as $t \to \infty$.

Next note that, if $\omega \in \Omega$, then there exists t_ω such that, if $t \geqslant t_\omega$, then $\varphi(\omega, t) < 1/2$. There is a neighborhood U_ω of ω in Ω such that, if $\bar{\omega} \in U_\omega$, then $\varphi(\bar{\omega}, t_\omega) < 1/2$. Choose $\omega_1, \ldots, \omega_r \in \Omega$ so that $\Omega = U_{\omega_1} \cup \cdots \cup U_{\omega_r}$. Write $U_i = U_{\omega_i}$, $t_i = t_{\omega_i}$ $(1 \leqslant i \leqslant r)$, and set $T = \max\{t_1, \ldots, t_r\}$. If $\bar{\omega} \in \Omega$ and $\bar{\omega} \in U_k$, then $\varphi(\bar{\omega}, t_j + t_k) = \varphi(\tau_{t_j}(\bar{\omega}), t_k)\varphi(\bar{\omega}, t_j) < (\frac{1}{2})^2$. This argument can be iterated to obtain the following: let $\alpha = \frac{\ln 2}{T}$ and $K = 2\sup\{\varphi(\omega, t) \mid \omega \in \Omega,\ 0 \leqslant t \leqslant T\}$; then $\varphi(\omega, t) \leqslant K\mathrm{e}^{-\alpha t}$ $(t \geqslant 0)$. We have shown that $\varphi(\omega, t)$ decays to zero in a uniform exponential way as $t \to \infty$. Now, the exponential decay has the consequence that, if $\omega \in \Omega$, then

$$\limsup_{t\to\infty} \frac{1}{t} \int_0^t f(\omega \cdot s)\,\mathrm{d}s \leqslant -\alpha. \tag{$*$}$$

We claim that condition $(*)$ is inconsistent with the hypothesis that $\int_\Omega f\,\mathrm{d}\nu_0 = 0$.

To see this, observe first that there exist ergodic measures ν_1, ν_2 on Ω such that $\int_\Omega f\,\mathrm{d}\nu_1 \leqslant 0$ and $\int_\Omega f\,\mathrm{d}\nu_2 \geqslant 0$ (these ergodic measures may coincide: $\nu_1 = \nu_2 = \nu$, in which case $\int_\Omega f\,\mathrm{d}\nu = 0$). This follows from the Choquet theory [131] applied to the weak-$*$ compact convex subset I of invariant measures on Ω. In fact, the invariant measure ν_0 can be expressed as an integral over I with respect to a measure H concentrated on the set E of ergodic measures, in such a way that $\int_\Omega f\,\mathrm{d}\nu_0 = \int_E (\int_\Omega f\,\mathrm{d}\nu)\,\mathrm{d}H(\nu)$. So if, say, $\int_\Omega f\,\mathrm{d}\nu > 0$ for all ergodic measures ν, then $\int_\Omega f\,\mathrm{d}\nu_0 > 0$, and similarly if $\int_\Omega f\,\mathrm{d}\nu < 0$ for all ergodic measures ν. This proves the existence of the measures ν_1, ν_2.

Now, if $\int_\Omega f\,\mathrm{d}\nu_2 \geqslant 0$, then there exists a set $\Omega_2 \subset \Omega$ with $\nu_2(\Omega_2) = 1$ such that, if $\omega_2 \in \Omega_2$, then

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f(\omega_2 \cdot s)\,\mathrm{d}s \geqslant 0,$$

in contradiction with $(*)$.

Recapitulating, we have shown that, if condition (i) does not hold, then there does not exist a constant c such that $\int_0^t f(\omega \cdot s)\,\mathrm{d}s \leqslant c$ for all $\omega \in \Omega$, $t \geqslant 0$. We now set $\Omega_N = \{\omega \in \Omega \mid \int_0^t f(\omega \cdot s)\,\mathrm{d}s \leqslant N$ for all $t \geqslant 0\}$. Then Ω_N is closed $(N = 1, 2, \ldots)$. If for some N, Ω_N contains a set U which is open in Ω, then there are times $t_1 > 0, \ldots, t_r > 0$ such that $\tau_{t_1}(U) \cup \cdots \cup \tau_{t_r}(U) = \Omega$. It follows that $\int_0^t f(\omega \cdot s)\,\mathrm{d}s$ is bounded above by a uniform constant for all $\omega \in \Omega$, $t \geqslant 0$. But we know that this is not possible. We conclude that $\bigcup_{N=1}^\infty \Omega_N$ has first category in Ω.

Let $\Omega_1 = \Omega \setminus \bigcup_{N=1}^\infty \Omega_N$. If ω is in the residual set Ω_1, then $\limsup_{t\to\infty} \int_0^t f(\omega \cdot s)\,\mathrm{d}s = \infty$. We apply reasoning similar to the above to obtain residual sets $\Omega_2, \Omega_3, \Omega_4 \subset \Omega$ for which the remaining oscillation conditions hold. Then the residual set $\Omega_* = \bigcap_{i=1}^4 \Omega_i$ satisfies the oscillation conditions (ii). □

One of the first proofs of the equivalence of condition (i) and the existence of a point ω_0 for which $\int_0^t f(\omega \cdot s)\,\mathrm{d}s$ is bounded was published in ([63], Theorem 14.11); we have repeated this proof. The fact that condition (ii) follows from the hypotheses of Lemma 3.1 was proved in [82]. There the condition that $\int_\Omega f\,\mathrm{d}\nu_0 = 0$ for some invariant measure ν_0 is substituted by the equivalent condition that, for some point $\omega \in \Omega$ and some sequence t_k such that $|t_k| \to \infty$, one has $\lim_{k\to\infty} \frac{1}{t_k}\int_0^{t_k} f(\omega \cdot s)\,\mathrm{d}s = 0$.

At several points in the discussion, we will apply a result proved by Cameron [25] in 1936 for almost periodic linear differential systems. It was later generalized to linear differential system with Birkhoff recurrent coefficients [41]. More recently, special cases of Cameron's theorem have been rediscovered by several authors.

THEOREM 3.2 (Cameron). *Suppose that* $(\Omega, \mathbb{R})$ *is minimal and that* $n \geqslant 1$. *Let* $A : \Omega \to \mathbb{M}_n$ *be a continuous function. Suppose that each* $\omega \in \Omega$ *has the property that all solutions of Eq.* (3.1_ω):

$$x' = A(\omega \cdot t)x$$

are bounded. Then there is a continuous and continuously invertible map $L : \Omega \to \mathrm{GL}(n, \mathbb{R})$ *with the following properties.*

- *L has symmetric values*: $L(\omega)^T = L(\omega)$ *for all* $\omega \in \Omega$.
- *The map* $\omega \to \frac{\mathrm{d}}{\mathrm{d}t} L(\omega \cdot t)|_{t=0}$ *is well-defined and continuous.*
- *The change of variables* $x = L(\omega \cdot t)y$ *takes* (3.1_ω) *to the form*

  $$y' = B(\omega \cdot t)y$$

  where $B : \Omega \to \mathbb{M}_n$ *is continuous and has antisymmetric values*: $B(\omega)^T = -B(\omega)$ *for all* $\omega \in \Omega$.
- *If A takes values in* $\mathrm{sl}(2, \mathbb{R})$, *then L takes values in* $\mathrm{SL}(2, \mathbb{R})$.

For a proof of this theorem see [41]. It makes use of Cameron's original idea and some results of Ellis concerning the structure of distal extensions of minimal flows.

We return to the case $n = 2$. Let $(\Omega, \mathbb{R})$ be a minimal flow. In what follows we will usually assume in addition that $(\Omega, \mathbb{R})$ is almost periodic. Let $A : \Omega \to \mathrm{sl}(2, \mathbb{R})$ be a continuous map. Our goal is to analyze the minimal subsets of the projective flow $(\Sigma, \mathbb{R})$ defined by Eqs. (3.1_ω). The first step is to count them. We begin with a strengthened version of a theorem of Sacker and Sell [139].

PROPOSITION 3.3. *Let* $(\Omega, \mathbb{R})$ *be a minimal almost periodic flow, let* $A : \Omega \to \mathrm{sl}(2, \mathbb{R})$ *be continuous, and let* $(\Sigma, \mathbb{R})$ *be the projective flow defined by Eqs.* (3.1_ω). *Suppose that* Σ *contains* $k \geqslant 3$ *distinct minimal sets. Then there is a continuous, continuously invertible* *matrix function* $L:\Omega \to \mathrm{SL}(2,\mathbb{R})$ *such that the change of variables* $y = L(\omega\cdot t)x$ *takes Eqs.* (3.1_ω) *into the form*

$$y' = \begin{pmatrix} 0 & -b_0 \\ b_0 & 0 \end{pmatrix} y \quad (\omega \in \Omega)$$

where $b_0 \in \mathbb{R}$ *is a constant.*

We include a proof of this proposition since it illustrates several simple but useful techniques.

PROOF. We first show that, for each $\omega \in \Omega$, the fundamental matrix solution $\Phi_\omega(t)$ of (3.1_ω) is bounded. For this, let M_1, M_2, M_3 be three distinct minimal subsets of Σ. Fix $\omega \in \Omega$, and let $l_1, l_2, l_3 \in \mathbb{P}$ be points such that $(\omega, l_i) \in M_i$, $1 \leqslant i \leqslant 3$. Such points can be found because $(\Omega, \mathbb{R})$ is minimal, hence the projection $\pi(M_i)$ equals Ω $(1 \leqslant i \leqslant 3)$.

Let d be a metric on $\mathbb{P}$. Note that, if $i \neq j$, then (ω, l_i) and (ω, l_j) form a distal pair in the sense that $\inf_{t\in\mathbb{R}} d(\Phi_\omega(t)l_i, \Phi_\omega(t)l_j) > 0$. This is because the orbits $\{\hat{\tau}_t(\omega, l_i) \mid t \in \mathbb{R}\}$ lie in the mutually disjoint minimal sets M_i $(1 \leqslant i \leqslant 3)$. Let $\bar{x}_i \in l_i$ be a vector of norm 1, and set $x_i(t) = \Phi_\omega(t)\bar{x}_i$ $(1 \leqslant i \leqslant 3)$.

Suppose for contradiction that $x_1(t)$ is unbounded. Then there exists a sequence $(t_k) \subset \mathbb{R}$ such that $|t_k| \to \infty$ and $|x_1(t_k)| \geqslant k$. Let $\Psi(t)$ be the matrix solution of (3.1_ω) whose columns are $x_1(t)$ and $x_2(t)$. By Liouville's formula, $\det \Psi(t) = c = \mathit{constant} \neq 0$, and it follows from the distallity condition on (ω, l_1) and (ω, l_2) that $\lim_{k\to\infty} x_2(t_k) = 0$. In a similar way, one proves that $\lim_{k\to\infty} x_3(t_k) = 0$.

Now, however, we let $\Psi_*(t)$ be the matrix solution of (3.1_ω) whose columns are $x_2(t)$ and $x_3(t)$. Then $\det \Psi_*(t) = c_* = \text{constant} \neq 0$ for all $t \in \mathbb{R}$. But $\lim_{k\to\infty} \Psi_*(t_k) = 0$. This contradiction proves that $x_1(t)$ is bounded. In a similar way, one proves that $x_2(t)$ is bounded (and also $x_3(t)$ for that matter). Since a generic solution of (3.1_ω) is a linear combination of $x_1(t)$ and $x_2(t)$, we see that $\Phi_\omega(t)$ is bounded.

Let us apply Theorem 3.2 to find a continuous, continuously invertible function $L:\Omega \to \mathrm{SL}(2,\mathbb{R})$ such that $\omega \mapsto \frac{\mathrm{d}}{\mathrm{d}t} L(\omega\cdot t)|_{t=0}$ is well-defined and continuous, and such that the change of variables $x = L(\omega\cdot t)y$ brings Eq. (3.1_ω) to the form

$$y' = \begin{pmatrix} 0 & -b(\omega\cdot t) \\ b(\omega\cdot t) & 0 \end{pmatrix} y \tag{3.3_ω}$$

where $b:\Omega \to \mathbb{R}$ is continuous. Let ν be an ergodic measure on Ω, and let $b_0 = \int_\Omega b(\omega)\,\mathrm{d}\nu(\omega)$. For each $\omega \in \Omega$, consider the quantity

$$\int_0^t \big(b(\omega\cdot s) - b_0\big)\,\mathrm{d}s = \int_0^t b(\omega\cdot s)\,\mathrm{d}s - b_0 t.$$

According to Lemma 3.1 this quantity satisfies exactly one of the following conditions.

(α) There exists a continuous function $B:\Omega \to \mathbb{R}$ such that $B(\omega\cdot t) - B(\omega) = \int_0^t (b(\omega\cdot s) - b_0)\,\mathrm{d}s$ for all $\omega \in \Omega$ and $t \in \mathbb{R}$.

(β) There is a residual set $\Omega_* \subset \Omega$ such that the oscillation conditions (ii) of Lemma 3.1 hold.

Suppose first that (α) holds. Make the change of variable

$$y = \begin{pmatrix} \cos B(\omega \cdot t) & -\sin B(\omega \cdot t) \\ \sin B(\omega \cdot t) & \cos B(\omega \cdot t) \end{pmatrix} z$$

in Eq. (3.3_ω), and note that, in the z-variable, Eqs. (3.1_ω) take the form

$$z' = \begin{pmatrix} 0 & -b_0 \\ b_0 & 0 \end{pmatrix} z.$$

So in this case Proposition 3.3 is proved.

Suppose the second possibility (β) holds. It turns out that, in this case, we cannot have $k \geqslant 3$. We indicate how this can be proved. Note that the circle group $\mathbb{T} = \mathbb{R}/\pi\mathbb{Z}$ acts on $\mathbb{P} = \mathbb{R}/\pi\mathbb{Z}$ by addition. Hence $\mathbb{T}$ acts on Σ in the following way: if $(\omega, l) \in \Sigma$ and $g \in \mathbb{T}$, then $g \cdot (\omega, l) = (\omega, g + l)$.

Let $(\Sigma, \{\hat{\tau}_t\})$ be the projective flow defined by Eqs. (3.3_ω). Then $\mathbb{T}$ commutes with $\{\hat{\tau}_t\}$ in the sense that, if $g \in \mathbb{T}$, then $g \cdot \hat{\tau}_t(\omega, l) = \hat{\tau}_t(g \cdot (\omega, l))$. Now, if $M \subset \Sigma$ is minimal, then $g \cdot M = \{g \cdot m \mid m \in M\}$ either coincides with M or is disjoint from M. Consider the subgroup G_0 of $\mathbb{T}$ defined as follows: $G_0 = \{g \in \mathbb{T} \mid g \cdot M = M\}$. Since $k > 1$, G_0 must be finite; let G_0 have r elements. Then M is an r-cover of Ω. It can be shown that $(M, \{\tau_t\})$ is an almost periodic minimal set [42,139].

Now however, let $\pi_2 : \Sigma \to \mathbb{P}$ be the projection on the second factor. Using almost periodicity of $(M, \mathbb{R})$, we see that the function

$$t \mapsto -b_0 t + \int_0^t b(\omega \cdot s)\,\mathrm{d}s \mod \pi$$

is almost periodic. Using an old theorem of Bohr (see Fink [54], Proposition 6.7), we see that $\int_0^t (b(\omega \cdot s) - b_0)\,\mathrm{d}s$ is an almost periodic function, hence is bounded. This last condition contradicts (β), so the proof of Proposition 3.3 is complete. In fact we have shown that, if (β) holds, then $k = 1$ for Eqs. (3.3_ω). That is, $(\Sigma, \mathbb{R})$ is minimal. □

REMARKS 3.4. If $(\Omega, \mathbb{R})$ is assumed to be minimal but not necessarily almost periodic, then we can only prove a weaker version of Proposition 3.3. We indicate what can be shown. Suppose that $k \geqslant 3$. We can still apply Cameron's theorem and reduce Eqs. (3.1_ω) to the skew-symmetric form (3.3_ω). The group $\mathbb{T} = \mathbb{R}/\pi\mathbb{Z}$ commutes with the flow on Σ defined by Eqs. (3.3_ω), so if $M \subset \Sigma$ is minimal, and if $G_0 = \{g \in \mathbb{T} \mid g \cdot M = M\}$, then G_0 is a finite subgroup of $\mathbb{T}$. This shows, incidentally, that $k = \infty$ because two points $g_1, g_2 \in \mathbb{T}$ with $g_1 - g_2 \notin G_0$ give rise to distinct minimal sets $M_i = g_i \cdot M$ $(i = 1, 2)$. Thus one has the so-called 1–2–∞ theorem [139].

Suppose now that G_0 contains r elements. Then each minimal subset $M \subset \Sigma$ is an r-cover of Ω, and in fact a finite Abelian group extension of Ω. In these circumstances, it seems reasonable to look for a change of variable $y = L_1(\omega \cdot t)z$ which takes Eqs. (3.3_ω)

to constant-coefficient form. However, we have not been able to prove that such a change of variables always exists.

Let us now turn to the case $k = 2$, where k is the number of distinct minimal subsets of Σ. When the coefficient matrix $A(\cdot)$ in (3.1) is 2π-periodic, the condition $k = 2$ is equivalent to the statement that the corresponding family of Eqs. (3.1_ω) has an exponential dichotomy. In this situation, the projective flow is isomorphic to that determined by the system $x' = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix} x$. Here the nonzero constant a is interpreted as a 2π-periodic function. We will see that other possibilities are present when (3.1) has an almost periodic coefficient $A(\cdot)$.

Unless otherwise stated, $(\Omega, \mathbb{R})$ will denote a minimal almost periodic flow. Let $A : \Omega \to \mathrm{sl}(2, \mathbb{R})$ be continuous, and let $(\Sigma, \mathbb{R})$ be the projective flow defined by the corresponding family of Eqs. (3.1_ω). Suppose that $(\Sigma, \mathbb{R})$ contains exactly two minimal sets.

Let us first note that, in this case, each Eq. (3.1_ω) admits an unbounded solution. For suppose that this is not the case, and that, for some $\omega \in \Omega$, the fundamental matrix solution $\Phi_\omega(t)$ of Eq. (3.1_ω) is bounded: $|\Phi_\omega(t)| \leqslant c < \infty$ for all $t \in \mathbb{R}$. If $\bar{\omega} \in \Omega$, let $(t_k) \subset \mathbb{R}$ be a sequence such that $\omega \cdot t_k \to \bar{\omega}$. Passing to a subsequence, we can assume that $\Phi_\omega(t + t_k)$ converges, uniformly on compact subsets of $\mathbb{R}$, to a matrix solution $\Psi(t)$ of Eq. $(1_{\bar{\omega}})$. Since $\det \Psi(t) = 1$ for all $t \in \mathbb{R}$, $\Psi(t)$ is an invertible matrix solution of $(1_{\bar{\omega}})$, and moreover $|\Psi(t)| \leqslant c$. It follows that all solutions of $(1_{\bar{\omega}})$ are bounded, as well.

Now however, we can apply Cameron's theorem to reduce Eqs. (3.1_ω) to the form (3.3_ω). Following the analysis of the family (3.3_ω) which was carried out in the proof of Proposition 3.3, we see that either $k = 1$ or $k = \infty$. So indeed each Eq. (3.1_ω) admits an unbounded solution.

It is of course possible that the family (3.1_ω) admits an exponential dichotomy. In this case, all solutions of all Eqs. (3.1_ω) are unbounded. However, we will encounter other, more interesting families of Eqs. (3.1_ω) for which $k = 2$.

First we consider some a priori properties of the minimal subsets of Σ when $(\Omega, \mathbb{R})$ is minimal and almost periodic, and $k = 2$. We begin with some terminology.

DEFINITIONS 3.5. Let Ω be a compact metric space with distance function d, and let $(\Omega, \mathbb{R})$ be a flow. Say that two points $\omega_1, \omega_2 \in \Omega$ form a *distal pair* if there exists $\delta > 0$ such that $d(\omega_1 \cdot t, \omega_2 \cdot t) \geqslant \delta$ for all $t \in \mathbb{R}$. Say that ω_1, ω_2 form a *proximal pair* if $\omega_1 \neq \omega_2$ and there exists a sequence $(t_k) \subset \mathbb{R}$ such that $\lim_{k \to \infty} d(\omega_1 \cdot t_k, \omega_2 \cdot t_k) = 0$. Say that $(\Omega, \mathbb{R})$ is *distal* if every pair $\omega_1, \omega_2 \in \Omega$ with $\omega_1 \neq \omega_2$ is distal, and say that $(\Omega, \mathbb{R})$ is *proximal* is every pair $\omega_1, \omega_2 \in \Omega$ with $\omega_1 \neq \omega_2$ is proximal. Let $(\Sigma, \mathbb{R})$ be another flow where Σ is a metric space, and let $\pi : \Sigma \to \Omega$ be a flow homomorphism. Say that Σ is a *distal extension* of Ω if whenever $\pi(\sigma_1) = \pi(\sigma_2)$ and $\sigma_1 \neq \sigma_2$, the pair (σ_1, σ_2) is distal with respect to the flow $(\Sigma, \mathbb{R})$. Similarly, say that Σ is a *proximal extension* of Ω if whenever $\pi(\sigma_1) = \pi(\sigma_2)$ and $\sigma_1 \neq \sigma_2$, the pair (σ_1, σ_2) is proximal with respect to the flow $(\Sigma, \mathbb{R})$.

In the context of systems of ordinary differential equations, the distallity concept is related to the well-known Favard property [52,158] and to properties of separatedness of the solutions. See [24,139].

PROPOSITION 3.6. *Let* $(\Omega, \mathbb{R})$ *be an almost periodic minimal flow, and let* $A : \Omega \to \mathrm{sl}(2, \mathbb{R})$ *be continuous. Suppose that the projective flow* $(\Sigma, \mathbb{R})$ *defined by Eqs.* (3.1_ω)

admits $k = 2$ minimal sets. Then each such minimal set $M \subset \Sigma$ is an almost automorphic extension of Ω.

For the definition of almost automorphic extension, see Section 2. It is of course understood that the flow homomorphism defining the extension is the projection $\pi : \Sigma \to \Omega$. Proposition 3.6 implies that there is a residual subset $\Omega_0 \subset \Omega$ such that, if $\omega \in \Omega_0$, then $\pi^{-1}(\omega) \cap M$ contains exactly one point. Of course, the possibility that $\Omega_0 = \Omega$ is not excluded.

PROOF. Let M_1 and M_2 be the minimal subsets of Σ. Let us first suppose that, for some $\omega \in \Omega$, the fiber $\pi^{-1}(\omega) \cap M_1$ contains two distinct points m_0, m_1. Further, let $m_2 \in M_2$ be a point such that $\pi(m_2) = \omega$. Let $m_i = (\omega, l_i)$, then choose $\bar{x}_i \in l_i$ with $|\bar{x}_i| = 1$, and set $x_i(t) = \Phi_\omega(t)\bar{x}_i$ $(i = 0, 1, 2)$. Note that (m_0, m_2) and (m_1, m_2) are distal pairs in Σ because M_1 and M_2 are compact and disjoint.

Suppose that $x_2(t)$ is unbounded, so that there is a sequence $(t_k) \subset \mathbb{R}$ such that $\lim_{k\to\infty} |x_2(t_k)| = \infty$. Let $\Psi_0(t)$ be the matrix solution of (3.1_ω) whose columns are $x_0(t)$ and $x_2(t)$. By Liouville's formula, $\det \Psi_0(t) = \det \Psi_0(0) \neq 0$ for all $t \in \mathbb{R}$. Since (m_0, m_2) is a distal pair, it must be the case that $\lim_{k\to\infty} x_0(t_k) = 0$. In an analogous way, we can conclude that $\lim_{k\to\infty} x_1(t_k) = 0$. Now, however, if $\Psi(t)$ is the matrix solution of (3.1_ω) whose columns are $x_0(t)$ and $x_1(t)$, then $0 \neq \det \Psi(0) = \lim_{k\to\infty} \det \Psi(t_k) = 0$, a contradiction. We conclude that $x_2(t)$ is bounded.

If $x_0(t)$ is also bounded, then the matrix solution $\Psi_0(t)$ is bounded. However, we have seen that this possibility is inconsistent with the hypothesis that $k = 2$. Therefore $x_0(t)$ is unbounded. Thus there is a sequence $(t_k) \subset \mathbb{R}$ such that $\lim_{k\to\infty} |x_0(t_k)| = \infty$. It follows that $x_2(t_k) \to 0$ as $k \to \infty$.

Suppose for contradiction that there is a point $m_3 \neq m_2$ in M_2 such that $\pi(m_3) = \omega$. Let $m_3 = (\omega, l_3)$, choose $\bar{x}_3 \in l_3$ such that $|\bar{x}_3| = 1$, and set $x_3(t) = \Phi_\omega(t)\bar{x}_3$. Reasoning as in the first two paragraphs of the proof, we see that $x_3(t_k) \to 0$ as $k \to \infty$. However, if we form the matrix solution of (3.1_ω) whose columns are $x_2(t)$ and $x_3(t)$, and if we apply Liouville's formula along the sequence (t_k), we obtain a contradiction. We conclude that m_2 is the unique point in the fiber $\pi^{-1}(\omega) \cap M_2$. This implies that M_2 is an almost automorphic extension of Ω.

We must still show that M_1 is an almost automorphic extension of Ω. If this is not the case, then for each $\omega \in \Omega$, the fiber $\pi^{-1}(\omega) \cap M_1$ contains at least two points. Arguing as above, we see that $\pi : M_2 \to \Omega$ is injective (actually a homeomorphism). Let $m_2 = (\omega, l_2) \in M_2$, and let $\bar{x}_2 \in l_2$ satisfy $|\bar{x}_2| = 1$. Carrying out the reasoning of the first two paragraphs of the proof, we see that $x_2(t) = \Phi_\omega(t)\bar{x}_2$ is bounded.

At this point, let us return to Eq. (3.3_ω) in Section 2. Letting $f : \Sigma \to \mathbb{R}$ be the function in that equation, and restricting f to M_2, we see that

$$\ln\left|x_2(t)\right| = \int_0^t f\left(\hat{\tau}(m_2)\right) \mathrm{d}s$$

for each $m_2 \in M_2$. Now, $(M_2, \mathbb{R})$ is a minimal flow and, according to the previous paragraph, $\int_0^t f(\hat{\tau}(m_2))\,\mathrm{d}s$ is bounded above for all $m_2 \in M_2$. It follows from the oscillation

Lemma 3.1 that $\int_0^t f(\hat{\tau}(m_2))\,\mathrm{d}s$ is uniformly bounded (above and below) for $m_2 \in M_2$, $t \in \mathbb{R}$. In particular, $x_2(t)$ is bounded away from zero for all $m_2 \in M_2$.

Again let $\omega \in \Omega$, let $m_0, m_1 \in M_1 \cap \pi^{-1}(\omega)$, and let $\Psi_0(t)$ be the matrix solution of (3.1_ω) whose columns are $x_0(t)$ and $x_2(t)$. Using the distallity of the pair (m_0, m_2), we see that $\Psi_0(t)$ is bounded. This is a contradiction, so we see that M_1 is an almost automorphic extension of Ω.

Until now, we have worked under the hypothesis that at least one fiber of M_1 contains two (or more) distinct points. To complete the proof of Proposition 3.6, we must examine the case that all fibers $\pi^{-1}(\omega) \cap M_1$ contain just one point. However, if this is true, and if M_2 is not an almost automorphic extension of Ω, then each fiber $\pi^{-1}(\omega) \cap M_2$ contains two or more points. We can thus repeat the reasoning just given; we obtain a contradiction; we then conclude that M_2 is an almost automorphic extension of Ω, after all. This conclude the proof of Proposition 3.6. □

We consider some examples in which $k = 2$. We will see that interesting phenomena are compatible with this hypothesis, which have no analogue in the periodic case. We begin with a class of simple equations which still have some interesting features. Let $(\Omega, \mathbb{R})$ be an almost periodic minimal flow, and let ν be the unique ergodic measure on Ω. Let $a : \Omega \to \mathbb{R}$ be a continuous function with mean value zero, that is to say $\int_\Omega a\,\mathrm{d}\nu = 0$. Let $C_0(\Omega)$ be the class of all such continuous functions on Ω. Next, let $a \in C_0(\Omega)$ be a function such that for some (hence all) $\omega \in \Omega$, the integral $\int_0^t a(\omega \cdot s)\,\mathrm{d}s$ is unbounded.

Consider the family of linear differential systems

$$x' = \begin{pmatrix} a(\omega \cdot t) & 0 \\ 0 & -a(\omega \cdot t) \end{pmatrix} x, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2. \tag{3.4_ω}$$

Letting $l_i \in \mathbb{P}$ be the x_i-axis $(i = 1, 2)$, we see that there are two minimal sets in Σ, namely $M_i = \{(\omega, l_i) \mid \omega \in \Omega\}$, $i = 1, 2$. There are no other minimal subsets of Σ because, if there were, then all solutions of Eqs. (3.4_ω) would be bounded, and that is not true. It is clear that this example shows that the hypothesis $k = 2$ is compatible with the vanishing of the Lyapunov exponent of the family (3.4_ω).

Actually, this class of examples can be divided into two subclasses, as follows. We first consider two subsets $\mathcal{C}_1$ and $\mathcal{C}_2$ of $C_0(\Omega)$ defined as follows: $\mathcal{C}_1 = \{a \in C_0(\Omega) \mid$ there exists a ν-measurable but noncontinuous function $A : \Omega \to \mathbb{R}$ such that $A(\omega \cdot t) - A(\omega) = \int_0^t a(\omega \cdot s)\,\mathrm{d}s$ for all $\omega \in \Omega,\ t \in \mathbb{R}\}$ and $\mathcal{C}_2 = \{a \in C_0(\Omega) \mid \int_0^t a(\omega \cdot s)\,\mathrm{d}s$ in unbounded for some $\omega \in \Omega$, and there is no measurable function $A : \Omega \to \mathbb{R}$ such that $A(\omega \cdot t) - A(\omega) = \int_0^t a(\omega \cdot s)\,\mathrm{d}s$ a.e. in $\omega\}$. It can be shown that both $\mathcal{C}_1$ and $\mathcal{C}_2$ are dense in $C_0(\Omega)$, though $\mathcal{C}_1$ is "bigger" in the sense that it is residual [74].

Next we note that the projective flow $(\Sigma, \mathbb{R})$ of the family (3.4_ω) can be explicitly described because, if $\theta(t)$ is the polar angle of a solution $x(t)$ of (3.4_ω), and if $u(t) = \cot\theta(t)$, then

$$u' = 2a(\omega \cdot t)u.$$

Suppose now that $a \in \mathcal{C}_1$. For each $\theta \in (0, \pi)$, the set $\{(\omega, \theta_\omega) \mid \cot\theta_\omega = \cot \mathrm{e}^{2A(\omega)}, \omega \in \Omega\}$ defines a measurable invariant section of Σ. This implies that, for each $\theta \in (0, \pi)$, there is an ergodic measure μ_θ on Σ defined as follows:

$$\int_\Sigma h\, \mathrm{d}\mu_\theta = \int_\Omega h(\omega, \theta_\omega)\, \mathrm{d}\nu(\omega), \qquad h \in C(\Sigma).$$

That is, the flow $(\Sigma, \mathbb{R})$ laminates into measurable invariant sections which define ergodic measures on Σ.

On the other hand, if $a \in \mathcal{C}_2$, then it can be shown that there are only two ergodic measures μ_1 and μ_2, which are supported on M_1 and M_2 respectively. We leave the proof as an exercise for the interested reader.

Let us now consider a class of examples of almost periodic Eqs. (3.1_ω) for which $k = 2$, and for which one of the minimal sets $M \subset \Sigma$ is almost automorphic but not almost periodic. We will see that examples exhibiting such minimal sets permit one to answer certain questions regarding linear nonhomogeneous almost periodic differential equations which do not satisfy the Favard separation property.

Let $(\Omega, \mathbb{R})$ be an almost periodic minimal flow, and let $a, b : \Omega \to \mathbb{R}$ be continuous functions. Consider the family of linear equations

$$x' = \begin{pmatrix} \frac{1}{2}a(\omega \cdot t) & b(\omega \cdot t) \\ 0 & -\frac{1}{2}a(\omega \cdot t) \end{pmatrix} x, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2. \tag{3.5_ω}$$

Introducing the polar angle θ in $\mathbb{R}^2$, and setting $u = \cot\theta$, we find that

$$u' = a(\omega \cdot t)u + b(\omega \cdot t). \tag{3.6_ω}$$

Let $(\Sigma, \mathbb{R})$ be the projective flow defined by Eqs. (3.5_ω). Note that, if $l_1 \in \mathbb{P}$ denotes the x_1-axis, then the set $M_0 = \{(\omega, l_1) \mid \omega \in \Omega\}$ is a minimal subset of Σ. The complement $\Sigma \setminus M_0$ is parameterized by points (ω, θ) where $0 < \theta < \pi$, or equivalently by points (ω, u) where $-\infty < u < \infty$. Thus we can study the projective flow $(\Sigma, \mathbb{R})$ by studying solutions of the family (3.6_ω) of linear nonhomogeneous equations, and vice-versa.

Let ν be the unique invariant measure on Ω, and let $a_0 = \int_\Omega a\, \mathrm{d}\nu$. It is well-known that, if $a_0 \neq 0$, then each Eq. (3.6_ω) admits a unique almost periodic solution. For example, if $a_0 > 0$, then

$$u_\omega(t) = \int_t^\infty b(\omega \cdot s) \mathrm{e}^{\int_s^t a(\omega \cdot r)\, \mathrm{d}r}\, \mathrm{d}s$$

is the unique almost periodic solution of (3.6_ω). Of course, we are dealing here with a simple case of a general theory regarding the existence, uniqueness, and recurrence properties of bounded solutions of a linear nonhomogeneous differential system whose linear part admits an exponential dichotomy.

Our interest here is in the case when $a_0 = 0$. It is convenient to consider two subcases. First, suppose that $\int_0^t a(\omega \cdot s)\, \mathrm{d}s$ is bounded for some (hence all) $\omega \in \Omega$. One then checks that, if some equation $(3.6_{\bar{\omega}})$ admits a bounded solution, then for every $\omega \in \Omega$, all solutions of (3.6_{ω}) are almost periodic. One can put this trivial observation in the context of the general Favard theory concerning bounded solutions of linear nonhomogeneous almost periodic differential systems in dimension n [52]. To see this, note that the family (3.6_{ω}) satisfies the Favard separation condition: if $\omega \in \Omega$ and if $u_1(t), u_2(t)$ are two bounded solutions of (3.6_{ω}), then $\inf_t |u_1(t) - u_2(t)| > 0$. The Favard theory then states that, if one Eq. $(6_{\bar{\omega}})$ admits a bounded solution, then every Eq. (3.6_{ω}) admits an almost periodic solution. In the one-dimensional case $n = 1$, this implies that all solutions of (3.6_{ω}) are almost periodic $(\omega \in \Omega)$.

We are going to consider the second case, which is when $\int_0^t a(\omega \cdot s)\,\mathrm{d}s$ is unbounded for some (hence all) $\omega \in \Omega$. In this situation, Eqs. (3.6_{ω}) do not all satisfy the Favard separation property (this follows from Lemma 3.1). For a long time it was not known if the existence of a bounded solution of some Eq. $(6_{\bar{\omega}})$ implies that Eqs. (3.6_{ω}) admit almost periodic solutions. The question was settled in the negative by examples of [158] and [73]. It turn out that, if some Eq. $(6_{\bar{\omega}})$ is assumed to have a bounded solution, then some Eq. (3.6_{ω}) (perhaps not the original equation) admits an almost automorphic solution; however this almost automorphic solution need not to be almost periodic. This follows from the discussion and the example in [73].

In a recent paper [126], Ortega and Tarallo carried out a systematic analysis of examples like those in [158] and [73]. They developed techniques which permit one to construct examples of "non-Favard type" for all almost periodic flows $(\Omega, \mathbb{R})$. In fact, they prove a result ([126], Theorem 2) which, for the one-dimensional family (3.6_{ω}), has the following consequence.

THEOREM 3.7. *Suppose that, for some $\bar{\omega} \in \Omega$, there holds*

$$\lim_{t \to -\infty} \int_0^t a(\bar{\omega} \cdot s)\,\mathrm{d}s = \lim_{t \to \infty} \int_0^t a(\bar{\omega} \cdot s)\,\mathrm{d}s = -\infty.$$

Then there exists a continuous function $b : \Omega \to \mathbb{R}$ such that each Eq. (3.6_{ω}) admits a bounded solution, and for which no Eq. (3.6_{ω}) admits an almost periodic solution.

We will not prove this theorem here, instead we refer to ([126], Section 6-7). There a particular concept of weak solution of an n-dimensional linear nonhomogeneous almost periodic system is developed. For related considerations see [127].

The $a(\cdot)$-term of the example in [158] and [73] satisfies the hypothesis of Theorem 3.7 for some $\bar{\omega} \in \Omega$. In fact, [158] makes use of a classical quasi-periodic example of Bohr [21], while [73] applies a limit periodic example of Conley and Miller [27].

We now extend Theorem 3.7 in that we discuss the recurrence properties of the bounded solutions of Eqs. (3.6_{ω}). We follow the discussion in [73]. We continue to assume that $k = 2$.

PROPOSITION 3.8. *Suppose that $\int_{\Omega} a\,\mathrm{d}\nu = 0$, and suppose that $\int_0^t a(\omega \cdot s)\,\mathrm{d}s$ is unbounded for some (hence all) $\omega \in \Omega$. Suppose further that some Eq. (3.6_{ω}) admits a bounded solution. Then there is a residual set $\Omega_* \subset \Omega$ such that, if $\omega \in \Omega_*$, then (3.6_{ω}) admits a unique bounded solution which is almost automorphic.*

Of course it is not excluded that the bounded solutions in this proposition are almost periodic.

PROOF. Let $\bar{\omega} \in \Omega$ be a point such that Eq. ($3.6_{\bar{\omega}}$) admits a bounded solution $\bar{u}(t)$. Let $\bar{\theta}(t) = \cot^{-1} u(t)$, and note that $\{(\tau_t(\bar{\omega}), \bar{\theta}(t)) \mid t \in \mathbb{R}\}$ is an orbit of the projective flow defined by Eqs. (3.5_ω). The orbit closure $\text{cls}\{(\tau_t(\bar{\omega}), \bar{\theta}(t)) \mid t \in \mathbb{R}\}$ is a compact invariant subset of Σ which is disjoint from the minimal set $M_0 = \{(\omega, l_1) \mid \omega \in \Omega\}$. It therefore contains a minimal set M_1, such that $M_1 \cap M_0 = \emptyset$.

Now, it is clear that, each Eq. (3.5_ω) admits an unbounded solution. In fact there exists $\omega \in \Omega$ for which each solution $x(t) = \binom{x_1(t)}{x_2(t)}$ satisfying $x_2(t) \equiv 0$ is unbounded, and therefore every Eq. (3.5_ω) admits an unbounded solution. Taking account of Proposition 3.3, we see that M_0 and M_1 are the only minimal subsets of Σ. By Proposition 3.6, M_1 is an almost automorphic extension of Ω. This means that there is a residual subset $\Omega_1 \subset \Omega$ such that, if $\omega_1 \in \Omega_1$, then the fiber $\pi^{-1}(\omega_1) \cap M_1$ contains exactly one point $(\omega_1, \bar{\theta}_1)$. One checks that, if $\hat{\tau}_t(\omega_1, \bar{\theta}_1) = (\tau_t(\omega), \theta_1(t))$, then $\theta_1(t)$ is an almost automorphic function. Hence $u_1(t) = \cot \theta_1(t)$ is an almost automorphic solution of (3.6_{ω_1}).

It follows from the oscillation Lemma 3.1 that there is a residual set $\Omega_2 \subset \Omega$ such that, if $\omega \in \Omega_2$, Eq. (3.6_ω) admits a unique bounded solution. Set $\Omega_* = \Omega_1 \cap \Omega_2$; then Ω_* satisfies the condition of Proposition 3.8. This completes the proof. □

Proposition 3.8 is of course applicable to all one-dimensional examples of Ortega–Tarallo type, and hence to the examples in [158] and [73]. Let us note that almost automorphicity is a more specific property than that of Besicovich almost periodicity. It was noted in [158] that a bounded solution of (3.6_ω) satisfies this last property; see that reference for the definition of almost periodicity in the Besicovich sense. The nature of Besicovich almost periodic functions has recently been discussed by Andres and his collaborators [6].

Examples of Ortega–Tarallo type correspond to two-dimensional families (3.5_ω) whose projective flows $(\Sigma, \mathbb{R})$ admit exactly two minimal subsets, one of which, say M_1, is an almost automorphic, non-almost periodic extension of Ω. There is a question concerning the structure of this minimal set M_1, which, so far as we know, has not been answered. In fact, it is not known (to us) if there exists a family (3.5_ω) for which the minimal set M_1 supports more than one invariant measure. In other words, we ask if there is a family (3.5_ω) for which the projection $\Omega_* = \{\omega \in \Omega \mid \pi^{-1}(\omega) \cap M_1$ contains just one point$\}$ of the set of all automorphic points of M_1 has zero ν-measure: $\nu(\Omega_*) = 0$. It is known that the minimal set of the family (3.5_ω) in [73] supports only one invariant measure; see [76].

Another question – probably more difficult – also awaits an answer. Namely, one would like to find a family of type (3.1_ω) for which $k = 2$, and for which both minimal subsets of Σ are almost automorphic, non-almost periodic extensions of Ω. Such a family (3.1_ω) cannot have the form (3.5_ω), because for families of the latter type the minimal set $M_0 = \{(\omega, l_1) \mid \omega \in \Omega\} \subset \Sigma$ is flow isomorphic to Ω via the projection π.

We continue to assume that $(\Omega, \mathbb{R})$ is an almost periodic minimal flow, and turn to the case when the projective flow $(\Sigma, \mathbb{R})$ defined by Eq. (3.1_ω) contains just one minimal set, i.e., $k = 1$. It will turn out that the projective flow can exhibit an ample range of interesting structures, most of which have no analogues in the periodic case. We divide the discussion into three subsections, according to the dynamical behavior of the solutions of Eq. (3.1_ω).

For want of better terminology, we label these subsections as follows: (A) the strongly elliptic case; (B) the weakly elliptic case; (C) the weakly hyperbolic case. There is actually a fourth case, that in which Eqs. (3.1_ω) admit an exponential dichotomy. In this case, the two minimal sets are defined by the "traces" in $\Omega \times \mathbb{P}$ of the Sacker–Sell subbundles $W_\pm$ which were discussed in Section 2. See [78] for the details. We will not discuss this case further.

(A) Strongly elliptic case

The condition we impose here is that all solutions of all Eqs. (3.1_ω) are bounded. This condition is consistent with the requirement that $k = 1$. We examine the various flows $(\Sigma, \mathbb{R})$ which can arise. We will see that Σ is itself a minimal set. However it can exhibit various structures which were already understood by Furstenberg about fifty years ago [55].

First of all, we apply Cameron's theorem 3.2 to find a continuous, continuously invertible function $L : \Omega \to \mathrm{SL}(2, \mathbb{R})$ such that $\omega \mapsto \frac{\mathrm{d}}{\mathrm{d}t} L(\omega \cdot t)|_{t=0}$ is continuous and such that the change of variables $x = L(\omega \cdot t)y$ brings Eqs. (3.1_ω) to the form

$$y' = \begin{pmatrix} 0 & -b(\omega \cdot t) \\ b(\omega \cdot t) & 0 \end{pmatrix} y, \tag{3.7_ω}$$

Let ν be the unique ergodic measure on Ω, and let $b_0 = \int_\Omega b \, \mathrm{d}\nu$. We consider two possibilities: (1) $\int_0^t (b(\omega \cdot t) - b_0) \, \mathrm{d}s$ is bounded for some (hence all) $\omega \in \Omega$; (2) $\int_0^t (b(\omega \cdot t) - b_0) \, \mathrm{d}s$ is unbounded for some (hence all) $\omega \in \Omega$.

If the first possibility is realized, there is a continuous function $B : \Omega \to \mathbb{R}$ such that $B(\omega \cdot t) - B(\omega) = \int_0^t (b(\omega \cdot t) - b_0) \, \mathrm{d}s$, and, as in the proof of Proposition 3.3, the further change of variables $y = \left(\begin{smallmatrix} \cos B(\omega \cdot t) & -\sin B(\omega \cdot t) \\ \sin B(\omega \cdot t) & \cos B(\omega \cdot t) \end{smallmatrix}\right) z$ brings Eqs. (3.7_ω) to the form

$$z' = \begin{pmatrix} 0 & -b_0 \\ b_0 & 0 \end{pmatrix} z. \tag{3.8_ω}$$

The hypothesis that $k = 1$ imposes a condition on b_0 relative to the frequency module $\mathcal{M}_\Omega$ of $(\Omega, \mathbb{R})$, namely that there does not exist a rational number r such that $r b_0 \in \mathcal{M}_\Omega$. When this condition holds, the flow $(\Sigma, \mathbb{R})$ defined by Eqs. (3.8_ω) is minimal and almost periodic. If $(\Omega, \mathbb{R})$ is quasi-periodic, then $(\Sigma, \mathbb{R})$ is a Kronecker-type minimal flow.

If the second possibility is realized one has a less simple situation, which however can be understood thanks to [55]. First of all, it turns out that $(\Sigma, \mathbb{R})$ is minimal but not almost periodic. Further information can be obtained by distinguishing two subcases, as follows. For simplicity, set $b_0 = 0$.

(1) Suppose that there exists a ν-measurable but noncontinuous function $B : \Omega \to \mathbb{R}$ such that $B(\omega \cdot t) - B(\omega) = \int_0^t b(\omega \cdot s) \, \mathrm{d}s$ for all $t \in \mathbb{R}$, $\omega \in \Omega$. We saw earlier that the class $\mathcal{C}_1$ of functions $b \in C_0(\Omega)$ which have this property is dense in $C_0(\Omega)$. Let $\theta \in \mathbb{R}$ be an angular coordinate mod π on $\mathbb{P}$. For each fixed value $\theta \in [0, \pi)$, set $\theta_\omega = (\theta + B(\omega)) \mathrm{mod}\, \pi$, then

note that the set $\{(\omega,\theta_\omega) \mid \omega \in \Omega\}$ is a measurable invariant section of Σ. There is an ergodic measure μ_θ on Σ defined as follows:

$$\int_\Sigma h\,\mathrm{d}\mu_\theta = \int_\Omega h(\omega,\theta_\omega)\,\mathrm{d}\nu(\omega) \quad \bigl(h \in C(\Sigma)\bigr).$$

That is, the flow $(\Sigma,\mathbb{R})$ laminates into measurable invariant sections which define ergodic measures μ_θ on Σ. A similar situation was considered when $k = 2$.

(2) There exists no measurable function B with the above property. As we saw before, the class $\mathcal{C}_2 \subset C_0(\Omega)$ of such functions b is residual in $C_0(\Omega)$. When $b \in \mathcal{C}_2$, the flow $(\Sigma,\mathbb{R})$ admits a unique ergodic measure μ (see [55], Lemma 2.1). It can be verified that μ is equivalent to the product of ν with the normalized Lebesgue measure m on $\mathbb{P}$.

Before considering the situations envisioned in subsections (B) and (C), we state a lemma which gives some a priori information concerning the structure of the projective flow in these cases.

LEMMA 3.9. *Let* $(\Omega,\mathbb{R})$ *be an almost periodic minimal flow, let* $A:\Omega \to \mathrm{sl}(2,\mathbb{R})$ *be a continuous function, and let* $(\Sigma,\mathbb{R})$ *be the projective flow defined by Eqs.* (3.1_ω). *Suppose that* Σ *contains exactly one minimal set* M, *and that some Eq.* (3.1_ω) *admits an unbounded solution.*

Then either M *is a proximal extension of* Ω, *or* M *is a proximal extension of an almost periodic two-cover of* Ω. *In the latter situation,* M *is an almost automorphic extension of an almost periodic two-cover of* Ω.

See Definitions 3.5 for the meaning of the term “proximal extension”. A proof of Lemma 3.9 is given in ([83], Section 7). We only sketch it here.

To begin, suppose that $(M,\mathbb{R})$ is not a proximal extension of $(\Omega,\mathbb{R})$. Then there is a point $\bar\omega \in \Omega$ together with a distal pair $\bar m_1 = (\bar\omega,\bar l_1), \bar m_2 = (\bar\omega,\bar l_2)$ of points in M. Letting d be a metric on $\mathbb{P}$, and writing $\hat\tau_t(\bar\omega,\bar l_i) = (\tau_t(\bar\omega), l_i(t))$ $(i = 1,2)$, we can find a number $\delta > 0$ such that $d(l_1(t),l_2(t)) \geqslant \delta$ for all $t \in \mathbb{R}$. Passing to the closure of one of the orbits $\{\hat\tau_t(\bar m_i) \mid t \in \mathbb{R}\}$ and using the minimality of M, one shows that, for each $\omega \in \Omega$, the fiber $\pi^{-1}(\omega) \cap M$ contains a distal pair (m_1,m_2). In particular, each fiber $\pi^{-1}(\omega)$ contains at least two points.

Using Lemma 3.1 together with by-now familiar reasoning using matrix solutions of Eqs. (3.1_ω), one can show that there is a residual subset $\Omega_* \subset \Omega$ such that, if $\omega \in \Omega_*$, then $\pi^{-1}(\omega) \cap M$ contains exactly two points. At this point we introduce the proximal relation $\mathcal{P} = \{(m_1,m_2) \in M \times M \mid (m_1,m_2)$ is a proximal pair$\}$. It is clear that $\mathcal{P}$ is invariant in the sense that, if $(m_1,m_2) \in \mathcal{P}$, then $(\hat\tau_t(m_1,m_2)) \in \mathcal{P}$ for all $t \in \mathbb{R}$. It is also clear that, if $(m_1,m_2) \in \mathcal{P}$, then $\pi(m_1) = \pi(m_2)$. Now, it can be shown that $\mathcal{P}$ is a *closed* subset of $M \times M$. According to a general result of Ellis ([40], Lemma 5.17), this implies that $\mathcal{P}$ is an equivalence relation on M. Define Ω_2 to be the topological quotient $M/\mathcal{P}$. One checks that Ω_2 is a compact metrizable space, that it inherits a flow $\{\hat\tau_t^2\}$ from $\{\hat\tau_t\}$, and that $(\Omega_2,\{\hat\tau_t^2\})$ is an almost periodic 2-cover of Ω. One also checks that M is a proximal extension of Ω.

The statement of Lemma 3.9 follows from these considerations. Let us now turn to a discussion of the weakly elliptic case.

(B) Weakly elliptic case

Let $(\Omega, \mathbb{R})$ be an almost periodic minimal flow, and let $A \to \mathrm{sl}(2, \mathbb{R})$ be a continuous function. We impose the following conditions on Eqs. (3.1_ω):

(i) one Eq. (3.1_ω) admits an unbounded solution, hence all Eqs. (3.1_ω) have this property;

(ii) the Lyapunov exponent β of the family (3.1_ω) is zero.

Here the Lyapunov exponent is that with respect to the ergodic measure ν on Ω; see Section 2. The condition (ii) implies that $\frac{1}{t}\ln|\Phi_\omega(t)x_0|$ tends to zero uniformly in $\omega \in \Omega$ and $x_0 \in \mathbb{R}^2$ with $|x_0| = 1$ as $|t| \to \infty$; see Proposition 2.5.

We mention two types of minimal subset $M \subset \Sigma$ which are encountered when conditions (i) and (ii) hold.

First, there are weakly elliptic families (3.1_ω) for which the projective flow $(\Sigma, \mathbb{R})$ contains a (unique) minimal set M which is a genuine 2-cover of Ω. This is a remarkable phenomenon which has no analogue in the periodic case. The first example illustrating it seems to be in [88]; another is given in [123]. These examples both have the property that $(\Omega, \mathbb{R})$ is of limit periodic type. Apparently no example of quasi-periodic type has been constructed. It would be interesting to have a larger body of examples of families (3.1_ω) for which $(\Sigma, \mathbb{R})$ exhibits such a minimal set.

Second, there exist weakly elliptic families (3.1_ω) such that the projective flow $(\Sigma, \mathbb{R})$ is minimal and admits a unique invariant measure. According to Lemma 3.9 and its proof, the minimality of $(\Sigma, \mathbb{R})$ implies that it is a proximal extension of $(\Omega, \mathbb{R})$. Examples of this type were constructed in a more general context by Glasner and Weiss [61]; they used ideas of Fathi and Herman [51] and Anosov and Katok [7]. Subsequently Nerurkar and Sussmann [114,115] have developed very general methods which permit in particular to obtain examples of the present type for various particular classes of coefficients matrices $A(\cdot)$. They can treat the case – very special from their point of view – when $A(\cdot)$ has the form $\begin{pmatrix} 0 & 1 \\ q(\omega) & 0 \end{pmatrix}$, which corresponds to Eqs. (3.1_ω) of Hill–Schrödinger type; they can find functions $q(\cdot)$ for which the family (3.1_ω) gives rise to a minimal, uniquely ergodic projective flow Σ which is a proximal extension of Ω. It seems possible that such projective flows arise “frequently” even in suitable parameterized families of quasi-periodic equations of Schrödinger type with analytic potential q and Diophantine frequency vector; see, e.g., Eliasson [38,39], Puig [134,133].

We do not know of examples of weakly elliptic families (3.1_ω) whose projective flows have properties qualitatively different from those illustrated above. In this regard, it would be interesting to know if there is a weakly elliptic family for which the minimal set $M \subset \Sigma$ is an almost automorphic, non-almost periodic extension of an almost periodic 2-cover Ω_2 of Ω. It would also be satisfying to decide whether or not there are examples for which M is a proximal non-almost automorphic extension of Ω but $M \neq \Sigma$.

We turn to the

(C) Weakly hyperbolic case

Once again let $(\Omega, \mathbb{R})$ be a minimal almost periodic flow, and let $A : \Omega \to \mathrm{sl}(2, \mathbb{R})$ be a continuous function. We impose the following conditions:

(i) the family (3.1_ω) does not admit an exponential dichotomy;

(ii) the Lyapunov exponent β of the family (3.1_ω) is positive: $\beta > 0$.

We say that weakly hyperbolic families (3.1_ω) are of "M-V type", in recognition of the papers of Millionščikov [108] and Vinograd [152]. Those authors constructed the first limit periodic resp. quasi-periodic examples of families (3.1_ω) of weakly hyperbolic type. They did not explicitly study the properties of the projective flows corresponding to their families. That was done later in [78,83] and other papers.

Generally speaking, the projective flow $(\Sigma, \mathbb{R})$ of a weakly hyperbolic family (3.1_ω) has remarkable properties. We will amplify this remark in the succeeding discussion. Let us begin by recalling from Section 2 that, in the weakly hyperbolic case, the projective flow $(\Sigma, \mathbb{R})$ admits exactly two ergodic measures μ_1 and μ_2. See Proposition 2.5 and the following discussion. These are supported on the measurable invariant sections of Σ which are defined by the "traces" in Σ of the Oseledets measurable invariant subbundles; see Section 2. Moreover, according to the discussion in the present Section 3, the projective flow $(\Sigma, \mathbb{R})$ contains exactly one minimal set $(M, \mathbb{R})$. It is an exercise to show that the two ergodic measures μ_1, μ_2 are actually supported on M. According to Proposition 3.9, $(M, \mathbb{R})$ is a proximal extension of $(\Omega, \mathbb{R})$.

It seems important to us to discuss the ideas underlying the construction of the examples in [108] and [152]. This is because those ideas, albeit in somewhat amplified form, have been applied by many authors since. So let us consider a fixed Bohr almost periodic function $A : \mathbb{R} \to \mathrm{sl}(2, \mathbb{R})$ together with the corresponding two-dimensional linear system

$$x' = A(t)x. \tag{3.9}$$

If x_1, x_2 are nonzero vectors in $\mathbb{R}^2$, let $\Theta(x_1, x_2)$ denote the counterclockwise angle between x_1 and x_2. Suppose that there exist a number $\delta > 0$, sequences $s_k \to -\infty$ and $t_k \to \infty$, and solutions $x_1^{(k)}(t), x_2^{(k)}(t)$ of (4.71) such that the following conditions hold:

$$\begin{aligned}
&\bigl|x_1^{(k)}(0)\bigr| = 1 \quad \text{and} \quad \frac{1}{t_k}\ln\bigl|x_1^{(k)}(t_k)\bigr| \leqslant -\delta;\\
&\bigl|x_2^{(k)}(0)\bigr| = 1 \quad \text{and} \quad \frac{1}{s_k}\ln\bigl|x_2^{(k)}(s_k)\bigr| \geqslant \delta;\\
&\Theta\bigl(x_1^{(k)}(0), x_2^{(k)}(0)\bigr) < \frac{1}{k}.
\end{aligned} \tag{$*$}$$

Let us note that Proposition 2.5 together with either of the first two conditions in $(*)$ imply that the family (3.1_ω) obtained from (3.9) via the Bebutov construction can be neither elliptic nor weakly elliptic. There remain just two possibilities: either the family (3.1_ω) has an exponential dichotomy, or it is weakly hyperbolic. Suppose for contradiction that it has an exponential dichotomy. Then it can be verified that Eqs. (4.71) admit solutions $x_1(t), x_2(t)$, with $|x_1(0)| = |x_2(0)| = 1$, such that the following conditions are valid:

$$\lim_{|t|\to\infty} \frac{1}{t}\ln\bigl|x_1(t)\bigr| \leqslant -\delta, \qquad \lim_{|t|\to\infty} \frac{1}{t}\ln\bigl|x_2(t)\bigr| \geqslant \delta. \tag{$**$}$$

We can write $x_i^{(k)}(0) = (\cos\theta_i^{(k)}, \sin\theta_i^{(k)})$ and $x_i(0) = (\cos\theta_i, \sin\theta_i)$ $(i = 1, 2)$. Passing to subsequences, we can assume that $\theta_i^{(k)} \to \bar{\theta}_i \in \mathbb{R}$ $(i = 1, 2)$. Let us show that $\bar{\theta}_1 =$

$\theta_1 \bmod \pi$. Suppose this is not true. Let $\Psi_k(t)$ be the fundamental matrix solution of (3.9) whose columns are $x_1^{(k)}(t)$ and $x_1(t)$. Then $|\sin(\theta_1^{(k)} - \theta_1)| = |\det \Psi_k(0)| = |\det \Psi_k(t)|$ is bounded away from zero for all sufficiently large k. However $\lim_{k\to\infty} \det \Psi_k(t_k) = 0$ by $(*)$ and $(**)$, so one has a contradiction.

In a similar way, one shows that $\bar{\theta}_2 = \theta_2 \bmod \pi$ (use the existence of the sequence (s_k)). Now, however, $\theta_2 = \theta_1 \bmod \pi$, because of the relation in the third line of $(*)$. We conclude that Eq. (3.9) gives rise to a family which is weakly hyperbolic.

Two points should be made concerning the above construction. First of all, Millionščikov's example [108] satisfies the condition $(*)$. Also, Vinograd's example [152] satisfies a variant of $(*)$ wherein one must translate the origin from $t = 0$ and make use of the stronger relations which substitute for $(**)$

$$\limsup_{|t-s|\to\infty} \frac{1}{t-s} \ln \frac{|x_1(t)|}{|x_1(s)|} \leqslant -\delta,$$

$$\liminf_{|t-s|\to\infty} \frac{1}{t-s} \ln \frac{|x_2(t)|}{|x_2(s)|} \geqslant \delta.$$

Second, conditions $(*)$ give a concrete meaning to the intuitive hypothesis that "the Oseledets bundles of the family (3.1_ω) are not bounded apart". This intuitive hypothesis has been rendered explicit and applied by later authors in ways similar to conditions $(*)$; see, e.g., [19,118].

Still another point to be made is that, for discrete two-dimensional system, a subharmonic trick due to Herman [68] allows one to construct discrete weakly hyperbolic cocycles with values in $\mathrm{SL}(2,\mathbb{R})$. Herman's method and generalizations thereof have been fruitfully applied in the study of the discrete Schrödinger operator, as evidenced by a large literature which we do not attempt to review.

We now show that examples of M-V type illustrate various interesting phenomena related to the behavior of solutions of linear, two-dimensional differential systems.

We first consider a problem posed by Marcus and Moore [107] concerning disconjugate solutions of the almost periodic Hill–Schrödinger equation. Consider the second-order equation

$$-\varphi'' + q(t)\varphi = E\varphi \tag{3.10}$$

where $q(\cdot)$ is a Bohr almost periodic function, and E is a real parameter. It is well-known that there is a number E_0 such that, if $E \leqslant E_0$, then every nonzero solution of (3.10) admits at most one zero on $[0,\infty)$. That is to say, Eq. (3.10) is disconjugate on $[0,\infty)$. On the other hand, if $E > E_0$, then every nonzero solution of (3.10) has infinitely many zeroes which cluster at $t = \infty$. One says that (3.10) is oscillatory at $t = \infty$. See, e.g., [28,67]. The concept of disconjugacy has connections to the calculus of variations which we do not discuss here.

If Eq. (3.10) is disconjugate, then it turns out that it admits a "smallest", or *principal* solution $\varphi_0(t)$. In the quasi periodic case (and more generally in the case when $q(\cdot)$ is

Birkhoff recurrent), the principal solution is positive on all of $\mathbb{R}$, and is uniquely defined by the following property: if $\varphi_1(t)$ is a solution of (3.10) which is independent of $\varphi_0(t)$, then

$$\frac{\varphi_0(t)}{\varphi_1(t)} \to 0 \quad \text{as } t \to \infty.$$

This condition makes sense because $\varphi_1(t)$ is nonzero for all large t. An equivalent property is the following: φ_0 is, up to a constant multiple, the unique solution of (3.10) which satisfies

$$\int^{\infty} \frac{\mathrm{d}t}{\varphi_0(t)^2} = \infty.$$

Marcus and Moore posed the following problem. Let $\varphi_0(t)$ be the principal solution of Eq. (3.10); is it the case that the logarithmic derivative $\frac{\mathrm{d}}{\mathrm{d}t} \ln \varphi_0(t)$ is almost periodic? If q is a periodic function, then $\frac{\varphi_0'(t)}{\varphi_0(t)}$ is indeed periodic with the same period as q. Also, if q is almost periodic and $\frac{\varphi_0'(t)}{\varphi_0(t)}$ is almost periodic, then the frequency module of the latter is contained in that of the former [107]. Still another general fact is that, if $E < E_0$, then the Marcus–Moore question has a positive answer. So, in this regard, E_0 is the only problematic value of E.

It turns out that the Marcus–Moore question in general has a negative answer. We proceed to explain in outline how counterexamples can be constructed.

It is convenient to rephrase the Marcus–Moore problem by introducing the Bebutov hull $(\Omega, \mathbb{R})$ of the function q. Let $Q : \Omega \to \mathbb{R}$ be a continuous function and let $\omega_0 \in \Omega$ such that $Q(\omega_0 \cdot t) = q(t)$. We introduce the family of equations

$$-\varphi'' + Q(\omega \cdot t)\varphi = E\varphi \tag{3.10_ω}$$

and the corresponding family of differential systems

$$x' = \begin{pmatrix} 0 & 1 \\ -E + Q(\omega \cdot t) & 0 \end{pmatrix} x, \quad x = \begin{pmatrix} \varphi \\ \varphi' \end{pmatrix} \in \mathbb{R}^2. \tag{3.11_ω}$$

For each $E \in \mathbb{R}$, Eqs. (3.11_ω) determine a projective flow $(\Sigma, \{\tau_t\})$.

We state without proof some general facts; see, e.g., [86] for more discussion. To begin, if $E < E_0$, then Eqs. (3.11_ω) admit an exponential dichotomy. For each $\omega \in \Omega$, let $l_\omega^+ \in \mathbb{P}$ resp. $l_\omega^- \in \mathbb{P}$ be the fiber of the stable resp. unstable bundle. Thus in the notation of Definitions 2.2., l_ω^+ is the image of the projection $P(\omega)$, while l_ω^- is its kernel. If $\Phi_\omega(t)$ is the fundamental matrix solution of Eq. (3.11_ω) and if $x_0 \in l_\omega^\pm$, then $\Phi_\omega(t)x_0 \to 0$ exponentially as $t \to \pm\infty$.

Let θ be a π-periodic angular coordinate on $\mathbb{P}$, so that $\theta = -\pi/2$ and $\theta = \pi/2$ parameterize the φ'-axis in the $\binom{\varphi}{\varphi'}$-space, and other elements $l \in \mathbb{P}$ have parameter values in $(-\pi/2, \pi/2)$. Using the particular way in which Eqs. (3.11_ω) depend on E, one can prove that, if $\omega \in \Omega$ and E is increased from $-\infty$ to E_0 then $l_\omega^+ \equiv l_\omega^+(E)$ moves counterclockwise on $\mathbb{P}$ starting near $\theta = -\pi/2$, while l_ω^- moves clockwise on $\mathbb{P}$ starting from $\theta = \pi/2$. Letting E tend to E_0 from the left, we see that, at $E = E_0$, there are two invariant sections $s^\pm$ of the projective bundle Σ, defined by $s^\pm(\omega) = \lim_{E \to E_0^-} l_\omega^\pm(E)$. These sections are semicontinuous in ω, because they can be expressed as pointwise limits of monotone sequences of continuous functions. See again [86] for a discussion of these facts. Furthermore, it can be shown [92] that, for each $\omega \in \Omega$, the line $s^+(\omega) \in \mathbb{P}$ determines the principal solution of Eqs. (3.10_ω) in the sense that, if $0 \neq \bar{x} = \binom{\bar{\varphi}}{\bar{\varphi}'} \in s^+(\omega)$, then the solution $\varphi_0(t)$ of (3.10_ω) with $\varphi_0(0) = \bar{\varphi}$, $\varphi_0'(0) = \bar{\varphi}'$ is principal; the converse statement also holds.

Let $S_\pm = \operatorname{cls}\{(\omega, s^\pm(\omega)) \mid \omega \in \Omega\} \subset \Sigma$, and let $M_\pm$ be a minimal subset of $S_\pm$. Let us observe that, by the semicontinuity of the sections $s_\pm$, there is a residual subset $\Omega_1 \subset \Omega$ such that, if $\omega \in \Omega_1$, then the fiber $\pi^{-1}(\omega) \cap S_\pm$ contains exactly one point, and hence $\pi^{-1}(\omega) \cap M_\pm$ also contains exactly one point, for each choice of sign $\pm$.

We claim that $M_+ = M_-$. For suppose that this is not true. Then M_+ and M_- are disjoint. Let $f : \Sigma \to \mathbb{R}$ be the function of Eqs. (2.3_ω), Section 2. Then if $\omega \in \Omega$, if $x(t)$ is a solution of (3.11_ω) with $|x_0(0)| = 1$, and if $l \in \mathbb{P}$ contains $x(0)$, then

$$\ln\bigl|x(t)\bigr| = \int_0^t f\bigl(\hat{\tau}_s(\omega, l)\bigr)\,\mathrm{d}s.$$

Let $m_1 \in M_+$ and let $\pi(m_1) = \omega \in \Omega$. Suppose that $\int_0^t f(\hat{\tau}_s(m_1))\,\mathrm{d}s$ is bounded (above and below) for all $t \in \mathbb{R}$. Using a familiar argument involving a matrix solution of Eq. (3.11_ω) (see the proofs of Proposition 2.5 and Proposition 3.3), one shows that, if $m_2 \in M_-$ and $\pi(m_2) = \omega$, then $\int_0^t f(\hat{\tau}_s(m_2))\,\mathrm{d}s$ is also bounded. This implies that all solutions of (3.11_ω) are bounded. But it is then an exercise to show that Eq. (3.10_ω) cannot have a principal solution. We conclude that $\int_0^t f(\hat{\tau}_s(m_1))\,\mathrm{d}s$ is unbounded for some (hence all) $m_1 \in M_+$. The same unboundedness holds for all $m_2 \in M_-$. Now however, using the oscillation Lemma 3.1, we can determine a residual subset $\tilde{\Omega}$ of Ω such that, if $\omega \in \tilde{\Omega}$, then Eq. ($3.10_\omega$) has no principal solution. So we must conclude that $M_+ = M_-$.

Let us write M for the common value of M_+ and M_-. Some additional reasoning shows that: (i) $(M, \mathbb{R})$ is an almost automorphic extension of $(\Omega, \mathbb{R})$ with projection equal to $\pi : M \to \Omega$; (ii) the Marcus–Moore question has an affirmative answer if and only if the projection π defines a flow isomorphism of $(M, \mathbb{R})$ onto $(\Omega, \mathbb{R})$. Furthermore, (iii) if π is a flow isomorphism, then the Lyapunov exponent of the family (3.11_ω) is zero. See [86].

Now however, one can construct families (3.10_ω) for which the corresponding families (3.11_ω) are of M-V type. One such family was determined in [86], using a rotation method of the sort applied by Millionščikov. Later other examples were found, in which $(\Omega, \mathbb{R})$ is quasi-periodic with Diophantine frequency vector and $Q : \Omega \to \mathbb{R}$ is real analytic (e.g., [14]). For families (3.11_ω) of M-V type, the minimal set $M \subset \Sigma$ must be an almost automorphic extension of Ω which is not almost periodic. The Marcus–Moore question has a negative answer for each Eq. (3.10_ω) if the family (3.11_ω) is of M-V type; this follows from the positivity of the Lyapunov exponent β for families of M-V type.

If Eqs. (3.11_ω) are of M-V type, then the principal solutions of the corresponding Eqs. (3.10_ω) exhibit quite remarkable behavior. We justify this remark. First, a variant of Lemma 3.1 allows one to find a residual subset $M_0 \subset M$ such that, if $m_0 \in M_0$, then

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t f\big(\hat{\tau}_s(m_0)\big)\,\mathrm{d}s = \beta, \qquad \liminf_{t\to\infty} \frac{1}{t}\int_0^t f\big(\hat{\tau}_s(m_0)\big)\,\mathrm{d}s = -\beta.$$

Let $\Omega_0 = \pi(M_0)$; one checks that Ω_0 is residual in Ω. On the other hand, there is a residual subset $\Omega_1 \subset \Omega$ such that, if $\omega \in \Omega_1$, then $\pi^{-1}(\omega) \cap M$ contains exactly one point (ω, l). If $0 \neq \bar{x} \in l$ and $\bar{x} = \binom{\bar{\varphi}}{\bar{\varphi}'}$, then the solution $\varphi_0(t)$ of (3.10_ω) satisfying $\varphi_0(0) = \bar{\varphi}$ and $\varphi_0'(0) = \bar{\varphi}'$ is a principal solution. Now $\Omega_* = \Omega_0 \cap \Omega_1$ is a residual subset of Ω. If $\omega \in \Omega_*$, then the principal solution $\varphi_0(t)$ exhibits both exponential growth and decay as $t \to \infty$, in the sense that $\limsup_{t\to\infty} \frac{1}{t} \ln \varphi_0(t) = \beta$ and $\liminf_{t\to\infty} \frac{1}{t} \ln \varphi_0(t) = -\beta$. This is surely a curious phenomenon.

As another consequence of the M-V property, we see that the principal solution of (3.10_ω) is not a continuous function of ω. In fact, in a family (3.11_ω) of M-V type, the section $s_+(\cdot)$ cannot be continuous because if they were, the projection $\pi : (M, \mathbb{R}) \to (\Omega, \mathbb{R})$ would be a flow isomorphism. Thus the lines $l_+(\omega) \in \mathbb{P}$ defining the initial conditions of the principal solutions of Eqs. (3.10_ω) have points of discontinuity; in fact $\omega \mapsto l_+(\omega)$ is only semicontinuous.

A second problem which can be solved by constructing a differential system of M-V type is the following. Return to the family (3.11_ω). The Lyapunov exponent β of this family is a function of the parameter E. It is natural to ask if $\beta = \beta(E)$ is continuous. We are in effect asking if the Lyapunov exponent of the almost periodic Schrödinger equation (3.10_ω) is always continuous as a function of E.

It is easy to see that β is continuous for values of E for which Eqs. (3.11_ω) admit an exponential dichotomy. One can further show that β is continuous at a point $E_* \in \mathbb{R}$ if $\beta(E_*) = 0$. Thus, if β is discontinuous at a point E_*, then the family (3.11_ω) with $E = E_*$ must be of M-V type, i.e., weakly hyperbolic.

There are reasons to think that β might be continuous at all values of E. For, the rotation number α of the family (3.11_ω) does depend continuously on E; this follows from Proposition 2.7 or from a direct argument of Krylov–Bogoliubov type. Moreover, the complex combination $w(E) = -\beta(E) + \mathrm{i}\alpha(E)$ admits a holomorphic extension from the real E-axis to the upper half of the complex E-plane [89]. Thus β and α are conjugate harmonic functions of E for $\Im E > 0$. One might now hope to apply the arguments available to the effect that regularity properties of $\alpha(\cdot)$ imply regularity properties of $\beta(\cdot)$ for real values of the argument E. See, e.g., [62].

It turns out, however, that one can construct an almost periodic function q for which $\beta = \beta(E)$ is discontinuous at the right end-point E_0 of the disconjugacy set $\{E \in \mathbb{R} \mid \text{Eq. } (3.10_\omega) \text{ is disconjugate}\}$. Apparently the first example of this type was given in [80]. The family (3.11_ω) corresponding to $E = E_0$ is of M-V type. By repeating the considerations concerning the stable and unstable bundles which were made in the discussion of the Marcus–Moore problem, we see that the unique minimal subflow $(M, \mathbb{R})$ of the projective flow $(\Sigma, \mathbb{R})$ determined by Eqs. (3.11_ω) is an almost automorphic extension of $(\Omega, \mathbb{R})$.

It seems that few examples are known of almost periodic Schrödinger equations for which the Lyapunov exponent $\beta = \beta(E)$ has discontinuity points. Indeed, there are results concerning the discrete quasi-periodic Schrödinger equation which state, roughly speaking, that if the frequency vector is Diophantine and if the potential is highly regular, then the Lyapunov exponent β depends continuously on the energy parameter E. See, e.g., [23]. Assuming that such results are also valid for the continuous Schrödinger operator (3.10_ω), one is led to formulate questions concerning "minimal conditions" on $(\Omega, \mathbb{R})$ and on Q which imply that β is continuous in E. So far these questions appear to be unanswered.

We close the discussion by noting that there are very general results concerning Lyapunov exponents and so-called dominated splittings which, in the present case, yield discontinuity points $Q_* \in C(\Omega)$ of β when β is thought as a function of $Q \in C(\Omega)$. See [19] for an excellent discussion of this topic. However, it is not clear from these results how one might pick out families of functions $Q = Q(E_1, \dots, E_r)$, depending on finitely many parameters, for which $\beta = \beta(E_1, \dots, E_r)$ admits points of discontinuity.

A third theme which can be elaborated upon in a fruitful way using families (3.1_ω) of M-V type is that of strange nonchaotic attractors or SNAs. In the following lines, we will summarize a part of an excellent article by Jorba, Núñez, Obaya, and Tatjer [99]; the reader is referred to that paper for more information. The main point to be made is that certain families (3.1_ω) with the M-V property give rise to projective flows which contain SNAs.

Following [99], we discuss the concept of SNA. We work for a moment in the category of *discrete* flows: we recall that a discrete flow on a metric space Ω is defined by iterating a fixed homeomorphism $\tau : \Omega \to \Omega$ and its inverse τ^{-1}. Let $\Omega = \mathbb{T}^d$, the d-torus with $d \geqslant 1$. Let $\psi_1, \dots, \psi_d$ be angular coordinates mod 2π on $\mathbb{T}^d$, and let $\gamma_1, \dots, \gamma_d$ be real numbers such that the $d + 1$-tuple $1, \gamma_1, \dots, \gamma_d$ is linearly independent over the rational field $\mathbb{Q}$. Consider the homeomorphism $\tau : \Omega \to \Omega$ defined by

$$\tau(\omega) = (\psi_1 + \gamma_1, \dots, \psi_d + \gamma_d) \quad \text{if } \omega = (\psi_1, \dots, \psi_d) \in \Omega.$$

Of course the angular quantities are taken mod 2π. It is well-known that, for each $\omega \in \Omega$, the orbit $\{\tau^n(\omega) \mid -\infty < n < \infty\}$ is dense in Ω, so one can think of the discrete flow (Ω, τ) as a discrete version of a Kronecker winding on $\Omega = \mathbb{T}^d$.

Now we introduce a continuous map $\hat{\tau} : \Omega \times \mathbb{R} \to \Omega \times \mathbb{R}$ of the following type. Let $\varphi : \Omega \times \mathbb{R} \to \mathbb{R}$ be continuous. Assume that the partial derivative $\frac{\partial \varphi}{\partial x}(\omega, x)$ is well-defined and continuous on $\Omega \times \mathbb{R}$, and that

$$\frac{\partial \varphi}{\partial x}(\omega, x) > 0 \quad \text{for all } (\omega, x) \in \Omega \times \mathbb{R}.$$

Then of course one has the monotonicity condition $\varphi(\omega, x_1) < \varphi(\omega, x_2)$ whenever $\omega \in \Omega$ and $x_1 < x_2$. We set

$$\hat{\tau}(\omega, x) = \bigl(\tau(\omega), \varphi(\omega, x)\bigr).$$

Clearly $\hat{\tau}$ is of skew-product type in the sense that the first coordinate on the right-hand side depends on ω but not on x.

Say that a Lebesgue measurable map $c:\Omega \to \mathbb{R}$ is a τ-invariant curve if it is bounded and $\varphi(\omega, c(\omega)) = c(\tau(\omega))$ for all $\omega \in \Omega$. The corresponding $\hat{\tau}$-invariant graph is $C = \{(\omega, c(\omega)) \mid \omega \in \Omega\} \subset \Omega \times \mathbb{R}$. Let ν be the unique τ-invariant measure on Ω (which coincides with the normalized Lebesgue measure on $\Omega = \mathbb{T}^d$). We can define a $\hat{\tau}$-invariant measure μ_c on $\Omega \times \mathbb{R}$ by the formula

$$\int_{\Omega\times\mathbb{R}} h(\omega, x)\, \mathrm{d}\mu_c(\omega, x) = \int_\Omega h\bigl(\omega, c(\omega)\bigr)\, \mathrm{d}\nu(\omega)$$

for each continuous function $h : \Omega \times \mathbb{R} \to \mathbb{R}$.

We introduce the *Lyapunov exponent* β_s of c, given by

$$\beta_s(c) = \int_\Omega \ln \frac{\partial\varphi}{\partial x}\bigl(\omega, c(\omega)\bigr)\, \mathrm{d}\nu(\omega) = \int_{\Omega\times\mathbb{R}} \left(\ln \frac{\partial\varphi}{\partial x}\right) \mathrm{d}\mu_c.$$

If the Lyapunov exponent $\beta_s(c)$ is negative, it measures the "vertical" contraction along the fiber $\mathbb{R}$ towards $c(\omega)$, for ν-a.a. $\omega \in \Omega$.

Next we follow T. Jäger [70] and make the following

DEFINITION 3.10. A strange nonchaotic attractor or SNA (in $\Omega \times \mathbb{R}$) is the graph of a noncontinuous, τ-invariant curve c with a negative Lyapunov exponent.

As is pointed out in [99], it is of interest to consider subsets $M \subset \Omega \times \mathbb{R}$ which are nonempty, compact, $\hat{\tau}$-invariant in the sense that $\hat{\tau}(M) \subset M$, and minimal with respect to these properties. Such a set M is called minimal. It turns out that, when restricted to M, $\hat{\tau}$ is invertible (we did not assume that $\hat{\tau}$ is invertible on $\Omega \times \mathbb{R}$), so that $(M, \hat{\tau})$ is a discrete flow. One can show that each orbit $\{\hat{\tau}^n(m) \mid -\infty < n < \infty\}$ is dense in M, and one can further show that $(M, \hat{\tau})$ is an almost automorphic extension of (Ω, τ). That is, if $\pi : \Omega \times \mathbb{R} \to \Omega$ is the projection, then there is a residual subset $\Omega_* \subset \Omega$ such that, if $\omega_* \in \Omega_*$, then $\pi^{-1}(\omega_*) \cap M$ contains exactly one point. Such minimal sets are analyzed in [99] and we refer to that paper for more information.

Whenever there exists a τ-invariant section c as above, the boundedness of the graph C ensures the existence of a minimal set M. Such an M need not contain an SNA. However, it is explained in [99] how a quasi-periodic family of Schrödinger equations with the M-V property at $E = E_0$ determines a minimal set which contains an SNA. We now discuss this matter.

Let us return momentarily to the framework of continuous (real) flows. Let $\Omega_c = \mathbb{T}^{d+1}$ be the $(d + 1)$-torus $(d \geqslant 1)$, and let $(\psi_0, \psi_1, \dots, \psi_d)$ be angular coordinates mod 2π on Ω_c. Let $\gamma_1, \dots, \gamma_d$ be real numbers such that $1, \gamma_1, \dots, \gamma_d$ are linearly independent over the rational field $\mathbb{Q}$. If $\omega_c = (\psi_0, \psi_1, \dots, \psi_d)$ mod 2π, set $\tau_t(\omega_c) = (\psi_0 + t, \psi_1 + \gamma_1 t, \dots, \psi_d + \gamma_d t)$ mod 2π. Let $Q : \Omega_c \to \mathbb{R}$ be a continuous function such that Eqs. (10_{ω_c}) have the M-V property at $E = E_0$. Let $(\Sigma, \{\hat{\tau}_t\})$ be the projective flow defined by Eqs. (10_{ω_c}).

Next let θ be the polar angle in $\mathbb{R}^2$, which we regard as a π-periodic angular coordinate on $\mathbb{P}$. It follows from the description given earlier of the projective flow that $(\Sigma, \mathbb{R})$

contains a unique minimal subflow $(M_c, \mathbb{R})$, which is an almost automorphic extension of $(\Omega_c, \mathbb{R})$. Moreover, if $(\omega, l) \in M_c$ and θ parameterizes l, then

$$-\pi/2 < \theta < \pi/2$$

for an appropriate branch of θ.

Next, consider the differential equation (4_{ω_c}) for θ of Section 2. We have

$$\theta' = g\big(\hat{\tau}_t(\omega_c, \theta)\big). \tag{3.12}$$

Let $(\omega_c, \bar{\theta}) \in \Omega_c \times \mathbb{R}$. Let $\theta(t) = \bar{\theta} + \int_0^t g(\hat{\tau}_s(\omega_c, \bar{\theta}))\, \mathrm{d}s$ be the solution of (3.12) satisfying $\theta(0) = \bar{\theta}$. We abuse notation and write $\hat{\tau}_t(\omega_c, \bar{\theta}) = (\tau_t(\omega_0), \theta(t))$. Then $\{\hat{\tau}_t\}$ defines a real flow on $\Omega \times \mathbb{R}$. If values of θ are identified mod π, then $\{\hat{\tau}_t\}$ becomes a flow on Σ defined by Eqs. (3.10_{ω_c}). We identify M_c in the natural way with a subset of $\Omega_c \times (-\pi/2, \pi/2)$; we call this subset M_c, as well.

Let us identify the d-torus $\Omega = \mathbb{T}^d$ with $\{\omega_c = (0, \omega) \mid \omega = (\psi_1, \ldots, \psi_d) \bmod 2\pi\} \subset \Omega_c$. Thus Ω is the section $\psi_0 = 0$ of Ω_c. We define $\tau : \Omega \to \Omega$: $\tau(\omega) = \tau_1(0, \omega)$, where τ_1 is the time-1 map of $\{\tau_t\}$. Then (Ω, τ) is a minimal discrete flow. We also define $\hat{\tau} : \Omega \times \mathbb{R} \to \Omega \times \mathbb{R}$ by $\hat{\tau}(\omega, \bar{\theta}) = \hat{\tau}_1(\omega, \bar{\theta})$. Let

$$M = M_c \cap \big\{(0, \omega, \bar{\theta}) \mid \omega \in \Omega, -\pi/2 < \bar{\theta} < \pi/2\big\}.$$

Then M is invariant with respect to $\hat{\tau}$, and it can be checked that $(M, \hat{\tau})$ is a minimal almost automorphic extension of (Ω, τ).

Now we return to the measurable invariant sections $s^\pm = s^\pm(\omega_c)$ of Σ which we encountered in our discussion of the Marcus–Moore problem. Let $\theta^\pm(\omega_c) \in (-\pi/2, \pi/2)$ be the corresponding values of the angular variable θ. Let $\beta > 0$ be the Lyapunov exponent of Eqs. (3.10_{ω_c}). Let us write $c^\pm(\omega) = \theta^\pm((0, \omega))$ for $\omega \in \Omega$. It can be shown that $\beta_s(c^\pm) = \mp\beta$. It follows that $C^+ = \{(\omega, c^+(\omega)) \mid \omega \in \Omega\}$ satisfies the condition of Definition 3.10, so C^+ is an SNA contained in M. See [99] for a detailed discussion and for further examples of SNAs obtained from quasiperiodic families (3.10_ω), and from quasiperiodic families of the more general form (3.1_ω) which have the M-V property.

We turn to a fourth issue which can be resolved by the construction of an almost periodic family (3.1_ω) which has the M-V property, namely a certain question concerning the behavior of the projective flow when Eqs. (3.1_ω) have unbounded mean motion. As usual, let $(\Omega, \mathbb{R})$ be an almost periodic minimal flow with unique invariant measure ν, let $A : \Omega \to \mathrm{sl}(2, \mathbb{R})$ be a continuous function, and let $(\Sigma, \{\hat{\tau}_t\})$ be the projective flow of the corresponding family (3.1_ω).

In Section 2 we defined the rotation number α of the family (3.1_ω) with respect to the measure ν. We recall the construction. If $\omega \in \Omega$ and $\bar{l} \in \mathbb{P}$, let $\hat{\tau}_t(\omega, \bar{l}) = (\tau_t(\omega), l(t))$. Introduce the polar angle θ in $\mathbb{R}^2$, and view θ as an angular coordinate mod π on $\mathbb{P}$. Letting $\theta(t)$ be a continuous parametrization of $l(t)$, we have Eq. (2.4_ω) of Section 2:

$$\theta' = g\big(\hat{\tau}_t(\omega, \bar{\theta})\big).$$

The rotation number α is by definition $\alpha = \lim_{t\to\infty}\frac{\theta(t)}{t}$. It turns out that the limit determining α is well-defined in the following sense: if ν is an ergodic measure on Ω, then there is a set $\Omega_1 \subset \Omega$ with $\nu(\Omega_1) = 1$ such that, if $(\omega, \bar{\theta}) \in \Omega_1 \times \mathbb{P}$, then the limit exists and does not depend on $(\omega, \bar{\theta})$. See [79]. In the present case, the flow $(\Omega, \mathbb{R})$ supports just one ergodic measure ν. Then, as noted in Proposition 2.7, the following stronger statement holds:

$$\alpha = \lim_{|t|\to\infty} \frac{1}{t} \int_0^t g\big(\hat{\tau}_s(\sigma)\big)\, \mathrm{d}s$$

where the limit is uniform in $\sigma \in \Sigma$.

It is natural to compare αt and $\int_0^t g(\hat{\tau}_s(\sigma))\, \mathrm{d}s$.

DEFINITION 3.11. Say that Eqs. (3.1_ω) have *bounded mean motion* if, for some $\sigma \in \Sigma$, there holds

$$\sup_{t\in\mathbb{R}} \left| \int_0^t g\big(\hat{\tau}_s(\sigma)\big)\, \mathrm{d}s - \alpha t \right| < \infty. \tag{$*$}$$

We now follow the discussion of Yi ([154]; also [16,71]). Yi works with general nonlinear skew-product flows on a circle bundle over Ω, so our case is special from his point of view. First of all, if Eqs. (3.1_ω) have bounded mean motion in the sense of Definition 3.11, then it turns out that condition $(*)$ holds for all $\sigma \in \Sigma$. See ([154], Proposition 3.1).

Second, if Eqs. (3.1_ω) have bounded mean motion, then one can draw deep conclusions about the structure of a minimal subflow $(M, \mathbb{R})$ of the projective flow $(\Sigma, \mathbb{R})$. We refer to [154] for precise statements and proofs, and only summarize the situation here. Assume for simplicity that $(\Omega, \mathbb{R})$ is quasiperiodic. Let $\mathcal{M}_\Omega$ be the frequency module of $(\Omega, \mathbb{R})$; we can write $\mathcal{M}_\Omega = \{n_1\gamma_1 + \cdots + n_d\gamma_d \mid n_1, \ldots, n_d \in \mathbb{Z}\}$ where $\gamma_1, \ldots, \gamma_d \in \mathbb{R}$ are linearly independent over $\mathbb{Q}$. Let $\mathcal{M}_1 = \{n_0\alpha + n_1\gamma_1 + \cdots + n_d\gamma_d \mid n_0, n_1, \ldots, n_d \in \mathbb{Z}\}$; then $\mathcal{M}_1$ is the frequency module of a quasiperiodic flow $(\Omega_1, \mathbb{R})$. The relation between $(\Omega_1, \mathbb{R})$ and $(\Omega, \mathbb{R})$ can be described in terms of the existence and nature of integral relations between $\alpha, \gamma_1, \ldots, \gamma_d$ [154].

It turns out that $(M, \mathbb{R})$ is an almost periodic minimal flow $(\Omega_*, \mathbb{R})$ which is "intermediate" between $(\Omega_1, \mathbb{R})$ and $(\Omega, \mathbb{R})$. If $n_0\alpha \notin \mathcal{M}_\Omega$ for all integers n_0, then $(\Omega_*, \mathbb{R}) = (\Omega_1, \mathbb{R})$. One thus has a general description of the minimal subsets of Σ when bounded mean motion is present. Yi's theory can be extended to the case of a general almost periodic minimal flow $(\Omega, \mathbb{R})$. One needs to define $\mathcal{M}_*$ to be the smallest subgroup of $\mathbb{R}$ which contains α and $\mathcal{M}_\Omega$.

The case when Eqs. (3.1_ω) have unbounded mean motion is less well-understood. Suppose for example that there does not exist an integer $r \neq 0$ such that $r\alpha \in \mathcal{M}_\Omega$. We say that α is irrationally related to $\mathcal{M}_\Omega$. Suppose in addition that Eqs. (3.1_ω) have unbounded mean motion. It has been asked in the literature if these conditions imply that $(\Sigma, \mathbb{R})$ is minimal.

The point we wish to make is that the answer to this question is no. In fact, as Yi notes, an example of [75] has the following properties: (i) there is no integer r such that $r\alpha \in \mathcal{M}_\Omega$;

(ii) Eqs. (3.1_ω) have the M-V property; (iii) the unique minimal set $M \subset \Sigma$ is an almost automorphic extension of Ω. In particular, $M \neq \Sigma$.

So far, we have seen that almost periodic families (3.1_ω) with the M-V property can be constructed which illustrate Marcus–Moore disconjugacy, a discontinuous Lyapunov exponent, and strange nonchaotic attractors. We have also seen that they may give rise to interesting minimal subsets $M \subset \Sigma$ when the bounded mean motion property is absent. All the minimal sets $M \subset \Sigma$ which have arisen have been almost automorphic extensions of Ω.

We note in passing that an almost automorphic minimal subset M of the projective flow $(\Sigma, \mathbb{R})$ determined by an almost periodic flow $(\Omega, \mathbb{R})$ may have interesting topological properties. For instance, the original Vinograd example [152] gives rise to an almost automorphic minimal set $M \subset \Sigma$ which is connected, locally connected at some points, but not locally connected at all points [84]. See [128] for other examples exhibiting this behavior which are invariant sets for 2-dimensional (nonlinear) differential systems. As another instance, we note that the minimal set M mentioned above in connection with mean motion property has "more" Čech cohomology than Ω in a precise sense, despite the fact that M is an almost automorphic extension of Ω. See [17].

We consider a final issue which can be resolved via the construction of an appropriate almost periodic family (3.1_ω) with the M-V property. The issue is that of the existence of minimal flows which have the property of Li–Yorke chaos, but which do not exhibit such traditional indicators of complicated orbital behavior as weak mixing or positive topological entropy. We refer to [18] for a general discussion of flows which do/do not exhibit Li–Yorke chaos. The concept itself derives from a classical paper by Li and Yorke [103].

Let us begin our discussion with the

DEFINITION 3.12. Let Ω be a metric space with metric d and let $(\Omega, \mathbb{R})$ be a flow. A pair $\{\omega_1, \omega_2\} \subset \Omega$ is a *Li–Yorke pair* if $\sup_{t>0}\{d(\omega_1 \cdot t, \omega_2 \cdot t)\} > 0$ and $\inf_{t>0}\{d(\omega_1 \cdot t, \omega_2 \cdot t)\} = 0$. A set $S \subset \Omega$ is called *scrambled* if each pair $\{\omega_1, \omega_2\} \subset S$ is a Li–Yorke pair. The flow $(\Omega, \mathbb{R})$ is *chaotic* in the *sense of Li and Yorke* if Ω contains an uncountable scrambled set.

Let us now focus on the problem of determining almost periodic families (3.1_ω) for which the projective flow $(\Sigma, \mathbb{R})$ admits a minimal subflow $(M, \mathbb{R})$ which exhibits Li–Yorke chaos. Let us note that such a minimal set cannot be weakly mixing. This is because a weakly mixing flow $(X, \mathbb{R})$ has by definition the property that there is a pair $(x_1, x_2) \in X \times X$ with $x_1 \neq x_2$ such that the set $\{(x_1 \cdot t, x_2 \cdot t) \mid t \in \mathbb{R}\}$ is dense in $X \times X$. If a pair $(m_1, m_2) \in M \times M$ had this property, then the pair $(\omega_1, \omega_2) \in \Omega \times \Omega$ of images $\omega_i = \pi(m_i)$ $(i = 1, 2)$ would have the property that $\{(\omega_1 \cdot t, \omega_2 \cdot t) \mid t \in \mathbb{R}\}$ is dense in $\Omega \times \Omega$. However, no such pair exists.

We now follow the remarkable thesis of Bjerklöv ([13]; see also [14]). Let Ω be the 2-torus $\mathbb{T}^2$ with angular coordinates (ψ_1, ψ_2) mod 1 Let γ be an irrational number, and set $\tau_t(\psi_1, \psi_2) = (\psi_1 + t, \psi_2 + t)$ mod 1. Let $Q : \Omega \to \mathbb{R}$ be a C^3 function which has a unique minimum on $\Omega = \mathbb{T}^2$. Suppose further that Q is quadratic in a neighborhood of the point ω_0 where it assumes its minimum value. Thus, assuming for convenience that Q attains its minimum value at $(\psi_1, \psi_2) = (0, 0)$, it is required that $Q(\psi_1, \psi_2) = B(\psi_1, \psi_2) + \mathrm{O}((\psi_1^2 + \psi_2^2)^{3/2})$ where B is a positive definite bilinear form.

Let us consider the following family of linear differential systems of Schrödinger type:

$$x' = \begin{pmatrix} 0 & 1 \\ -E + \lambda Q\bigl(\tau_t(\omega)\bigr) & 0 \end{pmatrix} x. \tag{3.14$_\omega$}$$

Here λ is a real parameter. Among other results, Bjerklöv shows that it is possible to find a (large) λ to which there correspond a (Diophantine) frequency γ and a (large) number E such that the corresponding family (3.14$_\omega$) is of M-V type, and determines a projective flow $(\Sigma, \mathbb{R})$ which is minimal. Such a minimal projective flow differs essentially those of Glasner–Weiss type which we encountered in the weakly elliptic case, because (3.14$_\omega$) has a positive Lyapunov exponent and hence $(\Sigma, \mathbb{R})$ supports exactly two ergodic measures.

Now, using an old triangularization technique (see [83], Section 5), one can show that there is a residual subset $\Omega_* \subset \Omega$ with the following property: if $\omega \in \Omega_*$, then $\pi^{-1}(\omega)$ is a scrambled subset of Σ. In fact, if $l_1, l_2 \in \pi^{-1}(\omega)$, let $\Theta(l_1, l_2)$ be the counterclockwise angle from l_1 to l_2. Let $\Phi_\omega(t)$ be the fundamental matrix solution of (3.14$_\omega$), and set $l_i(t) = \Phi_\omega(t) l_i$ $(i = 1, 2)$. Then one can show that

$$\inf_{t\in\mathbb{R}} \Theta\bigl(l_1(t), l_2(t)\bigr) = 0 \quad \text{and} \quad \sup_{t\in\mathbb{R}} \Theta\bigl(l_1(t), l_2(t)\bigr) = \pi.$$

Thus these examples have the Li–Yorke property. They have zero topological entropy because of a general result of Bowen [22]. See [17] for details. We note that it turns out that the Glasner–Weiss examples also have the Li–Yorke property; see [17].

We close the section by posing a question concerning the projective flow determined by almost periodic families (3.1$_\omega$) of linear differential systems. Namely, can one determine a family (3.1$_\omega$) for which $(\Sigma, \mathbb{R})$ admits a minimal subset M with the property that, for a residual set of $\omega \in \Omega$, the fiber $\pi^{-1}(\omega) \cap M$ is perfect and nowhere dense in $\mathbb{P}$? That is, does there exist an almost periodic, $\mathrm{SL}(2, \mathbb{R})$-valued "Denjoy cocycle"? According to a private communication of T. Jäger, this cannot happen when Ω is a 2-torus. Can other phase spaces Ω be excluded as well?

4. Algebro-geometric Sturm–Liouville coefficients

This section is devoted to the discussion of an inverse spectral problem for the classical Sturm–Liouville operator. The present discussion has interest for various reasons, which we want briefly to explain. First, we exploit methods from the theory of nonautonomous differential systems in dimension 2, by using techniques from ergodic theory, dynamical systems and other fields as well. It is our opinion that such an approach to the study of the inverse spectral properties of the Sturm–Liouville operator allows us to retrieve more information than is present in other sources. Second, it is well known that certain Sturm–Liouville operators have the property that their isospectral classes are preserved by the solutions of a corresponding nonlinear evolution equation, and indeed by the solution of a whole commuting hierarchy of such equations. A prototypical and extensively studied example of this phenomenon is the following. Let $u_t(x) = u(t,x)$ be the solution of the Korteweg–de Vries (K-dV) equation

$$\frac{\partial u}{\partial t} = 3u\frac{\partial u}{\partial x} - \frac{1}{2}\frac{\partial^3 u}{\partial x^3},$$

with appropriate initial condition $u_0(x) = u(0,x)$; then the spectrum $\Sigma(L_t)$ of the Schrödinger operator

$$L_t = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + u_t(x)$$

acting in $L^2(\mathbb{R})$ does not depend on t [56]. This fact permits one to explicitly solve the K-dV equation for an ample class of initial data u_0, as a vast literature testifies. The K-dV equation is one of an infinite family of commuting nonlinear evolution equations whose solutions preserve isospectral classes of the Schrödinger operator.

Another example is furnished by the Camassa–Holm equation

$$4\frac{\partial u}{\partial t} - \frac{\partial^3 u}{\partial x^2 \partial t} - 2u\frac{\partial^3 u}{\partial x^3} - 4\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x^2} + 24u\frac{\partial u}{\partial x} = 0.$$

In this case, the corresponding operator is of Sturm–Liouville type with a nonconstant density function. Actually more than one operator can be used. Following [11], we choose $-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + 1$; introducing a positive density function y, the relevant spectral problem is

$$-\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} + \varphi = \lambda y(x)\varphi. \tag{4.1}$$

It turns out that, if $u_t(x) = u(t,x)$ is an appropriate solution of the Camassa–Holm equation, then there is a t-dependent density $y_t = y_t(x)$ such that the spectrum of (4.1) does not depend on t.

We mention that there are also nonlinear evolution equations which are related to isospectral classes of ordinary differential operators which are not of Sturm–Liouville type. As an example, we mention the (nonfocusing) nonlinear Schrödinger (NLS) equation:

$$-\mathrm{i}\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 2|u|^2 u,$$

together with the AKNS operator

$$L = J^{-1}\left[\frac{\mathrm{d}}{\mathrm{d}x} - \begin{pmatrix} a & b \\ b & -a \end{pmatrix}\right], \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

acting on $L^2(\mathbb{R}, \mathbb{C}^2)$. Here, a and b are real-valued functions of x. For appropriate solutions $u_t(x) = u(t,x)$ of the NLS equation, there is a transformation $u_t \mapsto (a_t, b_t)$ such that the spectrum of

$$L_t = J^{-1}\left[\frac{\mathrm{d}}{\mathrm{d}x} - \left(\begin{pmatrix} a_t & b_t \\ b_t & -a_t \end{pmatrix}\right)\right]$$

does not depend on t. See [1] and much additional literature for more information.

Roughly speaking, there are two classical methods which are used to study the isospectral classes of an operator of Sturm–Liouville type: the first is that of the "inverse scattering method" (see, for instance [11,132]), the second is that in which methods of the theory of algebraic curves are used [35,118,34,96,95,155].

These facts provide a wonderful demonstration of the interplay between the theory of nonautonomous differential systems, algebraic curves and partial differential equations.

In the last two decades, many papers concerning the inverse spectral theory of the Schrödinger operator have been published (see, among others, [11,35,118,132]). We take inspiration from these works, and extend some of those well-known results to a general one dimensional Sturm–Liouville operator L, namely

$$L(\varphi) = (p\varphi')' - q\varphi = -\lambda y\varphi, \quad ' = \frac{\mathrm{d}}{\mathrm{d}t} \tag{4.2}$$

acting on $L^2(\mathbb{R})$.

Here is a more detailed formulation of the problem we study. Equation (4.2) can be written in system form as

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}' = \begin{pmatrix} 0 & \frac{1}{p(t)} \\ q(t) - \lambda y(t) & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{4.3}$$

Here p, q, y are uniformly continuous functions, p, y are strictly positive, and $p \in C^1(\mathbb{R})$ has bounded uniformly continuous derivative. This system has, in the nonautonomous case, a maximal Lyapunov exponent $\beta = \beta(\lambda)$. We assume that:

- (H1) The spectrum $\Sigma(L)$ of (4.2) is a finite union of intervals: $\Sigma(L) = [\lambda_0, \lambda_1] \cup [\lambda_2, \lambda_3] \cup \cdots \cup [\lambda_{2g}, \infty)$, where $-\infty < \lambda_0 < \lambda_1 < \cdots < \lambda_{2g}$;
- (H2) The Lyapunov exponent β vanishes on $\Sigma(L)$: $\beta(\lambda) = 0$ for all $\lambda \in \Sigma(L)$.

As we will see, these mild-looking assumptions are sufficient to permit the reconstruction of p, q, and y in terms of algebro-geometric data. In particular, we will see that $\sqrt{py}$ can be described as the restriction of a meromorphic function defined on a generalized Jacobian of the Riemann surface $\mathcal{R}$ of the algebraic relation

$$w^2 = -\prod_{i=0}^{2g}(\lambda - \lambda_i)$$

to a t-motion we will describe in detail and which is general nonlinear (unless $y/p = const$). Some words concerning the assumptions on the spectrum (H1) and (H2): in general, the spectrum $\Sigma(L)$ of a Sturm–Liouville operator L can have a very variegated structure. It can contain isolated points, bands, Cantor parts, or even more complicated sets. One can say that, since L (at least when $p(t)$ is of class $C^1(\mathbb{R})$) defines a self-adjoint operator on the weighted space $L^2(\mathbb{R}, \sqrt{y})$, the spectrum $\Sigma(L)$ is contained in $\mathbb{R}$, is bounded below by a number λ_0 and the set $\mathbb{R} \setminus \Sigma(L)$ is at most a countable union of disjoint open intervals $(\lambda_{2i-1}, \lambda_{2i})$ ($i \in \mathbb{N}$): $\mathbb{R} \setminus \Sigma(L) = (-\infty, \lambda_0) \bigcup_{i=1}^{\infty} (\lambda_{2i-1}, \lambda_{2i})$. However, it is fairly clear that it can be very difficult to study the properties of the spectrum of a Sturm–Liouville operator, as well as to formulate a consistent setting for an inverse problem to be solved. For this reason it is necessary to give assumptions on the functions p, q, y in order to have spectral structures which are "not too wild": those assumptions are equivalent to requiring that the Green function $G(t, s, \lambda)$ associated to Eq. (4.2) behaves "well" at the endpoints of the spectral gaps. This happens, for instance, in the case when we consider potentials which are called "reflectionless", i.e. potentials for which the nontangential limit $\lim_{\varepsilon \to 0^+} \Re G(t, t, \lambda + i\varepsilon) = 0$ for all $t \in \mathbb{R}$ and for Lebesgue a.a. $\lambda \in \Sigma(L)$. This class of potentials contains periodic potentials [106], ergodic potentials [34], algebro-geometric potentials [35,96,95], Sato–Segal–Wilson potentials [141,85], Eliasson potentials [38] and others as well: of course the potentials we consider here are reflectionless. It is possible to extend the results we will obtain under the Hypotheses (H1) and (H2) to some of the more general classes of potentials listed above, by using certain limit procedures (see [97]): in fact, all such potentials are reflectionless and have a spectrum $\Sigma(L)$ which is a countably infinite union of closed intervals such that the length of the spectral gaps satisfy certain specific assumptions. It is not yet known if all the potentials obtained by the limit procedure in [97] are almost periodic (see [34] for a detailed discussion in the case of the Schrödinger operator).

We follow the approach taken in [31] and [57], where analogous inverse problems were posed and solved for the AKNS operator and for the Verblunsky operator arising in the theory of orthogonal polynomials on the unit circle [148].

The key element we use to study our inverse spectral problem is a "dynamical" definition of the Weyl m-functions $m_\pm(\lambda)$. An ample discussion concerning the Weyl m-functions is present in the following. We will sketch the main results of the Gilbert–Person theory for the Schrödinger operator, which can be easily adapted to our more general Sturm–Liouville operator.

We now give a brief summary of the content of this section. It is divided into four subsections.

In Section 4.1, we review some basic facts concerning the Sturm–Liouville operator. Although these facts are very well known, we will systematically use them, hence it could be helpful for the reader to have a quick guide in order to retrieve the necessary information.

We begin Section 4.2 by describing the ergodic-dynamical setting. We define a family $\mathcal{F}$ of Sturm–Liouville systems (4.3), where the functions p, q, y vary in the hull of a given triple (p_0, q_0, y_0) of continuous functions which satisfies certain assumptions. We will apply the concept of exponential dichotomy, and define the Weyl m-functions in a dynamical way.

In Section 4.3 we consider the projective flow in the Sturm–Liouville setting. In particular, we let Ω be a compact metric space, and $\{\tau_t\}$ be a flow on Ω: the projective flow is built in analogy to that of the preceding sections, as $(\Sigma = \Omega \times \mathbb{P}(\mathbb{C}), \{\hat{\tau}_t\})$. Then we introduce the Floquet exponent, the rotation number, the Lyapunov exponent and we extend Kotani's seminal work on the Schrödinger operator [101]. We use systematically the projective flow. We will draw many important conclusions about the behavior of the Weyl m-functions when λ crosses the real line. Moreover we prove that the spectra of the family of Sturm–Liouville equations $\mathcal{F}$ are invariant, i.e. that for a.e. equation in $\mathcal{F}$, the spectrum equals a fixed closed set with no isolated eigenvalues and coincides with the non-dichotomy set of the family $\mathcal{F}$.

In Section 4.4 we pose and solve the inverse problem. In particular, first we find relations occurring between the potentials p, q, y in terms of the finite poles of a meromorphic function which is directly linked to the Weyl m-functions, then we characterize all the ergodic potentials p, q, y satisfying (H1) and (H2).

4.1. *Basic results on the Sturm–Liouville operator*

This part mainly refers to Atkinson [10], Coddington and Levinson [32] and Weinberger [153]. Consider the following differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(p(t)\frac{\mathrm{d}f}{\mathrm{d}t}\right) - q(t)f = -\lambda y(t)f(t), \quad t \in [0, b], \tag{4.4}$$

where p, q and y are real valued functions such that $p(t)$ is positive and of class C^1, $y(t)$ is positive and continuous, and $q(t)$ is continuous, together with the boundary conditions

(4.4a) $f(0)\cos\alpha + p(0)f'(0)\sin\alpha = 0$,

(4.4b) $f(b)\cos\beta + p(b)f'(b)\sin\beta = 0$,

where $\alpha, \beta \in [0, \pi)$.

DEFINITION 4.1. The values of $\lambda \in \mathbb{C}$ such that Eq. (4.4) together with the boundary conditions (4.4a) and (4.4b) has a nontrivial solution are the *eigenvalues*. The corresponding solutions are called *eigenfunctions*.

PROPOSITION 4.2. *Eigenfunctions corresponding to different eigenvalues are orthogonal with respect the weight function y.*

All the eigenvalues are real and positive. The set of eigenvalues is at most countable, and has a unique cluster point at ∞.

Equation (4.4) can be written in the following matrix form

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1/p \\ q - \lambda y & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{4.5}$$

which is equivalent to the system

$$\begin{cases} p u_1' = u_2, \\ u_2' = (q - \lambda y)u_1. \end{cases} \tag{4.6}$$

Let $u(t,\lambda)=\binom{u_1(t,\lambda)}{u_2(t,\lambda)}$ be the solution of (4.5) satisfying some fixed initial condition. As $|\lambda|\to\infty$, both u_1 and u_2 behave as follows:

$$u_1(t,\lambda), u_2(t,\lambda)=\mathcal{O}\bigl\{\mathrm{e}^{c\sqrt{|\lambda|}}\bigr\}, \tag{4.7}$$

where c is a constant.

From Eq. (4.7) we obtain:

- (4.7a) the maps $\lambda\mapsto u_{1,2}(t,\lambda)$ are entire and of order at most $1/2$ for every fixed $t\in(0,b)$;
- (4.7b) $u_{1,2}(t,\lambda)$ are entire functions if $t=b$, hence the map $\lambda\mapsto u_1(b,\lambda)\cos\beta+u_2(b,\lambda)\sin\beta$ is entire; this implies that this function can have only countably many zeros, and of course these zeroes are the eigenvalues;
- (4.7c) $\sum_n\lambda_n^{-1/2-\varepsilon}<\infty$ for each $\varepsilon>0$.

For results concerning the exponential growth of the zeros of an entire function, i.e. to prove (4.7c), see, among others, [149, Section 8.22]. In addition the eigenvalues can cluster only at ∞.

Let us now consider Eq. (4.4) together with the boundary conditions

$$f(0)=f(b)=0. \tag{4.8}$$

Let λ_1 be the smallest eigenvalue of Eq. (4.4) with boundary conditions (4.8), and let χ_1 be the corresponding eigenfunction. Then

$$0\neq\lambda_1=\frac{\int_0^b(p(\chi_1')^2+q\chi_1^2)\,\mathrm{d}t}{\int_0^b y\chi_1^2\,\mathrm{d}t}.$$

It is easy to prove that χ_1 does not vanish in $(0,b)$. Since the eigenfunctions are orthogonal, all the other eigenfunctions *must* change sign in $(0,b)$: even more, if $\lambda<\rho$ are different eigenvalues, χ and ψ are the corresponding eigenfunctions and t_0,t_1 are two consecutive zeroes of χ, then there exists $\bar t\in(t_0,t_1)$ such that $\psi(\bar t)=0$. This implies that

PROPOSITION 4.3. *The kth eigenfunction χ_k of Eq.* (4.4) *together with the boundary conditions* (4.8) *has exactly $k-1$ zeroes in* $(0,b)$.

Starting from Eq. (4.4) (which is considered in the interval $[0,b]$, $b\in\mathbb{R}$), we now retrieve information concerning the solutions of the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(p(t)\frac{\mathrm{d}f}{\mathrm{d}t}\right)-q(t)f=-\lambda y(t)f(t),\quad t\in[0,\infty). \tag{4.9}$$

To this end we introduce the Weyl m-functions for the Sturm–Liouville operator.

Let φ_α and ψ_α two linearly independent solutions of (4.4) satisfying the following initial conditions at 0

- (4.9a) $\varphi_\alpha(0,\lambda)=\sin\alpha$, $\psi_\alpha(0,\lambda)=\cos\alpha$
- (4.9b) $p(0)\varphi_\alpha'(0,\lambda)=-\cos\alpha$, $p(0)\psi_\alpha'(0,\lambda)=\sin\alpha$,

where $0 \leqslant \alpha < \pi$. Note that φ satisfies the condition (4.4a). Any other solution χ of Eq. (4.4) (except for ψ) has the form

$$\chi(t) = \varphi_\alpha(t) + m(\alpha, \lambda)\psi_\alpha(t),$$

where $m(\alpha, \lambda) \in \mathbb{C}$. We omit for the moment the α-dependence of the functions involved. If we require that χ satisfies the boundary condition (4.4b), namely

$$\chi(b)\cos\beta + p(b)\chi'(b)\sin\beta = 0,$$

then m must satisfy

$$m = -\frac{\varphi(b,\lambda)\cot\beta + p(b)\varphi'(b,\lambda)}{\psi(b,\lambda)\cot\beta + p(b)\psi'(b,\lambda)}.$$

Clearly, now we have $m = m(\alpha, \beta, b, \lambda) \in \mathbb{C}$.

From the above observations, all the functions $\varphi, \varphi', \psi, \psi'$ are entire functions in λ, hence m is meromorphic in the variable λ. Let us set $z = \cot\beta$; then for fixed (λ, b) we obtain

$$m = -\frac{Az + B}{Cz + D} \tag{4.10}$$

where z varies on the real line, and A, B, C, D are determined by the expression above. The image of the real line is a circle $C_b(\lambda)$ in the m-plane. Thus χ satisfies the boundary conditions in b if and only if $m \in C_b(\lambda)$.

If we use the notation

$$[f, g](t) = p(t)\big(f(t)\bar{g}(t)' - f(t)'\bar{g}(t)\big), \quad f, g \in C^1(0, b)$$

than the equation of the circle $C_b(\lambda)$ is

$$[\chi, \chi](b) = 0, \tag{4.11}$$

that is

$$C_b(\lambda) = \big\{m \in \mathbb{C} \mid [\chi, \chi](b) = 0\big\}.$$

The center c and the radius r of $C_b(\lambda)$ are

$$c = -\frac{[\varphi, \psi](b)}{[\psi, \psi](b)} \tag{4.12}$$

$$r = \frac{1}{|[\psi, \psi](b)|}, \tag{4.13}$$

The interior of the circle $C_b(\lambda)$ is given by the relation

$$\frac{[\chi,\chi](b)}{[\psi,\psi](b)} < 0.$$

Moreover

$$[\psi,\psi](b) = 2\mathrm{i}\Im\lambda \int_0^b y|\psi|^2\,\mathrm{d}t,$$

and

$$[\chi,\chi](b) - [\chi,\chi](0) = 2\mathrm{i}\Im\lambda \int_0^b y|\chi|^2\,\mathrm{d}t.$$

Since $[\chi,\chi](0) = -2\mathrm{i}\Im m$, the interior of the circle is given by

$$\int_0^b y|\chi|^2\,\mathrm{d}t < \frac{\Im m}{\Im\lambda} \quad (\Im\lambda \neq 0).$$

Points m are on the circle if and only if

$$\int_0^b y|\chi|^2\,\mathrm{d}t = \frac{\Im m}{\Im\lambda} \quad (\Im\lambda \neq 0). \tag{4.14}$$

The radius r is given, if $\Im\lambda \neq 0$, by

$$\frac{1}{r} = 2|\Im\lambda| \int_0^b y|\psi|^2\,\mathrm{d}t.$$

Let $0 < a < b < \infty$. If m is inside or on the circle $C_b(\lambda)$, we have

$$\int_0^a y|\chi|^2\,\mathrm{d}t < \int_0^b y|\chi|^2\,\mathrm{d}t \leqslant \frac{\Im m}{\Im\lambda},$$

so m is inside or on the circle $C_a(\lambda)$ as well. This means that if $a < b$, then $C_a(\lambda)$ contains $C_b(\lambda)$. For fixed λ with $\Im\lambda \neq 0$, the map $b \mapsto r(b) = (2|\Im\lambda| \int_0^b y|\psi|^2\,\mathrm{d}t)^{-1}$ is monotone decreasing, and hence the limit $\lim_{b\to+\infty} r(b) := r_+$ exists. We have two possibilities: either $r_+ = 0$ or $r_+ > 0$. In the first case the circles $C_b(\lambda)$ tend to a point $m_+(\lambda)$, in the second case the $C_b(\lambda)$ tend to a circle $C_+(\lambda)$. In the first case, since $r(b) \to 0$, then $\psi(t) \notin L^2([0,\infty), y)$, hence there exists only one solution (up to a constant multiple) which belongs to $L^2([0,\infty), y)$, and this solution is surely χ. In the other case, we can easily see that $\psi \in L^2([0,\infty), y)$. If $\hat{m}$ is any point of $C_+(\lambda)$ then $\hat{m}$ is inside every $C_b(\lambda)$, for all $b > 0$, so

$$\int_0^b y|\chi|^2\,\mathrm{d}t \leqslant \frac{\Im\hat{m}}{\Im\lambda},$$

and if $b\to\infty$, this proves that every solution $\chi=\varphi+\hat{m}\psi$ is in $L^2([0,\infty),y)$.

REMARK 4.4. The discussion we made above shows that solutions of (4.9) which are in $L^2([0,\infty),y)$ always exist: in the limit circle case, every solution which satisfies the boundary conditions mentioned above is square integrable, while in the limit point case there is (up to a constant multiple) only one solution which is square integrable, if $\Im\lambda\neq 0$, and this solution is $\chi(t)$.

We recall

PROPOSITION 4.5. *Let p,q,y be real valued bounded continuous functions such that $p(t)$ is positive and of class C^1, and $y(t)$ is positive. Then Eq.* (4.4) *together with the boundary conditions* (4.9a) *and* (4.9b) *is in the limit point case at ∞.*

As a consequence of Proposition 4.5, for every fixed λ with $\Im\lambda\neq 0$ there exists only one square integrable solution of (4.9) (up to a constant multiple) which is in $L^2([0,\infty),y)$. It turns out that the limit $\lim_{b\to\infty}C_b(\lambda)$ does not depend on the particular choice of β in the boundary conditions. In particular, if we put $\beta=0$ we obtain

$$m_+(\alpha,\lambda)=-\lim_{b\to\infty}\frac{\varphi_\alpha(b,\lambda)}{\psi_\alpha(b,\lambda)}.$$

REMARK 4.6. From now on, unless explicitly stated otherwise, we only consider $\alpha=\frac{\pi}{2}$ and set $m_+(\lambda)=m_+(\pi/2,\lambda)$, to obtain the particularly important relation

$$m_+(\lambda)=\frac{p(0)\chi'(0)}{\chi(0)}, \tag{4.15}$$

where $\chi(t)$ is the only solution of (4.9) which is in $L^2([0,\infty),y)$.

One can retrieve $m_+(\alpha,\lambda)$ from $m_+(\lambda)$ and vice-versa, by using the following simple formulas:

$$m_+(\lambda)=-\frac{\cos\alpha-m_+(\alpha,\lambda)\sin\alpha}{\sin\alpha+m_+(\alpha,\lambda)\cos\alpha},$$

$$m_+(\alpha,\lambda)=\frac{\cos\alpha+m_+(\lambda)\sin\alpha}{\sin\alpha-m_+(\lambda)\cos\alpha}.$$

The function $m_+(\lambda)$ is the *Weyl m-function* at $+\infty$. As is easily seen, an analogous construction can be made for the problem in the negative half-line, giving a Weyl m-function at $-\infty$. We will denote this function with $m_-(\lambda)$. Let $\varphi_+(t,\lambda)$ and $\varphi_-(t,\lambda)$ be the (essentially unique) solutions of (4.9) which are square integrable on the positive and negative half lines respectively (obviously with respect to the weighted measure $y\,\mathrm{d}t$). From (4.15) we have the following important formulas

$$m_+(\lambda) = \frac{p(0)\varphi'_+(0,\lambda)}{\varphi_+(0,\lambda)}, \qquad m_-(\lambda) = \frac{p(0)\varphi'_-(0,\lambda)}{\varphi_-(0,\lambda)}. \tag{4.16}$$

We have

PROPOSITION 4.7. *If* $\Im(\lambda) \neq 0$, $m_\pm$ *are analytic functions of* λ. *In addition,* $\operatorname{Sgn}(\Im m_\pm \Im\lambda) = \pm 1$. *If* $m_\pm$ *have zeroes or poles, they are all real and simple.*

Consider the operator

$$L(\varphi) = \frac{1}{y}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(p\frac{\mathrm{d}}{\mathrm{d}t}\varphi\right) - q\varphi\right]. \tag{4.17}$$

We will denote by L the self-adjoint closure of operator (4.17) on $L^2(\mathbb{R}, y)$ and with $L^\pm$ the closures of (4.17) on $L^2([0,\infty), y)$ and $L^2((-\infty,0], y)$ respectively.

Consider Eq. (4.4) in the interval $[0,b]$, together with the boundary conditions

$$u(0) = 0, \qquad p(b)u'(b) = 1.$$

Let $(\lambda_{b_n}) \subset \mathbb{R}$ be the sequence of eigenvalues of Eq. (4.4), and let (u_{b_n}) be the corresponding sequence of eigenfunctions. If ψ is the solution of (4.4) which satisfies the second half of the conditions (4.9a) and (4.9b), then ψ satisfies $\psi(0) = 0$, hence we must have $u_{b_n}(t) = r_{b_n}\psi(t,\lambda_{b_n})$, where $r_{b_n} \in \mathbb{C}$.

Let $\rho_b : \mathbb{R} \to \mathbb{R}$ be the monotone nondecreasing step function having jumps of $|r_{b_n}|^2$ at each eigenvalue λ_{b_n}. We take ρ_b to be right-continuous, and $\rho_b(0) = 0$. The function ρ defines a measure, that we still denote by ρ_b, called the spectral measure for Eq. (4.4) in the interval $(0,b)$. If σ is a measure of Lebesgue–Stieltjes type on $\mathbb{R}$, let $L^2(\mathbb{R},\sigma)$ be the set of all the functions $h(\lambda)$ which are σ-measurable and such that

$$\int_{-\infty}^{\infty} \left|h(\lambda)\right|^2 \mathrm{d}\sigma(\lambda) < \infty.$$

We have the following important

THEOREM 4.8.

(i) *There is a monotone nondecreasing function* $\rho(\lambda)$ *defined on the real line, such that*

$$\rho(\lambda) - \rho(\mu) = \lim_{b\to\infty}\bigl(\rho_b(\lambda) - \rho_b(\mu)\bigr),$$

at every point of continuity λ, μ *of* ρ.

(ii) *If* $f \in L^2(0,\infty)$*, there exists a function* $g(\lambda) \in L^2(\mathbb{R},\rho)$ *such that*

$$\lim_{b\to\infty} \int_{-\infty}^{\infty} \left| g(\lambda) - \int_0^b f(t)y(t)\psi(t,\lambda)\,\mathrm{d}t \right|^2 \mathrm{d}\rho(\lambda) = 0$$

and

$$\int_0^\infty y(t)\left|f(t)\right|^2 \mathrm{d}t = \int_{-\infty}^{\infty} \left|g(\lambda)\right|^2 \mathrm{d}\rho(\lambda).$$

(iii) *The map*

$$t \mapsto \int_{-\infty}^{\infty} g(\lambda)\psi(t,\lambda)\,\mathrm{d}\rho(\lambda)$$

converges in $L^2(0,\infty)$ *to* fy.

(iv) *If* $m_+(\lambda)$ *is the Weyl function at* ∞*, then*

$$\frac{\rho(\mu+0^+)-\rho(\mu+0^-)}{2} - \frac{\rho(\lambda+0^+)-\rho(\lambda+0^-)}{2} + \int_\lambda^\mu \mathrm{d}\rho$$
$$= \lim_{\varepsilon\to 0^+} \frac{1}{\pi} \int_\mu^\lambda \Im m_+(\nu + \mathrm{i}\varepsilon)\,\mathrm{d}\nu.$$

(v) *If* $\Im\lambda_0, \Im\lambda_1 \neq 0$*, then*

$$m_+(\lambda_0) - m_+(\lambda_1) = \int_{-\infty}^{\infty} \left(\frac{1}{\lambda-\lambda_0} - \frac{1}{\lambda-\lambda_1} \right) \mathrm{d}\rho(\lambda).$$

Analogous relations hold for the operator L^- *and the Weyl* m*-function* m_- *at* $-\infty$.

The above theorem can be extended to Sturm–Liouville operators on the whole line. As before, we start from the problem in an interval $I = [a,b]$ with $a < 0 < b$, and the boundary conditions

$$\cos\alpha u(a) + \sin\alpha p(a)u'(a) = 0, \tag{4.18}$$

$$\cos\beta u(b) + \sin\beta p(b)u'(b) = 0. \tag{4.19}$$

There is a matrix $P_I = (\rho_{jk})$ (called the *spectral matrix*) associated with Eq. (4.4) in the interval I. Its entries are nondecreasing step functions satisfying

$$\rho_{jk}(\lambda_n+) - \rho_{jk}(\lambda_n-) = r_j \bar{r}_k,$$

and otherwise constant. If we require that $P_I(\lambda+) = P_I(\lambda)$ then P_I is Hermitian, and, if $\lambda > \mu$, $P_I(\lambda) - P_I(\mu)$ is positive semidefinite. If we let $I \to \mathbb{R}$, then $P_I(\lambda) \to P(\lambda)$, where $P(\lambda)$ is a hermitian, positive semidefinite matrix.

REMARK 4.9. Since the matrix $P(\lambda)$ is Hermitian and nondecreasing, if follows that, for every interval $I = [\lambda, \mu]$,

$$\left|\rho_{jk}(\mu) - \rho_{jk}(\lambda)\right|^2 \leqslant \left(\rho_{jj}(\mu) - \rho_{jj}(\lambda)\right)\left(\rho_{kk}(\mu) - \rho_{kk}(\lambda)\right).$$

From this we conclude that the points of nonconstancy of the spectral matrix $P(\lambda)$ are the same of the points of nonconstancy of the trace $\mathrm{Tr}(P(\lambda))$.

We use the notation $\Sigma(L)$, $R(L)$ for the spectrum and the resolvent set of the full-line operator L and $\Sigma^{\pm}(L)$, $R^{\pm}(L)$ for the spectrum and the resolvent set of the positive and negative Sturm–Liouville operators $L^{\pm}$ respectively.

REMARK 4.10. The sets $\Sigma(L)$, $\Sigma^{\pm}(L)$ are closed subsets of $\mathbb{R}$. The set of the points of growth of the spectral matrix $P(\lambda)$ (i.e., of the trace $\mathrm{Tr}(P(\lambda))$) equals the set $\Sigma(L)$.

We finish this subsection by giving a brief summary of the theory of Gilbert and Pearson [60] adapted to the Sturm–Liouville operator. The original paper of Gilbert and Pearson deals with the Schrödinger operator. Using the weighted measure $\kappa = y\,\mathrm{d}t$ instead of the Lebesgue measure on $\mathbb{R}$, we can prove the results stated in the following lines.

We define a set $\mathcal{M} \subset \mathbb{R}$ to be a minimal support for a measure ι on $\mathbb{R}$ if:

(1) $\iota(\mathbb{R} \setminus \mathcal{M}) = 0$;

(2) if $\mathcal{M}_0 \subset \mathcal{M}$ and $\iota(\mathcal{M}_0) = 0$, then $\ell(\mathcal{M}_0) = 0$, where ℓ is the Lebesgue measure on $\mathbb{R}$.

A minimal support of a measure ι determines where the measure is concentrated, and is defined up to a set of Lebesgue (and ι) 0 measure.

Let $E = \{\lambda \in \mathbb{R} \mid (\mathrm{d}\rho^+/\mathrm{d}\ell)(\lambda)\text{exists}\}$, where ρ^+ is the spectral measure for Eq. (4.9) and $\mathrm{d}\rho^+/\mathrm{d}\ell$ is the Radon–Nikodym derivative of ρ^+ with respect to the measure ℓ. Note that $\mathrm{d}\kappa/\mathrm{d}\ell = y > 0$. Then $\ell(\mathbb{R} \setminus E) = 0$ and the minimal support $\mathcal{M}$ of ρ^+ has the decomposition $\mathcal{M} = \mathcal{M}_{ac} \cup \mathcal{M}_s = \mathcal{M}_{ac} \cup \mathcal{M}_{sc} \cup \mathcal{M}_p$, where $\mathcal{M}_{ac}$ and $\mathcal{M}_s$ are the absolutely continuous part and the singular part of $\mathcal{M}$ respectively. In particular

(i) $\mathcal{M} = \{\lambda \in \mathbb{R} \mid 0 < (\mathrm{d}\rho^+/\mathrm{d}\ell)(\lambda) \leqslant \infty\}$,

(ii) $\mathcal{M}_{ac} = \{\lambda \in \mathbb{R} \mid 0 < (\mathrm{d}\rho^+/\mathrm{d}\ell)(\lambda) < \infty\}$,

(iii) $\mathcal{M}_s = \{\lambda \in \mathbb{R} \mid (\mathrm{d}\rho^+/\mathrm{d}\ell)(\lambda) = \infty\}$,

(iv) $\mathcal{M}_{sc} = \{\lambda \in \mathbb{R} \mid (\mathrm{d}\rho^+/\mathrm{d}\ell)(\lambda) = \infty, \lim_{\delta\to 0^+} \rho^+(\lambda+\delta) - \rho^+(\lambda-\delta) = 0\}$,

(v) $\mathcal{M}_p = \{\lambda \in \mathbb{R} \mid (\mathrm{d}\rho^+/\mathrm{d}\ell)(\lambda) = \infty, \lim_{\delta\to 0^+} \rho^+(\lambda+\delta) - \rho^+(\lambda-\delta) > 0\}$.

Let $\lambda \in \mathbb{C}$. A solution $u_s(t,\lambda)$ of Eq. (4.9) is called *subordinate* at infinity if for every other linearly independent solution $u(t,\lambda)$ of Eq. (4.9), one has

$$\lim_{N\to\infty} \frac{\|u_s(t,\lambda)\|_N}{\|u(t,\lambda)\|_N} = 0,$$

where $\|\cdot\|_N$ denotes the $L^2((0,N), y)$-norm. Let $\alpha \in [0,\pi)$. Consider Eq. (4.9) together with the boundary condition (4.18) at $t = 0$, i.e.,

$$u(0)\cos\alpha + p(0)u'(0)\sin\alpha = 0.$$

The principal results of Gilbert and Pearson can be summarized as follows (see [60]):

THEOREM 4.11. *There exists a set* $E_1 \subset E$ *of full Lebesgue measure and full* ρ^+ *measure such that the nontangential limit*

$$\lim_{\varepsilon\to 0} m_+(\lambda + \mathrm{i}\varepsilon) = m_+(\lambda) \tag{4.20}$$

exists for every $\lambda \in E$, *and*

$$\frac{\mathrm{d}\rho^+}{\mathrm{d}\ell}(\lambda) = \frac{1}{\pi}\Im m_+(\lambda). \tag{4.21}$$

Moreover, we have

- (i) $\mathcal{M} = \{\lambda \in \mathbb{R} \mid 0 < \Im m_+(\lambda) \leqslant \infty\}$,
- (ii) $\mathcal{M}_{ac} = \{\lambda \in \mathbb{R} \mid 0 < \Im m_+(\lambda) < \infty\}$,
- (iii) $\mathcal{M}_s = \{\lambda \in \mathbb{R} \mid \Im m_+(\lambda) = \infty\}$,
- (iv) $\mathcal{M}_{sc} = \{\lambda \in \mathbb{R} \mid \Im m_+(\lambda) = \infty, \lim_{\delta\to 0^+} \rho^+(\lambda+\delta) - \rho^+(\lambda-\delta) = 0\}$,
- (v) $\mathcal{M}_p = \{\lambda \in \mathbb{R} \mid \Im m_+(\lambda) = \infty, \lim_{\delta\to 0^+} \rho^+(\lambda+\delta) - \rho^+(\lambda-\delta) > 0\}$

and

- (i) $\mathcal{M}^c = \{\lambda \in \mathbb{R} \mid$ *a subordinate solution of Eq.* (4.9) *exists and does not satisfy the boundary condition* (4.18)$\}$,
- (ii) $\mathcal{M}_{ac} = \{\lambda \in \mathbb{R} \mid$ *no subordinate solution of Eq.* (4.9) *exists*$\}$,
- (iii) $\mathcal{M}_s = \{\lambda \in \mathbb{R} \mid$ *a subordinate solution of Eq.* (4.9) *exists and satisfies the boundary condition* (4.18)$\}$,
- (iv) $\mathcal{M}_{sc} = \{\lambda \in \mathbb{R} \mid$ *a subordinate solution of Eq.* (4.9) *exists and satisfies the boundary condition* (4.18), *but it is not in* $L^2([0,\infty), y)\}$,
- (v) $\mathcal{M}_p = \{\lambda \in \mathbb{R} \mid$ *a subordinate solution of Eq.* (4.9) *exists, satisfies the boundary condition* (4.18) *and is in* $L^2([0,\infty), y)\}$.

Finally,

$$\Sigma^+(L) = \mathcal{M},$$

up to a set of 0 Lebesgue (*and* ρ^+) *measure.*

The theorem is still valid for Eq. (4.9) on $(-\infty, 0]$, with the spectral measure $\mathrm{d}\rho^-$ and the function m_-, with the appropriate changes of signs. From this theorem we derive the following important observations:

COROLLARY 4.12.

- (i) *The spectral measures* $\rho^\pm(\lambda)$, $P(\lambda)$ *increase exactly in the points of the spectra* $\Sigma^\pm(L)$, $\Sigma(L)$ *respectively;*
- (ii) *the nontangential limit* (4.20) *exists for Lebesgue a.e.* $\lambda \in \mathbb{R}$, $\Im m_+(\lambda) > 0$ ($\Im m_-(\lambda) < 0$) *for Lebesgue a.e* $\lambda \in \Sigma^+(L)$ ($\Sigma^-(L)$), *and* $\Im m_\pm(\lambda) = 0$ *for Lebesgue a.e.* $\lambda \in R^\pm(L)$;
- (iii) $|m_\pm(\lambda + \mathrm{i}\varepsilon)| \to \infty$ *for* $\varepsilon \to 0$ *if and only if* λ *is an eigenvalue of* $L^\pm$.

PROOF. The only thing we have to prove is that the measure $P(\lambda)$ increases exactly in the points of the spectrum $\Sigma(L)$. From Remark 4.9, the points of nonconstancy of $P(\lambda)$ are exactly the points of nonconstancy of the trace $\mathrm{Tr}(P(\lambda))$, and these are exactly the points λ for which

$$\Im\left(\frac{1}{m_-(\lambda)-m_+(\lambda)}+\frac{m_-(\lambda)m_+(\lambda)}{m_-(\lambda)-m_+(\lambda)}\right) \tag{4.22}$$

is positive. We now rewrite (4.22) in the following way

$$\Im\left(\frac{(1+m_-m_+)(\overline{m_- - m_+})}{\|m_- - m_+\|^2}\right). \tag{4.23}$$

From Theorem 4.11, it easy to conclude that (4.23) is not 0 exactly in the points of the spectrum. □

4.2. *The Sturm–Liouville equation as a nonautonomous differential system*

Let Ω be a compact metric space, and let $(\Omega, \{\tau_t\})$ be a flow. We suppose that there is a $\{\tau_t\}$-ergodic measure ν on Ω such that $\operatorname{Supp}\nu = \Omega$. We recall that the topological support $\operatorname{Supp}\nu$ of the measure ν is the complement of the largest open subset $U \subset \Omega$ satisfying $\nu(U) = 0$.

Suppose that $p, q, y : \Omega \to \mathbb{R}$ are continuous functions, and that p and y are strictly positive. For each $\omega \in \Omega$, we introduce the matrix function

$$t \mapsto \begin{pmatrix} 0 & \frac{1}{p(\omega\cdot t)} \\ q(\omega\cdot t)-\lambda y(\omega\cdot t) & 0 \end{pmatrix}$$

which by abuse of notation we call $\omega(t)$. We study the family of differential equations

$$\begin{pmatrix} \varphi \\ \psi \end{pmatrix}' = \begin{pmatrix} 0 & \frac{1}{p(\omega\cdot t)} \\ q(\omega\cdot t)-\lambda y(\omega\cdot t) & 0 \end{pmatrix}\begin{pmatrix} \varphi \\ \psi \end{pmatrix} \tag{SL_ω}$$

where ω ranges over Ω and λ is a complex parameter. If the function $p' : \Omega \to \mathbb{R} : \omega \mapsto \frac{\mathrm{d}}{\mathrm{d}t} p(\omega\cdot t)|_{t=0}$ is well defined and continuous, then (SL_ω) is equivalent to

$$-\big(p(\omega\cdot t)\varphi(t)'\big)' + q(\omega\cdot t)\varphi(t) = \lambda y(\omega\cdot t)\varphi(t). \tag{4.24_ω}$$

We have the following

PROPOSITION 4.13. *Suppose that* $\operatorname{Supp}\nu = \Omega$. *There is a set* $\Omega' \subset \Omega$ *with* $\nu(\Omega') = 1$ *and such that, if* $\omega \in \Omega'$, *then the orbit* $\{\omega\cdot t \mid t \in \mathbb{R}\}$ *is dense in* Ω. *Moreover,* Ω' *can be chosen such that each* $\omega \in \Omega'$ *is both positively and negatively Poisson recurrent.*

Recall that $\omega \in \Omega$ is positively (resp. negatively) Poisson recurrent if there is a sequence $\{t_n\} \to \infty$ (resp. $\{t_n\} \to -\infty$) such that $\lim_{n\to\infty} \omega \cdot t_n = \omega$.

PROOF. The proof is an easy consequence of the Birkhoff ergodic theorem. For, let $V_1, \dots, V_n, \dots$ be a countable base for the topology of Ω. Let χ_i be the characteristic function corresponding to V_i. Using the Birkhoff ergodic theorem, one checks easily that the sets $\Omega_i = \{\omega \in \Omega \mid \lim_{t\to\infty} \frac{1}{t}\int_0^t \chi_i(\omega \cdot s)\,\mathrm{d}s = \nu(V_i) > 0\}$ have ν-measure 1. It follows that if $\omega \in \Omega_i$, then the orbit of ω intersects V_i ($i \in \mathbb{N}$). It is clear now that each $\omega \in \Omega_i$ has dense orbit (in fact, it has dense positive semi-orbit). The set $\Omega' = \bigcap_{i=1}^{\infty} \Omega_i$ is the desired set.

The set Ω' consists of those points with dense positive semi-orbit; such points are in particular positively Poisson recurrent.

Using the fact that $\lim_{t\to-\infty} \frac{1}{t}\int_0^t \chi_i(\omega \cdot s)\,\mathrm{d}s = \nu(V_i)$ for ν-a.a. $\omega \in \Omega$, we see that there is a subset of Ω' of ν-measure 1 consisting of points which are negatively Poisson recurrent as well; we abuse notation slightly and let Ω' denote this set. This completes the proof. □

Let $\Phi_\omega(t)$ be the fundamental matrix solution of (SL_ω) with $\Phi_\omega(0) = I$. Define $\Sigma^{\mathrm{ed}}(L) = \{\lambda \in \mathbb{R} \mid$ the family (SL_ω) does NOT have exponential dichotomy$\}$, and $R^{\mathrm{ed}}(L) = \mathbb{C} \setminus \Sigma^{\mathrm{ed}}(L)$.

We next give a dichotomy-theoretic interpretation of the Weyl m-functions $m_\pm = m_\pm(\omega, \lambda)$ of Eqs. (4.24_ω) ($\omega \in \Omega$).

To begin with, let $\omega \in \Omega$ and $\lambda \in \mathbb{C} \setminus \Sigma^{\mathrm{ed}}(L)$. Define $m_+(\omega,\lambda)$ to be the unique (extended) complex number such that $\operatorname{Im} P_\omega = \operatorname{Span}\binom{1}{m_+(\omega,\lambda)}$. Here $m_+(\omega,\lambda) = \infty$ if and only if $\operatorname{Im} P_\omega = \operatorname{Span}\binom{0}{1}$. One can show that $\Im m_+(\omega,\lambda)\Im\lambda > 0$ whenever $\Im\lambda \neq 0$. In a similar way, we define $\operatorname{Ker} P_\omega = \operatorname{Span}\binom{1}{m_-(\omega,\lambda)}$ where $\Im m_-(\omega,\lambda)\Im\lambda < 0$ if $\Im\lambda \neq 0$. For every $\omega \in \Omega$, if $\Im\lambda \neq 0$ the maps $\lambda \mapsto m_\pm(\omega,\lambda)$ coincide with the classical Weyl m-functions of problem (4.24_ω). In fact, let $\omega \in \Omega$ and $\lambda \in \mathbb{C}$ with $\Im\lambda \neq 0$. Let $0 \neq u = (u_1, u_2) \in \mathbb{C}^2$ such that $\frac{u_2}{u_1} = m_+(\omega,\lambda)$ and let $u(t)$ be the solution of (SL_ω) with $u(0) = u$. Such a solution decays exponentially as $t \to \infty$, hence $u(t) \in L^2(0,\infty)$. Since (SL_ω) is in the limit-point case at ∞, the only L^2 solution at ∞ is $u(t)$. It is easy to check that $\frac{u_2}{u_1} = \frac{p(\omega)u_1'(0)}{u_1(0)}$. The same holds for $m_-(\omega,\lambda)$. We showed that the two definitions of the Weyl m-functions coincide. It follows that for each $\omega \in \Omega$, the maps $\lambda \mapsto m_\pm(\omega,\lambda)$ are holomorphic on $\mathbb{C} \setminus \mathbb{R}$ and meromorphic on $\mathbb{C} \setminus \Sigma^{\mathrm{ed}}(L)$.

The next proposition provides a result concerning the behavior of the Weyl m-functions when ω varies along an orbit.

PROPOSITION 4.14. *Let $\omega \in \Omega$ have dense orbit. Let $\bar\omega \in \Omega$ and let $\{t_n\} \subset \mathbb{R}$ be a sequence with $\lim_{n\to\infty} \omega \cdot t_n = \bar\omega$. Then the Weyl m functions $m_\pm(\omega \cdot t_n, \lambda)$ are well defined and*

$$\lim_{n\to\infty} m_\pm(\omega \cdot t_n, \lambda) = m_\pm(\bar\omega, \lambda),$$

uniformly on compact subsets of the upper (*lower*) *half-plane.*

PROOF. Let $\chi(t,\lambda)$ be a solution of Eq. (4.4) (in the interval $[0,b]$) satisfying the boundary conditions (4.4b) at $t=b$, and let $C_b(\lambda)$ be the circle defined by this solution as in Section 4.1. Let $D_b(\lambda)$ be the closure of $C_b(\lambda)$. Since $D_{b'}(\lambda)\subsetneq D_b(\lambda)$ for every $b'>b$, one has that $m_+(\omega,\lambda)$ is contained in every $D_b(\lambda)$. In an analogous way, we define the (closed) discs $D_b^n(\lambda)$ corresponding to $\omega\cdot t_n$. Let $\varepsilon>0$, and choose $b\in\mathbb{R}$ such that $\operatorname{diam} D_b(\lambda)<\varepsilon$. Fix any $b'>b$: it is easy to check that $C_{b'}^n(\lambda)$ converges to $C_{b'}(\lambda)$ as $n\to\infty$. This immediately implies that $|m_+(\omega\cdot t_n,\lambda)-m_+(\omega,\lambda)|<\varepsilon$. □

Now we turn to the first basic result concerning the spectrum of the family (SL_ω), namely (see [79])

PROPOSITION 4.15. *Let $\omega\in\Omega$ have dense orbit. Then a complex number λ lies in $R(L_\omega)$ if and only if the family (SL_ω) admits an exponential dichotomy: $R(L_\omega)=R^{\mathrm{ed}}(L)$.*

PROOF. Let $\lambda\in R^{\mathrm{ed}}(L)$: then the family (SL_ω) admits an exponential dichotomy over Ω. An easy consequence of the definition of the exponential dichotomy is that for every $\omega\in\Omega$, there exist two solutions $\varphi_\pm(\cdot,\omega,\lambda)$ of (4.24_ω) which are square integrable on the positive and negative half-lines respectively. Indeed $\varphi_\pm(\cdot,\omega,\lambda)$ are exponentially decreasing at ∞ and $-\infty$ respectively, for every $\omega\in\Omega$. Since $\varphi_\pm(\cdot,\omega,\lambda)$ are linearly independent, the Wronskian $W(\varphi_+,\varphi_-)=p(\varphi_+\varphi_-'-\varphi_+'\varphi_-)\neq 0$. It follows that a Green's function

$$G(t,s,\omega,\lambda)=\begin{cases}\dfrac{\varphi_+(t)\varphi_-(s)}{W(\varphi_+,\varphi_-)}y(s) & t\geqslant s,\\[2ex] \dfrac{\varphi_-(t)\varphi_+(s)}{W(\varphi_+,\varphi_-)}y(s) & t\leqslant s\end{cases}$$

is well defined, hence $\lambda\in R(L_\omega)$.

Conversely, assume that $\lambda\in R(L_\omega)$. We must prove that the family (SL_ω) admits exponential dichotomy over Ω. From a result of Sacker–Sell [137] and Selgrade [142], it is sufficient to show that the only bounded solution of the family is the trivial one. Suppose for contradiction that there exists $\bar\omega\in\Omega$ such that the corresponding differential equation $(\mathrm{SL}_{\bar\omega})$ has a nonzero bounded solution (hence $\lambda\notin R(L_{\bar\omega})$). Let $\{t_n\}\to\infty$ be a real sequence such that $\lim_{n\to\infty}\omega\cdot t_n=\bar\omega$. The Weyl m-functions $m_\pm(\omega\cdot t_n,\lambda)$ are defined, and $m_\pm(\omega\cdot t_n,\lambda)\to m_\pm(\bar\omega,\lambda)$ uniformly on every compact subset of the upper complex half-plane as $n\to\infty$ (see Proposition 4.14). It's clear then that the measure $\operatorname{Tr}P_n(\omega,\lambda):=\operatorname{Tr}P(\omega\cdot t_n,\lambda)$ weakly converges to $\operatorname{Tr}P(\bar\omega,\lambda)$ (note that the spectral matrix P depends on $\omega\in\Omega$). This means that for every $f\in C_c(\mathbb{R})$ we have

$$\int_{\mathbb{R}} f(\lambda)\,\mathrm{d}\big(\operatorname{Tr}P_n(\omega,\lambda)\big)\to\int_{\mathbb{R}} f(\lambda)\,\mathrm{d}\big(\operatorname{Tr}P(\bar\omega,\lambda)\big).$$

Since the resolvent set $R(L_\omega)$ is open, there is an interval I such that $\lambda\in I\subset R(L_\omega)$. Now, the resolvent set is invariant under the flow $(\Omega,\{\tau_t\})$ (in fact, one can easily check that if a Green's function corresponding to $\omega\in\Omega$ is well defined, then a Green's function corresponding to $\omega\cdot t\in\Omega$ is well defined as well, for every $t\in\mathbb{R}$), hence $\operatorname{Tr}P_n$ is constant on I. From Remark 4.10, it is straightforward to show that $\operatorname{Tr} P$ is constant on I, so $\lambda \in R(L_{\bar{\omega}})$, a contradiction. □

It will be often useful to think at the functions $p, q, y, m_\pm$ as functions of t, as follows: for every fixed $\omega \in \Omega$, we consider the map $t \mapsto p(\omega \cdot t)$ obtained by "following the orbit through $\omega \in \Omega$". The same holds for all the functions involved; for instance, for fixed $\omega \in \Omega$, we consider the maps $m_\pm(\cdot, \omega, \lambda) : t \mapsto m_\pm(\omega \cdot t, \lambda)$. We will often suppress the dependence with respect to $\omega \in \Omega$, unless it is strictly necessary.

If $\omega \in \Omega$, then $m_\pm(t, \omega, \lambda)$ is obtained from $m_\pm(\omega, \lambda)$ by "translating" the (unique) L^2 solutions at $\pm\infty$ of Eq. (4.24_ω): if we denote these solutions with $\varphi_\pm(t, \omega, \lambda)$, then

$$m_\pm(t, \omega, \lambda) = \frac{p(\omega \cdot t)\varphi'_\pm(t, \omega, \lambda)}{\varphi_\pm(t, \omega, \lambda)}.$$

Both m_+ and m_- are analytic in the upper and lower λ-half-planes respectively. The following proposition is fundamental.

PROPOSITION 4.16. *Let $\omega \in \Omega$. For every fixed $\lambda \in \mathbb{C}$ with $\Im\lambda \neq 0$ the Weyl m-functions $m_\pm(t, \lambda) \equiv m_\pm(t, \omega, \lambda)$ satisfy the Riccati equation*

$$m'_\pm(t, \lambda) + \frac{1}{p(t)} m^2_\pm(t, \lambda) = q(t) - \lambda y(t). \tag{4.25}$$

PROOF. Since

$$m_\pm(t, \lambda) = \frac{p(t)\varphi'_\pm(t)}{\varphi_\pm(t)},$$

differentiating with respect to t, we have

$$m'_\pm = \frac{(p\varphi'_\pm)'\varphi_\pm - p(\varphi'_\pm)^2}{\varphi^2_\pm}.$$

Hence

$$m'_\pm + \frac{1}{p} m^2_\pm = \frac{(p\varphi'_\pm)'}{\varphi_\pm} = q - \lambda y.$$

□

REMARK 4.17. Note that, if $\Im\lambda \neq 0$ and $\omega \in \Omega$, then

$$\varphi_\pm(t, \lambda) = \exp\left\{ \int_0^t \frac{m_\pm(s, \lambda)}{p(s)} \, \mathrm{d}s \right\} \tag{4.26}$$

is a solution of (4.24_ω) satisfying $\varphi_\pm(0, \lambda) = 1$ and $p(\omega)\varphi'_\pm(0, \lambda) = m_\pm(\omega, \lambda)$.

4.3. *The projective flow in the Sturm–Liouville setting*

We have already noted the importance of the projective flow in the study of nonautonomous differential systems. This subsection provides a further demonstration of this fact, since from the properties of the projective flow associated to a family of Sturm–Liouville differential systems, we will retrieve many important results concerning both the spectrum of the family (SL_ω) and the Weyl m-functions. We will use the notation and the terminology of Sections 2 and 3. Some proofs given in those sections will be repeated here, when they will lead to particular facts concerning the Sturm–Liouville operator.

Let Ω be a compact metric space, let $(\Omega, \{\tau_t\})$ be a real flow and let ν be a $\{\tau_t\}$-ergodic measure on Ω with $\operatorname{Supp}\nu = \Omega$. Consider the family of two-dimensional nonautonomous differential systems (SL_ω) defined as before. Let $(\Sigma, \mathbb{R})$ be the projective flow associated to $(\Omega, \mathbb{R})$, i.e. $\Sigma = \Omega \times \mathbb{P}(\mathbb{C})$ and $\hat{\tau}_t(\omega, l) = (\omega \cdot t, \Phi_\omega(t)l)$.

Fix $\omega \in \Omega$ and $\lambda \in \mathbb{R}$. The flow $(\Omega, \{\hat{\tau}_t\})$ induces a flow, which we still denote by $\{\hat{\tau}_t\}$, on $\Omega \times \mathbb{P}(\mathbb{R})$ as well. Following the developments of Section 2, we express Eqs. (SL_ω) in polar coordinates (r, θ), to obtain

$$\begin{cases} \dfrac{r'}{r} = \left(\dfrac{1}{p(\omega \cdot t)} + q(\omega \cdot t) - \lambda y(\omega \cdot t)\right)\sin\theta\cos\theta = f\big(\hat{\tau}_t(\omega, \theta)\big), \\ \theta' = \big(q(\omega \cdot t) - \lambda y(\omega \cdot t)\big)\cos^2\theta - \dfrac{1}{p(\omega \cdot t)}\sin^2\theta = g\big(\hat{\tau}_t(\omega, \theta)\big). \end{cases} \tag{4.27}$$

It is understood that we identify θ with the line through the origin it parametrizes, hence θ is taken mod π. If $\theta_0 \in \mathbb{P}(\mathbb{R})$ parametrizes the line l_0, then $\hat{\tau}_t(\omega, l_0) = \hat{\tau}_t(\omega, \theta_0) = (\omega \cdot t, \theta(t))$, where $\theta(t)$ satisfies the θ-equation in (4.27) with $\theta(0) = \theta_0$.

Recall that for every fixed $\omega \in \Omega$, the *rotation number* $\alpha_\omega(\lambda)$ is defined as

$$\alpha_\omega(\lambda) = -\lim_{t\to\infty} \frac{1}{t}\big(\theta(t, \lambda) - \theta_0(\lambda)\big) \quad (\lambda \in \mathbb{R}). \tag{4.28}$$

In the above formula, we explicitly wrote the fact that $\theta(\cdot)$ depends on the fixed $\lambda \in \mathbb{R}$. The questions we will answer in the following lines concern: (i) the existence of the limit (4.28); (ii) the independence of (4.28) with respect to the initial condition (ω, θ_0); (iii) continuity and monotonicity of the map $\lambda \mapsto \alpha(\lambda)$; (iv) the constancy of $\alpha(\lambda)$ in every interval $I \subset R^{\mathrm{ed}}(L)$. We will need the following

LEMMA 4.18. *Let $F : \Sigma \to \mathbb{R}$ be a continuous function such that $\int_\Sigma F(\sigma)\,\mathrm{d}\mu = 0$ for every invariant measure μ on Σ. Then $\frac{1}{t}\int_0^t F(\hat{\tau}_s(\sigma))\,\mathrm{d}s \to 0$ as $t \to \infty$, for all $\sigma \in \Sigma$. Moreover, the convergence is uniform with respect to $\sigma \in \Sigma$.*

PROOF. The proof uses techniques from Proposition 2.5; see, e.g. [89]. The space $C(\Sigma) = \{F : \Sigma \to \mathbb{R} \mid F \text{ is continuous}\}$ is separable, hence it contains a dense subset $C_1(\Sigma)$ which is generated by a countable family $\mathcal{F}$ of functions. We may assume that $F \in \mathcal{F}$. By contradiction, suppose that there exist sequences $\{t_n\} \to \infty$ and $\{\sigma_n\} = \{(\omega_n, \theta_n)\} \subset \Sigma$ such that

$$\frac{1}{t_n} \int_0^{t_n} F\big(\hat{\tau}_s(\sigma_n)\big)\,\mathrm{d}s \to \Delta(F) \neq 0.$$

We can assume that $\{\sigma_n\} \to \sigma = (\omega, \theta) \in \Sigma$. Let $F_1 \in \mathcal{F}$, and choose subsequences $\{t_n^{(1)}\} \subset \{t_n\}$ and $\{\sigma_n^{(1)}\} \subset \{\sigma_n\}$ such that

$$\lim_{t_n^{(1)} \to \infty} \frac{1}{t_n^{(1)}} \int_0^{t_n^{(1)}} F_1\big(\hat{\tau}_s\big(\sigma_n^{(1)}\big)\big)\,\mathrm{d}s$$

exists. Call it $\Delta(F_1)$. Continuing this way, for every $F_k \in \mathcal{F}$, we can find subsequences $\{t_n^{(k)}\} \subset \{t_n^{(k-1)}\}$ and $\{\sigma_n^{(k)}\} \subset \{\sigma_n^{(k-1)}\}$ such that

$$\lim_{t_n^{(k)} \to \infty} \frac{1}{t_n^{(k)}} \int_0^{t_n^{(k)}} F_k\big(\hat{\tau}_s\big(\sigma_n^{(k)}\big)\big)\,\mathrm{d}s$$

exists. Call it $\Delta(F_k)$. Now we choose the diagonal sequences $\{t_n^{(n)}\}$ and $\{\sigma_n^{(n)}\}$. For every $G \in \mathcal{F}$, one has

$$\Delta(G) = \lim_{t_n^{(n)} \to \infty} \frac{1}{t_n^{(n)}} \int_0^{t_n^{(n)}} G\big(\hat{\tau}_s\big(\sigma_n^{(n)}\big)\big)\,\mathrm{d}s.$$

Note that $\Delta(G)$ is $\mathbb{Q}$-linear, $\Delta(G) \geqslant 1$ if $G \geqslant 0$, and $\Delta(1) = 1$. It follows that the above limit defines a linear functional $l_1 : C_1(\Sigma) \to \mathbb{R}$ with $|l_1| = 1$, hence it extends to a unique linear functional $l : C(\Sigma) \to \mathbb{R}$ (here $|\cdot|$ denotes the L^∞ norm). By the Riesz representation theorem, l defines a measure μ_l on Σ such that $l(G) = \int_\Sigma G\,\mathrm{d}\mu_l$ for every $G \in C(\Sigma)$. Further, l is invariant in the sense that, if $t \in \mathbb{R}$ and $G \in C(\Sigma)$, then $l(G(\hat{\tau}_{-t}(\cdot))) = l(G(\cdot))$; this is because $t_n^{(n)} \to \infty$. But now we have

$$0 \neq \Delta(F) = \lim_{t_n^{(n)} \to \infty} \frac{1}{t_n^{(n)}} \int_0^{t_n^{(n)}} F\big(\hat{\tau}_s\big(\sigma_n^{(n)}\big)\big)\,\mathrm{d}s = \int_\Sigma F d\mu_l = l(F) = 0,$$

a contradiction. □

We use Lemma 4.18 in the proof of the following important theorem, which answers the questions (i), (ii) and (iii) posed before (for an analogous result in the case of the Schrödinger operator, see [89]).

THEOREM 4.19.

(a) *There exists a set* $\Omega_\nu \subset \Omega$ *with* $\nu(\Omega_\nu) = 1$, *such that* $\alpha_\omega(\lambda)$ *exists for every* $\omega \in \Omega_\nu$ *and does not depend on the choice of* $(\omega, \theta) \in \Omega_\nu \times \mathbb{P}(\mathbb{R})$: $\alpha_\omega(\lambda) = \alpha(\lambda)$, *for every* $\omega \in \Omega_\nu$;

(b) *if* ν *is the only ergodic measure on* Ω (*for instance if* $(\Omega, \mathbb{R})$ *is minimal and almost periodic*), *then* $\Omega_\nu = \Omega$, *and the limit* (4.28) *is uniform in* $t \in \mathbb{R}$ *and* $\sigma \in \Sigma$;

(c) *for every* $\omega \in \Omega$, *the maps* $\lambda \mapsto \alpha_\omega(\lambda)$ *are continuous, monotone nondecreasing, and* $\lim_{\lambda\to\infty} \alpha_\omega(\lambda) = +\infty$.

PROOF. Recall that the angular variable θ is defined mod π. First of all, we claim that the limit (4.28) does not depend on the choice of $\theta \in \mathbb{P}(\mathbb{R})$. For, let $0 < \theta_1 - \theta_0 < \pi$, and let $\sigma_1 = (\omega, \theta_1), \sigma_0 = (\omega, \theta_0) \in \Sigma$. It follows that $0 < \theta(t, \sigma_1) - \theta(t, \sigma_0) < \pi$ for all $t \in \mathbb{R}$, otherwise there would exists $t_0 \in \mathbb{R}$ such that $\theta(t_0, \sigma_1) - \theta(t_0, \sigma_0) = \pi$, and hence, by uniqueness, $\theta(t, \sigma_1) - \theta(t, \sigma_0)$ for all $t \in \mathbb{R}$, a contradiction. The claim now is obvious, since $|\theta(t, \sigma_1) - \theta(t, \sigma_0)|$ is bounded, for every $t \in \mathbb{R}$ and $\sigma_0, \sigma_1 \in \Sigma$.

Now, let μ be a $\{\hat{\tau}_t\}$-ergodic measure on Σ. Using the Birkhoff ergodic theorem, one finds a set Σ_μ with $\mu(\Sigma_\mu) = 1$ and such that the limit (4.28) exists for all $\sigma \in \Sigma_\mu$. Moreover we have

$$\alpha_\omega(\lambda) = -\int_\Sigma g(\omega, \theta)\, \mathrm{d}\mu,$$

hence the limit (4.28) does not depend on the choice of $(\omega, \theta) \in \Sigma_\mu$. Note that if ν is any $\{\tau_t\}$-ergodic measure on Ω, then $\Omega_\nu := \pi(\Sigma_\mu)$ has ν-measure 1. From the observations above, the function $g(\cdot) : \Sigma \to \mathbb{R}$ can be viewed as a function of the single variable $\omega \in \Omega$, since, when it exists, the limit (4.28) is independent of the choice of $\theta \in \mathbb{P}(\mathbb{R})$. It follows that for every $\omega \in \Omega_\nu$, the limit (4.28) exists and does not depend on the choice of $\omega \in \Omega_\nu$. We have proved (a).

To prove (b), we use Lemma 4.18. Let μ be any invariant measure on Σ. Then, by the Birkhoff ergodic theorem we conclude that there exists a set Σ_μ with $\mu(\Sigma_\mu) = 1$, and such that the limit (4.28) exists for all $\sigma = (\omega, \theta) \in \Sigma_\mu$. This limit defines a function $g^* : \Sigma_\mu \to \mathbb{R}$, which is μ-integrable and such that

$$\int_\Sigma g^*(\sigma)\, \mathrm{d}\mu = \int_\Sigma g(\sigma)\, \mathrm{d}\mu.$$

Arguing as above, the function g^* can be viewed as a function of the single variable $\omega \in \Omega$. Now we use the unique ergodicity of Ω. Since g^* is invariant with respect to the flow $(\Omega, \{\tau_t\})$, we conclude that $g^* = const$ on a set $\Omega_\nu \subset \Omega$ with $\nu(\Omega_\nu) = 1$. By the Birkhoff ergodic theorem, this constant is exactly the rotation number $\alpha(\lambda)$. At this point, (b) follows easily, applying Lemma 4.18 to the function $g(\cdot) + \alpha$.

To prove (c), let $\{\lambda_j\} \subset \mathbb{R}$ be a sequence which converges to $\lambda_0 \in \mathbb{R}$. Let $g_j(\sigma) = g(\sigma, \lambda_j)$, $g_0(\sigma) = g(\sigma, \lambda_0)$, and let $(\Sigma, \{\hat{\tau}_t^j\})$ be the corresponding projective flows (depending on λ_j). Let μ_j be any $\{\hat{\tau}_t^j\}$-ergodic measure on Σ. Then

$$\alpha(\lambda_j) = -\int_\Sigma g_j(\sigma)\, \mathrm{d}\mu_j.$$

Without loss of generality, we may assume that μ_j weakly converges to $\tilde{\mu}$, where $\tilde{\mu}$ is a $\{\hat{\tau}_t\}$-invariant measure on Σ. It follows that

$$\int_\Sigma g_0(\sigma)\,\mathrm{d}\mu_j \to \int_\Sigma g_0(\sigma)\,\mathrm{d}\tilde{\mu}.$$

Moreover

$$\left|g_j(\sigma) - g_0(\sigma)\right| \leqslant \max_{\omega\in\Omega}\left|y(\omega)\right| \cdot |\lambda_j - \lambda_0| \to 0$$

as $j \to \infty$. Using this inequality, we obtain

$$\alpha(\lambda_j) = -\int_\Sigma g_j(\sigma)\,\mathrm{d}\mu_j = -\int_\Sigma (g_j - g_0)(\sigma)\,\mathrm{d}\mu_j - \int_\Sigma g_0(\sigma)\,\mathrm{d}\mu_j \to \alpha(\lambda)$$

as $j \to \infty$. The monotonicity property and the asymptotic behavior of the map $\lambda \mapsto \alpha(\lambda)$ follow directly from the definition of the rotation number. ☐

REMARK 4.20. Any zero of the solution φ of Eq. (4.24$_\omega$) corresponds to a value $\bar{t} \in \mathbb{R}$ such that $\theta(\bar{t}, \lambda) = \frac{\pi}{2} \bmod \pi$. By (4.27), one has $\theta'(\bar{t}) = -\frac{1}{p(\omega\cdot\bar{t})} < 0$. It follows that θ decreases at such zero: this shows that there is exactly one more zero every time θ is increased by π. Let $N(t, \lambda)$ be the number of the zeros of the solution φ of (4.24$_\omega$) in $[0, t]$. Then we must have

$$\alpha(\lambda) = \pi \lim_{t\to\infty} \frac{1}{t} N(t, \lambda). \tag{4.29}$$

There remains to discuss the assertion that $\alpha(\lambda)$ is constant in every open interval $I \subset R^{\mathrm{ed}}(L)$. Actually, this and other considerations will enable us to prove a key fact concerning the spectrum of the family (SL$_\omega$). In fact, we will see that the points of growth of $\alpha(\lambda)$ *coincide* with the points of the spectrum $\Sigma(L_\omega)$, for Lebesgue a.a. $\lambda \in \mathbb{R}$. Since $\alpha(\lambda)$ does not depend on ω, for ν-a.a. $\omega \in \Omega$, it follows that $\Sigma(L_\omega)$ equals a fixed closed set $\Sigma(L)$ for ν-a.a. $\omega \in \Omega$. Moreover $\Sigma(L) = \Sigma^{\mathrm{ed}}(L)$.

We summarize these last remarks in the following theorem. In the remainder of Section 4.3 we will discuss its proof together with various facts concerning the spectral theory of the family (SL$_\omega$). We will also illustrate several methods which have proved useful in the study of 2-dimensional, nonautonomous linear systems.

THEOREM 4.21.

(i) *For ν-a.a. $\omega \in \Omega$, say $\omega \in \Omega_\nu$, the spectrum of Eq.* (4.24$_\omega$) *equals a closed set $\Sigma(L)$ which does not depend on $\omega \in \Omega_\nu$. Also, $R^{\mathrm{ed}}(L) = R(L)$; i.e., $R(L) := \mathbb{C} \setminus \Sigma(L) =$ $\{\lambda \in \mathbb{C} \mid$ the family* (SL$_\omega$) *has an exponential dichotomy over* $\Omega\}$.

(ii) *The set $\Sigma(L)$ contains no isolated points.*

(iii) *If $I \subset \mathbb{R}$ is an open interval, then $I \subset R(L)$ if and only if the rotation number α is constant on I.*

First, we note that Propositions 4.13 and 4.15 imply that there is a set $\Omega_\nu \subset \Omega$ with $\nu(\Omega_\nu) = 1$ such that, if $\omega \in \Omega_\nu$, then the spectrum of Eq. (4.24$_\omega$) equals the complement in $\mathbb{R}$ of $R^{\mathrm{ed}}(L)$. Thus point (i) of Theorem 4.21 has already been proved.

Second, we turn to a proof of the above mentioned constancy property of the rotation number. We need some arguments of Schwarzmann [140]. We sketch some facts of his theory. Let $H(\Omega, S^1)$ be the set of all the homotopy classes of the maps $f : \Omega \to S^1 \subset \mathbb{C}$. Each class in $H(\Omega, S^1)$ contains a map f such that the map $f' : \Omega \to \mathbb{C} : \omega \mapsto \frac{\mathrm{d}}{\mathrm{d}t} f(\omega \cdot t)_{|t=0}$ is well defined and continuous. Let $h : H(\Omega, S^1) \to \mathbb{R}$ be defined as follows: $h[f] = \Im \int_\Omega \frac{f'(\omega)}{f(\omega)} \mathrm{d}\nu(\omega)$, where, as usual, ν is a $\{\tau_t\}$-ergodic measure on Ω. Using the Birkhoff ergodic theorem, one can show that for ν-a.a. $\omega \in \Omega$, $h[f]$ equals $\lim_{t\to\infty} \frac{1}{t} \arg f(\omega \cdot t)$. Schwarzmann proved that the map h induces a homomorphism between the groups $H(\Omega, S^1)$ and $(\mathbb{R}, +)$. For, let C be the subgroup of $H(\Omega, S^1)$ consisting on the homotopy classes of the maps $g(\omega) = \mathrm{e}^{2\mathrm{i}s(\omega)}$, where $s : \Omega \to \mathbb{R}$ is a given continuous function. For every $g \in C$, one has $h[g] = 0$, hence the map h induces a homomorphism from $\tilde{H}(\Omega, \mathbb{Z}) = H(\Omega, S^1)/C$ into $(\mathbb{R}, +)$. This homomorphism is called the "Schwarzmann homomorphism". We can now prove the following result:

THEOREM 4.22. *Let $I \subset R(L) \cap \mathbb{R}$ be an open interval. Then $\alpha(\lambda)$ is constant in I.*

PROOF. We use the Schwarzmann homomorphism. Let us temporarily identify the projective circle $\mathbb{P}(\mathbb{R})$ with the circle $S^1 \in \mathbb{C}$ by mapping the line l parametrized by $\theta \in \mathbb{R}$ to $\mathrm{e}^{2\mathrm{i}\theta} \in S^1$ (recall that θ is a π-periodic coordinate on $\mathbb{P}(\mathbb{R})$). Let $i : \mathbb{P}(\mathbb{R}) \to S^1$ denote this map. Next note that, for each fixed $\lambda \in I$, the map $\tilde{g}_\lambda : \Omega \to \mathbb{P}(\mathbb{R}) : \omega \mapsto m_+(\omega, \lambda)$ is a continuous function. The composition $g_\lambda = i \circ \tilde{g}_\lambda$ lies in $H(\Omega, S^1)$.

Let us now calculate the image of g_λ under the Schwarzmann homomorphism h. Let $\omega \in \Omega$, and let θ_0 be a parameter value of $m_+(\omega, \lambda)$. Let $\theta(t)$ be the solution of (4.27) with initial value θ_0. We then obtain for ν-a.a. $\omega \in \Omega$:

$$-h[g_\lambda] = -2 \lim_{t\to\infty} \frac{\theta(t)}{t} = 2\alpha(\lambda).$$

This shows that $2\alpha \in h(\tilde{H}(\Omega, \mathbb{Z}))$, and the continuity of $\alpha(\lambda)$ implies that $\alpha(\lambda)$ is constant on I. □

The other useful quantity in the study of nonautonomous differential systems is the upper Lyapunov exponent $\beta(\lambda)$. It has been already defined and studied in detail in Section 2. All the properties of $\beta(\lambda)$ which were stated in that section retain validity in our case.

It is convenient to extend the concepts of the rotation number and the Lyapunov exponent to complex λ. To do this we define the Floquet exponent. We write z instead of λ, and use λ to denote real numbers.

Fix $z \in \mathbb{C}$ with $\Im z \neq 0$, and recall that for the family (SL$_\omega$), the Weyl m-functions are defined. Such functions satisfy the Riccati equation (4.25).

Let $\omega \in \Omega$ be fixed. For $\Im z \neq 0$ we define the Floquet exponent $w(\omega, z)$ as follows

$$w(\omega, z) = \frac{1}{2} \lim_{t\to\infty} \frac{1}{t} \int_0^t \frac{1}{p(\omega \cdot s)} \big(m_+(\omega \cdot s, z) - m_-(\omega \cdot s, z)\big)\, \mathrm{d}s. \tag{4.30}$$

Using the Birkhoff ergodic theorem, one sees that, for a fixed ergodic measure ν on Ω,

$$w(\omega, z) = \frac{1}{2}\int_\Omega \frac{[m_+(\omega, z) - m_-(\omega, z)]}{p(\omega)}\, \mathrm{d}\nu(\omega) = w(z), \tag{4.31}$$

for ν-a.a. $\omega \in \Omega$. In particular, it follows that $w(\omega, z)$ does not depend on the choice of $\omega \in \Omega$, for ν-a.a $\omega \in \Omega$, and we will write simply $w(z)$ instead of $w(\omega, z)$.

We have the following

PROPOSITION 4.23. *For* $\Im z \neq 0$, *we have*

$$w(z) = \int_\Omega \frac{1}{p(\omega)} m_+(\omega, z)\, \mathrm{d}\mu = -\int_\Omega \frac{1}{p(\omega)} m_-(\omega, z)\, \mathrm{d}\mu. \tag{4.32}$$

PROOF. For fixed $\omega \in \Omega$, consider the map (where the notation has an obvious meaning)

$$f(t, z) = \frac{m_+(t, z) + m_-(t, z)}{p(t)} = \frac{\varphi_+'(t, z)\varphi_-(t, z) + \varphi_-'(t, z)\varphi_+(t, z)}{\varphi_+(t, z)\varphi_-(t, z)}.$$

It is obvious that f is the derivative with respect to t of the function

$$\ln\bigl(\varphi_+(t, z)\varphi_-(t, z)\bigr).$$

Integrating between 0 and t we obtain

$$\begin{aligned}\int_0^t \frac{1}{p(s)}\bigl[m_+(s, z) + m_-(s, z)\bigr]\, \mathrm{d}s = {} & \ln\bigl(\varphi_+(t, z)\varphi_-(t, z)\bigr) \\ & - \ln\bigl(\varphi_+(0, z)\varphi_-(0, z)\bigr).\end{aligned}$$

If we divide both sides by t and pass to the limit as $t \to \infty$, we obtain (4.32). □

From (4.26), the solutions $\varphi_\pm(t, \lambda)$ have the asymptotic form

$$\varphi_\pm(t, \lambda) \sim \mathrm{e}^{\pm(\Re w(\lambda)t + i\Im w(\lambda)t)}. \tag{4.33}$$

It is quite natural to define the Lyapunov exponent and the rotation number for complex z in the following way:

DEFINITION 4.24. Let $z \in \mathbb{C} \setminus \mathbb{R}$. We define

$$\tilde{\beta}(z) = \Re w(z) \quad \text{(Lyapunov exponent)}$$

and

$$\tilde{\alpha}(z) = \Im w(z) \quad \text{(Rotation number)}.$$

Now we prove important results of convergence. Recall that for Lebesgue a.a. $\lambda \in \mathbb{R}$ there exists the nontangential limit of $m_\pm(z)$ as $z \to \lambda$, and that $|m_\pm(z)| \to \infty$ as $z \to \lambda$ if and only if λ is an eigenvalue for the operator $L^\pm$ respectively (see again Theorem 4.11 and Corollary 4.12).

Moreover, if $z \in \mathbb{C}$ is such that $\Im z \neq 0$, then the projective flow $(\Sigma, \mathbb{R})$ is defined as $(\Sigma = \Omega \times \mathbb{P}(\mathbb{C}), \{\hat{\tau}_t\})$, where $\{\hat{\tau}_t\}$ is the skew-product flow induced by $\{\tau_t\}$ and the solutions of (SL_ω). We prove the following

PROPOSITION 4.25. *For every* $\lambda \in \mathbb{R}$,

$$\lim_{\varepsilon \to 0} \tilde{\beta}(\lambda + \mathrm{i}\varepsilon) = \beta(\lambda). \tag{4.34}$$

PROOF. Recall that we have only two possibilities for $\lambda \in \mathbb{R}$: namely, either $\beta(\lambda) = 0$ or $\beta(\lambda) < 0$.

For every z with $\Im z \neq 0$ the Weyl m-function $m_+(z)$ defines an invariant (minimal) section in Σ, hence an ergodic measure μ_z on Σ expressed as follows

$$\int_\Sigma f(\sigma)\,\mathrm{d}\mu_z = \int_\Omega f\bigl(\omega, m_+(\omega, z)\bigr)\,\mathrm{d}\nu(\omega), \tag{4.35}$$

for every continuous function $f : \Sigma \to \mathbb{R}$. It is easy to check that μ_z is indeed $\{\hat{\tau}_t\}$-ergodic and that it projects to ν. Next, for every $z \in \mathbb{C}$, we define a continuous function $f_z : \Sigma \to \mathbb{R}$ as follows:

$$f_z(\sigma) = \Re \frac{\langle [y(\omega)\left(\begin{smallmatrix} 0 & 0 \\ -z & 0 \end{smallmatrix}\right) + C(\omega)]u, u\rangle}{\langle u, u\rangle}, \tag{4.36}$$

where $\sigma = (\omega, l) \in \Sigma$, u is any nonzero vector in l, $\langle \cdot, \cdot \rangle$ is the usual inner product in $\mathbb{C}$, and

$$C(\omega) = \begin{pmatrix} 0 & 1/p(\omega) \\ q(\omega) & 0 \end{pmatrix}.$$

For every sequence $z_n \to \lambda$, f_{z_n} converges to f_λ uniformly on Σ. Moreover, if $u(t)$ is any nonzero solution of Eq. (SL_ω), then

$$\frac{1}{t} \ln \frac{|u(t)|}{|u(0)|} = \frac{1}{t} \int_0^t f_z\bigl(\hat{\tau}_s(\omega, l_u)\bigr)\,\mathrm{d}s, \tag{4.37}$$

where l_u is the line in $\mathbb{C}$ containing 0 and $u(0)$.

From the Birkhoff ergodic theorem, (4.37), the definition of μ_z, and (4.33), we have

$$\tilde{\beta}(z) = \int_\Sigma f_z(\sigma)\,\mathrm{d}\mu_z = \int_\Omega f_z\bigl(\omega, m_+(\omega, z)\bigr)\,\mathrm{d}\nu \quad (\Im z > 0). \tag{4.38}$$

Suppose first that $\beta(\lambda) = 0$. Simple applications of the Choquet theorem [131], the Oseledets theorem [8,129] and the Birkhoff ergodic theorem show that for every invariant measure μ on Σ, we have

$$\int_\Sigma f_\lambda(\sigma)\,\mathrm{d}\mu = 0. \tag{4.39}$$

In fact, suppose that there exists an invariant measure μ on Σ such that

$$\int_\Sigma f_\lambda(\sigma)\,\mathrm{d}\mu < 0.$$

Then we apply the Choquet theorem to the set Γ of the $\{\hat{\tau}_t\}$-invariant measures on Σ which project to ν (Γ is weak-$*$ compact and convex), to show that there exists an ergodic measure μ_1 on Σ such that $\int_\Sigma f\,\mathrm{d}\mu_1 = \eta < 0$. Now, use the Birkhoff ergodic theorem to find a set $\Sigma_1 \subset \Sigma$ with $\mu_1(\Sigma_1) = 1$, and such that

$$\lim_{|t|\to\infty} \frac{1}{t}\int_0^t f\big(\hat{\tau}_s(\sigma)\big)\,\mathrm{d}s = \eta, \quad \text{for all } \sigma \in \Sigma_1.$$

Since $\beta(\lambda) = 0$, we apply the Oseledets theorem to find a set Ω_* with $\nu(\Omega_*) = 1$ and such that, if θ_0 is any initial value, then

$$\lim_{|t|\to\infty} \frac{1}{t}\int_0^t f\big(\hat{\tau}_s(\omega,\theta_0)\big)\,\mathrm{d}s = 0, \quad \text{for all } a \in \Omega_*.$$

Let $\Omega_1 = \pi(\Sigma_1)$, and choose $\omega \in \Omega_1 \cap \Omega_*$. Take points $(\omega, l_1) \in \Sigma_1$, $(\omega, l_2) \in \Sigma_2 = \pi^{-1}(\Omega_*)$. Clearly $l_1 \neq l_2$. If we let x_1, x_2 be unit vectors in l_1 and l_2 respectively, then x_1 and x_2 are linearly independent. Let $x_i(t) = \Phi_\omega(t)x_i$ $(i = 1, 2)$. Let $\Psi(t)$ be the fundamental matrix solution of (SL_ω) whose columns are $x_1(t)$ and $x_2(t)$. By Liouville's formula, we argue that $\det\Psi(t) = \det\Psi(0) = \text{const.}$ for all $t \in \mathbb{R}$. But also $\lim_{t\to-\infty}\det\Psi(t) = 0$, a contradiction.

We claim now that, if $\beta(\lambda) = 0$, then $\tilde{\beta}(z)$ converges to $\beta(\lambda)$ as $z \to \lambda$, and this convergence need not be nontangential, i.e. is nonrestricted. Suppose for contradiction that there exists a sequence $\{z_n\}$ of complex numbers converging to λ and such that

$$\tilde{\beta}(z_n) \to \beta^* \neq \beta(\lambda).$$

Using weak compactness of measures, we can assume that $\mu_{z_n} \rightharpoonup \mu_\lambda$, where μ_λ is invariant with respect to the flow $(\Sigma, \{\hat{\tau}_t^{(\lambda)}\})$. So we have, using Lemma 4.18,

$$\tilde{\beta}(z_n) = \int_\Sigma f_{z_n}(\sigma)\,\mathrm{d}\mu_{z_n} \to \int_\Sigma f_{z_n}(\sigma)\,\mathrm{d}\mu_\lambda \to \int_\Sigma f_\lambda(\sigma)\,\mathrm{d}\mu_\lambda = 0,$$

a contradiction.

Now we consider the case when $\beta(\lambda) < 0$. From the Oseledets theorem 2.1, there is a set $\Omega_\nu \subset \Omega$ with $\nu(\Omega_\nu) = 1$ which has the following property: if $\omega \in \Omega_\nu$, there is a vector $0 \neq u \in \mathbb{R}^2$ such that

$$\lim_{t\to\infty} \ln\left|\Phi_\omega(t)u\right| = \beta(\lambda) < 0. \tag{4.40}$$

In particular, it follows that the solution $\Phi_\omega(t)u \in L^2(0,\infty)$. Using Liouville's formula, one checks that u is unique up to a constant multiple.

Let $l_\omega \in \mathbb{P}(\mathbb{R})$ be the line containing u. Then l_ω contains the unit vector $\left(\begin{smallmatrix}\cos\alpha\\ \sin\alpha\end{smallmatrix}\right)$ for exactly one number $\alpha \in [0,\pi)$. Consider the following spectral problem:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\left(p(\omega\cdot t)\dfrac{\mathrm{d}\varphi}{\mathrm{d}t}\right) - q(\omega\cdot t)\varphi = -zy(\omega\cdot t)\varphi, \quad t \in [0,\infty), \\ \varphi(0) = \cos\alpha. \\ p(0)\varphi'(0) = \sin\alpha. \end{cases} \tag{$*$}$$

Note that the number λ which we are considering is an eigenvalue of this spectral problem. As stated previously, a Weyl m-function $m_+(\alpha,\omega,z)$ is associated to the spectral problem $(*)$; see Proposition 4.5 and the discussion preceding it. We have

$$m_+(\omega,z) = m_+(\pi/2,\omega,z) \quad (\Im z > 0);$$

see Remark 4.6.

Let us temporarily write

$$m_\alpha(z) = m_+(\alpha,\omega,z), \qquad m(z) = m_+(\omega,z).$$

One can check that

$$m(z) = -\frac{\cos\alpha - m_\alpha(z)\sin\alpha}{\sin\alpha + m_\alpha(z)\cos\alpha} \quad (\Im z \neq 0)$$

(see Remark 4.6 again). Now, an analogue of the Gilbert–Pearson theory is valid for the spectral problem $(*)$; in particular, there holds a statement corresponding to point (iii) in Corollary 4.12. This means that

$$m_\alpha(z) \to \infty \quad \text{when } z \to \lambda \text{ nontangentially.}$$

This in turn means that $m(z) \to \tan\alpha$ as $z \to \lambda$ nontangentially. That is, if we identify $m(z)$ with the complex line it parametrizes, we have

$$m(z) \to l_\omega \quad \text{when } z \to \lambda \text{ nontangentially.}$$

We conclude that, if $z \to \lambda$ nontangentially, then $f_z(\omega, m_+(\omega,z)) \to f_\lambda(\omega, l_\omega)$ for all $\omega \in \Omega_\nu$. Using (4.38), we obtain

$$\tilde{\beta}(z) = \int_\Omega f_z\big(\omega, m_+(\omega, z)\big)\,\mathrm{d}\nu \to \int_\Omega f_\lambda(\omega, l_\omega)\,\mathrm{d}\nu = \beta(\lambda). \qquad \square$$

Now we turn to $\Im w(z)$, and prove an analogous proposition, i.e.

PROPOSITION 4.26. *For every $\lambda \in \mathbb{R}$, we have*

$$\lim_{\varepsilon\to 0} \tilde{\alpha}(\lambda + \mathrm{i}\varepsilon) = \alpha(\lambda). \tag{4.41}$$

PROOF. Before proving this proposition, we want to give an equivalent definition of the rotation number for each complex number z with $\Im z \neq 0$. Let $\omega \in \Omega$. For $\Im z > 0$, consider a solution $u(t)$ of (4.24_ω) such that

$$\Im W[u, \bar{u}] > 0, \tag{4.42}$$

where $W[u, v] = p(uv' - uv')$ is the Wronskian of u and v. Note that $W[u, \bar{u}]$ is not constant, since $\bar{u}(t)$ is not a solution of (4.24_ω).

Let

$$h(t, z) = -\frac{1}{t} \arg \frac{u(t, z)}{u(0, z)} = -\frac{1}{t} \int_0^t \Im \frac{u'(s, z)}{u(s, z)}\,\mathrm{d}s. \tag{4.43}$$

We define the rotation number as the limit

$$\lim_{t\to\infty} h(t, z). \tag{4.44}$$

Suppose for the moment that the limit (4.44) exists and does not depend on the particular choice of the solution. Then

$$\lim_{t\to\infty} h(t, z) = \lim_{t\to\infty} \frac{1}{t} \int_0^t \frac{\Im m_+(s, z)}{p(s)}\,\mathrm{d}s = -\lim_{t\to\infty} \frac{1}{t} \int_0^t \frac{\Im m_-(s, z)}{p(s)}\,\mathrm{d}s = \tilde{\alpha}(z), \tag{4.45}$$

and the given definitions turn out to be equivalent. To prove the existence of the limit (4.44), and its invariance with respect to the particular solution, we may proceed as follows. Since $\Im W[u, \bar{u}] = -2p|u|^2 \Im \frac{u'}{u}$, the assumption (4.42) is equivalent to

$$\Im \frac{u'}{u} < 0, \tag{4.46}$$

hence $h(t, z) > 0$. Now, let v any other solution of (4.24_ω) satisfying (4.42). We may assume that $v(0, z) = 1 = u(0, z)$ (the assumption (4.42) retains validity if we replace u by cu, where c is a non zero constant). We claim that

$$\left| \int_0^t \Im \left(\frac{u'}{u} - \frac{v'}{v} \right) \mathrm{d}s \right| < \pi. \tag{4.47}$$

If true, this proves that if the limit exists for a particular solution, then it exists for any other solution, and does not depend on the choice of the solution. Let us prove the claim. The assumption (4.42) implies that both $\Im u'(0,z)$ and $\Im v'(0,z)$ are negative. If our claim were false, then there would exist $t^* > 0$ such that $\arg(u(t^*,z) - v(t^*,z)) = \pm\pi$, and so there would exist a positive constant c such that

$$u(t^*,z) + cv(t^*,z) = 0.$$

But $u + cv$ is a solution of (4.24_ω) satisfying (4.42), which in turn implies that $u + cv$ has no zeros, a contradiction. The existence of the limit follows easily.

Now we turn to the proof of Proposition 4.26. The function $w(z)$ is analytic in the upper complex half plane and has positive imaginary part there: hence it is an Herglotz function. From the theory of Herglotz functions [36], there is a nondecreasing function $\rho(\lambda)$ such that

$$\Im w(z) = \Im z \int_{-\infty}^{\infty} \frac{\mathrm{d}\rho(\lambda)}{|\lambda - z|^2} + c,$$

where c is a constant (it is well known that the asymptotic behavior of m_+ as $\Im z \to \infty$ ensures that $c = 0$). In addition the following relation holds

$$\rho(\lambda_2) - \rho(\lambda_1) = \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{\lambda_1}^{\lambda_2} \Im w(\lambda + \mathrm{i}\varepsilon)\,\mathrm{d}\lambda$$

at all continuity points of ρ. If $\{w_n(z)\}$ is a sequence of analytic functions with $\Im w_n(z) > 0$ if $\Im z > 0$, and if $w_n(z) \to w(z)$ pointwise, then we have that the corresponding sequence of measures $\{\mathrm{d}\rho_n(\lambda)\}$ weakly converges to $\mathrm{d}\rho(\lambda)$ (with abuse of notation).

Let $\Im z > 0$, and put

$$h(t,z) = -\frac{1}{t} \int_0^t \Im \frac{\varphi_+'(s,z)}{\varphi_+(s,z)}\,\mathrm{d}s.$$

For every fixed $t \in \mathbb{R}$, the map $z \mapsto h(t,z)$ is analytic in the upper complex half plane, has positive imaginary part, and pointwise converges to $\Im w(z)$ as $t \to \infty$. It follows that the corresponding densities define measures which are weakly convergent to the measure $\mathrm{d}\rho(\lambda)$. If we denote these densities by $\rho(t,\lambda)$ we obtain easily

$$\rho(t,\lambda_2) - \rho(t,\lambda_1) = \frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} h(t,\lambda)\,\mathrm{d}\lambda,$$

since $h(t,\lambda)$ is continuous for $\Im z \geqslant 0$. Using the fact that, as $t \to \infty$, $\rho(t,\lambda)$ converges to $\rho(\lambda)$ at all points λ of continuity of ρ, we can write

$$\rho(\lambda_2) - \rho(\lambda_1) = \frac{1}{\pi} \lim_{t \to \infty} \int_{\lambda_1}^{\lambda_2} h(t,\lambda)\,\mathrm{d}\lambda$$

at all continuity points λ_1, λ_2 of ρ.

Recall that if $\lambda \in \mathbb{R}$, then $h(t,\lambda) \to \alpha(\lambda)$ as $t \to \infty$ (this is an easy consequence of Corollary 4.12). Using Sturm's comparison theorem we can also prove that this convergence is bounded on $[\lambda_1, \lambda_2]$, hence we can write

$$\rho(\lambda_2) - \rho(\lambda_1) = \frac{1}{\pi} \int_{\lambda_1}^{\lambda_2} \alpha(\lambda)\,\mathrm{d}\lambda.$$

From these facts we obtain

$$\Im w(z) = \tilde{\alpha}(z) = \frac{\Im z}{\pi} \int_{\mathbb{R}} \frac{\alpha(\lambda)\,\mathrm{d}\lambda}{|\lambda - z|^2}, \quad \Im z > 0. \tag{4.48}$$

Next we put $z = \eta + \mathrm{i}\varepsilon$ and compute the limit for $\varepsilon \to 0$: we have, using standard arguments,

$$\lim_{\varepsilon \to 0} \tilde{\alpha}(\eta + \mathrm{i}\varepsilon) = \lim_{\varepsilon \to 0} \frac{\varepsilon}{\pi} \int_{\mathbb{R}} \frac{\alpha(\lambda)\,\mathrm{d}\lambda}{(\lambda - \eta)^2 + \varepsilon^2} = \alpha(\eta)$$

for all $\eta \in \mathbb{R}$. □

REMARK 4.27. Using (4.48), we can obtain the following representation

$$\frac{w(z) - w(z_0)}{z - z_0} = \frac{1}{\pi} \int_{\mathbb{R}} \frac{\alpha(\lambda)\,\mathrm{d}\lambda}{(\lambda - z_0)(\lambda - z)} \quad (\Im z, \Im z_0 > 0), \tag{4.49}$$

which allows us to determine $w(z)$ up to an additive constant if $\alpha(\lambda)$ is known.

Now we go deeper in the examination of the properties of the sets $\Sigma(L)$ and $R(L)$. In particular, we will retrieve some useful information about the existence of isolated eigenvalues for the full-line operator, as well for the half-line operators. We will prove that the assumptions on Ω and on the functions p, q, y ensure that the spectrum $\Sigma(L)$ of the family (SL_ω) has no isolated points. Moreover, at most one of the Weyl m-functions can have a pole in a given open interval $J \subset R(L)$; that is, if $J \subset R(L)$ is an open interval, then $m_+(\omega,\lambda)$ and $m_-(\omega,\lambda)$ have at most one pole $P(\omega) \in J$, and such a point $P(\omega)$ satisfies $\lim_{\varepsilon\to 0} |\Im m_+(\omega, P(\omega) + \mathrm{i}\varepsilon)| = \infty$ and $\lim_{\varepsilon \to 0} |m_-(\omega, P(\omega) + \mathrm{i}\varepsilon)| < \infty$, or vice-versa.

Let $\omega \in \Omega$, and define $\mathrm{RE}^{\pm}(L_\omega) = R^{\pm}(L_\omega) \cup \{\lambda \in \Sigma(L) \mid \lambda \text{is an isolated point}\}$. We begin with the following proposition, whose proof is straightforward.

PROPOSITION 4.28. *Let $\omega \in \Omega$. If $J \subset R(L)$ is any open interval, then both $m_+(\omega,\cdot)$ and $m_-(\omega,\cdot)$ extend meromorphically through J. In particular $m_\pm(\omega,\cdot)$ has a pole at every point $\lambda_0 \in J \cap \{\mathrm{RE}^{\pm}(L_\omega) \setminus R^{\pm}(L_\omega)\}$.*

REMARK 4.29. For every $\lambda \in J \subset R(L)$, $m_+(\omega,\lambda)$ and $m_-(\omega,\lambda)$ are elements of $\mathbb{P}(\mathbb{R})$. Moreover, the maps $\lambda \mapsto m_\pm(\omega,\lambda)$ rotate without rest in opposite directions in $\mathbb{P}(\mathbb{R})$ as λ increases in J. In particular, $m_+(\omega,\cdot)$ rotates counterclockwise, and $m_-(\omega,\cdot)$ rotates clockwise. In fact, this is a simple restatement of the following elementary facts: (i) for $\Im\lambda = 0$, $\mathrm{d}m_-/\mathrm{d}\lambda < 0$ if $m_- \neq \infty$, and $\mathrm{d}(1/m_-)/\mathrm{d}\lambda > 0$ for $m_- \neq 0$; (ii) for $\Im\lambda = 0$, $\mathrm{d}m_+/\mathrm{d}\lambda > 0$ if $m_+ \neq \infty$, and $\mathrm{d}(1/m_+)/\mathrm{d}\lambda < 0$ for $m_+ \neq 0$.

We now consider point (ii) of Theorem 4.21.

PROPOSITION 4.30. *The spectrum $\Sigma(L)$ of the family (SL_ω) does not contain any isolated point.*

PROOF. We argue by contradiction. Suppose that $\lambda_0 \in \Sigma(L)$ is an isolated point. Let $D \subset \mathbb{C}$ be a disc centered in λ_0 such that $D \cap \Sigma(L) = \{\lambda_0\}$. Then $D \subset \mathrm{RE}^{\pm}(L_\omega)$ for every $\omega \in \Omega$, and both $m_+(\omega, \cdot)$ and $m_-(\omega, \cdot)$ extend meromorphically through $D \cap \mathbb{R} = J$, for ν a.a. $\omega \in \Omega$. Set $m(\omega) = m_+(\omega, \lambda_0)$. Consider the set $C = \{\omega \in \Omega \mid L_\omega \text{has no eigenvalue at} \lambda_0\}$. We claim that C is closed. To prove the claim, we can argue as follows: first, we note that the set C is $\{\tau_t\}$-invariant. Let $\bar{\omega} \in C$, and let $\{t_n\} \subset \mathbb{R}$ be a sequence such that $t_n \to \infty$ as $n \to \infty$. Let $\omega_0 = \lim_{n\to\infty} \bar{\omega} \cdot t_n$. Let $P_n(\bar{\omega})$ denote the spectral matrix corresponding to $\bar{\omega} \cdot t_n$, for every $n \in \mathbb{N}$. Since λ_0 is an isolated point, then the trace $\operatorname{Tr} P_n(\bar{\omega})$ is constant in an open interval I_0 centered at λ_0, for every $n \in \mathbb{N}$. Since $\operatorname{Tr} P_n(\bar{\omega})$ weakly converges to $\operatorname{Tr} P(\omega_0)$, it follows that $\operatorname{Tr} P(\omega_0)$ is constant on I_0 as well, and hence $\omega_0 \in C$ (because the points of growth of $\operatorname{Tr} P(\omega_0)$ coincide with the points in the spectrum of the operator L_{ω_0}). We have proved the claim.

The complement set $\Omega_1 = \Omega \setminus C$ is open and dense, since, if $\omega \in \Omega_1$ has dense orbit (such a point exists, since ν is ergodic), then $\omega \cdot t \in \Omega_1$, for all $t \in \mathbb{R}$. Consider the function $h_n : \Omega \to \mathbb{P}(\mathbb{R}) : \omega \mapsto m_+(\omega, \lambda_0 + 1/n)$. It follows that $m(\omega) = \lim_{n\to\infty} h_n(\omega)$ is a pointwise limit of continuous functions, hence it has a residual set Ω_2 of continuity points. It is easy to check that Ω_2 is $\{\tau_t\}$-invariant. Let Ω_3 be the set of all $\omega \in \Omega$ with dense orbit in Ω. Then Ω_3 is a residual set as well. Choose a point $\bar{\omega} \in \Omega_1 \cap \Omega_2 \cap \Omega_3$. Then λ_0 is an isolated eigenvalue of $L_{\bar{\omega}}$, the function $m(\omega)$ is continuous at $\bar{\omega}$, and $\{\omega \cdot t \mid t \in \mathbb{R}\}$ is dense in Ω. Let Σ_1 be the closure of the set

$$\big\{(\bar{\omega} \cdot t, m(\bar{\omega} \cdot t)) \mid t \in \mathbb{R}\big\} \subset \Omega \times \mathbb{P}(\mathbb{R}) \subset \Omega \times \mathbb{P}(\mathbb{C}) = \Sigma.$$

As we already noted, the flow $(\Sigma, \{\hat{\tau}_t\})$ induces a flow $\{\hat{\tau}_t\}$ in $\Omega \times \mathbb{P}(\mathbb{R})$ (abusing the notation slightly). It follows that Σ_1 is compact and $\{\hat{\tau}_t\}$-invariant . Moreover, it is not difficult to show that Σ_1 contains a residual set of points $\sigma = (\omega, l)$ such that $\{\hat{\tau}_t(\sigma) \mid t \in \mathbb{R}\}$ is dense in Σ_1; in fact, Σ_1 is the α-limit set of the point $(\bar{\omega}, m(\bar{\omega}))$. Let $\sigma_0 = (\omega, l_0)$ be such a point. Let $\Sigma_2 = \Sigma_1 \cap \{\Omega_1 \times \mathbb{P}(\mathbb{R})\}$. Then Σ_2 is an open, dense and $\{\hat{\tau}_t\}$-invariant subset of Σ_1. Let $\sigma = (\omega, l) \in \Sigma_2$ and let $f_{\lambda_0}(\sigma)$ as in Eq. (4.36). Let u_0 be a vector of norm 1 which is contained in l. Then

$$\ln\big|u(t)\big| = \int_0^t f_{\lambda_0}\big(\hat{\tau}_s(\sigma)\big)\,\mathrm{d}s \to -\infty$$

as $t \to \pm\infty$. By the Baire category theorem, there exists an integer N and a set $\Sigma_3 \subset \Sigma_2$ such that for every $\sigma \in \Sigma_3$,

$$\int_0^t f_{\lambda_0}\bigl(\hat{\tau}(\sigma)\bigr)\,\mathrm{d}s \leqslant N$$

for every large $t \in \mathbb{R}$. But then $\hat{\tau}_t(\sigma_0) \in \Sigma_3$ for every large t, a contradiction. $\square$

The following result is of independent interest.

PROPOSITION 4.31. *For ν a.a $\omega \in \Omega$, the following relations hold*:

$$\mathrm{RE}^+(L_\omega) = \mathrm{RE}^-(L_\omega) = R(L) = R^{\mathrm{ed}}(L).$$

PROOF. According to Proposition 4.13, there is a set $\Omega_\nu \subset \Omega$ with $\nu(\Omega_\nu) = 1$ such that, if $\omega \in \Omega_\nu$, then ω has dense orbit and is positive and negative Poisson recurrent in the sense that there are sequences $\{t_n\} \to \infty$ and $\{s_n\} \to -\infty$ such that $\lim_{n\to\infty} \omega \cdot t_n = \lim_{n\to\infty} \omega \cdot s_n = \omega$.

Let $T > 0$, and consider Eq. (4.24_ω) on the interval $[0, T]$ together with the boundary conditions $\varphi(0) = \varphi(T) = 0$. There is a corresponding self-adjoint differential operator L_ω^T on the weighted space $L^2[0, T]$ with density $y(t)$. Let $J = (\lambda', \lambda'') \subset \mathbb{R}$ be an open interval containing exactly $k \geqslant 0$ eigenvalues of the half-line operator L_ω^+. We will use a beautiful argument of Hartman [66] to show that J contains no more than $k+1$ eigenvalues of L_ω^T.

To prove this statement, let $\varphi(t, \lambda_1), \ldots, \varphi(t, \lambda_k)$ be eigenfunctions of L_ω^+ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_k \in J$. Suppose for contradiction that L_ω^T admits $N \geqslant k+2$ eigenvalues in J with corresponding eigenfunctions $\varphi_1(t), \ldots, \varphi_N(t)$; thus $\varphi_j(0) = \varphi_j(T) = 0$ $(1 \leqslant j \leqslant N)$. Write

$$z(t) = \sum_{j=1}^N c_j \varphi_j(t) \quad (0 \leqslant t \leqslant T), \tag{4.50}$$

then determine coefficients $c_1, \ldots, c_N$, not all zero, such that

$$\begin{aligned} &z'(T) = 0,\\ &\int_0^T z(t)\varphi(t, \lambda_j) y(t)\,\mathrm{d}t = 0 \quad (1 \leqslant j \leqslant k). \end{aligned} \tag{4.51}$$

This is possible because $N \geqslant k+2$, while (4.51) consists of only $k+1$ linear relations.

Next extend $z(t)$ to all of $[0, \infty)$ by setting $z(t) = 0$ for $t \geqslant T$. Then z is in the domain of L_ω^+. Let $\bar{\lambda} = \frac{\lambda' + \lambda''}{2}$, and set $w = L_\omega^+(z) - \bar{\lambda} z$. Using (4.50) together with the spectral theorem for self-adjoint operators, we see that

$$\int_0^\infty w^2(t) y(t)\,\mathrm{d}t \geqslant \frac{(\lambda'' - \lambda')^2}{4} \int_0^\infty z^2(t) y(t)\,\mathrm{d}t.$$

On the other hand, condition (4.50) together with the fact that $L^T_\omega(z) = L^+_\omega(z)$ for $0 \leqslant t < T$ imply that

$$\int_0^\infty w^2(t)y(t)\,\mathrm{d}t = \int_0^T w^2(t)y(t)\,\mathrm{d}t < \frac{(\lambda'' - \lambda')^2}{4}\int_0^T z^2(t)y(t)\,\mathrm{d}t$$
$$= \frac{(\lambda'' - \lambda')^2}{4}\int_0^\infty z^2(t)y(t)\,\mathrm{d}t.$$

These relations are inconsistent, so $N \leqslant k + 1$ as stated.

Continuing, we claim that J contains no more than $2k + 2$ eigenvalues of $L^+_{\omega\cdot T}$. It is easiest to prove this by passing to the $u = \binom{u_1}{u_2}$-plane. The proof is by contradiction. Let $u(t,\lambda) = \binom{u_1(t,\lambda)}{u_2(t,\lambda)}$ satisfy (SL_ω) and $u_1(0,\lambda) = 0, u_2(0,\lambda) = 1$ ($\lambda \in J, 0 \leqslant t \leqslant T$). Furthermore, write $m(t,\lambda) = m_+(\omega\cdot t,\lambda)$ ($\lambda \in J, 0 \leqslant t \leqslant T$). Let us parametrize $u(t,\lambda)$ in a continuous way using the polar coordinate θ in the u-plane, then identify $u(t,\lambda)$ with its θ-value; we carry out the same operation on $m(t,\lambda)$ ($\lambda \in J, 0 \leqslant t \leqslant T$). It is clear that the interval $(m(T,\lambda'), m(T,\lambda''))$ contains $2k + 3$ consecutive numbers of the form $\frac{\pi}{2} + j\pi$, $j \in \mathbb{Z}$. A moment's thought shows that the interval $(u(T,\lambda'), u(T,\lambda''))$ contains $k+2$ consecutive numbers of the form $\frac{\pi}{2} + l\pi$, $l \in \mathbb{Z}$. But then L^T_ω has at least $k+2$ eigenvalues in J, contradicting the previous result. So in fact J contains non more than $2k + 2$ eigenvalues of $L^+_{\omega\cdot T}$, for each $T > 0$.

Now, let $\omega_1 \in \Omega$, and let $\{t_n\} \subset \mathbb{R}$ be a sequence of positive real numbers such that $\lim_{n\to\infty} \omega\cdot t_n = \omega_1$. Let $\rho^+_n := \rho^+_{\omega\cdot t_n}$ be the spectral measure corresponding to the half-line operator $L^+_{\omega\cdot t_n}$. Then ρ^+_n weakly converges to ρ^+, where ρ^+ is the spectral measure associated to $L^+_{\omega_1}$. The eigenvalues of $L^+_{\omega_1}$ correspond to the discontinuity points of ρ^+. It follows that ρ^+ has no more than $2k+3$ discontinuity points, because the previous remarks apply to each t_n.

Now, consider Eq. ($4.24_{\omega\cdot(-T)}$) in the interval $[0, T]$, together with the boundary conditions $u(0) = u(T) = 0$ ($T > 0$). It follows again that J contains no more than $2k + 3$ eigenvalues of the operator $L^+_{\omega\cdot(-T)}$. If we consider Eq. (4.24_ω) with the boundary conditions

$$u(0) = u(-T) = 0$$

then the number of eigenvalues in J of the operator L^{-T}_ω is equal to the number of eigenvalues, always in J, of the operator $L^T_{\omega\cdot(-T)}$, hence it is $\leqslant 2k + 3$.

Let $\{s_n\} \subset \mathbb{R}$ be a sequence of positive real numbers such that $\lim_{n\to\infty} \omega\cdot(-s_n) = \omega_1$. For every $n \in \mathbb{N}$, consider the spectral measure, which we still denote by ρ^+_n, corresponding to Eq. ($4.24_{\omega\cdot(-s_n)}$) in the interval $[-s_n, 0]$, together with the boundary conditions $u(0) = u(-s_n) = 0$. Recall that ρ^+_n is piecewise constant, with a jump of discontinuity equal to

$$\left(\int_{-s_n}^0 y(t)\left\|\chi_i(t)\right\|^2 \mathrm{d}t\right)^{-1}$$

at every eigenvalue λ_i of Eq. (4.24$_{\omega\cdot(-s_n)}$), where $\chi_i(t)$ is the normalized eigenfunction corresponding to the eigenvalue λ_i. It can be shown that ρ_n^+ converges weakly to $\rho_{\omega_1}^-$, where $\rho_{\omega_1}^-$ is the spectral measure for the negative half-line operator $L_{\omega_1}^-$. As before, we can prove that $\rho_{\omega_1}^-$ has at most $2k+3$ discontinuity points in J. We conclude that both $m_+(\omega,\cdot)$ and $m_-(\omega,\cdot)$ extend meromorphically through J. It follows that the spectral matrix M of the full line operator extends meromorphically through J, hence the full line operator L_ω has at most isolated eigenvalues on J. However, the spectrum of L_ω contains no isolated eigenvalues and in fact $R(L_\omega) = R(L)$, which in turn equals $R^{\mathrm{ed}}(L)$ by Proposition 4.15. All this implies that $R^+(L_\omega) \cup R^-(L_\omega) \subset R^{\mathrm{ed}}(L)$. Since $R^{\mathrm{ed}}(L) \subset R^+(L_\omega) \cup R^-(L_\omega)$, we have proved Proposition 4.31. □

Let us note that the statement of Proposition 4.31 holds on the orbit closure of any point ω which is positive and negative Poisson recurrent.

The final result concerning the structure of the spectrum is stated in the following

THEOREM 4.32. *Let $\omega \in \Omega$ have dense orbit, and let J be an open interval with $J \subset \mathrm{RE}^+(L_\omega)$. Then the half-line operator L_ω^+ has at most one eigenvalue in J. If L_ω^+ has an eigenvalue in J, then L_ω^- has no eigenvalue in J, and vice-versa. In particular, if $m_+(\omega,\cdot)$ has a pole $\lambda_0 \in J$, then λ_0 is a regular point for $m_-(\omega,\cdot)$ and vice-versa.*

PROOF. By some results of Coppel [29], the bundles Ker P_ω and Im P_ω vary continuously with respect to λ. Since $J \subset \mathrm{RE}^+(L_\omega)$, we have that $m_+(\omega,\lambda) \neq m_-(\omega,\lambda)$ for all $\lambda \in J$. Moreover, the above observations ensure that only one of $m_+(\omega,\cdot)$ and $m_-(\omega,\cdot)$ can attain the value ∞, and such a value can be attained at most once. This concludes the proof, since, from Theorem 4.11, $m_+(\omega,\cdot)$ and $m_-(\omega,\cdot)$ attain the values ∞ only at those points $\lambda \in J$ which are isolated eigenvalues for L_ω^+ and L_ω^- respectively. □

Our next goal is to complete the proof of point (iii) of Theorem 4.21. We need a formula for the z-derivative of $w(z)$ which involves the potential $y(\omega)$ and the Weyl m-functions as well. This formula will turn to be very useful also in the following, when we will discuss a theory of Kotani-type for the general Sturm–Liouville operator. In [59,31,89], and also in [81], one can find very interesting (trace) formulas for $\mathrm{d}w/\mathrm{d}z$ and its representation in terms of the trace of the spectral matrix $P(\lambda)$ arising from the Green (matrix) function. In our case, it is not very difficult to compute directly such a derivative. Let $\mathcal{L} = \{f : \Omega \to \mathbb{R} \mid f \text{ is continuous and strictly positive}\}$. Then $\mathcal{L}$ can be regarded as an open subset of the Banach space $C(\Omega)$ of continuous real-valued functions on Ω. By a direct observation of the Riccati equation (4.25), letting $\tilde{z} = zy$, the Floquet exponent $w(\cdot)$ can be considered as a functional from $\mathcal{L}$ into $\mathbb{C}$, as follows: $w(z,\cdot) : \mathcal{L} \to \mathbb{C} : y \mapsto w(zy)$, for each fixed $z \in \mathbb{C}$. We now adapt to our Sturm–Liouville setting the developments in [89].

PROPOSITION 4.33. *For every fixed $z \in \mathbb{C} \setminus \mathbb{R}$, the functional $w(z,\cdot)$ defined above admits a Frechet derivative $D_f w(z,\cdot)$ along every direction $f \in \mathcal{L}$. Moreover, for every $z \in \mathbb{C}$ with $\Im z \neq 0$, and for every ergodic measure ν on Ω, we have*

$$\frac{\mathrm{d}}{\mathrm{d}z} w(z) = \int_\Omega \frac{y(\omega)}{m_-(\omega,z) - m_+(\omega,z)} \,\mathrm{d}\nu(\omega). \tag{4.52}$$

PROOF. Let $z \in \mathbb{C} \setminus \mathbb{R}$. Consider the functionals $m_\pm : \mathcal{L} \to \mathbb{C} : y \mapsto m_\pm(z, y)$, and suppose for the moment that $m_\pm$ are Frechet differentiable at $y \in \mathcal{L}$. Let $f \in \mathcal{L}$ and denote by $D_f m_\pm(z, y)$ the Frechet derivatives of $m_\pm(z, \cdot)$ computed at $y \in \mathcal{L}$ in the direction $f \in \mathcal{L}$. Then we have

$$D_f w(z, y) = \int_\Omega \frac{D_f m_+(z, y(\omega)) - D_f m_-(z, y(\omega))}{2p(\omega)} \, \mathrm{d}\nu, \tag{4.53}$$

where ν is a fixed ergodic measure on Ω. This shows that w is Frechet differentiable at $y \in \mathcal{L}$.

We compute $D_f m_\pm(z, y)$. From the Riccati equation (4.25) it follows that $D_f m_\pm(z, y)$ must satisfy

$$\big(D_f m_\pm(z, y)\big)' + \frac{2}{p} m_\pm D_f m_\pm(z, y) = -zf. \tag{4.54}$$

The solutions $\hat{D}_f m_\pm(z, y)$ of these first order differential equations are

$$\begin{aligned} \hat{D}_f m_-(z, y) &= \int_{-\infty}^{t} -zf(s) \exp\left\{2 \int_t^s \frac{m_-}{p} \, \mathrm{d}r\right\} \mathrm{d}s, \\ \hat{D}_f m_+(z, y) &= \int_t^{\infty} zf(s) \exp\left\{2 \int_t^s \frac{m_+}{p} \, \mathrm{d}r\right\} \mathrm{d}s \end{aligned} \tag{4.55}$$

hence

$$\hat{D}_f m_-(z, y) = \int_{-\infty}^{t} -zf(s) \frac{\varphi_-^2(s)}{\varphi_-^2(t)} \, \mathrm{d}s \tag{4.56}$$

and

$$\hat{D}_f m_+(z, y) = \int_t^{\infty} zf(s) \frac{\varphi_+^2(s)}{\varphi_+^2(t)} \, \mathrm{d}s. \tag{4.57}$$

Now we show that $D_f m_\pm(z, y)$ actually exist and equal the solutions $\hat{D}_f m_\pm(z, y)$ defined in (4.56) and (4.57). We prove this only for m_+. Let us simplify the notation and omit the dependence with respect to $z \in \mathbb{C} \setminus \mathbb{R}$. Let $\Delta m_+ = m_+(y + \varepsilon f) - m_+(y)$. Then the Riccati equation (4.25) ensures that

$$(\Delta m_+)' + \frac{1}{p}\big(m_+(y + \varepsilon f) + m_+(y)\big)\Delta m_+ = -z\varepsilon f. \tag{4.58}$$

Since $\Im z \neq 0$, the family (SL_ω) admits an exponential dichotomy, and an estimate of Coppel ([29], page 34) ensures that $m_+(y + \varepsilon f) = m_+(y) + \mathcal{O}(\varepsilon f)$. It follows that Eq. (4.58) admits a unique bounded solution which (obviously) must be Δm_+. Moreover

$$\Delta m_+(t) = \int_t^{\infty} z\varepsilon f(s) \exp\left\{\int_t^s \frac{2m_+(y) + \mathcal{O}(\varepsilon f)}{p} \, \mathrm{d}r\right\} \mathrm{d}s.$$

It is now easy to see that

$$\lim_{\varepsilon\to 0}\frac{|\Delta m_+ - \varepsilon\hat{D}_f m_+(z,y)|}{\varepsilon}=0,$$

where the limit is appropriately uniform in $f\in\mathcal{L}$. This shows that $D_f m_+(z,y)$ exists and equals $\hat{D}_f m_+(z,y)$ for every $f\in\mathcal{L}$. The same holds for m_-.

It remains to prove the formula (4.52). Let us consider the Frechet derivatives $D_y m_\pm(z,y)$ along the direction y. By a direct computation, we get

$$\begin{aligned}D_y w(z,y)&=\lim_{\varepsilon\to 0}\frac{w(z(y+\varepsilon y))-w(zy)}{\varepsilon}=\lim_{\varepsilon\to 0}\frac{w(zy(1+\varepsilon))-w(zy)}{\varepsilon}\\&=z\frac{\mathrm{d}}{\mathrm{d}z}w(z,y).\end{aligned}$$

We calculate $D_y w(z,y)$: we have

$$\begin{aligned}&\frac{D_y m_+(z,y)-D_y m_-(z,y)}{p}\\&\quad=\frac{z}{p}\left[\left(\int_t^\infty\frac{\varphi_+^2(s)}{\varphi_+^2(t)}y(s)\,\mathrm{d}s+\int_{-\infty}^t\frac{\varphi_-^2(s)}{\varphi_-^2(t)}y(s)\,\mathrm{d}s\right)\right].\end{aligned}\tag{4.59}$$

Next we observe that

$$\left(\frac{\varphi_+}{\varphi_-}\right)'=\frac{\varphi_+'\varphi_- - \varphi_+\varphi_-'}{\varphi_-^2}=-\frac{W}{p\varphi_-^2}\quad\text{and}\quad\left(\frac{\varphi_-}{\varphi_+}\right)'=\frac{\varphi_-'\varphi_+ - \varphi_-\varphi_+'}{\varphi_+^2}=\frac{W}{p\varphi_+^2}$$

where W is the Wronskian of φ_+ and φ_-. We obtain

$$\begin{aligned}&\frac{D_y m_+(z,y)-D_y m_-(z,y)}{2p}\\&\quad=\frac{z}{2W}\left[\left(\frac{\varphi_-}{\varphi_+}\right)'\int_t^\infty\varphi_+^2(s)y(s)\,\mathrm{d}s-\left(\frac{\varphi_+}{\varphi_-}\right)'\int_{-\infty}^t\varphi_-^2(s)y(s)\,\mathrm{d}s\right].\end{aligned}\tag{4.60}$$

It follows that

$$\begin{aligned}D_y w(z,y)=\lim_{T\to\infty}\frac{1}{T}\int_0^T\frac{z}{2W}\bigg[&\left(\frac{\varphi_-}{\varphi_+}\right)'\int_t^\infty\varphi_+^2(s)y(s)\,\mathrm{d}s\\&-\left(\frac{\varphi_+}{\varphi_-}\right)'\int_{-\infty}^t\varphi_-^2(s)y(s)\,\mathrm{d}s\bigg]\mathrm{d}t.\end{aligned}\tag{4.61}$$

To make the computation easier, we add in the integral the derivative with respect to t of the function

$$h(t)=\frac{z}{2W}\left[\left(-\frac{\varphi_-}{\varphi_+}\right)\int_t^\infty \varphi_+^2(s)y(s)\,\mathrm{d}s+\left(\frac{\varphi_+}{\varphi_-}\right)\int_{-\infty}^t \varphi_-^2(s)y(s)\,\mathrm{d}s\right]. \tag{4.62}$$

Note that $h'(t)$ gives no contribution to the integral (4.61). We obtain

$$\begin{aligned}&\lim_{T\to\infty}\frac{z}{T}\int_0^T\frac{1}{2W}\left[\left(\frac{\varphi_-}{\varphi_+}\right)'\int_t^\infty \varphi_+^2(s)y(s)\,\mathrm{d}s-\left(\frac{\varphi_+}{\varphi_-}\right)'\int_{-\infty}^t \varphi_-^2(s)\,\mathrm{d}s\right]\mathrm{d}t\\ &\quad=\lim_{T\to\infty}\frac{z}{T}\int_0^T\frac{1}{2W}\left[\left(\frac{\varphi_-}{\varphi_+}\right)'\int_t^\infty \varphi_+^2(s)y(s)\,\mathrm{d}s\right.\\ &\qquad\left.-\left(\frac{\varphi_+}{\varphi_-}\right)'\int_{-\infty}^t \varphi_-^2(s)y(s)\,\mathrm{d}s+h'(s)\right]\mathrm{d}t\\ &\quad=\lim_{T\to\infty}\frac{z}{T}\int_0^T\frac{y(t)(\varphi_+(t)+\varphi_-(t))}{2(m_-(\omega\cdot t,z)-m_+(\omega\cdot t,z))}\,\mathrm{d}t.\end{aligned} \tag{4.63}$$

Using the Birkhoff ergodic theorem we have

$$\frac{\mathrm{d}w}{\mathrm{d}z}=\int_\Omega\frac{y(\omega)(\varphi_+(0)+\varphi_-(0))}{2(m_-(\omega,z)-m_+(\omega,z))}\,\mathrm{d}\nu(\omega)=\int_\Omega\frac{y(\omega)}{m_-(\omega,z)-m_+(\omega,z)}\,\mathrm{d}\nu(\omega)$$

for every $\{\tau_t\}$-ergodic measure ν on Ω. The proof is complete. □

We can now return to point (iii) of Theorem 4.21. Arguing as in [89, Section 6], and using our formula (4.52) for $\frac{\mathrm{d}w}{\mathrm{d}z}$ together with (4.49), we can show that, if $\alpha(\lambda)$ is constant in an interval $I\subset\mathbb{R}$, then $I\subset R(L)$. We omit the details as they are similar to those in the cited paper [89].

In the rest of Section 4.3 we discuss a theory of Kotani-type for the Sturm–Liouville operator. We begin with the following lemmas.

LEMMA 4.34. *If* $\Im z\neq 0$, *then*

$$\Re w(z)=\mp\frac{\Im z}{2}\int_\Omega\frac{y(\omega)}{\Im m_\pm(\omega,z)}\,\mathrm{d}\nu. \tag{4.64}$$

In particular,

$$\Re w(z)\leqslant 0.$$

PROOF. We consider the imaginary parts in the Riccati equation for m_+, obtaining

$$\Im m_+'+\frac{2}{p}\Re m_+\Im m_+=-y\Im z.$$

Now, dividing by $\Im m_+$ on both sides of the equation, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\ln(\Im m_+) + \frac{2}{p}\Re m_+ = -\frac{y\Im z}{\Im m_+},$$

and hence the proof follows integrating both sides on Ω with respect to the ergodic measure ν, and using the Birkhoff ergodic theorem. □

Let us denote by $H^\pm$ the upper and lower complex half-planes respectively.

LEMMA 4.35. *Let $I \subset \mathbb{R}$ be an open interval and let $f : H^+ \to H^+$ be analytic.*

Suppose that $\lim_{\varepsilon\to 0^+}\Re(f(\lambda + \mathrm{i}\varepsilon)) = 0$ *for Lebesgue a.a. $\lambda \in I$. Then f extends analytically through I, and its extension $\bar{f}$ has positive imaginary part on I.*

Lemma 4.35 is a standard result from the theory of complex functions, hence we will not give the proof. Instead we refer to a standard text such as [64] for more details.

Recall that $\lim_{\varepsilon\to 0}\Re w(\lambda + \mathrm{i}\varepsilon) = \beta(\lambda)$ and $\lim_{\varepsilon\to 0}\Im w(\lambda + \mathrm{i}\varepsilon) = \alpha(\lambda)$, for every $\lambda \in \mathbb{R}$. Motivated by Kotani's seminal work on the Schrödinger operator [101], and using the above formulas (4.64) and (4.52), we obtain, after omitting to write the dependence with respect to z of the functions involved:

$$\begin{aligned}
&-4\left(\frac{\Re w}{\Im z} + \Im\frac{\mathrm{d}w}{\mathrm{d}z}\right) \\
&\quad = \int_\Omega \frac{y(\omega)(\Im m_-(\omega) - \Im m_+(\omega))}{\Im m_-(\omega)\Im m_+(\omega)}\left(1 + 4\frac{\Im m_-(\omega)\Im m_+(\omega)}{|m_-(\omega) - m_+(\omega)|^2}\right)\mathrm{d}\nu \\
&\quad = \int_\Omega y(\omega)\left(\frac{1}{\Im m_+(\omega)} - \frac{1}{\Im m_-(\omega)}\right) \\
&\qquad \times \left(\frac{[\Re(m_-(\omega) - m_+(\omega))]^2 + [\Im(m_-(\omega) + m_+(\omega))]^2}{|m_-(\omega) - m_+(\omega)|^2}\right)\mathrm{d}\nu.
\end{aligned}$$

Let $I \subset \mathbb{R}$ be an open interval such that $\beta(\lambda) = 0$ for Lebesgue a.a. $\lambda \in I$. Note that the function $-w(\cdot)$ is analytic in the upper half-plane H^+ and has positive real part there. In addition $-w(\cdot)$ is bounded in every compact subset of H^+. We apply the Schwarz reflection principle [136], to prove that $w(\cdot)$ admits an analytic extension from H^+ through I to the set $H = I \cup \{\lambda \in \mathbb{C} \mid \Im\lambda \neq 0\}$. Since $\beta(\lambda) = 0$ for Lebesgue a.a. $\lambda \in I$, we have $\Re w(\lambda) = 0$. It follows that

$$\Im\frac{\mathrm{d}w}{\mathrm{d}\lambda}(\lambda) = -\lim_{\varepsilon\to 0^+}\frac{\Re w(\lambda + \mathrm{i}\varepsilon)}{\varepsilon}.$$

Since the limits $\lim_{\varepsilon\to 0^+} m_\pm(\omega, \lambda + \mathrm{i}\varepsilon) := m_\pm(\omega, \lambda)$ exist for Lebesgue a.a. $\lambda \in I$ and for ν a.a. $\omega \in \Omega$, we can use Fatou's Lemma and the computation above to obtain that $\Re m_+(\omega, \lambda) = \Re m_-(\omega, \lambda)$ and $\Im m_+(\omega, \lambda) = -\Im m_-(\omega, \lambda)$, for Lebesgue a.a. $\lambda \in I$ and for ν a.a. $\omega \in \Omega$. We claim that the maps $(m_+ - m_-)(\omega, \cdot)$ and $\frac{m_- m_+}{m_- - m_+}(\omega, \cdot)$ extend analytically through I to H. To prove the claim, we first describe a general fact concerning functions which are analytic in H^+. Let $f(z)$ be an analytic function on H^+ with $\Im f(z) > 0$ for all $z \in H^+$. Put

$$F(z) = \ln f(z) = \ln\bigl|f(z)\bigr| + \mathrm{i} \arg f(z) = u(z) + \mathrm{i} v(z),$$

where $0 \leqslant \arg f(z) < \pi$. It follows that $u(z)$ is harmonic in H^+, and that $v(z)$ is harmonic and bounded in H^+. For $F(z)$ a generalized version of the Schwarz reflection principle holds [64]: namely, if $\Im F(\lambda) = 0$ for Lebesgue a.a $\lambda \in I$, then $F(z)$ extends analytically through I from $H \cap H^+$ to $H \cap H^-$. In our case, it follows that for ν-a.a $\omega \in \Omega$, the function $m(\omega, \cdot) := m_+(\omega, \cdot) - m_-(\omega, \cdot)$ has the properties that $\Im m(\omega, z) > 0$, for every $z \in H^+$, $\Re m(\omega, \lambda) = 0$ and $\Im m_+(\omega, \lambda) = -\Im m_-(\omega, \lambda)$, for Lebesgue a.a $\lambda \in I$. Letting $G(z) = \ln m(\omega, z)$, we have $G(z) = \ln|m(\omega, z)| + i\pi/2$. The function $F(z) = G(z) - \mathrm{i}\pi/2$ satisfies $\Im F(\lambda) = 0$ for Lebesgue a.a. $\lambda \in I$. From the above observations, it follows that $m_+(\omega, z) - m_-(\omega, z)$ extends analytically through I to H^-. The same holds for the function $\frac{m_- m_+}{m_- - m_+}$. Now, it is easy to prove that both m_+ and m_- extend analytically through I from $H \cap H^+$ to $H \cap H^-$. We proved the first part of the following important

THEOREM 4.36. *Let $I \subset \mathbb{R}$ be an open interval, and suppose that $\beta(\lambda) = 0$ for Lebesgue a.a. $\lambda \in I$. Then, for ν a.a. $\omega \in \Omega$, the Weyl m-functions $m_\pm(\omega, \cdot)$ admit analytic extensions through I from H^+ to H^-. If we denote these extensions with $h_\pm(\omega, \cdot)$, then*

$$h_+(\omega, \lambda) = \begin{cases} m_+(\omega, \lambda), & \Im\lambda > 0, \\ m_-(\omega, \lambda), & \Im\lambda < 0 \end{cases} \quad \textit{and} \quad h_-(\omega, \lambda) + \begin{cases} m_-(\omega, \lambda), & \Im\lambda > 0, \\ m_+(\omega, \lambda), & \Im\lambda < 0. \end{cases}$$

PROOF. We give the proof only for $h_+(\omega, \cdot)$. Let $\Omega_1 \subset \Omega$ such that both m_+ and m_- extend analytically through I. Let $\Omega_2 \subset \Omega$ be the set consisting on those points $\omega \in \Omega$ with dense orbit. Then $\nu(\Omega_1 \cap \Omega_2) = 1$. Choose $\bar{\omega} \in \Omega_1 \cap \Omega_2$. Let $\lambda \in \mathbb{C}$ with $\Im\lambda \neq 0$, and define $h_+(\lambda) := h_+(\bar{\omega}, \lambda)$. We write down the Riccati equation (4.25)

$$m' + \frac{1}{p} m^2 = q - \lambda y.$$

Suppose that $\Im\lambda < 0$. Then, for every initial condition $m_0 \in \mathbb{R}$ and for every $t > 0$, we have $\Im m(t, \lambda) > 0$ and $\Im n(t, \lambda) := \Im \frac{1}{m(t, \lambda)} < 0$, where $m(t, \lambda)$ is the solution of Eq. (4.25) satisfying $m(0, \lambda) = m_0 \in \mathbb{R}$. We define a flow $\{\tilde{\tau}_t\}$ on $\Omega \times \mathbb{P}(\mathbb{C}) = \Sigma$ by setting

$$\tilde{\tau}_t(\omega, m) = \bigl(\omega \cdot t, m(t)\bigr),$$

where $m(t)$ is the solution of Eq. (4.25) satisfying $m(0) = m \in \mathbb{C}$. Note that the flow $\{\tilde{\tau}_t\}$ can be defined in an equivalent manner by considering the action of the flow $\{\hat{\tau}_t\}$ on complex lines in $\mathbb{P}(\mathbb{C})$, i.e.,

$$\tilde{\tau}_t(\omega, m) = \hat{\tau}_t\left(\omega, \begin{pmatrix} 1 \\ m \end{pmatrix}\right).$$

From the above observations, it follows easily that for every fixed $\omega \in \Omega$, $\lambda \in H^-$ and $t > 0$, the map $\tilde{\tau}_t^\omega : \mathbb{P}(\mathbb{C}) \to \mathbb{P}(\mathbb{C}) : m \mapsto \tilde{\tau}_t(\omega, m) = (\omega \cdot t, m(t))$ maps the closure of the set $\{\infty\} \times \{z \in \mathbb{C} \mid \Im z \geqslant 0\}$ into H^+. Alternatively, identifying H^+ with a disc B contained in $\mathbb{P}(\mathbb{C})$, $\tilde{\tau}_t^\omega$ maps the closure of B into B, for every $\omega \in \Omega$, $\lambda \in H^-$ and $t > 0$.

These things being said, let $S^\pm(\lambda) = \{(\omega, m_\pm(\omega, \lambda)) \mid a \in \Omega\} \subset \Sigma$. Then $S^\pm(\lambda)$ are $\{\tilde{\tau}_t\}$-invariant sections of Σ. Recall that $S^\pm(\lambda)$ are minimal subsets of the projective flow $(\Sigma, \{\hat{\tau}_t\})$.

We claim that $S^\pm(\lambda)$ are the only $\{\tilde{\tau}_t\}$-invariant sections of Σ. To prove this, we use methods already exploited in Section 3. Suppose for contradiction that S is another invariant section of Σ. Let $\omega \in \Omega$ have dense orbit. Then there exists $l \in \mathbb{P}(\mathbb{C})$ such that $(\omega, l) \in S$. It follows that $(\omega, m_+(\omega, \lambda))$, $(\omega, m_-(\omega, \lambda))$ and (ω, l) form distal pairs in Σ. Choose unit vectors $u^+ \in m_+(\omega, \lambda)$, $u^- \in m_-(\omega, \lambda)$, $u \in (\omega, l)$. Then u^+, u^-, u are pairwise linearly independent. Let $\Phi_\omega(t)$ be the fundamental matrix solution of Eq. (SL_ω) with $\Phi_\omega(0) = I$. Let $u^\pm(t) = \Phi_\omega(t) u^\pm$, $u(t) = \Phi_\omega(t) u$. Since $\Im\lambda \neq 0$, the family (SL_ω) admits an exponential dichotomy over Ω, hence for every fixed $\omega \in \Omega$,

$$\lim_{t \to -\infty} \left|u^+(t)\right| = \infty, \qquad \lim_{t \to \infty} \left|u^-(t)\right| = \infty.$$

Suppose that $u(t)$ is unbounded at ∞. Let $\Psi(t)$ be the fundamental matrix solution of Eq. (SL_ω) whose columns are $u^-(t)$ and $u(t)$. By Liouville's formula, $\det \Psi(t) = \det \Psi(0) = \mathit{const} \neq 0$, and since $u^-(t)$ is unbounded at ∞, we must have $|u(t)| \to 0$, a contradiction. Hence $u(t)$ is bounded at ∞. We carry out the same reasoning at $-\infty$ to prove that $u(t)$ decays as $|t| \to \infty$. However, in this case, let $\Psi_1(t)$ be the fundamental matrix solution of Eq. (SL_ω) whose columns are $u^+(t)$ and $u(t)$. Then $\liminf_{t \to \infty} \det \Psi_1(t) = 0$, a contradiction. We proved that $S^\pm(\lambda)$ are the unique $\{\tilde{\tau}_t\}$-invariant minimal sections of Σ.

Now, let D_0 be a complex domain with $D_0 \supset (I \cup \{\lambda \mid \Im\lambda \neq 0\})$ and such that $h_+(\lambda)$ is analytic and has positive imaginary part in D_0. For every $t > 0$, we define a function $g_t : D_0 \to \mathbb{C}$, by setting

$$\bigl(\tilde{\tau}_t(\bar{\omega}), g_t(\lambda)\bigr) = \tilde{\tau}_t\bigl(\bar{\omega}, h_+(\lambda)\bigr),$$

i.e., $g_t(\lambda) = h_+(t, \lambda)$, where $h_+(t, \lambda)$ is the solution of Eq. (4.25) satisfying $h_+(0, \lambda) = h_+(\lambda)$. It follows that, for every $t > 0$, g_t maps D_0 into H^+. In fact, if $\Im\lambda > 0$, this follows from the invariance of the section $S^+(\lambda)$; if $\lambda \in I$, it follows from $\Im h_+ > 0$, and if $\Im\lambda < 0$, it follows from $\Im h_+ > 0$ and the first part of the proof. Moreover, for every $t > 0$, the map g_t in analytic in D_0, since $g_t(\lambda) = h_+(t, \lambda) = \Phi_{\bar{\omega}}(t)\binom{1}{h_+(\lambda)}$ and $\Phi_{\bar{\omega}}(t)$ is linear. It follows that the family $\{g_t(\lambda), t > 0\}$ is a normal family. Let $\omega \in \Omega$, and choose a sequence $\{t_n\} \to \infty$ such that $\bar{\omega} \cdot t_n \to \omega$ and such that $g_{t_n}(\lambda)$ converges to a holomorphic function $g_\omega(\lambda)$ on compact subsets of D_0 (if necessary, we pass to a subsequence). From the invariance and the continuity of the section $S^+(\lambda)$ with respect to λ, we conclude that $g_\omega(\lambda) = m_+(\omega, \lambda)$ for every λ with $\Im\lambda > 0$. Moreover, one easily sees that $g_{t_n}(\lambda)$ converges to $g_\omega(\lambda)$ on compact subsets of D_0, for every sequence $\{t_n\} \subset \mathbb{R}$ such that $\bar{\omega} \cdot t_n \to \omega$ as $n \to \infty$.

We can repeat this construction for every $\omega \in \Omega$. Now, using standard arguments about normal families [37], we prove that the section $S(\lambda) = \{(\omega, g_\omega(\lambda)) \mid \omega \in \Omega\}$ is continuous and $\{\tilde{\tau}_t\}$-invariant for every $\lambda \in \mathbb{C}$ with $\Im\lambda > 0$. Of course $S(\lambda) = S^-(\lambda)$ if $\Im\lambda < 0$, and the proof is complete (for m_- we use similar arguments). $\square$

We summarize the results we obtained about the behavior of the Weyl m-functions when crossing the real line, in the following

THEOREM 4.37. *For ν a.a. $\omega \in \Omega$, the Weyl m-functions extend meromorphically through every interval $J \subset R(L)$ from H^+ to H^-. Moreover, $m_\pm(\omega,\lambda) \in \mathbb{P}(\mathbb{R})$, for all $\lambda \in J$.*

If $I \subset \mathbb{R}$ is an open interval such that $\beta(\lambda) = 0$ for Lebesgue a.a. $\lambda \in I$, then $m_\pm$ extend analytically through I from H^+ to H^-. The extensions $h_\pm(\omega,\cdot)$ satisfy $h_\pm(\omega,z) = m_\mp(\omega,z)$ for all $z \in H^-$, and $\Re h_+(\omega,\lambda) = \Re h_-(\omega,\lambda)$ for Lebesgue a.a. $\lambda \in I$.

We finish this subsection by observing that actually $\beta(\lambda)$ can attain the value 0 only at points $\lambda \in \Sigma(L)$. In fact, if $\lambda \in R(L)$, it is a simple consequence of the Oseledets theorem that $\beta(\lambda) < 0$ (the Oseledets bundles are disjoint, and are contained in the bundles of exponential dichotomy, hence there are exactly two Lyapunov exponents, $\beta_1(\lambda) = -\beta_2(\lambda) \neq 0$).

4.4. *The inversion problem*

In this part, we pose and solve our inverse algebro-geometric problem for the Sturm–Liouville operator. We impose the following basic

HYPOTHESIS 4.38.

(H1) The spectrum of the family (SL_ω) is a finite union of closed intervals: $\Sigma(L) = [\lambda_0,\lambda_1] \cup [\lambda_2,\lambda_3] \cup \cdots \cup [\lambda_{2g},\infty)$ $(\lambda_0 > 0)$.

(H2) The Lyapunov exponent vanishes almost everywhere in the spectrum: $\beta(\lambda) = 0$ for Lebesgue a.a. $\lambda \in \Sigma(L)$.

We claim that the hypothesis that $\lambda_0 > 0$ is not restrictive, i.e., that if $\lambda_0 \leqslant 0$, then we can "translate" Eqs. (SL_ω) in such a way that the translated spectrum has $\tilde{\lambda}_0 > 0$, without any loss of regularity of the functions involved. For, suppose that $\lambda_0 \leqslant 0$, and choose $\eta > 0$. It is easy to check that the "translated" family of equations

$$(p\varphi')' - \big[q - (\lambda_0 - \eta)y\big]\varphi = -\lambda y\varphi \qquad (5.2^\eta_\omega)$$

has a spectrum Σ_η which equals the set $\Sigma(L) - \lambda_0 + \eta$. Moreover, the family (5.2^η_ω) satisfies Hypotheses 4.38. See ([95], Section 4) for more details concerning this fact.

Let $\mathcal{R}$ be the Riemann surface of the algebraic relation

$$w^2 = -(\lambda - \lambda_0)(\lambda - \lambda_1)\cdots(\lambda - \lambda_{2g}).$$

The surface $\mathcal{R}$ can be viewed as the union of two copies of the Riemann sphere $\mathbb{C} \cup \{\infty\}$ cut open along $\Sigma(L)$ and glued together in the standard way. Alternatively, $\mathcal{R}$ can be viewed as an analytic g torus, with holes obtained by cutting the Riemann sphere along (closed) paths c_i issuing from λ_{2i-1} $(i = 1, \dots, g)$ such that $\lambda_{2i} \in c_i$ for all $i = 1, \dots, g$. Let $\pi : \mathcal{R} \to \mathbb{C} \cup \{\infty\}$ be the projection of $\mathcal{R}$ to the Riemann sphere. Then π is 2-1 except at the ramification (or branch) points $\{\lambda_0, \lambda_1, \dots, \lambda_{2g}, \infty\}$, where it is 1-1. We abuse notation slightly, and write $\lambda_1, \lambda_2, \dots, \lambda_{2g}, \infty \in \mathcal{R}$. Let $\pi^{-1}(0) = \{0^+, 0^-\}$. Define a meromorphic function k on $\mathcal{R}$ by letting $k(0^+)$ be the positive square root of $\prod_{i=0}^{2g} \lambda_i$ and letting $k(P)$ be the appropriate square root of $-\prod_{i=0}^{2g}(\pi(P) - \lambda_i)$ for each $P \in \mathcal{R}$.

Let $\omega \in \Omega$, and define a function M_ω on $\mathcal{R} \setminus \{\lambda_0, \lambda_1, \dots, \lambda_{2g}\}$ as follows: put $M_\omega(0^+) = m_+(\omega, 0)$, $M_\omega(0^-) = m_-(\omega, 0)$. We define $M_\omega(P)$ by continuation along a curve in $\mathcal{R}$ joining 0^+ to P, for each $P \in \mathcal{R}$. Theorem 4.37 ensures that the function M_ω on $\mathcal{R} \setminus \{\lambda_0, \lambda_1, \dots, \lambda_{2g}\}$ defined in this way is meromorphic in a neighborhood of every point $P \in \mathcal{R}$ which is not a ramification point of $\mathcal{R}$.

Let us prove the following important

THEOREM 4.39. *Suppose that Hypotheses* (H1) *and* (H2) *hold, and let* $\omega \in \Omega$. *Then* $m_+(\omega, \cdot)$ *and* $m_-(\omega, \cdot)$ *define a single meromorphic function* $M_\omega : \mathcal{R} \to \mathbb{C} \cup \{\infty\}$. *Moreover, the map* $M : \Omega \times \mathcal{R} \to \mathbb{C} \cup \{\infty\} : (\omega, P) \mapsto M_\omega(P)$ *is jointly continuous.*

PROOF. To extend M_ω to a meromorphic function defined on all $\mathcal{R}$, we examine the behavior of the Weyl m-functions near a ramification point $\{\lambda_0, \lambda_1, \dots, \lambda_{2g}, \infty\}$. So, let $\lambda_* \in \{\lambda_0, \lambda_1, \dots, \lambda_{2g}\}$. Let $\omega \in \Omega$, and let $D_* \subset \mathbb{C}$ be a punctured disc centered at λ_* such that D_* has diameter equal to the union of an interval $I_* \subset \mathbb{R} \setminus \Sigma(L)$ and an interval contained in $\Sigma(L)$. If $m_+(\omega, \cdot)$ is continued around a circle in D_* centered in λ_*, it goes into $m_-(\omega, \cdot)$, and vice-versa. If we introduce the coordinate $z = \sqrt{\lambda - \lambda_*}$, then the inverse image of D_* in the z-plane is a punctured disc $D_\#$ centered at $z = 0$, and $m_\pm(\omega, \cdot)$ are branches of a meromorphic function M_ω in $D_\#$. Shrinking the disc $D_\#$ if necessary, we see that M_ω omits infinitely many (real) values in $D_\#$. By the Picard theorem [64], M_ω is meromorphic on the full disc $D_\# \cup \{0\}$. For the ramification point ∞, let D_∞ be a disc in the Riemann sphere centered at ∞ such that $D_\infty \cap \mathbb{R} \subset (-\infty, \lambda_0) \cup (\lambda_{2g}, \infty)$. Arguing as above, we observe that M_ω omits infinitely many real values in the punctured disc $\pi^{-1}(D_\infty)$, and applying the Picard theorem again, we conclude that M_ω extends meromorphically at ∞. Moreover, the bundles $\operatorname{Im} P_\omega$ and $\operatorname{Ker} P_\omega$ of the exponential dichotomy vary continuously with respect to $\omega \in \Omega$, ensuring that the map $M : \Omega \times \mathcal{R} \to \mathbb{C} \cup \{\infty\}$ is jointly continuous. ☐

Let $(\Omega, \{\tau_t\})$ be a topological flow, and let ν be a $\{\tau_t\}$-ergodic measure on Ω. We suppose that the topological support of ν equals Ω. Let p, q, and $y : \Omega \to \mathbb{R}$ be continuous functions such that p and y assume only positive values. We assume that the derivative $\omega \mapsto \frac{\mathrm{d}}{\mathrm{d}t} p(\omega \cdot t)|_{t=0}$ exists and is continuous. Again a word concerning the way we let the functions involved vary with $t \in \mathbb{R}$: if we fix $\omega \in \Omega$, we can consider p, q, y as functions of t: for instance, $p(t)$ is identified with $p(\omega \cdot t)$, and similarly for q, y. Suppose that Hypotheses 4.38 hold for these data.

Our goal is to determine all such functions p, q, y. Some of the following material follows closely the exposition in a recent paper in the journal Discrete and Continuous Dynamical Systems A **18**; see reference [96] below. We thank the editor Prof. Shouchuan Hu of that journal for permission to present this material here. Let $\omega \in \Omega$, $\lambda \in \mathbb{C} \setminus \Sigma(L)$, and let $m_\pm(t) = m_\pm(\omega \cdot t, \lambda)$ be the corresponding m-functions. They satisfy the Riccati equation

$$m' + \frac{1}{p}m^2 = q - \lambda y.$$

It is understood that this equation holds for each $t \in \mathbb{R}$.

We change the meaning of z again, and let $z^2 = -\lambda$ be a parameter near $\lambda = \infty$ on the Riemann surface $\mathcal{R}$. We view points $P \in \mathcal{R}$ near ∞ as being parameterized by z. Write $M'_\omega(z)$ for the derivative of the function $t \mapsto M_{\omega\cdot t}(z)$ in $t = 0$: thus $M'_\omega(z) = \frac{\mathrm{d}}{\mathrm{d}t}M_{\omega\cdot t}(z)|_{t=0}$ for each $\omega \in \Omega$. Then

$$M'_\omega(z) + \frac{1}{p}M^2_\omega(z) = q + z^2 y. \tag{4.65}$$

We expand the function $z \mapsto M_\omega(z)$ in a Laurent series near $z = \infty$, obtaining

$$M_\omega(z) = \alpha_1 z + \alpha_0 + \sum_{n=1}^{\infty} \alpha_{-n} z^{-n}.$$

Using the Riccati equation (4.65) one finds

$$\alpha_1^2 = yp, \quad \alpha_0 = -\frac{(py)'}{4y}, \quad \alpha_0' + \frac{\alpha_0^2}{p} + \frac{2\alpha_1\alpha_{-1}}{p} = q. \tag{4.66}$$

In particular, the above derivatives exist; in fact, by using the Cauchy integral formula along a small circle centered at $z = \infty$, one sees that α_{-1} exists and is continuous in $\omega \in \Omega$. It follows that the derivative α_0' exists and is continuous in $\omega \in \Omega$, so the same holds for y, and moreover py has second derivative which is continuous in $\omega \in \Omega$.

Let $\sigma : \mathcal{R} \to \mathcal{R}$ be the hyperelliptic involution $(\lambda, w) \mapsto (\lambda, -w)$ (i.e. the map which interchanges the sheets). It is sometimes convenient to think of M_ω as "the meromorphic extension of $m_+(\omega, \cdot)$ to $\mathcal{R}$". From this point of view, one has $m_-(\omega, \cdot) = M_\omega \circ \sigma(\cdot)$. We then have

$$\begin{aligned} M_\omega(z) = m_+(z) &= \sqrt{py}\, z - \frac{(py)'}{4y} + \sum_{n=1}^{\infty} \alpha_{-n} z^{-n}, \\ M_\omega(z) \circ \sigma(z) = m_-(z) &= -\sqrt{py}\, z - \frac{(py)'}{4y} + \sum_{n=1}^{\infty} (-1)^n \alpha_{-n} z^{-n}. \end{aligned} \tag{4.67}$$

We adopt the convention that $\sqrt{py} < 0$ if g is even, while $\sqrt{py} > 0$ if g is odd. The derivatives in (4.66) and (4.67) are calculated at $t = 0$. Thus, for example, $(py)'$ means $\frac{\mathrm{d}}{\mathrm{d}t}(py)(\omega \cdot t)|_{t=0}$. Let us now write P_i $(1 \leqslant i \leqslant g)$ for the finite poles of M_ω. Note that these poles P_i depend on $\omega \in \Omega$.

We know that $\pi(P_i) \in [\lambda_{2i-1}, \lambda_{2i}]$ for each $i \leqslant i \leqslant g$. Our aim in the next lines is to obtain an explicit formula for M_ω in terms of the poles and of p, q, y.

Let P denote a generic point of the Riemann surface $\mathcal{R}$. Abbreviating $m_\pm(\omega, P)$ to $m_\pm(P)$, we first note that $m_+ - m_- = M_\omega - M_\omega \circ \sigma$ is meromorphic on $\mathcal{R}$ and has zeroes $\{\lambda_0, \dots, \lambda_{2g}\} \subset \mathcal{R}$. It has poles at the inverse images $\pi^{-1}\pi(P_i)$ $(1 \leqslant i \leqslant g)$ and at ∞. Taking account of (4.67), we have

$$(m_+ - m_-)(P) = \frac{2\sqrt{py}\,k(P)}{H(\lambda)}, \tag{4.68}$$

where

$$H(\lambda) = \prod_{i=1}^{g} \bigl[\lambda - \pi(P_i)\bigr],$$

and $k(P)$ is defined by the conditions

$$k(P) = \sqrt{-\bigl(\pi(P) - \lambda_0\bigr)\cdots\bigl(\pi(P) - \lambda_{2g}\bigr)}, \quad k\bigl(0^+\bigr) > 0. \tag{4.69}$$

Next, for each fixed $\omega \in \Omega$, the function $m_+ + m_-$ is meromorphic on the Riemann sphere; i.e., it is a rational function. It has poles at $\pi(P_1), \dots, \pi(P_g)$ and moreover tends to $2\alpha_0$ as $\lambda \to \infty$. This implies that

$$(m_+ + m_-)(P) = \frac{2Q(\lambda)}{H(\lambda)}$$

where $\lambda = \pi(P)$ and $Q(\lambda) = q_g\lambda^g + \cdots + q_1\lambda + q_0$ is a polynomial of degree $\leqslant g$ in λ. It is of degree g whenever $\alpha_0 \neq 0$. Putting together all this information we obtain

$$M_\omega(P) = m_+(\omega, P) = \frac{Q(\lambda) + \sqrt{py}\,k(P)}{H(\lambda)},$$
$$M_\omega \circ \sigma(P) = m_-(\omega, P) = \frac{Q(\lambda) - \sqrt{py}\,k(P)}{H(\lambda)}.$$

Here M_ω is viewed as a meromorphic function of the generic point $P \in \mathcal{R}$. The polynomials Q and H are functions of $\lambda = \pi(P)$, and so can also be viewed as functions on $\mathcal{R}$, for each fixed $\omega \in \Omega$.

Let us now suppose for a moment that none of the poles are ramification points of $\mathcal{R}$. Then $m_-(P_i)$ is finite for each $1 \leqslant i \leqslant g$ and so we must have

$$Q(P_i) = \sqrt{py}\,k(P_i) \quad 1 \leqslant i \leqslant g. \tag{4.70}$$

The above formulas retain validity even if one or more poles P_i are ramification points of $\mathcal{R}$: in fact, if this were not true, then such a point P_i would not be a simple pole for M_ω

(to see this, suppose that $Q(P_i) \neq \sqrt{py}\,k(P_i)$, then compute the derivative of $M_\omega(\cdot)$ with respect to λ at $\lambda = P_i$, and use the local coordinate $z^2 = \lambda - P_i$.)

Equation (4.70) provides g linear relations for the $g+1$ coefficients of Q. To get another relation, note that $\lambda = 0$ is to the left of $\Sigma(L)$. We point out again the fact that we may assume without any loss of generality that $\lambda_0 > 0$. Hence $m_+(\omega, 0) = m_+^0$ and $m_-(\omega, 0) = m_-^0$ have real values. We therefore have

$$q_0 = Q(0) = \frac{(-1)^g [m_+^0 + m_-^0] \prod_{i=1}^g \pi(P_i)}{2}.$$

Now we can determine the coefficients of Q from the following van der Monde system:

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & \pi(P_1) & \dots & \pi(P_1)^g \\ \vdots & \vdots & & \vdots \\ 1 & \pi(P_g) & \dots & \pi(P_g)^g \end{pmatrix} \begin{pmatrix} q_0 \\ q_1 \\ \vdots \\ q_g \end{pmatrix} = \begin{pmatrix} Q(0) \\ \sqrt{py}\,k(P_1) \\ \vdots \\ \sqrt{py}\,k(P_g) \end{pmatrix}. \tag{4.71}$$

From now on, we will often abuse the notation, and write P_i instead of $\pi(P_i)$ ($i = 1, \dots, g$), when no confusion arises (i.e., we will omit π). In particular, we can determine the highest-order coefficient $q_g = \alpha_0 = -\frac{(py)'}{4y}$. In fact, letting Δ be the determinant of the matrix on the right-hand side of (4.71), we obtain

$$\begin{aligned} q_g \Delta = {} & \det \begin{pmatrix} P_1 & P_1^2 & \dots & P_1^{g-1} & \sqrt{py}\,k(P_1) \\ \vdots & \vdots & & \vdots & \vdots \\ P_g & P_g^2 & \dots & P_g^{g-1} & \sqrt{py}\,k(P_g) \end{pmatrix} \\ & + (-1)^g q_0 \det \begin{pmatrix} P_1 & P_1^2 & \dots & P_1^{g-1} \\ \vdots & \vdots & & \vdots \\ P_g & P_g^2 & \dots & P_g^{g-1} \end{pmatrix}. \end{aligned}$$

Working out the determinants, we get

$$q_g = \frac{(-1)^g q_0}{\prod_{i=1}^g P_i} + \sum_{i=1}^g \frac{\sqrt{py}\,k(P_i)}{P_i \prod_{r \neq i} (P_i - P_r)}. \tag{4.72}$$

Now we obtain a formula for α_{-1}. The Laurent expansions (4.67) for $m_\pm$ together with (4.68) give

$$2\sqrt{py}\,z + \frac{2\alpha_{-1}}{z} + 2\sum_{k=1}^{\infty} \frac{\alpha_{-2k-1}}{z^{2k+1}} = \frac{2\sqrt{py}\,k(z)}{\prod_{i=1}^g (-z^2 - P_i)}.$$

Multiplying both sides of this relation by $z\prod_{i=1}^{g}(z^2+P_i)$ and comparing the z^{2g}-coefficients , we obtain

$$\alpha_{-1}=\sqrt{py}\left(\frac{1}{2}\sum_{i=0}^{2g}\lambda_i-\sum_{i=1}^{g}P_i\right).$$

Let us now fix $\omega\in\Omega$. Now we will "follow the trajectory $\{\omega\cdot t\}$ through ω". As already remarked, this allows us to think at the functions involved as varying with $t\in\mathbb{R}$. Thus write $m_\pm(t,z)=m_\pm(\omega\cdot t,z)$, $p=p(t), q=q(t), y=y(t), P_i=P_i(t)\ (-\infty<t<\infty)$, $Q=Q(\lambda,t), H=H(\lambda,t)$. Using the formula for q in (4.66) we obtain

$$q(t)=y(t)\left(\sum_{i=0}^{2g}\lambda_i-2\sum_{i=1}^{g}P_i(t)\right)+q_g'(t)+\frac{q_g(t)^2}{p(t)}. \tag{4.73}$$

Here q_g is given by (4.72).

Now, consider

$$m_+'=\left(\frac{\sqrt{p(t)y(t)}\,k(P)+Q(\lambda,t)}{H(\lambda,t)}\right)',\qquad '=\frac{\mathrm{d}}{\mathrm{d}t}.$$

From the Riccati type equation for m_+ we obtain

$$m_+'=\left(\frac{\sqrt{p(t)y(t)}\,k(P)+Q(\lambda,t)}{H(\lambda,t)}\right)'=-\frac{1}{p}m_+^2+q-\lambda y,$$

hence,

$$m_+'=\frac{(\sqrt{py}\,k+Q)'H-H'(\sqrt{py}\,k+Q)}{H^2}=-\frac{1}{p}\frac{(\sqrt{py}\,k+q)^2}{H^2}+q-\lambda y,$$

and multiplying by H^2 we obtain

$$(\sqrt{py}\,k+Q)'H-(\sqrt{py}\,k+Q)H'+\frac{1}{p}(\sqrt{py}\,k+Q)^2=(q-\lambda y)H^2. \tag{4.74}$$

Let us write $k_r=k(P_r)$. If we compute (4.74) at a pole P_r we obtain, since $H(P_r)=H_r=0$ and $\sqrt{py}\,k(P_r)=\sqrt{py}\,k_r=Q(\pi(P_r))$,

$$-(\sqrt{py}\,k_r)H_r'+\frac{1}{p}(2\sqrt{py}\,k_r)^2=0, \tag{4.75}$$

hence we obtain the following important relation

$$\frac{H_r'(t)}{k_r(t)}=\frac{2}{p(t)}\sqrt{p(t)y(t)}. \tag{4.76}$$

What we want to do now is to express the above fraction (4.76) as a product of the poles $P_1(t), \dots, P_g(t)$ of the function $M_{\omega\cdot t}(P)$ To this aim, we compute m_- at $\lambda = 0$, and we have

$$m_-^0(t) = \frac{Q(0,t) - \sqrt{py}\,k(0^+)}{(-1)^g \prod_{i=1}^g P_i(t)},$$

hence

$$\sqrt{py} = \frac{(-1)^{g+1} m_-^0(t) \prod_{i=1}^g P_i(t) + Q(0,t)}{k(0^+)},$$

and (4.76) becomes

$$\frac{H_r'}{k_r} = \frac{2(-1)^{g+1} m_-^0(t) \prod_{i=1}^g P_i(t) + 2Q(0,t)}{p(t)k(0^+)}. \tag{4.77}$$

Hence we have

$$\frac{H_r'(t)}{k_r(t)} = \frac{(-1)^{g+1}[m_-^0(t) - m_+^0(t)] \prod_{i=1}^g P_i(t)}{p(t)k(0^+)}. \tag{4.78}$$

Finally

$$\frac{H_r'(t)}{P_r(t)k_r(t)} = \frac{(-1)^{g+1}[m_-^0(t) - m_+^0(t)] \prod_{s\neq r} P_s(t)}{p(t)k(0^+)}. \tag{4.79}$$

We proved the following

THEOREM 4.40. *Suppose that Hypotheses* 4.38 *hold. Then the functions p, q and y satisfy the following relations involving the t-motion of the poles of the meromorphic function $M_\omega(P)$ on $\mathcal{R}$,*

$$\frac{2\sqrt{p(t)y(t)}}{p(t)} = \frac{H_r'(t)}{k_r(t)} = \frac{(-1)^{g+1}[m_-^0(t) - m_+^0(t)]\prod_{i=1}^g P_i(t)}{p(t)k(0^+)},$$

and

$$q(t) = y(t)\left(\sum_{i=0}^{2g} \lambda_i - 2\sum_{i=1}^{g} P_i(t)\right) + q_g'(t) + \frac{q_g^2(t)}{p(t)},$$

where

$$q_g(t) = -\frac{(p(t)y(t))'}{4y(t)} = \frac{(-1)^g Q(0,t)}{\prod_{i=1}^g P_i(t)} + \sum_{j=1}^{g} \frac{\sqrt{p(t)y(t)}k(P_j(t))}{P_j(t)\prod_{r\neq j}(P_j(t)-P_r(t))}$$
$$= -\frac{p(t)}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\{\ln\left(\left[m_-^0(t) - m_+^0(t)\right]\prod_{i=1}^{g} P_i(t)\right)\right\}.$$

Now we study the nature of the t-dependence of p, q, y. Clearly this issue is related to that of the pole motion $t \mapsto \{P_1(t), \dots, P_g(t)\}$. We will see that a good deal of information about the pole motion can be obtained by introducing certain "nonstandard" Abel–Jacobi coordinates. In what follows, we will use the language of classical algebraic geometry. We will not pause to give much information concerning general facts of the theory of algebraic curves. We will only give a general outline of the classical theory, and spend a few more lines in the explanation of the main concepts occurring in the theory of the "singular Riemann surfaces". For more details, the reader is referred to such standard texts as [146,112]. For the theory of generalized Jacobian varieties, see [53].

Let $\mathcal{R}$ be the Riemann surface of the algebraic relation

$$w^2 - (\lambda - \lambda_0)\cdots(\lambda - \lambda_{2g}).$$

It was already pointed out that $\mathcal{R}$ is homeomorphic to an analytic g-torus. Let $\{\alpha_i, \beta_i\}$ ($i = 1, \dots, g$) be a homology basis of $\mathcal{R}$. There is one simple standard way to take such a basis: all α_i, β_i ($i = 1, \dots, g$) consist of a simple closed curves satisfying the intersection relations

$$\begin{aligned} &\alpha_i \cap \alpha_j = \beta_i \cap \beta_j = \emptyset, \quad 1 \leqslant i, j \leqslant g, \\ &\alpha_i \cap \beta_j = \begin{cases} \{A_i\} & i = j, \\ \emptyset, & i \neq j \end{cases} \end{aligned} \tag{4.80}$$

where $\{A_i\}$ denotes a point in $\mathcal{R}$ which varies with i, for every $i = 1, \dots, g$. Let $w_1, \dots, w_g$ be a standard basis of normalized differential forms of the first kind on $\mathcal{R}$: to be clearer, $w_1, \dots, w_g$ are holomorphic differentials such that

$$\int_{\alpha_i} w_j = \delta_{ij} \quad \text{and} \quad \int_{\beta_i} w_j = z_{ij},$$

where $z_{ij} \in \mathbb{C}$. The $g \times g$ matrix $Z = (z_{ij})$ is called the *period matrix* of the surface $\mathcal{R}$. This matrix is symmetric and has positive imaginary part [146]. It turns out that the $\mathbb{Z}$-lattice in $\mathbb{C}^g$ spanned by all the vectors of the form

$$\int_c (w_1, \dots, w_g), \quad c = \alpha_1, \dots, \alpha_g, \beta_1, \dots, \beta_g$$

is completely determined once the matrix $P = \binom{I}{Z}$ is given, where I is the $g \times g$ identity matrix. We denote this $\mathbb{Z}$-lattice with Λ^{2g}. Clearly the rank of Λ^{2g} is $2g$. We define the Jacobian variety as

$$J(\mathcal{R}) = \mathbb{C}^g / \Lambda^{2g}.$$

There is another way to describe the Jacobian variety. Any formal sum $\mathcal{D} = \sum_{i=1}^{p} k_i Q_i$, where each $k_i \in \mathbb{Z}$, and Q_i are points in $\mathcal{R}$, is called a *divisor*. The sum of two divisors $\mathcal{D} = \sum_{i=1}^{p} k_i Q_i$ and $\mathcal{D}_1 = \sum_{i=1}^{s} n_i R_i$ is the divisor

$$\mathcal{D} + \mathcal{D}_1 = \sum_{i=1}^{p} k_i Q_i + \sum_{i=1}^{s} n_i R_i.$$

The degree of a divisor $\mathcal{D} = \sum_{i=1}^{p} k_i Q_i$ is the integer $\sum_{i=1}^{p} k_i$. With the usual rules of the summation, the set of all the divisors of the Riemann surface $\mathcal{R}$ is an Abelian additive group. The inverse of a divisor $\mathcal{D} = \sum_{i=1}^{p} k_i Q_i$ is the divisor $-\mathcal{D} = \sum_{i=1}^{p} (-k_i) Q_i$.

Let $f : \mathcal{R} \to \mathbb{C}$ be any meromorphic function. Then f has the same number of zeroes and poles. If $Q_1, \dots, Q_p$ are the zeroes of f and $R_1, \dots, R_p$ are the poles of f (all counted with their multiplicities), then we can associate to f a divisor (f) of degree 0 defined as

$$(f) = \sum_{i=1}^{p} (k_i Q_i - n_i R_i) \quad \left(\sum_{i=1}^{p} (k_i - n_i) = 0 \right)$$

where $k_i, n_i \in \mathbb{Z}^+$ are the multiplicities of the points Q_i, R_i as zeroes or poles of f $(i = i, \dots, p)$. If f and g are two meromorphic functions on $\mathcal{R}$, then $(fg) = (f) + (g)$, $(fg^{-1}) = (f) - (g)$, and $(f) = (g)$ implies $f = cg$, where $c \in \mathbb{C}$ is a constant.

Every divisor of degree g can be thought as an element of $\mathrm{Symm}^g(\mathcal{R})$, the set of unordered g-tuples of points in $\mathcal{R}$. Given any divisor $\mathcal{D}$ of degree g (note that g is the genus of the surface $\mathcal{R}$) we form its class $[\mathcal{D}]$ consisting of all the divisors $\mathcal{D}_1$ of the same degree g such that $\mathcal{D} - \mathcal{D}_1$ is the divisor of a meromorphic function $f : \mathcal{R} \to \mathbb{C}$. Now, let P_* be any fixed point on $\mathcal{R}$ different from the ramification points. It is a classical fact that two divisors $\mathcal{D} = (Q_1, \dots, Q_g)$ and $\mathcal{D}_1 = (R_1, \dots, R_g)$ belong to the same class if and only if

$$\sum_{i=1}^{g} \int_{Q_i}^{R_i} (w_1, \dots, w_g) = Zt,$$

where t is a column vector in $\mathbb{Z}^g$.

To any divisor $\mathcal{D} = (Q_1, \dots, Q_g) \in \mathrm{Symm}^g(\mathcal{R})$ we associate the complex g-vector $(s_1, \dots, s_g) \in \mathbb{C}^g$, given by

$$(s_1, \dots, s_g) = \sum_{i=1}^{g} \int_{P_*}^{Q_i} (w_1, \dots, w_g).$$

We define a map $I : \mathrm{Symm}^g(\mathcal{R}) \to \mathbb{C}^g/\Lambda^{2g}$ as follows

$$I(Q_1, \dots, Q_g) = \sum_{i=1}^{g} \int_{P_*}^{Q_i} (w_1, \dots, w_g) \mod \Lambda^{2g} \in J(\mathcal{R}).$$

This map is called the *Abel map*. It is surjective and, when restricted to an appropriate open dense subset of $\mathrm{Symm}^g(\mathcal{R})$, is bijective. It is clear that $I(Q_1, \dots, Q_g) = I(R_1, \dots, R_g)$ if and only if $(Q_1, \dots, Q_g)$ and $(R_1, \dots, R_g)$ belong to the same class. It is easy to check that $J(\mathcal{R})$ can be identified with the set of the classes $[\mathcal{D}]$ of the divisors $\mathcal{D}$ on $\mathcal{R}$ of degree g. Equivalently, it can be proved that $J(\mathcal{R})$ can be identified as the quotient space of the set of all the divisors on $\mathcal{R}$ of degree 0, modulo the set of the divisors of meromorphic functions on $\mathcal{R}$ (see [146] for more details concerning this last remark).

Now we define the Riemann theta function. Let Z be any g-period matrix (i.e., any symmetric complex $g \times g$ matrix with positive imaginary part). Define $\Theta : \mathbb{C}^g \to \mathbb{C}$ as follows:

$$\Theta(s_1, \dots, s_g) = \Theta(s) = \sum_{t \in \mathbb{Z}^g} \exp\bigl\{\pi \langle t, Zt \rangle + 2\pi \mathrm{i} \langle t, s \rangle\bigr\},$$

where $(s_1, \dots, s_g) \in \mathbb{C}^g$. The function Θ has many important properties; we list some of them:

(i) $\Theta(-s) = \Theta(s)$;
(ii) $\Theta(s + t) = \Theta(s)$, for every integer column $t \in \mathbb{Z}^g$;
(iii) $\Theta(s + Zt) = \exp\{-\pi \langle t, Zt \rangle - 2\pi \mathrm{i} \langle t, s \rangle\}\Theta(s)$, for every column vector $t \in \mathbb{Z}^g$.

It was a Riemann's fundamental idea to compute the zeroes of the function Θ. To this aim, we define the so-called *vector of Riemann constants* $\Delta = (\Delta_1, \dots, \Delta_g) \in \mathbb{C}^g$, where

$$\Delta_k = -\frac{1}{2} z_{kk} + \sum_{j \neq k}^{g} \int_{\alpha_j} \left(\int_{P_*}^{P} w_k \right) w_j \quad (k = 1, \dots, g). \tag{4.81}$$

In order to simplify the notation, if $P \in \mathcal{R}$, we will write

$$w(P) = \int_{P_*}^{P} (w_1, \dots, w_g) \in \mathbb{C}^g.$$

THEOREM 4.41. *Let $s \in \mathbb{C}^g$ be fixed. If the map $F : P \mapsto \Theta(w(P) - s + \Delta)$ does not vanish identically, then it has exactly g zeroes $Q_1, \dots, Q_g \in \mathcal{R}$ such that*

$$s = \sum_{i=1}^{g} \int_{P_*}^{Q_i} (w_1, \dots, w_g) \mod \Lambda^{2g}.$$

Theorem 4.41 provides an answer to the Riemann inversion problem, which consists in associating to each $s \in \mathbb{C}^g$ a divisor $Q_1, \dots, Q_g$ such that $I(Q_1, \dots, Q_g) = s \bmod \Lambda^{2g}$.

In fact, when restricted to $J(\mathcal{R})$, the equation $\Theta(s, Z) = 0$ defines the 1-codimensional analytic subvariety Υ_g of the zero-locus of Θ, called the *Θ-divisor*. It can be shown that Υ_g admits a parametrization

$$\Upsilon_g = \left\{ s \in \mathbb{C}^g \;\middle|\; s = \sum_{j=1}^{g-1} \int_{P_*}^{X_j} (w_1, \dots, w_g) - \Delta \bmod \Lambda^{2g} \right\}, \tag{4.82}$$

where X_j $(j = 1, \dots, g-1)$ are arbitrarily chosen points on $\mathcal{R}$ which are different from the ramification points. If, for some $s \in \mathbb{C}^g$, the map $F : P \mapsto \Theta(w(P) - s + \Delta) \equiv 0$, then $s \in \Upsilon_g$ and there exist points $X_1, \dots, X_{g-1} \in \mathcal{R}$ satisfying (4.82). The g-tuple of points $P_*, X_1, \dots, X_g$ defines a divisor such that $I(P_*, X_1, \dots, X_{g-1}) = s \bmod \Lambda^{2g}$. See [146, Chapter 4, Vol. 2].

Now we go deeper in the examination of the zeroes of Θ. Let $A, B \in \mathcal{R}$, and consider the differential $w_{B,A}$ of the third kind on $\mathcal{R}$, having simple poles at A and B with residues 1 and -1 respectively. Let $s = I(Q_1, \dots, Q_g)$, where $Q_i \in \mathcal{R}$ $(i = 1, \dots, g)$ are different from the ramification points. Applying Theorem 4.41, we see that the zeroes of $F(P) = \Theta(w(P) - s + \Delta)$ are exactly the points Q_i $(i = 1, \dots, g)$. Next, we compute the sum of the residues of the differential $w_{B,A}\, \mathrm{d} \ln F(P)$, to obtain

$$\sum_{i=1}^{g} \int_{P_*}^{Q_i} w_{B,A} = -\frac{\Theta(s + \Delta - w(B))}{\Theta(s + \Delta - w(A))} + c.$$

Another important application of Theorem 4.41 and the parametrization (4.82) is the following fundamental theorem, which provides an answer to the following question: is it possible to express meromorphic functions defined on $\mathcal{R}$ in terms of functions defined on $J(\mathcal{R})$? The answer is positive: first, we use the map $w : P \mapsto \int_{P_*}^{P} (w_1, \dots, w_g)$ which sends the zeroes and the poles of f (which belong to $\mathcal{R}$) into complex g-vectors in $J(\mathcal{R})$, then we use the Riemann Θ-function to express f.

THEOREM 4.42. *Let f be a meromorphic function on $\mathcal{R}$. Let $\mathcal{D} = (f) = P_1 + \dots + P_m - Q_1 - \dots - Q_m$. Let $X_1, \dots, X_{g-1}$ be any $g-1$ points in $\mathcal{R}$ which are different from the ramification points of $\mathcal{R}$ and the points P_i, Q_i $(i = 1, \dots, g)$. Then*

$$f(P) = \delta \prod_{i=1}^{m} \frac{\Theta(w(P) - w(P_i) - \Delta + \sum_{j=1}^{g-1} w(X_j))}{\Theta(w(P) - w(Q_i) - \Delta + \sum_{j=1}^{g-1} w(X_j))},$$

where $\delta \in \mathbb{C}$ depends only on the choice of the points $X_1, \dots, X_g$.

An easy consequence of the above theorem is the following

THEOREM 4.43. *Let f be a meromorphic function defined on $\mathcal{R}$, with*

$$(f) = \sum_{i=1}^{m} (P_i - Q_i)$$

(*it is understood that the points* P_i *and* Q_i $(i = 1, \dots, m)$ *are the zeroes and the poles of* f *respectively*). *Let* $S_1, \dots, S_g$ *be* g *points on* $\mathcal{R}$ (*here some of the points* S_i *may coincide with some of the* P_i*'s or* Q_i*'s*). *Then*

$$\prod_{i=1}^{g} f(S_i) = \gamma \prod_{j=1}^{m} \frac{\Theta(I(S_1, \dots, S_g) - w(P_j) - \Delta)}{\Theta(I(S_1, \dots, S_g) - w(Q_j) - \Delta)},$$

where $\gamma \in \mathbb{C}$ *is a constant which does not depend on the choice of the points* $S_1, \dots, S_g$.

The theory we have just outlined is very useful if one has to face problems where only differentials of the first kind occur, but when we have to consider nonholomorphic differentials as well, things become a little more difficult.

We give a method, essentially due to Fay [53], to solve the Riemann inversion problem when in presence of one nonholomorphic differential. This method uses the concept of singular Riemann surface, which leads to the introduction of the generalized Jacobian variety, the generalized Abel map, and the generalized Riemann Θ-function as well. To begin the discussion, let D be the unit circle. We will denote points on D with z.

Following [53], we build a family $\mathfrak{F}$ of Riemann surfaces over D in the following way: let $A \neq B \in \mathcal{R}$. Let U_A and U_B be disjoint neighborhood of A and B respectively, and $z_A : U_A \to D$, $z_B : U_B \to D$ be local coordinates of A and B. Let

$$\begin{aligned} W = \big\{&(P, z) \mid z \in D,\ P \in \mathcal{R} \setminus (U_A \cup U_B) \text{ or } P \in U_A(U_B) \text{ and} \\ &\big|z_A(P)\big| > |z|\big(|z_B(P)| > |z|\big)\big\}. \end{aligned}$$

Let $\mathcal{S}$ be the surface

$$\mathcal{S}\colon \big\{XY = z \mid (X, Y, z) \in D^3\big\},$$

and define

$$\mathfrak{F} = W \cup \mathcal{S},$$

where the points $(P_A, z) \in (W \cap U_A) \times D$ are identified with $(z_A(P_A), z/z_A(P_A), z) \in \mathcal{S}$ and the points $(P_B, z) \in (W \cap U_B) \times D$ are identified with $(z_B(P_B), z/z_B(P_B), z) \in \mathcal{S}$ $(z \in D)$. We take

$$x = \frac{1}{2}(X + Y) \quad \text{and} \quad y = \frac{1}{2}(X - Y)$$

as coordinates on $\mathcal{S}$. For every fixed $z \in D$, the fiber $\mathcal{R}_z$ over z is a Riemann surface of genus $g + 1$ for which $\mathcal{R}_z \cap \mathcal{S}$ is the surface of $y = \sqrt{x^2 - z}$, having $\pm\sqrt{z}$ as ramification points. The fiber $\mathcal{R}_0$ is a curve of genus $g + 1$ with a double point corresponding to the identification of the points A and B: hence the branches of $\mathcal{R}_0 \cap \mathcal{S}$ are $y = x$ and $y = -x$, with local parameters $x = z_A/2$ and $x = z_B/2$ respectively. Now, choose a canonical basis $\alpha_1(z), \dots, \alpha_g(z), \alpha_{g+1}(z), \beta_1(z), \dots, \beta_g(z), \beta_{g+1}(z)$ for the homology group of

$\mathcal{R}_z$ as follows: $\alpha_1(z), \dots, \alpha_g(z), \beta_1(z), \dots, \beta_g(z)$ are the extensions on $\mathcal{R}_z$ of the canonical basis $\alpha_1, \dots, \alpha_g, \beta_1, \dots, \beta_g$ of the homology group of $\mathcal{R}$ (moreover, we may assume that $\alpha_i, \beta_i \in \mathcal{R} \setminus (U_A \cup U_B)$ $(i = 1, \dots, g)$): then define $\alpha_{g+1}(z) = \partial U_B \times \{z\}$ and, if $|z| < 1/4$, $\beta_{g+1}(z) = l \times \{z\} \cup l_{A,z} \cup l_{B,z} \subset W$, where l is a fixed path from $z_A^{-1}(1/2)$ to $z_B^{-1}(1/2)$ in the (canonically) dissected surface $\mathcal{R}$, $l_{A,z}$ and $l_{B,z}$ are paths from $z_A^{-1}(\sqrt{z})$ to $z_A^{-1}(1/2)$ and from $z_B^{-1}(1/2)$ to $z_B^{-1}(\sqrt{z})$ respectively, which vary continuously with respect to $z \in D$. This choice seems to be a little complicated, but its significance is clear if we note that $\beta_{g+1}(0)$ is simply a path from A to B lying in $\mathcal{R}$.

We have the following

PROPOSITION 4.44. *There are $g + 1$ linearly independent holomorphic 2-forms on $\mathfrak{F}$, having residues $w_1(z), \dots, w_{g+1}(z)$ on $\mathcal{R}_z$, for z in a sufficiently small disc D_ε centered in 0. These residues are a normalized basis of holomorphic differentials on $\mathcal{R}_z$, for $z \neq 0$. For $z = 0$, the differentials $w_1(0), \dots, w_g(0)$ form a normalized basis of holomorphic differentials on $\mathcal{R}$, and $w_{g+1}(0)$ is the normalized differential of the third kind $w_{B,A}$, having simple poles at A and B with residues $\pm 1/2\pi\mathrm{i}$ respectively.*

In particular,

$$w_i(z) = w_i + \zeta_i z\omega + \mathcal{O}(z^2), \quad i = 1, \dots, g,$$

where $\zeta_i \in \mathbb{C}$ is a constant $(i = 1, \dots, g)$, ω is the normalized differential of the second kind on $\mathcal{R}$ having double poles of residue zero at A and B, w_i is the i-th differential of the first kind in $\mathcal{R}$, and $\mathcal{O}(z^2)$ is an expression involving only the holomorphic differentials on $\mathcal{R}$.

Moreover

$$w_{g+1}(z) = w_{B,A} + z\tilde{\omega} + \mathcal{O}(z^2),$$

where $\tilde{\omega}$ is a differential of the second kind on $\mathcal{R}$ having as poles the points A and B, each of order 3.

PROOF. We use without explanations various concepts of algebraic geometry, which are amply discussed by Griffiths and Harris [65]. Let $\mathcal{W}_{\mathfrak{F}}$ and $\mathcal{W}_z$ be the sheafs of holomorphic 2-forms on $\mathfrak{F}$ and on $\mathcal{R}_z$ respectively. Under the residue map $r_z : \mathcal{W}_{\mathfrak{F}} \to \mathcal{W}_z$, the quotient space $\mathcal{W}_{\mathfrak{F}}/\mathcal{R}_z$ is the sheaf consisting, for $z = 0$, of the holomorphic differentials on $\mathcal{R}$, plus the differentials having simple poles at A and B with opposite residues. It turns out that, for every $z \in D$, the cohomology group $H^0(\mathcal{R}_z, \mathcal{W}_{\mathfrak{F}}/\mathcal{R}_z)$ has dimension $g + 1$. Using methods from Fay ([53], Proposition 3.1, page 37), this fact implies that there is a sufficiently small disc D_ε centered at $z = 0$, and differential 2-forms $W_1(z), \dots, W_{g+1}(z)$ on $\mathfrak{F}$, whose first g residues along $z = 0$ give a normalized basis $w_1, \dots, w_g$ of holomorphic differentials on $\mathcal{R}$. For $z = 0$, the $(g + 1)$-th residue gives the differential of the third kind $w_{B,A}$ on $\mathcal{R}$, having simple poles at A and B with residues $\pm 1/2\pi\mathrm{i}$ respectively. By continuity, for $z \in D_\varepsilon$, the residue map computed at the differential forms $W_i(z)$ $(i = 1, \dots, g + 1)$, provides a normalized basis $w_1(z), \dots, w_{g+1}(z)$ of differentials of the first kind on $\mathcal{R}_z$.

Now, let x be the coordinate in a neighborhood $U \subset \mathcal{R}_z \cap \mathcal{S}$ of the double point; then, for every $i = 1, \dots, g+1$, set

$$w_i(z) = \sum_{j=0}^{\infty} a_{i,j}(z) x^j \,\mathrm{d}x + \sum_{n=0}^{\infty} b_{i,n}(z) \frac{x^n}{\sqrt{x^2 - z}} \,\mathrm{d}x,$$

where $a_{i,j}(z)$ and $b_{i,n}(z)$ are holomorphic functions near $z = 0$ for every $j, n \in \mathbb{N}$ and $i = 1, \dots, g+1$. Then, for $i = 1, \dots, g$,

$$w_i = w_i(0) = \sum_{j=0}^{\infty} a_{i,j}(0) x^j \,\mathrm{d}x \pm \sum_{n=0}^{\infty} b_{i,n}(0) x^{n-1} \,\mathrm{d}x,$$

where the sign of the second summand on the right-hand side depends on whether we move in a neighborhood of A or B (where, in fact, the local coordinate is x or $-x$ respectively). If we let $w_i(z) = w_i(x, z) dx$, for every $i = 1, \dots, g+1$, then $b_{i,0}(0) = 0$, $w_i(A, 0) = a_{i,0}(0) + b_{i,1}(0)$, $w_i(B, 0) = a_{i,0}(0) - b_{i,1}(0)$, etc. Moreover

$$\lim_{z \to 0} \frac{w_i(z) - w_i}{z} = \sum_{j \geqslant 0} a'_{i,j}(0) x^j \,\mathrm{d}x \pm \sum_{n \geqslant 0} \left(\frac{1}{2} b_{i,n+1}(0) + x b'_{i,n}(0) \right) x^{n-2} \,\mathrm{d}x$$

is a differential of the second kind on $\mathcal{R}$ whose only double poles are A and B, with opposite principal parts. For $i = g+1$, we have $w_{g+1}(0) = w_{B,A}$, hence $b_{g+1,0}(0) = -1/2\pi \mathrm{i}$, and thus the limit

$$\begin{aligned} \lim_{z \to 0} \frac{w_{g+1}(z) - w_{B,A}}{z} &= \sum_{j \geqslant 0} a'_{g+1,j}(0) x^j \,\mathrm{d}x \\ &\quad \pm \sum_{n \geqslant 0} \left(\frac{b_{g+1,n}(0)}{2x} + x b'_{g+1,n}(0) \right) x^{n-2} \,\mathrm{d}x \end{aligned}$$

is a normalized differential of the second kind on $\mathcal{R}$, having only poles at A and B, each of order 3. The proof follows easily, if we put $\zeta_i = \frac{w_i(A,0) - w_i(B,0)}{4}$. □

The following theorem is fundamental.

THEOREM 4.45. *With the $g+1$-tuple of the normalized basis of holomorphic differentials on $\mathcal{R}_z$ ordered as $(w_{g+1}(z), w_1(z), \dots, w_g(z))$, the period matrix for the Riemann surface $\mathcal{R}_z$ is the $(g+1) \times (g+1)$ matrix*

$$Z_z = \begin{pmatrix} \alpha_i + z\sigma_{gi} & z_{ij} + z\sigma_{ij} \\ \ln z + c_1 + c_2 z & \alpha_j + z\sigma_{jg} \end{pmatrix}_{1 \leqslant i,j \leqslant g} + \mathcal{O}(z^2),$$

where c_1 and c_2 are constants, $\alpha_i = \int_A^B w_i$, σ_{ij} are functions involving the holomorphic differentials on $\mathcal{R}$ together with their derivatives, and $Z = (z_{ij})$ is the $g \times g$ period matrix for $\mathcal{R}$.

PROOF. The proof follows easily from Proposition 4.44, except for the entry $z_{g+1,1}$ of Z_z. However, this follows directly from Proposition 4.44 as well, since $z_{g+1,1} - \ln z$ is a well defined holomorphic function of the variable z in the punctured disc $D_\varepsilon \setminus \{0\}$, and it must be holomorphic in the disc D_ε as well, otherwise $\Im Z_z$ would not be positive definite as $z \to 0$. $\square$

Following Fay again [53], let Λ_z^{2g+2} be the $\mathbb{Z}$-lattice spanned by all vectors of the form

$$\int_c \bigl(w_{g+1}(z), w_1(z), \dots, w_g(z)\bigr) \quad \bigl(c = \alpha_1(z), \dots, \alpha_{g+1}(z), \beta_1(z), \dots, \beta_{g+1}(z)\bigr).$$

From the above theorem, one can easily prove that two vectors $s, s_1 \in \mathbb{C}^{g+1}$ differ by a vector in Λ_z^{2g+2} if and only if, as $z \to 0$, they differ by a vector in the $\mathbb{Z}$-lattice Λ_0^{2g+1} of rank $2g+1$, generated by the rows of the matrix $\left(\begin{smallmatrix} I \\ Z_0 \end{smallmatrix}\right)$, where I is the $(g+1) \times (g+1)$ identity matrix, $Z_0 = \left(\begin{smallmatrix} C & Z \\ 0 & 0 \end{smallmatrix}\right)$, and C is the transpose of the vector $\int_A^B (w_1, \dots, w_g)$. Equivalently, two vectors $s, s_1 \in \mathbb{C}^{g+1}$ satisfy $s - s_1 = \Lambda_z^{2g+2} t$ $(t \in \mathbb{Z}^{2g+2})$, if and only if there exists $t_0 \in \mathbb{Z}^{2g+1}$ such that $s - s_1 = \Lambda_0^{2g+1} t_0 + \mathcal{O}(z)$.

Now, we let $\mathcal{G}$ be the family of the $(g+1)$-dimensional manifolds over the disc D and denote by $\tilde{\pi}$ the canonical map $\tilde{\pi} : \mathcal{G} \to D$. If $z \in D \setminus \{0\}$, the fiber $\tilde{\pi}^{-1}(z)$ is the Jacobian variety $J(\mathcal{R}_z) = \mathbb{C}^{g+1}/\Lambda_z^{2g+2}$ of the Riemann surface $\mathcal{R}_z$, while $\tilde{\pi}^{-1}(0)$ is $J_0(\mathcal{R}) = \mathbb{C}^{g+1}/\Lambda_0^{2g+1}$.

In [72] it is shown that $J_0(\mathcal{R})$ is a manifold, and that the canonical map π is analytic. We want to describe $J_0(\mathcal{R})$.

Let P be the double point of $\mathcal{R}_0$; as already observed, P coincides with the identification of the points $A, B \in \mathcal{R}$. Every divisor $\mathcal{D} = \sum_{Q_i \neq P} n_i Q_i + nP$ on $\mathcal{R}_0$ can be lifted to a divisor $\mathcal{D} = \sum_{Q_i \neq P} n_i Q_i + n(A+B)$ on $\mathcal{R}$ (we abuse slightly the notation). Let $\mathcal{B} - \mathcal{D}$ is the divisor of a meromorphic function f on $\mathcal{R}_0$; if we lift $\mathcal{B} - \mathcal{D}$ on $\mathcal{R}$, then we must have $f(B)/f(A) = 1$, and since f is meromorphic on $\mathcal{R}$ as well, Abel's Theorem gives (see [53])

$$\frac{f(B)}{f(A)} = \exp\left\{\int_{\mathcal{D}}^{\mathcal{B}} v_{B,A} - \sum_{i=1}^{g} \int_A^B k_i w_i\right\} \in \mathbb{C}^*, \tag{4.83}$$

where

$$(k_1, \dots, k_g) = \int_{\mathcal{D}}^{\mathcal{B}} (w_1, \dots, w_g) \in \mathbb{C}^g,$$

and we used the notation $\int_{\mathcal{D}}^{\mathcal{B}} (\cdot) = \sum_{i=1}^{p} \int_{D_i}^{B_i} (\cdot)$. The relation (4.83) holds even if f has a zero or a pole at A: in fact, from the identification of A with B, f must have the same zero or pole at B. Now it is quite natural to associate to a divisor $\mathcal{B}-\mathcal{D}$ of degree 0 on $\mathcal{R}_0$, the equivalence class of the vector

$$\int_{\mathcal{D}}^{\mathcal{B}} (w_{B,A}, w_1, \dots, w_g) \in \mathbb{C}^{g+1} \quad \operatorname{mod} \Lambda_0^{2g+1},$$

and $J_0(\mathcal{R})$ is the group of the equivalence classes of the divisors of degree 0 on $\mathcal{R}_0$, with two divisors $\mathcal{D}$ and $\mathcal{D}'$ identified if $\mathcal{D}-\mathcal{D}'$ is the divisor of a meromorphic function on $\mathcal{R}_0$, i.e. a function satisfying $f(B)/f(A)=1$.

The exact sequence of groups

$$0 \longrightarrow \mathbb{C}^* \xrightarrow{\Phi} J_0(\mathcal{R}) \xrightarrow{\Psi} J(\mathcal{R}) \longrightarrow 0$$

where, for every $r \in \mathbb{C}^*$, $\Phi(r)$ is the class of the divisor of any meromorphic function on $\mathcal{R}$ with $f(B)/f(A)=r$, and Ψ is the projection $J_0(\mathcal{R}) \to J(\mathcal{R})$, gives the fundamental identification

$$J_0(\mathcal{R}) \cong \mathbb{C}^* \times J(\mathcal{R}).$$

We call $J_0(\mathcal{R})$ the *generalized Jacobian variety of* $\mathcal{R}$. If $\tilde{s} = (z_0, z_1, \dots, z_g) \in \mathbb{C}^{g+1}$ is a point in the universal cover of $J_0(\mathcal{R})$, then $\Phi(r)$ is the class of $(\ln r, 0, \dots, 0)$ (mod Λ_0^{2g+1}), and $\Psi(\tilde{s})$ is the class of $(z_1, \dots, z_g) \in \mathbb{C}^g$ in $J(\mathcal{R})$.

The generalized Jacobian $J_0(\mathcal{R})$ is the natural setting for the solution of the Riemann inversion problem when in the presence of a differential of the third kind, as stated in the following

THEOREM 4.46. *Let A and B be two distinct points on $\mathcal{R}$, and let $\mathcal{D}$ be a divisor of degree $g+1$ not containing A and B. Then for every $r \in \mathbb{C}^*$ and $z \in \mathbb{C}^g$, there is a unique divisor $\mathcal{B}$ of degree $g+1$ (not containing A and B) such that*

$$\int_{\mathcal{D}}^{\mathcal{B}} (w_{B,A}, w_1, \dots, w_g) = (\ln r, z_1, \dots, z_g) \in \mathbb{C}^{g+1}/\Lambda_0^{2g+1}. \tag{4.84}$$

Let P_* be a fixed point on $\mathcal{R}_0$ different from the ramification points and the double point of $\mathcal{R}_0$. The map $I_0 : \mathrm{Symm}^{g+1}(\mathcal{R}_0) \to J_0(\mathcal{R})$ defined by

$$I_0(P_0, \dots, P_g) = \sum_{i=0}^{g} \int_{P_*}^{P_i} (w_{B,A}, w_1, \dots, w_g) \quad \operatorname{mod} \Lambda_0^{2g+1}$$

is the generalized Abel map with base point P_*.

We now want to describe a method which uses Θ-functions to express meromorphic functions on the generalized Jacobian. The standard theory must be suitably adapted for our purpose.

Let $\tilde{s} = (z, s) \in \mathbb{C} \times \mathbb{C}^g$. The generalized Θ-function is defined as the map

$$\Theta_0 : \mathbb{C}^{g+1} \to \mathbb{C} : \tilde{s} \mapsto \mathrm{e}^{z/2}\Theta\left(s + \frac{1}{2}q\right) + \mathrm{e}^{-z/2}\Theta\left(s - \frac{1}{2}q\right), \tag{4.85}$$

where Θ is the Riemann Θ-function on $J(\mathcal{R})$, and

$$q = \int_A^B (w_1, \dots, w_g).$$

It is worth noting that Θ_0 has the property that if we shift the vector $\tilde{s}$ by a period vector in the lattice Λ_0^{2g+1} (i.e. by a vector $Z_0 t$, where t is a $\mathbb{Z}^{g+1}$-column), then Θ_0 is multiplied by a simple exponential.

Now, take a point $P \in \mathcal{R}$ different from the ramification points and the points A, B, then consider the function

$$F(P) = \Theta_0\bigl(\tilde{w}(P) - v\bigr),$$

where $v \in \mathbb{C}^{g+1}$,

$$\tilde{w}(P) = \int_{P_*}^P (w_{B,A}, w_1, \dots, w_g),$$

and P_* is any fixed base point. The following holds:

PROPOSITION 4.47. *If the function $F(P)$ does not vanish identically on $\mathcal{R}_0$, then it has exactly $g+1$ zeros* (*possibly with multiplicities*).

PROOF. As in the case of the classical Θ-function, we canonically dissect $\mathcal{R}_0$, and compute

$$\begin{aligned}&\frac{1}{2\pi\mathrm{i}}\int_{\partial\mathcal{R}_0} \mathrm{d}\ln F(P)\\ &\quad= \frac{1}{2\pi\mathrm{i}}\int_{\partial\mathcal{R}} \mathrm{d}\ln F(P) + \operatorname{Res}\mathrm{d}\ln F(B) + \operatorname{Res}\mathrm{d}\ln F(A).\end{aligned} \tag{4.86}$$

The first integral on the right side of (4.86) is exactly g, while, as a simple computation shows,

$$\operatorname{Res}\mathrm{d}\ln F(B) = \operatorname{Res}\mathrm{d}\ln F(A) = \frac{1}{2},$$

and the proof follows easily. □

The above proposition suggests what has to be done: we have to introduce a generalized Riemann vector $\Delta_0 \in \mathbb{C}^{g+1}$, and prove results analogous to those of the classical theory.

We define Δ_0 in the following way:

$$\Delta_0 = (\Delta_*, \Delta_1) = \left(\sum_{i=1}^{g} \int_{\alpha_i} \left(\int_{P_*}^{P} w_i \right) w_{B,A}, \Delta - \frac{w(B) + w(A)}{2} \right) \in \mathbb{C} \times \mathbb{C}^g,$$

where $\Delta \in \mathbb{C}^g$ is the Riemann vector in $J(\mathcal{R})$, whose coordinates are defined in (4.81). Note that Δ_0 depends only on the choice of the base point P_* and the canonical dissection of $\mathcal{R}_0$.

The natural extension of Theorem 4.41 to the setting of singular Riemann surfaces is the following (for a proof, see [53]):

THEOREM 4.48. *For every fixed* $\tilde{s} \in \mathbb{C}^{g+1}$, *if the function* $\Theta_0(\tilde{w}(P) - \tilde{s} + \Delta_0)$ *does not vanish identically, then it has exactly* $g + 1$ *zeros* $P_0, \dots, P_g$ *such that*

$$\tilde{s} = I_0(P_0, \dots, P_g).$$

The analytic noncompact subvariety Υ_0 of $J_0(\mathcal{R})$ defined by the equation $\Theta_0(\tilde{s}) = 0$ is the Θ-divisor of the generalized Jacobian variety. In analogy to the case of the Riemann surface $\mathcal{R}$, Υ_0 admits a parametrization which is very similar to that of Eq. (4.82), i.e.,

$$\Upsilon_0\colon \left\{ \tilde{s} = \sum_{j=1}^{g} \tilde{w}(X_j) - \Delta_0 \bmod \Lambda_0^{2g+1} \right\},$$

where X_j $(j = 1, \dots, g)$ are arbitrarily chosen points on $\mathcal{R}$ different from the ramification points and A, B. It is clear that the Θ-divisor is the set of the solutions of the equation

$$e^{z/2}\Theta\left(s + \frac{1}{2}q\right) - e^{-z/2}\Theta\left(s - \frac{1}{2}q\right) = 0. \tag{4.87}$$

It is possible to retrieve the following parametric form for points on the Θ-divisor:

$$\begin{cases} s = I(P_1, \dots, P_g) - \Delta_1, \\ z = -\ln \dfrac{\Theta(s - \Delta - w(A))}{\Theta(s - \Delta - w(B))} + c = -\ln \dfrac{\Theta(s + (w(B) - w(A))/2)}{\Theta(s + (w(A) - w(B))/2)} + c. \end{cases} \tag{4.88}$$

It is clear that (4.88) follows from (4.87), up to an additive constant. We can formulate a generalized version of Theorem 4.43, as follows.

THEOREM 4.49. *Let* f *be a meromorphic function on* $\mathcal{R}_0$ *such that* $(f) = \mathcal{D} - \mathcal{B}$, *where* $\mathcal{D} = Q_1 + \cdots + Q_m$ *and* $\mathcal{B} = R_1 + \cdots + R_m$ *are two divisors of the same degree* m. *If*

$X_1 + \cdots + X_g$ is a divisor of degree g such that every X_j is different from the points Q_i and R_i and the ramification points of $\mathcal{R}_0$, then

$$f(P) = \delta \prod_{i=1}^{m} \frac{\Theta_0(\tilde{w}(P) - \tilde{w}(Q_i) - \Delta_0 + \sum_{j=1}^{g} \tilde{w}(X_j))}{\Theta_0(\tilde{w}(P) - \tilde{w}(R_i) - \Delta_0 + \sum_{j=1}^{g} \tilde{w}(X_j))},$$

where δ is a constant which depends on the choice of the points $X_1, \dots, X_g \in \mathcal{R}$.

Now, using methods from [146] together with Theorems 4.48 and 4.49, we can prove a generalization of Theorem 4.43, i.e.

THEOREM 4.50. *Let $P_0, \dots, P_g$ be any $(g+1)$-tuple of points of $\mathcal{R}_0$ different from the ramification points and A, B. Let f be a meromorphic function on $\mathcal{R}_0$, and let $(f) = \mathcal{D} - \mathcal{B} = Q_1 + \cdots + Q_m - R_1 - \cdots - R_m$ be the divisor of f. Then*

$$f(P_0) f(P_1) \cdots f(P_g) = \gamma \prod_{i=1}^{m} \frac{\Theta_0(I_0(P_0, \dots, P_g) - \tilde{w}(Q_i) - \Delta_0)}{\Theta_0(I_0(P_0, \dots, P_g) - \tilde{w}(R_i) - \Delta_0)}, \tag{4.89}$$

where γ does not depend on the choice of the points $P_0, \dots, P_g \in \mathcal{R}$.

We have given a general outline of basic facts concerning singular hyperelliptic Riemann surfaces. Although we have considered only surfaces with only one singular point, it is worth noting that all the above reasoning extend to the case of surfaces having more than one singular point as well. Obviously, the definitions must be suitably adapted; but see [53] for a good discussion of this topic.

Now we apply the above results to our particular case. We review developments in [96]. Let us introduce the following differential forms on the Riemann surface $\mathcal{R}$:

$$v_{r-1} = \frac{\lambda^{r-2} \mathrm{d}\lambda}{k(P)}, \quad \lambda = \pi(P),\ 1 \leqslant r \leqslant g+1.$$

Note that $v_1, \dots, v_g$ are holomorphic differentials, while $v_0 = \frac{\mathrm{d}\lambda}{\lambda k(P)}$ is not holomorphic, but is a differential of the third kind with simple poles at 0^+ and 0^- with residues

$$\frac{1}{k(0^{\pm})} = \frac{\pm 1}{\sqrt{\prod_{i=0}^{2g} \lambda_i}}$$

respectively.

Let $\omega \in \Omega$, and let $P_1(t), \dots, P_g(t)$ be the finite poles of $M_{\omega \cdot t}$. Abusing notation slightly, write

$$v_{r-1}(t) = \sum_{i=1}^{g} \int_{P_*}^{P_i(t)} v_{r-1}, \quad 1 \leqslant r \leqslant g+1. \tag{4.90}$$

Here P_* is a fixed point of $\mathcal{R}$ different from 0^+ and 0^-.

Differentiating, and simplifying the notation, we obtain

$$v'_{r-1}(t) = \sum_{j=1}^{g} \frac{P_j^{r-2} P'_j}{k(P_j)} = \sum_{j=1}^{g} \frac{P_j^{r-1} P'_j}{P_j k(P_j)}.$$

We have

$$H'(P_r) = \left(\prod_{j=1}^{g} (\lambda - P_j) \right)' \Bigg|_{\lambda = P_r} = -P'_r \prod_{s \neq r} (P_r - P_s). \tag{4.91}$$

Note that

$$\prod_{s \neq r} (P_r - P_s) = \sum_{j=0}^{g-1} \theta^{(r)}_{g-j-1} P_r^j \tag{4.92}$$

where $\theta^{(r)}_{g-j-1}$ is $(-1)^{g-j-1}$ times the $(g-j-1)$-th elementary symmetric function in $(P_1, \ldots, \hat{P}_r, \ldots, P_g)$.

If we let

$$A = \begin{pmatrix} 1 & \theta_1^{(1)} & \cdots & \theta_{g-1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \theta_1^{(g)} & \cdots & \theta_{g-1}^{(g)} \end{pmatrix},$$

then we can rewrite (4.92) as follows

$$A \begin{pmatrix} \frac{P_1^{g-1} P'_1}{k(P_1) P_1} & \cdots & \frac{P_g^{g-1} P'_g}{k(P_g) P_g} \\ \vdots & \ddots & \vdots \\ \frac{P'_1}{k(P_1) P_1} & \cdots & \frac{P'_g}{k(P_g) P_g} \end{pmatrix} = \mathrm{diag} \left(\frac{(-1)^g [m_-^0 - m_+^0] \prod_{r \neq s} P_s}{p(t) k(0^+)} \right)_{r=1 \ldots g}.$$

Multiplying both sides on the right by $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ we obtain

$$A \begin{pmatrix} v'_{g-1} \\ \vdots \\ v'_0 \end{pmatrix} = \begin{pmatrix} \frac{-[m_-^0(t) - m_+^0(t)] \theta_{g-1}^{(1)}}{p(t) k(0^+)} \\ \vdots \\ \frac{-[m_-^0(t) - m_+^0(t)] \theta_{g-1}^{(g)}}{p(t) k(0^+)} \end{pmatrix}.$$

Multiplying by A^{-1} we have

$$\begin{pmatrix} v'_{g-1} \\ \vdots \\ v'_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ -\frac{[m_-^0(t)-m_+^0(t)]}{p(t)k(0^+)} \end{pmatrix},$$

hence

$$v'_0(t) = -\frac{[m_-^0(t) - m_+^0(t)]}{p(t)k(0^+)}. \tag{4.93}$$

This implies that

$$\begin{aligned} v_{r-1}(t) &= \sum_{j=1}^{g} \int_{P_*}^{P_j(t)} v_{r-1} \\ &= \begin{cases} c_{r-1}, & r = 2, \dots, g, \\ c_0 - \displaystyle\int_0^x \frac{[m_-^0(s) - m_+^0(s)]}{p(s)k(0^+)}\,\mathrm{d}s, & r = 1. \end{cases} \end{aligned} \tag{4.94}$$

From (4.91) and (4.78), we get the following important formula describing the pole motion

$$P'_r(t) = \frac{(-1)^g k(P_r(t))[m_-^0(t) - m_+^0(t)]\prod_{i=1}^g P_i(t)}{p(t)k(0^+)\prod_{s\neq r}(P_r(t) - P_s(t))}, \quad 1 \leqslant r \leqslant q. \tag{4.95}$$

Note that (4.95) states that a pole $P_j(t)$ $(1 \leqslant j \leqslant g)$ has 0 derivative if and only if it attains one of the values $\lambda_{2j-1}, \lambda_{2j}$ in the resolvent interval containing it: when it reaches such a value, then it jumps from one sheet to the other in the Riemann surface $\mathcal{R}$, and begins moving in the opposite direction in the interval.

We see that the quantities $v_0(t), \dots, v_{g-1}(t)$ are known once $p(t) = p(\omega \cdot t)$, $m_+^0(t)$, and $m_-^0(t)$ are given, while $v_g(t)$ is as yet undetermined. At this point we propose to study the pole motion by introducing an appropriate Abel map which sends the pole divisor $\{P_1(t), \dots, P_g(t)\}$ into a generalized Jacobian variety. It will turn out that the lack of an explicit formula for $v_g(t)$ is related to the fact the Abel map sends the pole divisor into Υ_0. Thus, roughly speaking, $v_g(t)$ is determined implicitly by $v_0(t), \dots, v_{g-1}(t)$.

Let $\alpha_1, \dots, \alpha_g, \beta_1, \dots, \beta_g$ be a homology basis on $\mathcal{R}$, satisfying the intersection relations (4.80). Since v_0 is a differential of the third kind on $\mathcal{R}$, the natural setting for studying the pole motion is the surface $\mathcal{R}_0$ obtained, as before, by building the family $\{\mathcal{R}_z, z \in D \setminus \{0\}\}$ of Riemann surfaces, each with double points $0^+, 0^- \in \mathcal{R}$, and then letting $z \to 0$. The surface $\mathcal{R}_0$ has a double point corresponding to the identification of the points $0^\pm \in \mathcal{R}$.

Now, the differentials $v_0, v_1, \dots, v_g$ are not normalized. In order to use instruments from the theory described above, we introduce the normalizing invertible $(g+1)\times(g+1)$ complex matrix

$$C_0 = \begin{pmatrix} c_{00} & c_{01} & \dots & c_{0g} \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix},$$

such that

$$\begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_g \end{pmatrix} = C_0 \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_g \end{pmatrix}.$$

Here the differentials $w_0, w_1, \dots, w_g$ are the normalized differentials considered before, hence $w_1, \dots, w_g$ are a basis of holomorphic differentials on $\mathcal{R}$; also, w_0 is a differential of the third kind on $\mathcal{R}$, having poles at $0^{\pm}$ with residue $\pm 1/2\pi\mathrm{i}$ respectively, and such that $\int_{\alpha_i} w_0 = 0$, for every $i = 1, \dots, g$.

We are interested in the restriction of the map I_0 (which by abuse of terminology we again call I_0) to the set $\mathrm{Symm}^g(\mathcal{R}_0)$ of unordered g-tuples $\{P_1, \dots, P_g\}$ of points in $\mathcal{R} \setminus \{0^+, 0^-\}$, where $\mathrm{Symm}^g(\mathcal{R}_0)$ is viewed as a subset of $\mathrm{Symm}^{g+1}(\mathcal{R}_0)$ by setting $P_0 = P_*$. In particular, we wish to describe the image $I_0(\mathrm{Symm}^g(\mathcal{R}_0)) \subset J_0(\mathcal{R})$.

Let Θ_0 be the generalized Riemann Θ-function. It is a simple observation that, if $\Theta_0(\tilde{s}) = 0$, then $\Theta_0(\tilde{s} + C_0^{-1}\lambda) = 0$ for all $\lambda \in \Lambda_0$. Define

$$\Upsilon = \left\{ \sum_{i=1}^{g} \int_{P_*}^{P_i} (v_0, v_1, \dots, v_g) \mid P_1, \dots, P_g \in \mathcal{R} \setminus \left\{0^+, 0^-\right\} \right\} \subset \mathbb{C}^{g+1}.$$

Then $\varsigma \in \Upsilon$ if and only if $\Theta_0(C_0^{-1}\varsigma - \Delta_0) = 0$. It follows easily that

$$\Upsilon_0 = \left\{ (\varsigma - C_0\Delta_0) \bmod \Lambda_0 \mid \varsigma \in \Upsilon \right\}.$$

Thus the pole motion $\{P_1(t), \dots, P_g(t)\} \mapsto I_0(P_1(t), \dots, P_g(t)\}$ takes place in the affine translate Υ of Υ_0, and also in Υ_0 itself via the relation $\{P_1(t), \dots, P_g(t)\} \mapsto C_0^{-1}[I_0(P_1(t), \dots, P_g(t))] - \Delta_0$.

Now recall the expression for $2\sqrt{py}$ as in (4.78), i.e.,

$$2\sqrt{p(t)y(t)} = \frac{(-1)^{g+1}[m_-^0(t) - m_+^0(t)] \prod_{i=1}^{g} P_i(t)}{k(0^+)}.$$

If we put

$$\delta(t) = \frac{(-1)^{g+1}[m_-^0(t) - m_+^0(t)]}{2k(0^+)},$$

then

$$\sqrt{p(t)y(t)} = \delta(t)\prod_{i=1}^{g} P_i(t). \tag{4.96}$$

Next, consider the function $P \mapsto \pi(P)$; it is a meromorphic function having a double pole at ∞ and zeros at 0^+ and 0^-. Let $P_0 = P_*$. It follows that (4.96) is the restriction of a well defined meromorphic function defined on $J_0(\mathcal{R})$, to Υ_0. Note that in general this restriction is not any more a meromorphic function on Υ_0. These observations, together with Theorem 4.50, provide the proof of the following important

THEOREM 4.51. *$\sqrt{p(t)y(t)}$ coincides with the restriction of a meromorphic function, defined on $J_0(\mathcal{R})$, to Υ_0.*

In particular

$$\begin{aligned}\sqrt{p(t)y(t)} &= \frac{(-1)^{g+1}[m_-^0(t) - m_+^0(t)]\prod_{i=1}^{g}\pi(P_i(t))}{2k(0^+)}\\ &= \delta(t)\gamma\frac{\Theta_0(C_0^{-1}[I_0(P_1(t),\dots,P_g(t))] - \int_{P_*}^{0^+}(w_0,w_1,\dots,w_g) - \Delta_0)}{\Theta_0^2(C_0^{-1}[I_0(P_1(t),\dots,P_g(t))] - \int_{P_*}^{\infty}(w_0,w_1,\dots,w_g) - \Delta_0)}\\ &\quad\times\frac{\Theta_0(C_0^{-1}[I_0(P_1(t),\dots,P_g(t))] + \int_{P_*}^{0^+}(w_0,w_1,\dots,w_g) - \Delta_0)}{\Theta_0^2(C_0^{-1}[I_0(P_1(t),\dots,P_g(t))] - \int_{P_*}^{\infty}(w_0,w_1,\dots,w_g) - \Delta_0)}.\end{aligned} \tag{4.97}$$

Here γ does not depend on the choice of $P_j(t)$ $(1 \leqslant j \leqslant g)$.

The following implicit relation holds between the components of $(w_0(t),\dots,w_g(t)) = C_0^{-1}[I_0(P_1(t),\dots,P_g(t))] = \sum_{i=1}^{g}\int_{P_*}^{P_i(t)}(w_0,\dots,w_g)$:

$$\Theta_0\left(\sum_{i=1}^{g}\int_{P_*}^{P_i(t)}(w_0,\dots,w_g) - \Delta_0\right) = 0.$$

Another way of retrieving formula (4.97) is presented in [156], Section 8. In that paper, formulas for the expression on $J_0(\mathcal{R})$ of all the elementary symmetric functions of g points on $\mathcal{R}$ are obtained, depending on the logarithmic derivative of a well-known Θ-quotient.

We now discuss another approach to the study of the pole motion which is based on the Baker–Akhiezer function $\varphi(t,P)$ of Eqs. (SL_ω). Fix $\omega \in \Omega$. For each $\lambda \in \mathbb{C}$ which is not a pole of $m_+(0,\lambda)$, let $\varphi_+(t,\lambda)$ be the solution of (4.24_ω) which satisfies $\varphi_+(0,\lambda) = 1, \varphi'(0,\lambda) = \frac{m_+(0,\lambda)}{p(0)}$. Similarly, define $\varphi_-(t,\lambda)$ for each $\lambda \in \mathbb{C}$ which is not a pole of $m_-(0,\lambda)$. It turns out that, for each $t \in \mathbb{R}$, these two functions glue together to form a single meromorphic function $\varphi(t,z)$ on the punctured Riemann surface $\mathcal{R} \setminus \{\infty\}$. This function has an essential singularity at $\infty \in \mathcal{R}$. If $z = \frac{1}{\sqrt{-\lambda}}$ is a local parameter at ∞, then by (4.67):

$$\varphi(t, z) = \exp\left(\frac{-1}{z} \int_0^t \sqrt{\frac{y(s)}{p(s)}}\, \mathrm{d}s\right) f(t, z)$$

where $f(t, \cdot)$ is holomorphic on some disc in the z-plane centered at $z = 0$.

Unfortunately, one cannot apply the Abel Theorem to $\varphi(t, \cdot)$, because it is not meromorphic at ∞. However, following [35], we analyze the differential form $w = d \ln \varphi$. We note the following interesting features of w:

(a) w has a simple pole at $P_i(0)$ with residue -1 $(1 \leqslant i \leqslant g)$.

(b) w has a simple pole at $P_i(t)$ with residue 1 $(1 \leqslant i \leqslant g)$.

(c) w has a pole of order 2 at ∞, and

$$w = \left(\int_0^t \sqrt{\frac{y(s)}{p(s)}}\, \mathrm{d}s\right) \frac{\mathrm{d}z}{z^2} + \cdots$$

near $z = 0$.

(d) φ is single-valued on $\mathcal{R} \setminus \{\infty\}$, and hence $\int_{\alpha_j} w = 2\pi \mathrm{i} m_j$ and $\int_{\beta_j} w = 2\pi \mathrm{i} n_j$ where $m_j, n_j \in \mathbb{Z}$ $(1 \leqslant j \leqslant g)$.

Following the developments in [35], we let η be the differential of the second kind on $\mathcal{R}$ which has a pole of order 2 at ∞ and which is normalized so that $\int_{\alpha_j} \eta = 0$ $(1 \leqslant j \leqslant g)$. As usual, let $w_1, \ldots, w_g$ be the normalized basis of holomorphic differentials on $\mathcal{R}$. Let $w_{P_i(0), P_i(t)}$ be the differential of the third kind on $\mathcal{R}$ having simple poles at $P_i(0)$ and $P_i(t)$ with residues -1 and $+1$ respectively $(1 \leqslant i \leqslant g)$, and normalized in such a way that

$$\int_{\alpha_j} w_{P_i(0), P_i(t)} = 0 \quad (1 \leqslant i, j \leqslant g).$$

This normalization and Riemann's bilinear relation [112] ensure that

$$\int_{\beta_j} w_{P_i(0), P_i(t)} = \int_{P_i(0)}^{P_i(t)} w_j \mod \Lambda^{2g} \quad (1 \leqslant i, j \leqslant g).$$

It is clear now that

$$w = \int_0^t \sqrt{\frac{y(s)}{p(s)}}\, \mathrm{d}s \eta + \sum_{i=1}^{g} w_{P_i(0), P_i(t)} + \sum_{j=1}^{g} \gamma_j w_j,$$

where $\gamma_1, \dots, \gamma_g$ are complex constants. Since $\int_{\alpha_j} w = 2\pi \mathrm{i} m_j$, we see that $\gamma_j = 2\pi \mathrm{i} m_j$ $(1 \leqslant j \leqslant g)$. Since $\int_{\beta_k} w = 2\pi \mathrm{i} n_k$, we see that

$$2\pi \mathrm{i} n_k = -\int_0^t \sqrt{\frac{y(s)}{p(s)}}\, \mathrm{d}s U_k + \sum_{i=1}^{g} \int_{P_i(0)}^{P_i(t)} w_k + \sum_{i=1}^{g} 2\pi \mathrm{i} m_j z_{jk},$$

where $U_k = -\int_{\beta_k} \eta$ $(1 \leqslant k \leqslant g)$. We thus find

$$\sum_{i=1}^{g} \int_{P_i(0)}^{P_i(t)} w_k = 2\pi \mathrm{i}\left(n_k - \sum_{j=1}^{g} m_j z_{jk}\right) + \left(\int_0^t \sqrt{\frac{y(s)}{p(s)}}\, \mathrm{d}s\right) U_k \quad (1 \leqslant k \leqslant g).$$

Let us consider the standard Abel map $I : \mathrm{Symm}^g(\mathcal{R}) \to J(\mathcal{R})$. We see that the pole motion $t \to \{P_1(t), \dots, P_g(t)\}$ is mapped into $J(\mathcal{R})$, and assumes a very suggestive form, i.e.,

$$I\big(P_1(t), \dots, P_g(t)\big) = c + \left(\int_0^t \sqrt{\frac{y(s)}{p(s)}}\, \mathrm{d}s\right) U \mod \Lambda^{2g} \tag{4.98}$$

where c is a constant vector and $U = (U_1, \dots, U_g)$.

This approach, however, does not provide a solution of our difficulty concerning the implicit nature of the pole motion. In fact, from (4.79) and Theorem 4.40, we see that the function $\sqrt{\frac{y(t)}{p(t)}}$ depends on the product $\prod_{i=1}^{g} P_i(t)$. Thus (4.98) gives the pole motion implicitly. At this point one can use the change of variable

$$\frac{\mathrm{d}\tilde{t}}{\mathrm{d}t} = \prod_{i=1}^{g} P_i(t),$$

as in [2]. With respect to this new variable, the motion $\tilde{t} \to \{P_1(\tilde{t}), \dots, P_g(\tilde{t})\}$ defines a motion along a straight line in $J(\mathcal{R})$ whose velocity is given by

$$\frac{(-1)^{g+1}[m_-^0(\tilde{t}) - m_+^0(\tilde{t})]}{2p(\tilde{t})k(0^+)}.$$

Again, the implicit relation seen earlier between v_g and $(v_0, \dots, v_{g-1})$ is reflected here in the relation between t and $\tilde{t}$.

Theorem 4.40 ensures that, if Hypotheses 4.38 hold, then there exist relations of algebro-geometric type for the coefficients $p(t), q(t), y(t)$ and for the Weyl m-functions $m_\pm$ satisfying the Riccati equation (4.25). Going backwards, we seek to determine if a fixed algebro-geometric configuration determines the coefficients of a Sturm–Liouville operator.

To this end, we fix the Riemann surface $\mathcal{R}$ of genus g of the algebraic relation

$$w^2 = -(\lambda - \lambda_0)(\lambda - \lambda_1)\cdots(\lambda - \lambda_{2g})$$

where $\lambda_i \in \mathbb{R}^+$ $(0 \leqslant i \leqslant 2g)$ and are all distinct. Let $\tilde{P}_1, \ldots, \tilde{P}_g$ be points on $\mathcal{R}$ such that $\lambda_{2i-1} \leqslant \tilde{P}_i \leqslant \lambda_{2i}$ $(1 \leqslant i \leqslant g)$. Define

$$H_0(\lambda) = \prod_{i=1}^{g} \bigl(\lambda - \pi(\tilde{P}_i)\bigr).$$

Let $\tilde{\Omega}$ be a compact metric space, let $\{\tilde{\tau}_t \mid t \in \mathbb{R}\}$ be a continuous one-parameter group of homeomorphisms of $\tilde{\Omega}$, and let $\tilde{\nu}$ be a $\{\tilde{\tau}_t\}$-ergodic measure on $\tilde{\Omega}$ such that $\tilde{\Omega}$ is the topological support of $\tilde{\nu}$. Let $p : \tilde{\Omega} \to \mathbb{R}$ and $\mathcal{M}_1 : \tilde{\Omega} \to \mathbb{R}$ be positive bounded continuous functions. Suppose further that the function $\tilde{\omega} \mapsto \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}_1(\tilde{\tau}(\tilde{\omega}))|_{t=0} : \tilde{\Omega} \to \mathbb{R}$ is defined and continuous. Fix $\tilde{\omega} \in \tilde{\Omega}$, and write $p(t) = p(\tilde{\tau}_t(\tilde{\omega}))$, $\mathcal{M}_1(t) = \mathcal{M}_1(\tilde{\tau}_t(\tilde{\omega}))$.

Let $(P_1(t), \ldots, P_g(t))$ be the solution of the following system of first-order differential equations

$$\frac{-P_r'(t)\prod_{s\neq r}(P_r(t) - P_s(t))}{k(P_r(t))} = \frac{(-1)^{g+1}\mathcal{M}_1(t)\prod_{i=1}^{g} P_i(t)}{p(t)k(0^+)} \qquad (1 \leqslant r \leqslant g) \tag{4.99}$$

with initial condition $P_r(0) = \tilde{P}_r$ for every $r = 1, \ldots, g$ (i.e. the analogue of (4.95)), and $k(P)$ is the meromorphic function on $\mathcal{R}$ defined in (4.69). Note that (4.99) implies that the motion of the points P_r $(1 \leqslant r \leqslant g)$ has the following properties:

(1) $P_r'(t)$ attains the value 0 only at the points t such that $P_r(t)$ is a ramification point;

(2) when $P_r(t)$ reaches a ramification point, say at $t = \tilde{t}$, then $P_r(t)$ jumps from a sheet to the other in the Riemann surface $\mathcal{R}$, and begins to move towards the opposite ramification point in the resolvent interval containing it (this can be seen by introducing a parameter $z = \sqrt{\lambda - P_r(\tilde{t})}$ in a neighborhood of the ramification point $P_r(\tilde{t})$ of $\mathcal{R}$).

As before, let $c_i = \pi^{-1}[\lambda_{2i-1}, \lambda_{2i}] \subset \mathcal{R}$ $(1 \leqslant i \leqslant g)$. The product $\tilde{D} = c_1 \times c_2 \times \cdots \times c_g$ is a real analytic g-torus which we view as an embedded submanifold of $\mathrm{Symm}^g(\mathcal{R}_0)$. The differential equations (4.99) define a one-parameter group of homeomorphisms $\{\tau_t\}$ on $\tilde{\Omega} \times \tilde{D}$, as follows: if $\tilde{\omega} \in \tilde{\Omega}$ and $\tilde{d} = (\tilde{P}_1, \ldots, \tilde{P}_g) \in \tilde{D}$, then

$$\tau_t(\tilde{\omega}, \tilde{d}) = \bigl(\tilde{\tau}_t(\tilde{\omega}), d(t)\bigr),$$

where $d(t) = \{P_1(t), \ldots, P_g(t)\}$ is obtained by solving (4.99) with initial value $\tilde{d}$.

Fix $\tilde{\omega} \in \tilde{\Omega}$, $\tilde{d} \in \tilde{D}$. Define a (positive) function $y(t)$ such that

$$2\sqrt{p(t)y(t)} = \frac{(-1)^{g+1}\mathcal{M}_1(t)\prod_{i=1}^{g} P_i(t)}{k(0^+)}. \tag{4.100}$$

Put

$$H(t,\lambda)=\prod_{i=1}^{g}\bigl[\lambda-\pi\bigl(P_i(t)\bigr)\bigr].$$

Consider the nonstandard Abel Jacobi coordinates (4.90). We repeat the computations after (4.90) with $\mathcal{M}_1(t)$ instead of $m_-^0(t)-m_+^0(t)$. The implicit dependence of $v_g(t)$ with respect to $v_0(t),\dots,v_{g-1}(t)$ is a general fact and does not depend on the particular choice of the surface and the points on it. Thus $v_g(t)$ is a transcendental function of $v_0(t),v_1(t)\dots,v_{g-1}(t)$ such that

$$\varsigma(t)=\sum_{i=1}^{g}\int_{P_*}^{P_i(t)}(v_0,v_1,\dots,v_g)\in\Upsilon,$$

or equivalently, $\Theta_0(C_0^{-1}\varsigma(t)-\Delta_0)=0$.

Let $D=I_0(\tilde{D})\subset\Upsilon$. We can identify $\tilde{\Omega}\times\tilde{D}$ with the product space $\tilde{\Omega}\times D\subset\tilde{\Omega}\times\Upsilon$. Let $(\tilde{\omega},\tilde{d})\in\tilde{\Omega}\times\tilde{D}$, then let

$$\Omega=\mathrm{cls}\left\{\left(\tilde{\tau}_t(\tilde{\omega}),I_0(0)+\left(\sum_{i=1}^{g}\int_{P_*}^{P_i(t)}(v_0,v_1,\dots,v_g)\right)\right)\,\middle|\,t\in\mathbb{R}\right\},$$

where $I_0(0)=(\sum_{i=1}^{g}\int_{P_*}^{P_i(0)}(v_0,v_1,\dots,v_g))$. The flow $(\tilde{\Omega}\times\tilde{D},\{\tau_t\})$ induces a flow on Ω which by abuse of notation we denote by $(\Omega,\{\tau_t\})$. Let $\tilde{\omega}\in\tilde{\Omega}$ be a point such that the orbit $\{\tilde{\tau}_t(\tilde{\omega})\mid\tilde{\omega}\in\tilde{\Omega}\}$ is dense in $\tilde{\Omega}$ (this holds for $\tilde{\nu}$-a.a. $\tilde{\omega}\in\tilde{\Omega}$), then let $\tilde{\pi}:\tilde{\Omega}\times\tilde{D}\to\tilde{\Omega}$ be the projection. The flow $(\Omega,\{\tau_t\})$ is an extension of the flow $(\tilde{\Omega},\{\tilde{\tau}_t\})$, in the sense that the projection $\tilde{\pi}$ restricted to Ω is a flow homeomorphism of Ω onto $\tilde{\Omega}$. There exists a $\{\tau_t\}$-ergodic measure ν on Ω which is the lift of the $\{\tilde{\tau}_t\}$-ergodic measure $\tilde{\nu}$ on $\tilde{\Omega}$, that is, $\tilde{\pi}(\nu)=\tilde{\nu}$. We give a short proof of this fact. First, we use the Birkhoff ergodic theorem together with a construction of Krylov–Bogoliubov type similar to that used in the proof of Lemma 4.18, to prove that the set $\mathcal{S}$ of all the $\{\tau_t\}$-invariant measures ι on Ω such that $\tilde{\pi}(\iota)=\tilde{\nu}$ is nonempty. Moreover, one can see that $\mathcal{S}$ is compact and convex in the weak-* topology. By the Krein–Mil'man theorem, $\mathcal{S}$ is the hull of its extreme points. It is easy to show that every extreme point ν of $\mathcal{S}$ is $\{\tau_t\}$-ergodic and projects into $\tilde{\nu}$.

If necessary, redefine Ω to be topological support of the measure ν: $\Omega=\operatorname{Supp}\nu$. Fix $\omega=(\tilde{\omega},d)\in\Omega$. Put

$$k_r(t)=k\bigl(P_r(t)\bigr),$$

and define a polynomial $Q(t,\lambda)$ of degree g in λ and with coefficients depending on t, such that

$$Q_i(t)=Q\bigl(t,\pi\bigl(P_i(t)\bigr)\bigr)=\sqrt{p(t)y(t)}\,k_i(t)\quad(i=1,\dots,g).\tag{4.101}$$

We also require that the highest term coefficient q_g is

$$q_g(t) = -\frac{(p(t)y(t))'}{4y(t)} \tag{4.102}$$

Define

$$M\big(d(t), P\big) = \frac{\sqrt{p(t)y(t)}\,k(P) + Q(t,\lambda)}{H(t,\lambda)}, \quad \lambda = \pi(P). \tag{4.103}$$

It is clear that

$$M\big(d(t), \sigma(P)\big) = \frac{Q(t,\lambda) - \sqrt{p(t)y(t)}\,k(P)}{H(t,\lambda)},$$

where σ is the hyperelliptic involution.

Define

$$h_+(t,P) = M\big(d(t), P\big) \quad \text{and} \quad h_-(t,P) = M\big(d(t), \sigma(P)\big), \tag{4.104}$$

and note that

$$h_+(t,0) - h_-(t,0) = -\mathcal{M}_1(t). \tag{4.105}$$

Put

$$\mathcal{M}_2(t) = h_+(t,0) + h_-(t,0) = \frac{2Q(t,0)}{H(t,0)}. \tag{4.106}$$

Finally, let

$$q(t) = y(t)\left(\sum_{i=0}^{2g} \lambda_i - 2\sum_{i=1}^{g} P_i(t)\right) + \frac{q_g^2(t)}{p(t)} + q_g'(t). \tag{4.107}$$

Corresponding to every $\omega \in \Omega \subset \tilde{\Omega} \times \Upsilon$, there is a one parameter family of matrices

$$\begin{pmatrix} 0 & 1/p(t) \\ q(t) - \lambda y(t) & 0 \end{pmatrix},$$

where y and q are defined as in (4.100) and (4.107) respectively.

We are ready to prove the following important result

THEOREM 4.52. *Let $\omega \in \Omega$, $\lambda = \pi^{-1}(P) \in \mathbb{C}$ with $\Im\lambda \neq 0$ and let $p(t), y(t)$ and $q(t)$ be continuous functions such that the relations* (4.100) *and* (4.107) *hold. Then $h_\pm(t,P)$ satisfy the Riccati equation*

$$h_\pm' + \frac{1}{p}h_\pm^2 = q - \lambda y. \tag{4.108}$$

PROOF. We will give the proof only for h_+. We want to prove that, differentiating (4.103) with respect to t, we obtain

$$\begin{aligned}&\big((\sqrt{py})'k + Q'\big)H - (\sqrt{py}\,k + Q)H' \\ &\quad + \frac{1}{p}(\sqrt{py}\,k + Q)^2 - (q - \lambda y)H^2 = 0. \end{aligned}\tag{4.109}$$

We rearrange (4.109) in the following way

$$\begin{aligned}&Q'H - QH' + yk^2 + \frac{1}{p}Q^2 - (q - \lambda y)H^2 \\ &\quad = k\bigg(\sqrt{py}H' - (\sqrt{py})'H - \frac{2}{p}\sqrt{py}Q\bigg). \end{aligned}\tag{4.110}$$

We claim that

$$\begin{aligned}F(t,\lambda) = \sqrt{p(t)y(t)}\,H'(t,\lambda) - \big(\sqrt{p(t)y(t)}\big)'H(t,\lambda) \\ - \frac{2}{p(t)}\sqrt{p(t)y(t)}\,Q(t,\lambda) = 0, \end{aligned}\tag{4.111}$$

Such a F is a polynomial in λ of degree g. It is easy to observe that $F(t,\lambda)$ has the points $\pi(P_i(t))$ as roots for every $i = 1,\dots,g$, hence it has g roots. Moreover the λ^g coefficient of $F(t,\lambda)$ is

$$-\big(\sqrt{p(t)y(t)}\big)' + \frac{2}{p(t)}\sqrt{p(t)y(t)}\,\frac{(p(t)y(t))'}{4y(t)} = 0,$$

hence $F(t,\lambda) = 0$. To conclude the proof, we examine the remaining term in (4.110), i.e.

$$\begin{aligned}L(t,\lambda) = Q'(t,\lambda)H(t,\lambda) - Q(t,\lambda)H'(t,\lambda) + y(t)k^2(P) \\ + \frac{1}{p(t)}Q^2(t,\lambda) - \big(q(t) - \lambda y(t)\big)H^2(t,\lambda), \quad \pi(P) = \lambda. \end{aligned}\tag{4.112}$$

$L(t,\lambda)$ is polynomial in λ of degree at most $2g+1$. It is easy to observe that for every $i = 1,\dots,g$, $\pi(P_i(t))$ is a root of $L(t,\lambda)$. As before we compute the coefficient α of λ^{2g+1} and β of λ^{2g}: we have

$$\alpha = y - y = 0$$

$$\beta = y(t)\Big(\sum\lambda_i - 2\sum P_i(t)\Big) + \frac{1}{p(t)}q_g^2(t) + q_g'(t) - q(t) = 0,$$

hence $L(t,\lambda)$ has degree at most $2g-1$. If we show that every $\pi(P_i(t))$ is a double root of $L(t,\lambda)$, then the proof is complete.

We observe that $H^2(t,\lambda)$ certainly has all the points $\pi(P_i(t))$ as double zeros, so we consider only

$$\begin{aligned} V(t,\lambda) &= Q'(t,\lambda)H(t,\lambda) - Q(t,\lambda)H'(t,\lambda) + y(t)k^2(P) \\ &\quad + \frac{1}{p(t)}Q^2(t,\lambda) = H(t,\lambda)N(t,\lambda), \quad \pi(P)=\lambda \end{aligned} \tag{4.113}$$

where

$$N(t,\lambda) = Q'(t,\lambda) - Q(t,\lambda)\frac{H'(t,\lambda)}{H(t,\lambda)} + y(t)\frac{k^2(P)}{H(t,\lambda)} + \frac{Q^2(t,\lambda)}{p(t)H(t,\lambda)}. \tag{4.114}$$

In view of the fact that $F(t,\lambda)=0$, we can write

$$H'(t,\lambda) = \frac{(p(t)y(t))'}{2p(t)y(t)}H(t,\lambda) + \frac{2}{p(t)}Q(t,\lambda) \tag{4.115}$$

and substituting this expression in (4.113) we have

$$\begin{aligned} V(t,\lambda) &= Q'(t,\lambda)H(t,\lambda) - Q(t,\lambda)\left(\frac{(p(t)y(t))'}{2p(t)y(t)}H(t,\lambda) + \frac{1}{p(t)}Q(t,\lambda)\right) \\ &\quad + y(t)k^2(P). \end{aligned} \tag{4.116}$$

Differentiating V with respect to t we have

$$\begin{aligned} V'(t,\lambda) &= Q''(t,\lambda)H(t,\lambda) + Q'(t,\lambda)H'(t,\lambda) \\ &\quad - Q'(t,\lambda)\left(\frac{(p(t)y(t))'}{2p(t)y(t)}H(t,\lambda) + \frac{1}{p(t)}Q(t,\lambda)\right) \\ &\quad - Q(t,\lambda)\left(\frac{(p(t)y(t))'}{2p(t)y(t)}H(t,\lambda) + \frac{1}{p(t)}Q(t,\lambda)\right)' + y'(t)k^2(P), \end{aligned} \tag{4.117}$$

and

$$\begin{aligned} V'(t,P_r) &= Q_r'(t)H_r'(t) - \frac{2}{p(t)}Q_r(t)Q_r'(t) - \frac{(p(t)y(t))'}{2p(t)y(t)}H_r'(t)Q_r(t) \\ &\quad + \frac{p'(t)}{p^2(t)}Q_r^2(t) + y'(t)k_r^2(t), \end{aligned} \tag{4.118}$$

since $H_r(t) = H(t,\pi(P_r(t))) = 0$.

Substituting in (4.118) the formulas for H_r' and Q_r obtained from (4.99), (4.100) and (4.101), we see that $V'(t,P_r)=0$, and we have proved that H divides $V' = HN' + H'N$, and so H divides $H'N$. Now, $H'(t, P_r) = -P_r'(t)\prod_{s\neq r}(P_r(t) - P_s(t))$, and $H_r'(t) = 0$ if and only if $P_r'(t) = 0$. If $P_r' \neq 0$ for every r then H divides N too, and the proof is complete.

Now, suppose that there is an index j and a value $\tilde{t} \in \mathbb{R}$ such that $P_j'(\tilde{t}) = 0$. In this case we must have necessarily $\sigma(P_j(\tilde{t})) = P_j(\tilde{t})$. In fact, if $\sigma(P_r(\tilde{t})) \neq P_r(\tilde{t})$ for every r, then we see that $P_r'(\tilde{t}) \neq 0$, since the (restricted) Abel map I_0 is a diffeomorphism of $\mathrm{Symm}^g(\mathcal{R}_0)$ onto Υ and π is invertible in a neighborhood of each P_r. If then $\sigma(P_j(\tilde{t})) = P_j(\tilde{t})$ for some $j \in \{1, \dots, g\}$ and $\tilde{t}$, we may approximate the divisor $d(\tilde{t})$ with an arbitrarily near divisor $d_\varepsilon(\tilde{t}) = P_{1,\varepsilon}(\tilde{t}) + \cdots + P_{g,\varepsilon}(\tilde{t})$ such that $\sigma(P_{r,\varepsilon}(\tilde{t})) \neq P_{r,\varepsilon}(\tilde{t})$ for every r. For this divisor, the conclusion of the proof above holds, and, passing to the limit, the same conclusion holds for $d(\tilde{t})$. The proof is complete. □

From Theorem 4.52 we obtain this important

THEOREM 4.53. *The real flow $(\Omega, \{\tau_t\})$ has the property that the spectrum of the family of two-dimensional nonautonomous differential systems of Sturm–Liouville type*

$$\begin{aligned}\begin{pmatrix} u \\ v \end{pmatrix}' &= \begin{pmatrix} 0 & \frac{1}{p(\omega\cdot t)} \\ q(\omega\cdot t) - \lambda y(\omega\cdot t) & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} \\ &= \omega(t,\lambda)\begin{pmatrix} u \\ v \end{pmatrix} \quad (\omega \in \Omega) \end{aligned} \tag{SL$_\omega$}$$

does not depend on the choice of $\omega \in \Omega$ and has the form

$$\Sigma(L) = [\lambda_0, \lambda_1] \cup [\lambda_2, \lambda_3] \cup \cdots \cup [\lambda_{2g}, \infty). \tag{4.119}$$

Moreover, $\beta(\lambda) = 0$, for all $\lambda \in \Sigma(L)$.

PROOF. For $\lambda \in \mathbb{C}$, define a flow $\{\hat{\tau}_t\}$ on $\Omega \times \mathbb{C}^2$ by setting

$$\hat{\tau}_t(\omega, u) = \big(\omega\cdot t, \Phi_\omega(t)u\big),$$

where, for every $\omega \in \Omega$, $\Phi_\omega(t)$ is the fundamental matrix solution for Eq. (SL$_\omega$) such that $\Phi_\omega(0) = \mathrm{Id}$. Clearly, this flow is induced on $\Omega \times \mathbb{P}(\mathbb{C})$ as well.

Given $\lambda \in \mathbb{C}$ with $\Im\lambda \neq 0$, let $P \in \pi^{-1}(\lambda)$ such that $\Im\lambda\Im k(P) > 0$; since $\omega = (\tilde{\omega}, d) \in \tilde{\Omega} \times \Upsilon$, define $m_+(\tilde{\omega}, d, \lambda) = M(d, P)$ and $m_-(\tilde{\omega}, d, \lambda) = M(d, \sigma(P))$.

By Theorem 4.52, both m_+ and m_- satisfy the Riccati equation (4.25), hence the sections

$$S^\pm(\lambda) = \big\{\big(\lambda, m_\pm(\tilde{\omega}, d, \lambda)\big) \mid (\tilde{\omega}, d) \in \Omega\big\} \subset \Omega \times \mathbb{P}(\mathbb{C})$$

are invariant under the flow $(\Sigma, \{\hat{\tau}_t\})$ for every λ with $\Im\lambda \neq 0$. By continuity, we can extend $S^\pm(\lambda)$ to the real line, keeping them still invariant, and hence we can define $m_\pm$ for every λ with $\Im\lambda = 0$.

Now, take $\lambda \in (\lambda_0, \lambda_1) \cup \cdots \cup (\lambda_{2g}, \infty)$; then it easy to observe that $\Im m_+(\tilde{\omega}, d, \lambda) > 0 > \Im m_-(\tilde{\omega}, d, \lambda)$. It follows that $S^+(\lambda)$ and $S^-(\lambda)$ are distinct. Clearly $\Omega \times \mathbb{P}(\mathbb{R})$ is invariant under the flow $(\Sigma, \{\hat{\tau}_t\})$. By using Liouville's formula, it is easy to show that every solution of (SL_ω) is bounded, for every $\omega \in \Omega$. Thus $\beta(\lambda) = 0$ and λ belongs to the spectrum of the full-line operator (SL_ω). On the other hand, if $\lambda \notin \Sigma(L)$, then the sections $S^{\pm}(\lambda)$ are real, hence the rotation number $\alpha(\lambda)$ of (SL_ω) is constant. It is clear that the only possibility is that λ belongs to the resolvent set (in fact equations (SL_ω) admit an exponential dichotomy; see Theorem 4.21(iii)).

Thus we proved that the spectrum has the form $\Sigma(L) = [\lambda_0, \lambda_1] \cup \cdots \cup [\lambda_{2g}, \infty)$ for every $\omega \in \Omega$, and that $\beta(\lambda) = 0$ for all $\lambda \in \Sigma(L)$. □

A more detailed study of the flow $(\Omega, \{\tau_t\})$ has been carried out in [95]. It is possible to obtain more information concerning the structure of the flow $(\Omega, \{\tau_t\})$ if the set of endpoints of the intervals which determine the spectrum $\Sigma(L)$ satisfy a certain condition, namely that if $\mathcal{R}$ is the Riemann surface of the algebraic relation $w^2 = -(\lambda - \lambda_0)(\lambda - \lambda_1) \cdots (\lambda - \lambda_{2g})$ and $c_i = \pi^{-1}([\lambda_{2i-1}, \lambda_{2i}])$ $(i = 1, \dots, g)$, then the vectors $\zeta_i = \int_{c_i} (v_0, \dots, v_{g-1})$ $(i = 1, \dots, g)$ are linearly independent in $\mathbb{R}^g$. This condition is satisfied in an open dense subset of the space of endpoints, and there are examples for which the above condition is not fulfilled, at least when g is odd. In any case, if the above condition is satisfied, then one can show that for "almost" every pair of functions $(p(t), \mathcal{M}(t))$, the flow $(\Omega, \{\tau_t\})$ generated by $(p, \mathcal{M})$ laminates into closed subsets Ω_h, where h varies in a quotient group $\mathcal{H}$ of the g-torus. Moreover, each set Ω_h is the topological support of an ergodic measure ν_h on Ω. It turns out that, in general, it is difficult to actually determine $\mathcal{H}$, and hence an exact parametrization of the sets Ω_h. However, if $p(t)$ and $\mathcal{M}(t)$ are Bohr almost periodic, then it is possible to determine the "size" of the quotient group $\mathcal{H}$ by looking at the frequency module of the hull of the pair $(p, \mathcal{M})$ (see [95] for more information on this topic).

5. Genericity of exponential dichotomy

As was stated in Section 1, we will consider the question of the genericity of the exponential dichotomy property for cocycles generated by quasi -periodic, two-dimensional linear systems. Thus, we will consider minimal quasi-periodic flows of Kronecker-type generated on the torus $\mathbb{T}^k$ by a vector of frequencies γ, where $\gamma \in \Gamma \subset \mathbb{R}^k$, the set of frequencies whose components are rationally independent. To be precise we study the minimal flow $\{\tau_t\}$ on $\mathbb{T}^k$ given by $\tau_t(\psi) = \psi + \gamma t$ where $\psi = (\psi_1, \dots, \psi_k) \in \mathbb{T}^k$ and $\gamma = (\gamma_1, \dots, \gamma_k) \in \Gamma$. Let us agree to abbreviate exponential dichotomy by E.D. We have that, by the robustness properties of the E.D. (see [29,130]), the set $\mathcal{C}_{\mathrm{ED}} = \{A \in C_0 \mid$ the family of systems $x' = A(\tau_t(\psi))x$ $(\psi \in \mathbb{T}^k)$ admits an E.D.$\}$ is open in the corresponding topology. Here recall that C_0 is the set defined by $C_0 = \{A : \mathbb{T}^k \to \mathrm{sl}(2, \mathbb{R}) \mid A$ is continuous$\}$ in which we consider the sup norm $\|A\| = \sup_{\psi \in \mathbb{T}^k} |A(\psi)|$, where in turn $|\cdot|$ is the sup norm on $\mathrm{sl}(2, \mathbb{R})$ given by $|B| = \sup\{|Bx|_{\mathbb{R}^2}$ with $|x|_{\mathbb{R}^2} = 1\}$. What we will show is the density of the set $\mathcal{C}_{\mathrm{ED}}$ for a residual subset of frequencies $\gamma \in \Gamma$. It will be apparent from the proof that these frequencies are of "Liouville-type". By this we mean rather vaguely that the frequencies are "sufficiently close" to rational numbers; the result we discuss below gives no estimate on the degree of "closeness".

Let $\mathcal{D} = \{(\gamma, A) \in \mathbb{R}^k \times C_0 \mid \text{ the system } x' = A(\tau_t(\psi))x \text{ admits an E.D. over } \mathbb{T}^k\}$. We will prove that the set $\mathcal{D}$ is open and dense in $\mathbb{R}^k \times C_0$. We have already observed that the set $\mathcal{C}_{\mathrm{ED}}$ is open. That $\mathcal{D}$ is open is a consequence of a theorem of Sacker and Sell [137] which is a variant of the results concerning roughness of exponential dichotomies mentioned above [29,130]. What is of interest here is the density result. If we fix a frequency $\gamma \in \Gamma$, we could ask if the set $\mathcal{D}_\gamma = \{A \in C_0 \mid \text{ the system } x' = A(\tau_t(\psi))x \text{ admits an E.D. over } \mathbb{T}^k\}$ is open and dense in C_0. While the openness follows from [29] (also [130,137]), we have that the density of $\mathcal{D}_\gamma$ cannot be proved for all $\gamma \in \Gamma$, but only for a residual set of frequencies.

We will state and prove the exact result below; before doing so, recall that in the Introduction we already noted that if we consider as a compact metric space Ω the circle and as the flow the periodic flow given by the rigid rotation on the circle, we have that the density fails to hold. There are also some other examples in which the density property fails, depending on the topology of the compact metric space Ω and on the nature of the flow (see [75]).

The main result of this section is the following:

THEOREM 5.1. *Consider the set* $\mathcal{D} = \{(\gamma, A) \in \mathbb{R}^k \times C_0 \mid$ *the system* $x' = A(\tau_t(\psi))x$ *admits an E.D. over* $\mathbb{T}^k\}$. *Then the set* $\mathcal{D}$ *is open and dense in* $\mathbb{R}^k \times C_0$.

We have said that the openness of the set $\mathcal{D}$ in $\mathbb{R}^k \times C_0$ follows from a theorem of Sacker and Sell [137] which we will state in a form convenient for our purposes in Theorem 5.3 below. To prove the density we will proceed by contradiction: supposing that there is an open set $V \subset \mathbb{R}^k \times C_0$ such that, if $(\gamma, A) \in V$, then the corresponding family of systems does not have an E.D. over $\mathbb{T}^k$, we will arrive at a contradiction.

We will give the scheme of the proof of this fact. Before doing so, we recall here some general facts about the properties of the rotation number and the Lyapunov exponent of the systems considered. Consider then the family of equations

$$x' = A\bigl(\tau_t(\psi)\bigr)x, \quad x \in \mathbb{R}^2 \tag{5.1}$$

where $t \in \mathbb{R}$, $\psi \in \mathbb{T}^k$, $\tau_t(\psi) = \psi + \gamma t$, $\gamma \in \mathbb{R}^k$ and $A \in C_0$. We can write the coefficient matrix as

$$A = \begin{pmatrix} a & -b+c \\ b+c & -a \end{pmatrix}$$

where $a, b, c : \mathbb{T}^k \to \mathbb{R}$ are continuous functions. As already explained in Section 2, after passing to polar coordinates (r, θ), we can define the rotation number $\alpha = \alpha(\psi, \gamma, A)$ as the following limit

$$\alpha = \lim_{t \to \infty} \frac{\theta(t)}{t}$$

where θ is the solution of the equation for the polar angle given by

$$\theta' = b(\psi \cdot t) - a(\psi \cdot t)\sin 2\theta + c(\psi \cdot t)\cos 2\theta$$

where we set $\psi \cdot t = \psi + \gamma t$. If we consider a vector of frequencies $\gamma \in \Gamma$, we have that the limit defining α exists and depends neither on the choice of the initial value $\theta(0)$ nor on the choice of the element $\psi \in \mathbb{T}^k$ ([89] and Theorem 4.19). On the other hand, if we take a frequency vector $\gamma \notin \Gamma$, that is a vector whose components satisfy linear relations with coefficients in $\mathbb{Q}$, then the rotation number does not depend on the initial value $\theta(0)$, but may depend on $\psi \in \mathbb{T}^k$. In this case we have that the rotation number is constant on the closure of the orbit passing through ψ, for each $\psi \in \mathbb{T}^k$.

As already mentioned the rotation number has good continuity properties which we now discuss and which will be useful in the proof of the result. If we fix a vector of frequencies $\gamma \in \mathbb{R}^k$, then using the basic Krylov–Bogoliubov method for the construction of invariant measures, one can prove that α is "very" continuous when the matrix A varies. In fact, consider a topological space $\mathcal{E}$ and a continuous map $\mathcal{E} \to C_0$: $E \mapsto A_E$. We have then that $(\psi, E) \mapsto \alpha(\psi, A_E) : \mathbb{T}^k \times \mathcal{E} \to \mathbb{R}$ is jointly continuous with respect to $\psi \in \mathbb{T}^k$ and $E \in \mathcal{E}$. Moreover, there is a sort of continuity with respect to the frequency: suppose that $\gamma \in \Gamma$, and consider a sequence of frequencies $\{\gamma_n\} \subset \mathbb{R}^k$ such that $\gamma_n \to \gamma$ as $n \to \infty$. Observe that it is not required that γ_n belongs to Γ. Then one has, considering a continuous map $E \mapsto A_E$: $\mathcal{E} \to C_0$, that

$$\alpha(\psi, \gamma_n, A_E) \to \alpha(\psi, \gamma, A_E) \equiv \alpha(\psi, A_E) \tag{5.2}$$

when the convergence is uniform on compact subsets of $\mathbb{T}^k \times \mathcal{E}$. Observe that the last equality follows from the fact that, when $\gamma \in \Gamma$, α does not depend on the frequency. The convergence (5.2) can be proved using the Krylov–Bogoliubov method already mentioned. See [113] and [89].

The following theorem relates the presence of an E.D. for a (real) perturbation of the family (5.1) to the monotonicity of the rotation number.

THEOREM 5.2. *Let $A \in C_0$ and consider the family of equations*

$$x' = \left[A\left(\tau_t(\psi)\right) + JE\right]x \tag{5.3}$$

where $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $E \in \mathbb{R}$ and $\psi \in \mathbb{T}^k$. Let $I = (E_1, E_2) \subset \mathbb{R}$ be an open interval. Suppose that $\gamma \in \Gamma$, then the family of Eqs. (5.3) *admits an E.D. over $\mathbb{T}^k$ for all $E \in I$ if and only if the rotation number $\alpha(E)$ is constant on I.*

The proof of the theorem can be found in [59]; see Theorem 4.21. Remember that the rotation number also depends on the ergodic measure defined in our compact metric space, which in the present case is the normalized Lebesgue measure on $\mathbb{T}^k$.

The rotation number of Eq. (5.3), for E fixed and $\psi \in \mathbb{T}^k$, defines a function $E \to \alpha(E)$ which is a monotone, nondecreasing nonnegative function which increases exactly on the spectrum of the corresponding linear operator $L_\psi := J^{-1}[\frac{\mathrm{d}}{\mathrm{d}t} - A(\tau_t(\psi))]$ for each $\psi \in \mathbb{T}^k$. This operator, called the AKNS operator, is viewed as un unbounded self-adjoint operator on the Hilbert space $L^2(\mathbb{R}, \mathbb{C}^2)$. There is an important relation between the dichotomy property and the spectral properties of the operator L_ψ: a complex number E belongs to the resolvent of the operator L_ψ if and only if the corresponding differential system

$x' = [A(\tau_t(\psi)) + JE]x$ admits an E.D. See [79,59] and Proposition 4.15 together with the succeeding discussion.

In the proof of Theorem 5.1, we will perturb our family of equations and arrive at a family of periodic systems. For this reason we recall here some facts and properties of periodic ordinary differential operators which we will use later.

Consider a summable function $a:[0,T) \to \mathrm{sl}(2,\mathbb{R})$; we can view $a(\cdot)$ as a T-periodic function on the whole real axis by extending it in the natural way. Consider then the corresponding periodic differential operator $L = J^{-1}[\frac{\mathrm{d}}{\mathrm{d}t} - a(t)]$ on the space $L^2(\mathbb{R},\mathbb{C}^2)$. Using Floquet theory and extending the arguments of [104], one can obtain the following facts:

1) The spectrum Σ of the self-adjoint operator L is a union (finite or countably infinite) of nondegenerate closed intervals.
2) Consider the rotation number $\alpha(E)$ of the periodic system $x' = [a(t) + JE]x$. Let $I \subset \mathbb{R}$ be a resolvent interval of L; i.e., I is a maximal subinterval of $\mathbb{R} \setminus \Sigma$. Then there exists a unique integer l such that $E \in \operatorname{cls} I$ if and only if $\alpha(E) = \frac{\pi l}{T}$, in particular $\alpha(E)$ is a integer multiple of $\frac{\pi}{T}$.
3) Let $\Phi_E(T)$ be the period matrix of the system $x' = [a(t) + JE]x$ and let $I \subset \mathbb{R}$ be a resolvent interval of L. We have then that for each $E \in I$, there are two linearly independent real eigenvectors $v_\pm$ of $\Phi_E(T)$. (We refer to this as the hyperbolic case). If $\alpha(E) = \frac{2\pi m}{T}$, that is, l is even, with $l = 2m$, $m \in \mathbb{Z}$, then the corresponding eigenvalues $B_\pm$ verify $0 < B_- < 1 < B_+$.

   If E is an endpoint of the interval I, then-up to constant multiple- $\Phi_E(T)$ admits a unique eigenvector v.The corresponding eigenvalue is -1 if the integer l is odd, and is $+1$ if l is even. We refer to this as the parabolic case.
4) Consider E in the interior of the spectrum Σ such that $\alpha(E) = \frac{\pi l}{T}$ for $l \in \mathbb{Z}$. Then we have that $\Phi_E(T) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ if l is odd, while $\Phi_E(T) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ if $l = 2m$ is even. In either case we say that E defines a *closed gap* of L. We refer to this as the elliptic case.
5) Consider a sequence $\{a_n\}$ of T-periodic, $\mathrm{sl}(2,\mathbb{R})$-valued functions such that $\int_0^T |a_n(s) - a(s)|\,\mathrm{d}s \to 0$ as $n \to \infty$. Suppose also that I_n and I are resolvent intervals of the corresponding periodic operators L_n and L where $L_n = J^{-1}[\frac{\mathrm{d}}{\mathrm{d}t} - a_n(t)]$ and $L = J^{-1}[\frac{\mathrm{d}}{\mathrm{d}t} - a(t)]$. Let α_n be the rotation number associated to the operator L_n and suppose that $\alpha(E) = \frac{\pi l}{T}$ if $E \in I$ and $\alpha_n(E) = \frac{\pi l}{T}$ if $E \in I_n$ where l is a fixed integer and $n = 1, 2, \ldots$. Then $I_n \to I$, in the sense that the sequences of endpoints of I_n converge to the endpoints of I, and moreover one has the following result. Let $I_n = (b_n, c_n)$ be a resolvent interval for each $n \geqslant 1$, and suppose that $\overline{E}$ defines a closed gap of the operator L. Suppose also that $\alpha(\overline{E}) = \frac{\pi l}{T} = \alpha_n(E)$ where $E \in I_n, n = 1, 2, \ldots$. Then $b_n \to \overline{E}$ and $c_n \to \overline{E}$ as $n \to \infty$.

Before passing to the proof of Theorem 5.1, we recall some facts about the upper Lyapunov exponent $\beta = \beta(\psi, A)$ of the family of equations $x' = A(\psi + \gamma t)x$. This object is defined by

$$\beta = \lim_{t\to\infty} \frac{1}{t} \log \bigl\| \Phi_\psi(t) \bigr\| \tag{5.4}$$

where $\Phi_\psi(t)$ is the fundamental matrix of the equation defined by $\psi \in \mathbb{T}^k$. It is known that, if $\gamma \in \Gamma$, the limit (5.4) exists and takes on a constant value for Lebesgue-a.a. $\psi \in \mathbb{T}^k$. For fixed $A \in C_0$, if the vector of frequencies $\gamma \in \mathbb{Q}^k$, then the limit defining β exists for all $\psi \in \mathbb{T}^k$ and the function $\psi \to \beta(\psi)$ is continuous. This follows from Floquet theory for periodic systems (remember that $\gamma \in \mathbb{Q}^k$). On the other hand, if $\gamma \notin \mathbb{Q}^k$, the limit (5.4) need not exist for *all* $\psi \in \mathbb{T}^k$. So, when we consider the frequency vector $\gamma \in \Gamma$, we refer to this a.e.-constant limit as the upper Lyapunov exponent of the family $x' = A(\psi + \gamma t)x$. Moreover, β is an upper semi-continuous function. More precisely, let $\mathcal{E}$ be a topological space, and let $E \to A_E; \mathcal{E} \to C_0$ be a continuous function. Let $\gamma \in \Gamma$; then the function $E \to \beta(A_E)$ is upper semi-continuous, but is not in general continuous. In our case $A(\cdot) \in \mathrm{sl}(2,\mathbb{R})$, and therefore we have two "ordinary" Lyapunov exponents $\beta_1 = -\beta_2$ with $\beta_2 \geqslant 0$, and $\beta = \beta_2$.

We recall now the robustness properties of the E.D. in a form convenient for our case. Consider the space $\mathcal{L} = L^\infty(\mathbb{R}, \mathrm{sl}(2,\mathbb{R}))$ and consider the weak-* topology in $\mathcal{L}$: a generalized sequence $\{A_n\}$ converges to A in $\mathcal{L}$ if and only if

$$\int_{-\infty}^{\infty} A_n(t)\varphi(t)\,\mathrm{d}t \to \int_{-\infty}^{\infty} A(t)\varphi(t), \quad \forall \varphi \in L^1(\mathbb{R}).$$

By the Alaoglu theorem, norm-bounded subsets of the space $\mathcal{L}$ are precompact in the weak-* topology (they are also metrizable), so, when we consider the hull of $A \in \mathcal{L}$ given by $Y = \mathrm{cls}\{A_s(\cdot) \mid s \in \mathbb{R}\}$ where $A_s(\cdot) = A(\cdot + s)$ is the translation with respect to the time, we obtain a compact and time-translation invariant subset of $\mathcal{L}$. Observe that there is no need to restrict attention to weak-* compact, translation invariant sets Y which are the hull of some $A \in \mathcal{L}$, but rather we can consider Y to be any weak-* compact, translation invariant subset of $\mathcal{L}$.

Now we state the robustness properties of the E.D. concept as follows. Suppose that the family of equations

$$x' = y(t)x, \quad y \in Y \tag{5.5}$$

admits an E.D. over the compact set Y. According the definition this means that there exist constants $c, \eta > 0$ and a projection valued function $P = P_y : Y \to \mathrm{Proj}(\mathbb{R}^2)$ where $\mathrm{Proj}(\mathbb{R}^2) = \{Q \in \mathbb{M}_2 \mid Q^2 = Q\}$, such that, if $\Phi_y(t)$ is the fundamental matrix solution of (5.5), then the following inequalities hold:

$$\begin{aligned}
&\left\|\Phi_y(t)P_y\Phi_y(s)^{-1}\right\| \leqslant c\mathrm{e}^{-\eta(t-s)}, \quad t \geqslant s, \\
&\left\|\Phi_y(t)(I - P_y)\Phi_y(s)^{-1}\right\| \leqslant c\mathrm{e}^{\eta(t-s)}, \quad t \leqslant s.
\end{aligned}$$

The perturbation theorem for the E.D. [137] states the following

THEOREM 5.3. *Let $Y_0 \subset \mathcal{L}$ a fixed norm-bounded weak-* compact, translation invariant set. Let $Y \subset Y_0$ be a weak-* compact, translation invariant set such that Eqs.* (5.5) *admit* *an E.D. over Y. There exists a weak-* neighborhood U of Y in Y_0 so that, if $\tilde{Y} \subset U$ is another weak-* compact and translation invariant set, then the corresponding Eqs.* (5.5) *have an E.D. over $\tilde{Y}$.*

Observe that we consider neighborhoods with respect to a distance defined by a metric compatible with the weak-* topology considered in the space $Y_0 \subset \mathcal{L}$.

From Theorem 5.3 there follows the openness in $\mathbb{R}^k \times C_0$ of the set $\{(\gamma, A) \in \mathbb{R}^k \times C_0 \mid x' = A(\psi + \gamma t)x$ has an E.D. over $\mathbb{T}^k\}$. See [44] for more details.

5.1. *Proof of Theorem 5.1 (Density result)*

In this subsection we will sketch the proof of Theorem 5.1; as already mentioned, the proof is a fairly lengthy argument by contradiction. We start from the following

HYPOTHESIS 5.4. There exists an open subset $V \subset \mathbb{R}^k \times C_0$ such that, if the pair $(\gamma, A) \in V$, then the corresponding differential system

$$x' = A(\psi + \gamma t)x$$

does not have an E.D. over $\mathbb{T}^k$.

Remember that, if the frequency vector $\gamma \in \Gamma$, Hypothesis 5.4 means that E.D. fails for *all* $\psi \in \mathbb{T}^k$. If the vector $\gamma \in \mathbb{R}^k \setminus \Gamma$, Hypothesis 5.4 means that for at least one point $\psi \in \mathbb{T}^k$ the E. D. property does not hold. Further, if there is no E.D. for (at least) one point $\psi \in \mathbb{T}^k$, then E.D. fails to hold for each point in the orbit closure $\mathrm{cls}\{\tau_t(\psi) \mid t \in \mathbb{R}\}$ of the point as well.

We can consider w.l.g. an open set $V \subset \mathbb{R}^k \times C_0$ of the following type: $V = V_1 \times V_2$ where $V_1 \subset \mathbb{R}^k$ and $V_2 \subset C_0$ are open sets. Consider now $\Gamma_1 = \Gamma \cap V_1$. Using a result of [44], see Theorem 5.5 below, we have that there exists a residual subset W_* of $\Gamma_1 \times C_0$ such that, if the pair $(\gamma, A) \in W_*$, then the Lyapunov exponent $\beta(\gamma, A)$ of the corresponding family $x' = A(\psi + \gamma t)x$ equals zero. Let us note that there are recent results of Bochi and Viana [19] which generalize this statement; for our purposes the result of [44] is quite sufficient.

Recall here that the presence of an E.D. for a family of equations implies that the corresponding Lyapunov exponent is positive, but that the positivity of the Lyapunov exponent is not enough to guarantee the presence of an E.D. The exact result (here we are considering the continuous case) proved in [44] is the following

THEOREM 5.5. *Suppose that* $0 \leqslant r < 1$ *where* $r \in \mathbb{R}$ *and let* $C^r = C^r(\mathbb{T}^k, \mathrm{sl}(2, \mathbb{R}))$ *be the set of C^r- mappings from $\mathbb{T}^k$ to* $\mathrm{sl}(2, \mathbb{R})$ *with the usual topology. There is a residual subset W_* of $\Gamma \times C^r$ with the following property*: *if $w = (\gamma, A) \in W_*$, then either equations $x' = A(\psi + \gamma t)x$ admit an E.D. or the Lyapunov exponent $\beta(w)$ equals zero.*

According to Theorem 5.5, we can then find a frequency $\gamma_* \in \Gamma \cap V_1$ and a function $A_* \in V_2$, such that the corresponding Lyapunov exponent $\beta(\gamma_*, A_*)$ equals zero. From now on we fix γ_* and A_*. What we are going to do now is to perturb the frequencies of our system of equations to obtain a certain rational frequency vector and a corresponding family of periodic linear systems. It is a fact that the Lyapunov exponent $\beta = \beta(\gamma, A)$ is a semicontinuous function with respect to the coefficient matrix A and with respect to various topologies; in the proof we apply the following result. For a proof we refer to [44].

LEMMA 5.6. *Consider the sequence* $\{\gamma_n\} \subset \mathbb{Q}^k$ *of frequency vectors with rationally dependent components such that* $\gamma_n \to \gamma_*$. *Let* $A_n \to A_*$ *as* $n \to \infty$ *in* C_0. *Indicate with* $\beta(\psi, \gamma_n, A_n)$ *the Lyapunov exponent of the periodic system* $x' = A_n(\psi + \gamma_n t)x$. *Then* $\beta(\psi, \gamma_n, A_n) \to \beta(\gamma_*, A_*) = 0$ *where the convergence is uniform with respect to* $\psi \in \mathbb{T}^k$.

After these considerations and remarks we can proceed with the proof of Theorem 5.1; we are considering a pair $(\gamma_*, A_*) \in \Gamma_1 \times V_2$ such that the corresponding system does not have an E.D. and the related Lyapunov exponent $\beta(\gamma_*, A_*) = 0$. Consider now the family of equations

$$x' = \left[A_*(\psi + \gamma_* t) + JE\right]x, \quad E \in \mathbb{R},$$

and its corresponding rotation number $\alpha_*(E)$; we have then that the function $E \to \alpha_*(E)$ is continuous and it strictly increases at $E = 0$ by the monotonicity of the rotation number and by the fact that for $E = 0$ there is no E.D. There exists $E_0 > 0$ such that, if $|E| < 2E_0$, then $A_* + JE \in V_2$. The rotation number is strictly increasing at $E = 0$: precisely, for each $\delta > 0$, $\alpha(\delta) - \alpha(0) > 0$ and $\alpha(0) - \alpha(-\delta) > 0$.

Now, let $\gamma_n \in \mathbb{Q}^k$ be a sequence of frequency vectors with rational coefficients such that $\gamma_n \to \gamma_*$. We choose $\gamma_n = (\frac{p_1^{(n)}}{q_1^{(n)}}, \dots, \frac{p_k^{(n)}}{q_k^{(n)}})$ and require that each pair $q_i^{(n)}$, $q_j^{(n)}$ of denominators be relatively prime for $1 \leqslant i < j \leqslant k$. This condition implies that, if $T_n = q_1^{(n)} q_2^{(n)} \cdots q_k^{(n)}$, then the flow $\psi + \gamma_n t$ is T_n-periodic. Let $\rho > 0$ a positive number. As proved in [77] (see also [44] and [45]) it is the case that there exists a number $N_0 \geqslant 1$ such that, to each $n \geqslant N_0$, there correspond an integer m and a continuous function $E_n : \mathbb{T}^k \to \mathbb{R}$ such that

(i) The function E_n is constant along each γ_n-orbit, that is

$$E_n(\psi) = E_n(\psi + \gamma_n t) \quad (\psi \in \mathbb{T}^k, 0 \leqslant t \leqslant T_n).$$

(ii) $\|E_n\| < \rho$.

(iii) The rotation number $\alpha_n(\psi)$ of the periodic system

$$x' = \left[A_*(\psi + \gamma_n t) + JE_n(\psi)\right]x \tag{5.6}$$

satisfies

$$\alpha_n(\psi) = \frac{2\pi m}{T_n} \quad (\psi \in \mathbb{T}^k).$$

We obtain E_n by using the monotonicity and continuity properties of the rotation number to arrive to a family of T_n-periodic systems for which the rotation number is an even multiple of $\frac{\pi}{T_n}$. In this case, according to the discussion made in the previous subsection for the periodic differential operator L associated to a T-periodic function with value in the algebra $\mathrm{sl}(2, \mathbb{R})$, only two situations may occur (remember that we are imposing Hypothesis 5.4, so E.D. does not hold over $\mathbb{T}^k$ in the open set $V \subset \mathbb{R}^k \times C_0$).

- If the (periodic) operator $L_\psi = J^{-1}[\frac{\mathrm{d}}{\mathrm{d}t} - A_*(\psi + \gamma_n t)]$ admits a resolvent interval I_ψ wherein the rotation number $\alpha(\psi, \gamma_n, A_* + JE)$ has value $\frac{2\pi m}{T_n}$, then define $E_n(\psi)$ to be the center point of I_ψ.
- If on the other hand there is a number $E_0 \in \mathbb{R}$ which defines a closed gap of L_ψ with $\alpha(\psi, \gamma_n, A_* + JE_0) = \frac{2\pi m}{T_n}$, then set $E_n(\psi) = E_0$. One verifies that E_n is continuous.

Now we continue discussing the proof of Theorem 5.1. Choose a positive number ρ in such a way that the ball in C_0 of center A_* and radius ρ is contained in the open set V_2. Moreover determine N_0, $n \geqslant N_0$, m and E_n as above (for the construction of them see the discussion in [44]) and let $\psi^0 = (0, 0, \dots, 0)$ be the origin of $\mathbb{T}^k$ with respect to the angular coordinates $(\psi_1, \dots, \psi_k)$. Write $\bar{A}(\psi) = A_*(\psi) + JE_n(\psi)$ $(\psi \in \mathbb{T}^k)$, so that $\bar{A}(\psi) \in V_2$ and for the corresponding periodic system of equations $x' = \bar{A}(\psi + \gamma_n t)x$ E.D. does not hold over $\mathbb{T}^k$, and the corresponding rotation number is given by $\alpha(\psi, \gamma_n, \bar{A}) = \frac{2\pi m}{T_n}$. Adapting a method of Moser [110], as done in [45], we find an $A_0 \in C_0$ with $\|A_0 - A_*\| < \rho$ and such that the corresponding system along the orbit of ψ^0 with frequency γ_n has an E.D. (e.g. such that the system $x' = A_0(\psi^0 + \gamma_n t)x$ admits an E.D.).

We discuss in more detail the method of Moser. Observe that the condition $\alpha(\psi, \gamma_n, \bar{A}) = \frac{2\pi m}{T_n}$ induces a restriction on the behavior of the period matrix $\Phi_\psi(T_n)$ of the (periodic) system $x' = \bar{A}(\psi + \gamma_n t)x$. We reformulate this restriction geometrically: we have already listed, in the previous section , the relations between the rotation number and the spectral properties of a T-periodic differential operator (see (1), …, (5) on page 253). The facts that the rotation number $\alpha(\psi, \gamma_n, \bar{A}) = \frac{2\pi m}{T_n}$ and the absence of an E.D. for the corresponding periodic system $x' = \bar{A}(\psi + \gamma_n t)x$, imply that $\Phi_\psi(T_n)$ verifies one of the following conditions:

(i) $\Phi_\psi(T_n) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, the identity matrix. This is the elliptic case.

(ii) $\Phi_\psi(T_n)$ has eigenvalue 1 and a unique one-dimensional eigenspace (parabolic case).

We cannot have the hyperbolic case because the system does not admit an E.D. Geometrically these conditions mean that the logarithm of the periodic matrix solution $\Phi_\psi(T_n)$ lies on the surface of a certain cone C whose construction was made and used in [111] for the study of the quasi-periodic Schrödinger operator. Introduce then the following cone C: observe that each matrix $A \in \mathrm{sl}(2, \mathbb{R})$ can be written in the form $\left(\begin{smallmatrix} a & b-c \\ b+c & -a \end{smallmatrix}\right)$ where a, b, c are real numbers. Let C defined by the set $\{(a, b, c) \in \mathbb{R}^3 \mid c^2 = a^2 + b^2\}$, then C is just the set of matrices A which admit zero as an eigenvalue. We of course identify the triple $(a, b, c) \in \mathbb{R}^3$ with the matrix $\left(\begin{smallmatrix} a & b-c \\ b+c & -a \end{smallmatrix}\right)$. Now, let $D_* \subset \mathrm{SL}(2, \mathbb{R})$ be the set of all matrices Φ which are parabolic, or hyperbolic and have positive eigenvalues. There is an open set $D \subset \mathrm{SL}(2, \mathbb{R})$ which contains D_* and a single-valued, real-analytic branch $\ln : D \to \mathrm{sl}(2, \mathbb{R})$ of the matrix logarithm such that $\ln I = \ln\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Moreover $\ln(D)$ contains the cone C and its exterior $\{(a, b, c) \mid a^2 + b^2 > c^2\}$. Writing this fact in terms of the T-period matrix solution $\Phi_\psi(T)$ one has immediately that:

$\Phi_\psi(T)$ is elliptic if and only if $\ln\Phi_\psi(T) = (0,0,0)$;

$\Phi_\psi(T)$ is parabolic if and only if $\ln\Phi_\psi(T)$ lies on the surface of the cone C but is not the vertex $(0,0,0)$;

$\Phi_\psi(T)$ is hyperbolic if and only if $\ln\Phi_\psi(T)$ lies outside the cone C.

We have thus that in our case the logarithm $\ln\Phi_\psi(T_n)$ of the period matrix solution $\Phi_\psi(T_n)$ of the periodic system $x' = \bar{A}(\psi + \gamma_n t)x$ lies on the surface of the cone C. It is now intuitively reasonable that we can perturb A to a matrix function $A_0 \in C_0$ such that $x' = A_0(\psi^0 + \gamma_n t)x$ has an E.D. See [43] and [44] for details.

Let us write $\gamma = \gamma_n$ and then consider the rational frequency vector $\gamma = (\frac{p_1}{q_1}, \dots, \frac{p_k}{q_k})$ for which the properties (i)–(iii) hold. Recall also that each pair q_i, q_j is relatively prime ($1 \leqslant i < j \leqslant k$). The proof of Theorem 5.1 now proceeds as follows. We extend A_0 along the various faces of a "quasi-section" $S \subset \mathbb{T}^k$ to the linear flow on $\mathbb{T}^k$ with frequency vector γ. In the end we obtain a function $A \in C_0$ with $\|A - A_*\| < \rho$ and such that $x' = A(\psi + \gamma t)x$ has an E.D. for all $\psi \in S$ and hence – by the choice of S – for all $\psi \in \mathbb{T}^k$. The presence of an E.D. contradicts the Hypothesis 5.4 and thus Theorem 5.1 is proved.

Let us explain briefly the construction of the "quasi-section" S. Let S be the following set: $S = \{\psi = (\psi_1, \dots, \psi_k) \in \mathbb{T}^k \mid \psi_k = 0 \text{ and } 0 \leqslant \psi_i \leqslant \frac{2\pi}{q_i} (1 \leqslant i \leqslant k-1)\}$. (Recall that $\psi_1, \dots, \psi_k$ are angular coordinates mod 2π on $\mathbb{T}^k$). Observe that the S lies in the hyperplane $\{\psi \in \mathbb{T}^k \mid \psi_k = 0\}$ and that $\psi^0 = (0, 0, \dots, 0)$ is one of its vertices. Moreover $S = \bigcup_{p=0}^{k-1} S_p$ where S_p is the union of all p-dimensional faces of S where each face is a p-dimensional rectangle.

If $k = 2$, the quasi-section S is simply the segment

$$S = \left\{ (\psi_1, \psi_2) \in \mathbb{T}^2 \;\middle|\; 0 \leqslant \psi_1 \leqslant \frac{2\pi}{q_1}, \psi_2 = 0 \right\}$$

and for $k = 3$ the quasi-section S is the rectangle

$$S = \left\{ (\psi_1, \psi_2, \psi_3) \in \mathbb{T}^3 \;\middle|\; 0 \leqslant \psi_1 \leqslant \frac{2\pi}{q_1}, 0 \leqslant \psi_2 \leqslant \frac{2\pi}{q_2}, \psi_3 = 0 \right\}$$

in the ψ_1–ψ_2 plane. We can thus consider $\psi^0 = (0,0,0)$ as the lower-left hand vertex of the rectangle S.

Moreover observe that, if $\psi \in \mathbb{T}^k$, then its orbit $\{\psi + \gamma t \mid 0 \leqslant t \leqslant T_n\}$ intersects S. If an orbit intersects S in an interior point (i.e., a point $\psi \in S \setminus \bigcup_{p=0}^{k-2} S_p$), then ψ is the unique point of intersection between the orbit and S. On the other hand, if an orbit intersects S in a noninterior point, then the orbit intersects S in other noninterior points as well. Observe that we have in $\mathbb{T}^k$ the flow defined by the rational frequency γ which permutes the various p faces whose union is S_p. We can then consider noninterior points of S as "boundary" points. Denote the set of such boundary points of S as ∂S.

In the detailed proof carried out in [44] the desired function A for which there is an E.D. for the corresponding system is obtained altering A_* many times to obtain a contradiction with Hypothesis 5.4. First a map from the boundary ∂S into the unit circle S^1 is defined, then, after proving that this map is homotopic to a constant for a suitable perturbation of A_*, this map is extended to the whole of S. Then, using Tietze's theorem, the map is extended to the space $\mathbb{T}^k$. In the end one obtains a continuous function defined on the entire space $\mathbb{T}^k$ and belonging to the open set V_2 for which the corresponding family of systems admits an E.D.

We include here the details for $k = 3$. In this case the quasi-section is the rectangle

$$S = \left\{ (\psi_1, \psi_2, \psi_3) \in \mathbb{T}^3 \,\middle|\, 0 \leqslant \psi_1 \leqslant \frac{2\pi}{q_1}, 0 \leqslant \psi_2 \leqslant \frac{2\pi}{q_2}, \psi_3 = 0 \right\}$$

in the ψ_1–ψ_2 plane. We already noticed that we can consider $\psi^0 = (0, 0, 0)$ as the lower-left hand vertex of the rectangle S.

We choose $\gamma_n = \gamma$ and A_0 such that the system $x' = A_0(\psi^0 + \gamma t)x$ has an E.D. and we can also assume (by choosing $\gamma = \gamma_n$ even closer to γ_*) that, for each $\psi \in S$, the resolvent interval $I_m(\psi)$ (if such an interval exists) corresponding to the (constant) value $\frac{2\pi m}{T_n}$ of the rotation number has length so small that, if $E : \mathbb{T}^k \to \mathbb{R}$ is a continuous function with $E(\psi) \in I_m(\psi)$ for any $\psi \in \mathbb{T}^k$, then the function $A_0 + JE$ belongs to V_2.

Let ψ^1, ψ^2, ψ^3 be the other vertices of the rectangle S in the $\psi_1 - \psi_2$ plane, where $\psi^1 = (0, \frac{2\pi}{q_2}, 0)$, $\psi^2 = (\frac{2\pi}{q_1}, 0, 0)$. We have then that $S = S_0 \cup S_1 \cup S_2$ where $S_0 = \{\psi^0, \psi^1, \psi^2, \psi^3\}$; $S_1 = \overline{\psi^0\psi^1} \cup \overline{\psi^1\psi^3} \cup \overline{\psi^2\psi^3} \cup \overline{\psi^0\psi^2}$ and S_2 is the interior of the rectangle which has vertices $\{\psi^0, \psi^1, \psi^2, \psi^3\}$. Observe that, when we consider the flow on $\mathbb{T}^3$ generated by the rational frequency γ, the segment $\overline{\psi^0\psi^1}$ is mapped to $\overline{\psi^2\psi^3}$ and the segment $\overline{\psi^0\psi^2}$ to the segment $\overline{\psi^1\psi^3}$. Let $t_1, t_2 \in [0, T_n)$ be the unique times such that $\tau_{t_i}(\psi^0) = \psi^0 + \gamma t_i = \psi^i$ $(i = 1, 2)$. We have then that $\psi^1 = \tau_{t_1}(\psi^0)$, $\psi^2 = \tau_{t_2}(\psi^0)$, and the other vertex $\psi^3 = \tau_{t_1+t_2}(\psi^0)$.

Before defining the desired function A (which will contradict Hypothesis 5.4), we make a digression to introduce the technique we will use. Let C_T be the set of T-periodic, continuous functions $\mathcal{A} : \mathbb{R} \to \mathrm{sl}(2, \mathbb{R})$. If $\mathcal{A} \in C_T$, consider the system

$$x' = \big[\mathcal{A}(t) + JE\big]x \tag{5.7}$$

where $E \in \mathbb{R}$ and $J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Write $\mathcal{A} = \begin{pmatrix} a & -b+c \\ b+c & -a \end{pmatrix}$ and after considering the polar angle $\theta = \operatorname{Arg} x$, we have that θ verifies the equation

$$\theta' = \big(E + b(t)\big) - a(t)\sin 2\theta + c(t)\cos 2\theta. \tag{5.8}$$

If $\phi \in \mathbb{R}$, let $v_\phi = \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$. If $x(t)$ is the solution of Eq. (5.7) with the initial condition $x(0) = v_\phi$, then the solution $\theta(t)$ of Eq. (5.8) such that $\theta(0) = \phi$ is a continuous determination of $\arg x(t)$. Let m a fixed integer. It can be shown that there is a unique $E_m \in \mathbb{R}$ such that the solution $\theta(t, E_m)$ of Eq. (5.8) with $E = E_m$ and verifying the initial condition $\theta(0, E_m) = \phi$ satisfies

$$\theta(T, E_m) - \phi = 2\pi m.$$

Consider now the corresponding periodic operator $L = L_m = J^{-1}[\frac{\mathrm{d}}{\mathrm{d}t} - \mathcal{A}(t)]$ acting on $L^2(\mathbb{R}, \mathbb{C}^2)$. According to the discussion carried out on page 253, one has just two possibilities: either E_m is an element of the closure of a resolvent interval I_m of L, in which case the rotation number of the operator equals $\frac{2\pi m}{T}$ (parabolic case), or E_m defines a closed gap (elliptic case).

Define now $F^{\phi}(\mathcal{A}) = \mathcal{A} + JE_m$. Then F^{ϕ} is a nonlinear transformation from C_T into C_T which is continuous in the topology of the uniform convergence on $[0, T]$. Observe that, for each $\tilde{A} \in C_T$, v_{ϕ} is an eigenvector of the period matrix of the system $x' = F^{\phi}(\mathcal{A})(t)x$ (which is by definition $x' = [\mathcal{A}(t) + JE_m]x$). Moreover, if the rotation number of the system $x' = \mathcal{A}(t)x$ is already equal to $\frac{2\pi m}{T}$, then the transformation $F^{\phi}(\mathcal{A})$ is obtained from $\mathcal{A}$ by adding a translate JE_m such that $E_m = 0$ if the real number 0 defines a closed gap of L, and $|E_m|$ is no greater than the length of I_m if 0 is in the resolvent interval I_m of L.

We have finally that, for each $\phi \in \mathbb{R}$, the family (depending on the real parameter s) $F_s^{\phi} : C_T \to C_T$ defined by

$$F_s^{\phi}(\mathcal{A}) = \mathcal{A} + sJE_m \quad (0 \leqslant s \leqslant 1) \tag{5.9}$$

is an homotopy joining F^{ϕ} to the identity transformation on C_T. We will obtain the desired A – for which we will have an E.D. – using the functional F^{ϕ}.

Let r be a real parameter varying in the interval $[0, t_1 + t_2] = [0, t_1] \cup [t_1, t_1 + t_2]$. Let P_r be the projection defined by the presence of an E.D. for the translated equation $x' = A_0(\psi^0 + \gamma(t + r))x$ for $r \in [0, t_1 + t_2]$ for the translation invariance of the E.D. property. Extend the parameter r by letting it decrease from $t_1 + t_2$ to t_2, then from t_2 to 0. Consider a parametrization of the boundary ∂S of the section S given by the following map $\psi = \psi(r)$ with $0 \leqslant r \leqslant t_1 + t_2$.

$$\partial S := \begin{cases} \psi\colon\ [0, t_1] \to [\psi^0, \psi^1] \text{ is linear with } \psi(0) = \psi^0 \text{ and } \psi(t_1) = \psi^1, \\ \psi : [t_1, t_1 + t_2] \to [\psi^1, \psi^3] \text{ is linear with } \psi(t_1) = \psi^1 \\ \quad \text{and } \psi(t_1 + t_2) = \psi^3, \\ \psi : [t_1 + t_2, t_2] \to [\psi^3, \psi^2] \text{ is linear with } \psi(t_1 + t_2) = \psi^3, \\ \quad \psi(t_2) = \psi^2, \\ \psi : [t_2, 0] \to [\psi^2, \psi^0] \text{ is linear with } \psi(t_2) = \psi^2 \text{ and } \psi(0) = \psi^0. \end{cases}$$

Of course we have put $[\psi^0, \psi^1] = \overline{\psi^0\psi^1}, \ldots, [\psi^0, \psi^2] = \overline{\psi^0\psi^2}$ for the segments whose union is the boundary ∂S.

Now, we associate to each value of the parameter r a continuously determined argument $\theta(r)$ of the line in the plane $\mathbb{R}^2$ which is equal to the image of the projection P_r given by the presence of an E.D. Recall that when there is an E.D. the corresponding projection P has a one-dimensional range and a one dimensional kernel which are lines in the plane $\mathbb{R}^2$. For each $\psi \in \partial S$, let $v(\psi)$ be the unit vector in the plane $\mathbb{R}^2$ with polar angle $\theta(r)$ where r corresponds to the point ψ on the boundary ∂S. We have then that the map $v : \partial S \to S^1 : \psi \to v(\psi)$ is homotopic to a constant map. Here we indicate with S^1 the unit circle on the plane.

We define $A_1(\psi) = A_0(\psi) + JE(\psi)$ for $\psi \in \partial S$ as follows: if $\psi \in [\psi^0, \psi^1] \cup [\psi^1, \psi^3]$, let $A'_{\psi}(t) = A_0(\psi + \gamma t)$. Then let $A_{\psi}^{(1)}(t) = F^{\theta(r)}(A'_{\psi})(t)$ where r is the parameter which corresponds to the point ψ and $0 \leqslant t \leqslant T_n$. Finally set $A_1(\psi + \gamma t) = A_{\psi}^{(1)}(t)$ for $0 \leqslant t \leqslant$

T_n. We observe that A_1 is defined on a subset of the torus $\mathbb{T}^3$ containing the whole boundary ∂S. We also have that, if $\psi \in [\psi^0, \psi^1] \cup [\psi^1, \psi^3]$, then the solution $x_\psi(t) := x(t)$ of the corresponding (periodic) system $x' = A_\psi^{(1)}(t)x$ with initial condition $x_\psi(0) = v(\psi)$ has the property that $x(T_n)$ is a positive constant multiple of $v(\psi)$.

Next define $v_1(\psi) = v(\psi)$ if $\psi \in [\psi^0, \psi^1] \cup [\psi^1, \psi^3]$. If we consider $\psi \in [\psi^2, \psi^3]$, it follows that, by the properties of the flow generated on $\mathbb{T}^3$ by the rational frequency γ, there exists a unique $\tilde{\psi} \in [\psi^0, \psi^1]$ such that $\psi = \tilde{\psi} + \gamma t_2$. Define $v_1(\psi) = \frac{x_{\tilde{\psi}}(t_2)}{|x_{\tilde{\psi}}(t_2)|}$. Observe that the map $v_1(\psi)$ has values in the unit circle S^1. Moreover, if $\psi \in [\psi^0, \psi^2]$, there exists a unique $\tilde{\psi} \in [\psi^1, \psi^3]$ such that $\psi = \tilde{\psi} - \gamma t_1$ and, as before, we put $v_1(\psi) = \frac{x_{\tilde{\psi}}(-t_1)}{|x_{\tilde{\psi}}(-t_1)|}$. We have thus that the map v_1 is defined on the whole boundary ∂S with values in the unit circle S^1.

Consider now the map $f : \partial S \to C_T$ defined by $f(\psi)(t) = A_0(\psi + \gamma(t + r))$ where r is the parameter value corresponding to the point ψ on the boundary ∂S. Moreover we consider $g : \partial S \to C_T$ defined by $g(\psi)(t) = A_1(\psi + \gamma t)$. Using the contractibility of the space C_T, we have then that the maps f and g are homotopic via a homotopy $H : [0, 1] \times \partial S \to C_T : (s, \psi) \to H(s, \psi)$. This homotopy can be modified in a convenient way. First, using the transformation F^θ where $\phi = \theta$, we can arrange the homotopy in such a way that, for each $0 \leqslant s \leqslant 1$ and each $\psi \in [\psi^0, \psi^1] \cup [\psi^1, \psi^3]$, the corresponding system of equations $x' = H(s, \psi)(t)x$ has a solution $x_\psi(t)$ with $x_\psi(0) = v(\psi)$ and $x_\psi(T_n) = cv(\psi)$ where c is a positive number. Second, we can also impose some translation invariance relations.

(i) If $\psi \in [\psi^2, \psi^3]$ and $\tilde{\psi} \in [\psi^0, \psi^1]$ is the unique point such that $\psi = \tilde{\psi} + \gamma t_2$, then

$$H(s, \psi)(t) = H(s, \tilde{\psi})(t + t_2).$$

(ii) If $\psi \in [\psi^0, \psi^2]$ and if $\tilde{\psi} \in [\psi^1, \psi^3]$ is the unique point such that $\psi = \tilde{\psi} - \gamma t_1$, we arrange that

$$H(s, \psi)(t) = H(s, \tilde{\psi})(t - t_1).$$

Now, using the translation relations one can show that v and v_1 are homotopic as maps from the boundary ∂S to the unit circle S^1. We have hence that v_1 is homotopic to a constant map too and admits a continuous extension to all of S. Indicate with $v_1 : S \to S^1$ such an extension. When $\psi \in \operatorname{int} S$, let θ_ψ be a continuous determination of the polar angle of the unit vector $v_1(\psi)$. We now extend, using Tietze's theorem, the function A_1 to a continuous map (again called A_1) defined on all the torus. Then we redefine $A_1(\psi + \gamma t)$ to be $F^{\theta(\psi)}(A_\psi^{(1)})(t)$ $(\psi \in S,\ 0 \leqslant t \leqslant T_n)$. One can arrange that $A_1 \in V_2$. For each $\psi \in S$ (and hence for each ψ in the torus) one has a unit vector $v_1(\psi) \in \mathbb{R}^2$ varying continuously with ψ such that, if $x_\psi(t)$ is the solution of the system $x' = A_1(\psi + \gamma t)x$ with $x_\psi(0) = v_1(\psi)$, then $x_\psi(T_n) = cv_1(\psi)$ where $c > 0$. Now, if $c > 1$ for all $\psi \in \mathbb{T}^3$, then the corresponding system of equations has an E.D. over $\mathbb{T}^3$.

The last thing to do is alter the function A_1 for the last time, using a triangularization technique [50], to arrange that the corresponding Lyapunov exponent $\beta(\psi) > 0$ for all

$\psi \in \mathbb{T}^3$. Equivalently, $c > 1$ for all $\psi \in \mathbb{T}^3$. Now, putting everything together, we have arrived at a contradiction with Hypothesis 5.4 and then Theorem 5.1 is finally proved.

For $k > 3$, an analogous construction can be carried out. One constructs A starting from A_0 in an inductive way, modifying A_0 first on S_0, then on $S_1, S_2, \ldots, S_{k-1}$. There is no essential difficulty in passing from S_p to S_{p+1} for $p \geqslant 2$, for the following reason. Suppose that A has been defined on S_p. One wishes to define A along one of the $p+1$-faces $\mathcal{F}$ making up S_{p+1}. The boundary $\partial\mathcal{F}$ of $\mathcal{F}$ is a topological p-sphere which is a union of p-faces. When one carries out the appropriate analogue of the construction outlined above, one arrives at a crucial step in which one must extend a continuous map $v_1 : \partial\mathcal{F} \to S^1$ to a continuous map (again called v_1) from $\mathcal{F}$ to S^1. If $p \geqslant 2$ this is always possible because v_1 is homotopic to a constant map; see, e.g. Dugundji [37].

As one can see from the arguments given above, it seems quite difficult to obtain the same result considering the Hölder space $C^r(\mathbb{T}^k)$ with $r \geqslant 1$ instead of C_0. In view of the results obtained and the techniques used in the paper [45], we conjecture that Theorem 5.1 can be generalized to the case of $0 < r < 1$ but not to the case $r \geqslant 1$. We have already stated the first result on page 255 (Theorem 5.5), the second is the following [45]:

THEOREM 5.7. *Let $C^r = C^r(\mathbb{T}^k, \mathrm{sl}(2,\mathbb{R}))$ There is a dense subset $P_r \subset W_r = \Gamma \times C^r$ with the following property: if $w = (\gamma, A) \in P_r$, then the Lyapunov exponent $\beta = \beta(w)$ of the corresponding family of equations $x' = A(\psi + \gamma t)x$ is strictly positive.*

Both of these results hold for $0 \leqslant r < 1$, and this, together with the nature of their proofs, leads one to suspect that Theorem 5.1 also holds when $0 \leqslant r < 1$.

Observe that, using the Baire category theorem, we can give a "fixed-frequency" version of Theorem 5.1:

COROLLARY 5.8. *There is a residual subset $\Gamma_* \subset \Gamma$ such that, if $\gamma \in \Gamma_*$, then the set $\mathcal{D}_\gamma = \{A \in C_0 \mid x' = A(\psi + \gamma t)x$ has an E.D. over $\mathbb{T}^k\}$ is open and dense in C_0.*

Note first that the set Γ is a Baire space and thus it admits a metric with respect to which it is a complete metric space.

PROOF. The openness of the set $\mathcal{D}_\gamma$ in C_0 follows from Theorem 5.3 for all $\gamma \in \mathbb{R}^k$. It remains to prove the density result. Take a countable dense subset of C_0 given by $\{A_n \mid n = 1, 2, \ldots\}$ and indicate with $B_{j,n}$ the open ball in C_0 centered in A_n with radius $1/j$ where $j = 1, 2, \ldots$. Let $\Gamma_{n,j} = \{\gamma \in \Gamma \mid \text{for each} A \in B_{n,j}$ the system $x' = A(\psi + \gamma t)$ does not have an E.D. over $\mathbb{T}^k\}$. For Theorem 5.3, it follows that each $\Gamma_{n,j}$ is closed in Γ. Moreover Theorem 5.1 implies that each $\Gamma_{n,j}$ has no interior in Γ. If we put $\Gamma_* = \Gamma \setminus \bigcup\{\Gamma_{n,j} \mid n, j \geqslant 1\}$ we have that Γ_* is residual in Γ and it verifies the condition of corollary. □

References

[1] M. Ablowitz, D. Kaup, A. Newell and H. Segur, *The inverse scattering transform: Fourier analysis for non-linear systems*, Stud. Appl. Math. **53** (1974), 249–315.

[2] M. Alber and Y. Fedorov, *Algebraic geometrical solutions for certain evolution equations and Hamiltonian flows on nonlinear subvarieties of generalized Jacobians*, Inverse Problems **17** (2001), 1017–1042.

[3] M.Alber and Y. Fedorov, *Wave solutions of evolution equations and Hamiltonian flows on nonlinear subvarieties of generalized Jacobians*, J. Phys. A: Math. Gen. **33** (2000), 8409–8425.

[4] A. Alonso, C. Núñez and R. Obaya, *Complete guiding sets for a class of almost periodic differential equations*, J. Differential Equations **208** (2005), 124–146.

[5] L. Amerio, *Soluzioni quasi-periodiche, o limitate, di sistemi differenziali nonlineari quasi-periodici, o limitati*, Ann. Math. Pura Appl. **39** (1955), 97–119.

[6] J. Andres, A. Bersani and F. Grande, *Hierarchy of almost periodic function spaces*, Rend. Mat., in press.

[7] D. Anosov and A. Katok, *New examples in smooth ergodic theory*, Trans. Moscow Math. Soc. **23** (1970), 1–35.

[8] L. Arnold, *Random Dynamical Systems*, Springer-Verlag, Berlin (1998).

[9] L. Arnold, Dinh Cong Nguyen and V. Oseledets, *Jordan normal form for linear cocycles*, Random Oper. Stochastic Equations **7** (1999), 303–358.

[10] F.V. Atkinson, *Discrete and Continuous Boundary Value Problems*, Academic Press Inc., New York (1964).

[11] R. Beals, D. Sattinger and J. Szmigielski, *Multipeakons and the classical moment problem*, Advances in Math. **154**, 229–257.

[12] M. Bebutov, *On dynamical systems in the space of continuous functions*, Boll. Moscov. Univ. Matematica **2** (1941), 1–52.

[13] K. Bjerklöv, *Dynamical properties of quasi-periodic Schrödinger equations*, Ph.D. Thesis, Royal Institute of Technology (2003).

[14] K. Bjerklöv, *Positive Lyapunov exponent and minimality for a class of one-dimensional quasi-periodic Schrödinger equations*, Ergodic Theory Dynam. Systems **25** (2005), 1015–1045.

[15] K. Bjerklöv, D.Damanik and R. Johnson, *Lyapunov exponents of continuous Schrödinger cocycles over irrational rotations*, Ann. Mat. Pura Appl. **187** (2008), 1–6.

[16] K. Bjerklöv and T. Jäger, *Rotation numbers for quasi-periodically forced interval maps – mode locking vs. strict monotonicity*, Preprint (2006).

[17] K. Bjerklöv and R. Johnson, *Minimal subsets of projective flows*, Discrete Contin. Dynam. Systems, in press.

[18] F. Blanchard, E. Glasner, S. Kolyada and A. Maass, *On Li–Yorke pairs*, J. Reine Angew. Math. **547** (2002), 51–68.

[19] J. Bochi and M. Viana, *The Lyapunov exponents of generic volume preserving and symplectic systems*, Ann. Math. **161** (2005), 1423–1485.

[20] S. Bochner, *A new approach to almost periodicity*, Proc. Nat. Acad. Sci. USA **48** (1962), 2039–2043.

[21] H. Bohr, *Fastperiodische Funktionen*, Springer-Verlag, Berlin (1932).

[22] R. Bowen, *Entropy for group endomorphisms ond homogeneous spaces*, Trans. Amer. Math. Soc. **153** (1971), 401–414.

[23] J. Bourgain and S. Jitomirskaya, *Continuity of the Liapunov exponent for quasiperiodic operators with analytic potentials*, J. Statist. Phys. **108** (2002), 1203–1218.

[24] I. Bronstein, *Extensions of Minimal Transformation Groups*, Sijthoff and Noosdhoff, Alphen aan den Rijn (1979).

[25] R. Cameron, *Almost periodic properties of bounded solutions of linear differential equations with almost periodic coefficients*, J. Math. Phys. **15** (1936), 73–81.

[26] V. Chulaevsky and Y. Sinai, *Anderson localization for the* $1-D$ *discrete Schrödinger operator with two-frequency potential*, Comm. Math. Phys. **125** (1989), 91–112.

[27] C. Conley and M. Miller, *Asymptotic stability without uniform stability*, J. Differential Equations **1** (1965), 333–336.

[28] W.A. Coppel, *Disconjugacy*, Lecture Notes in Math., Vol. 220, Springer-Verlag, Berlin (1971).

[29] W.A. Coppel, *Dichotomies in Stability Theory*, Lecture Notes in Math., Vol. 629, Springer-Verlag, Berlin (1978).

[30] D. Damanik and D. Lenz, *A condition of Boshernitzan and uniform convergence in the multiplicative ergodic theorem*, Duke Math. J. **133** (2006), 95–123.

[31] C. De Concini and R. Johnson, *The algebraic-geometric AKNS potentials*, Ergodic Theory Dynam. Systems **7** (1987), 1–24.

[32] E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill (1955).

[33] F. Colonius and W. Kliemann, *The Dynamics of Control*, Birkhäuser, Basel (2000).

[34] W. Craig, *The trace formula for Schrödinger operators on the line*, Comm. Math. Phys. **126** (1989), 379–407.

[35] B. Dubrovin, V. Matveev and S. Novikov, *Non-linear equations of Korteweg–de Vries type, finite-zone linear operators, and Abelian varieties*, Russ. Math. Surveys **31** (1976), 59–146.

[36] N. Dunford and J. Schwartz, *Linear Operators*, Vols. 1,2,3, Interscience Publishers (1958).

[37] P. Duren, *Theory of* H^p *Spaces*, Academic Press, New York (1970).

[38] L.H. Eliasson, *Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation*, Comm. Math. Phys. **146** (1992), 447–482.

[39] L.H. Eliasson, *Almost reducibility of linear quasi-periodic systems*, Proc. Sympos. Pure Math. **69** (2001), 679–705.

[40] R. Ellis, *Lectures on Topological Dynamics*, Benjamin, New York, 1969.

[41] R. Ellis and R. Johnson, *Topological dynamics and linear differential systems*, J. Differential Equations **44** (1982), 21–39.

[42] R. Ellis, S. Glasner and L. Shapino, *Proximal-isometric (PI) flows*, Adv. in Math. **17** (1975), 213–260.

[43] R. Fabbri, Genericità dell 'Iperbolicità nei sistemi differenziali lineari di dimensione due, Ph. D. Thesis, Università di Firenze (1997).

[44] R. Fabbri and R. Johnson, *Genericity of exponential dichotomy for two-dimensional quasi-periodic systems*, Ann. Math. Pura Appl. **178** (2000), 175–193.

[45] R. Fabbri and R. Johnson, *On the Lyapunov exponent of certain* $\mathrm{SL}(2,\mathbb{R})$*-valued cocycles*, Differential Equations and Dynamical Systems **7** (3) (July 1999), 349–370

[46] R. Fabbri, R. Johnson and C. Núñez, *Disconjugacy and the rotation number for linear, non-autonomous Hamiltonian systems*, Ann. Mat. Pura Appl., in press.

[47] R. Fabbri, R. Johnson and C. Núñez, *On the Yakubovich Frequency Theorem for linear non-autonomous control processes*, Discrete Contin. Dynam. Systems **9** (2003), 677–704.

[48] R. Fabbri, R. Johnson and C. Núñez, *The rotation number for non-autonomous linear Hamiltonian systems I: basic properties*, Z. Angew. Math. Phys. **54** (2003), 484–502.

[49] R. Fabbri, R. Johnson and C. Núñez, *The rotation number for non-autonomous linear Hamiltonian systems II: the Floquet coefficient*, Z. Angew. Math. Phys. **54** (2003), 652–676.

[50] R. Fabbri, R. Johnson and R. Pavani, *On the nature of the spectrum of the quasi-periodic Schrödinger operator*, Nonlin. Anal.: Real World Problems **3** (2002), 37–59.

[51] A. Fathi and M. Herman, *Existence de difféomorphismes minimaux*, Soc. Math. de France Astérisque **49** (1977), 37–59.

[52] J. Favard, *Sur les équations différentielles a coefficients presque-périodiques*, Acta Math. **51** (1927), 31–81.

[53] J.D. Fay, *Theta Functions on Riemann Surfaces*, Lecture Notes, Vol. 352, Springer-Verlag, Berlin.

[54] A. Fink, *Almost Periodic Differential Equations*, Lecture Notes in Math., Vol. 377, Springer-Verlag, Berlin (1974).

[55] H. Furstenberg, *Strict ergodicity and transformations on the torus*, Amer. J. Math. **83** (1961), 573–601.

[56] C. Gardner, J. Greene, M. Kruskal and R. Miura, *Methods for solving the Korteweg–de Vries equation*, Phys. Rev. Lett. **19** (1967), 1095-1097.

[57] J. S. Geronimo and R. Johnson, *Rotation number associated with difference equations satisfied by polynomials orthogonal on the unit circle*, J. Differential Equations **132** (1996), 140–178.

[58] F. Gesztesy and H. Holden, *Algebro-geometric solutions of the Camassa–Holm hierarchy*, Rev. Math. Iberoamericano **19** (2003), 73–142.

[59] R. Giachetti and R. Johnson, *The Floquet exponent for two-dimensional linear systems with bounded coefficients*, J. Math. Pures Appl. **65** (1986), 93–117.

[60] D.J. Gilbert and D.B. Pearson, *On subordinacy and analysis of the spectrum of one-dimensional Schrödinger operator*, J. Math. Anal. Appl. **128** (1987), 30–56.

[61] S. Glasner and B. Weiss, *On the construction of minimal skew-products*, Israel J. Math. **34** (1979), 321–336.

[62] M. Goldstein and W. Schlag, *Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions*, Ann. Math. **154** (2001), 155–203.

[63] W. Gottschalk and G. Hedlund, *Topological Dynamics*, AMS Colloquium Publications, Vol. 36, Amer. Math. Soc., Providence, RI (1955).

[64] R. Greene and S. Krantz, *Function Theory of One Complex Variable*, second ed., Graduate Stud. Math., Vol. 40, Amer. Math. Soc. (2001).

[65] P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, Reprint of the 1978 original, John Wiley & Sons, Inc., New York (1994).

[66] P. Hartman, *A characterization of the spectra of one-dimensional wave equations*, Amer. J. Math. **71** (1949), 915–920.

[67] P. Hartman, *Ordinary Differential Equations*, John Wiley and Sons, New York (1964).

[68] M. Herman, *Une méthod pour minorer les exponents teh Lyapunov et quelques exemples monstrant le caractère local d'un theorèm d'Arnold e de Moser sur le tore di dimension 2*, Comment. Math. Helv. **58** (1983), 453–502.

[69] P. Imkeller and C. Lederer, *The cohomology of stochastic and random differential equations, and local linearization of stochastic flows*, Stoch. Dynam. **2** (2002), 131–159.

[70] T. Jäger, *Quasiperiodically forced interval maps with negative Schwarzian derivative*, Nonlinearity **16** (2003), 1239–1255.

[71] T. Jäger and J. Stark, *Towards a classification of quasiperiodically forced circle homeomorphisms*, J. London Math. Soc. **73** (2006), 727–744.

[72] T. Jambois, *Seminar on Degeneration of Algebraic Varieties*, (I.A.S.) Fall 1970.

[73] R. Johnson, *A linear almost periodic equation with an almost automorphic solution*, Proc. Amer. Math. Soc. **82** (1981), 417–422.

[74] R. Johnson, *Almost periodic functions with unbounded integral*, Pacific J. Math. **87** (1980), 347–362.

[75] R. Johnson, *An example concerning the geometric significance of the rotation number-integrated density of states*, Lecture Notes in Math., Vol. 1186, Springer-Verlag, Berlin (1986), 216–226.

[76] R. Johnson, *Bounded solutions of scalar, almost periodic linear systems*, Illinois J. Math. **25** (1981), 632–643.

[77] R. Johnson, *Cantor spectrum for the quasi-periodic Schrödinger equation*, J. Differential Equations **91** (1991), 88–110.

[78] R. Johnson, *Ergodic theory and linear differential equations*, J. Differential Equations **28** (1978), 23–34.

[79] R. Johnson, *Exponential dichotomy, rotation number, and linear differential operators with bounded coefficients*, J. Differential Equations **61** (1986), 54–78.

[80] R. Johnson, *Lyapunov exponents for the almost periodic Schrödinger equation*, Illinois J. Math. **28** (1984), 397–419.

[81] R.Johnson, *m-functions and Floquet exponents for linear differential systems*. Ann. Mat. Pura Appl. **147** (1987), 211–248.

[82] R. Johnson, *Minimal functions with unbounded integral*, Israel J. Math. **31** (1978), 133–141.

[83] R. Johnson, *On a Floquet theory for two-dimensional, almost periodic linear systems*, J. Differential Equations **37** (1980), 184–205.

[84] R. Johnson, *On almost periodic linear systems of Millionscikov and Vinograd*, J. Math. Anal. Appl. **85** (1982), 452–460.

[85] R. Johnson, *On the Sato–Segal–Wilson solutions of the K-dV equation*, Pacific J. Math. **132** (1988), 343–355.

[86] R. Johnson, *The recurrent Hill's equation*, J. Differential Equations **46** (1982), 165–194.

[87] R. Johnson, *Two-dimensional, almost periodic systems with proximal and recurrent behavior*, Proc. Amer. Math. Soc. **82** (1981), 417–422.

[88] R. Johnson, *Un'equazione lineare quasi-periodica con una proprietá inconsueta*, Boll. UMI **6** (1983), 115–121.

[89] R. Johnson and J. Moser, *The rotation number for almost periodic potentials*, Commun. Math. Phys. **84** (1982), 403–438.

[90] R. Johnson and M. Nerurkar, *Exponential dichotomy and rotation number for linear Hamiltonian systems*, J. Differential Equations **108** (1994), 201–216.

[91] R. Johnson and M. Nerurkar, *Stabilization and random linear regulator problem for random linear control processes*, J. Math. Anal. Appl. **197** (1996), 608–629.

[92] R. Johnson, S. Novo and R. Obaya, *An ergodic and topological approach to disconjugate linear Hamiltonian systems*, Illinois J. Math. **45** (2001), 1045–1079.

[93] R. Johnson, S. Novo and R. Obaya, *Ergodic properties and Weyl M-functions for linear Hamiltonian systems*, Proc. Roy. Soc. Edinburgh Sect. A **130** (2000), 1045–1079.

[94] R. Johnson, K. Palmer and G. Sell, *Ergodic properties of linear dynamical systems*, SIAM J. Math. Anal. **18** (1987), 1–33.

[95] R. Johnson and L. Zampogni, *Description of the algebro-geometric Sturm–Liouville coefficient*, J. Differential Equations, in press.

[96] R. Johnson and L. Zampogni, *On the inverse Sturm–Liouville problem*, Discrete Contin. Dyn. Systems **18** (2007), 405–428.

[97] R. Johnson and L. Zampogni, *Uniform limits of algebro-geometric Sturm–Liouville coefficients and infinite order hierarchies of K-dV and CH type*, preprint.

[98] À. Jorba and C. Simó, *On the reducibility of linear differential equations with quasiperiodic coefficients*, J. Differential Equations **98** (1992), 111–124.

[99] À. Jorba, C. Núñez, R. Obaya and C. Tatjèr, *Old and new results for SNAs on the real line*, Int. J. Bifur. Chaos, in press.

[100] O. Knill, *Positive Lyapunov exponents for a dense set of bounded measurable $SL(2,\mathbb{R})$-cocycles*, Ergodic Theory Dynam. Systems **12** (1992), 319–331.

[101] S. Kotani, *Lyapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators*, Taniguchi Symp. SA, Katata (1982), 225–247.

[102] R. Krikorian, *Reducibility, differentiable rigidity and Lyapunov exponents for quasi-periodic cocycles on $\mathbb{T}\times SL(2,\mathbb{R})$*, preprint.

[103] T. Li and J. Yorke, *Period three implies chaos*, Amer. Math. Monthly **82** (1975), 985–992.

[104] W. Magnus and S. Winkler, *Hill's Equations*, Interscience Publishers, New York–London–Sidney (1966).

[105] R. Mañe, *Quasi-Asonov diffeomorphisms and hyperbolic manifolds*, Trans. Amer. Math. Soc. **229** (1977), 351–370.

[106] V. Marchenko and V. Ostrovsky, *Approximation of periodic by finite-zone potentials*, Selecta Math. Sovietica **6** (1987), 101–136.

[107] L. Marcus and R. Moore, *Oscillation and disconjugacy for linear differential equations with almost periodic coefficients*, Acta Math. **96** (1956), 99–123.

[108] V. Millionščikov, *Proof of the existence of irregular systems of linear differential equations with almost periodic coefficients*, Differ. Uravn. **5** (1969), 1979–1983.

[109] V. Millionščikov, *On the typicality of almost reducible systems with almost periodic coefficients* (Russian), Differ. Uravn. **14** (1978), 634–636.

[110] J. Moser, *An example of a Schrödinger operator with almost periodic potential and nowhere dense spectrum*, Helv. Math. Acta **56** (1981), 198–224.

[111] J. Moser and J. Pöschel, *An extension of a result by Dinaburg and Sinai on quasi-periodic potentials*, Comment. Math. Helv. **59** (1984), 39–85.

[112] D. Mumford, *Tata Lectures on Theta*, Vols. 1,2,3, Birkhäuser (1983).

[113] V. Nemytskii and V. Stepanov, *Qualitative Theory of Differential Equations*, Princeton University Press, Princeton, NJ (1960).

[114] M. Nerurkar, *On the construction of smooth ergodic skew-products*, Ergodic Theory Dynam. Systems **8** (1988), 311–326.

[115] M. Nerurkar and H. Sussmann, *Construction of minimal cocycles arising from specific differential equations*, Israel J. Math. **100** (1997), 309–326.

[116] Nguyen Dinh Cong, *Topological Dynamics of Random Dynamical Systems*, Clarendon Press, Oxford (1997).

[117] Nguyen Dinh Cong, *A generic bounded linear cocycle has simple Lyapunov spectrum*, Ergodic Theory Dynam. Systems **25** (2005), 1775–1797.

[118] S. Novikov, *The periodic Korteweg–de Vries problem*, Func. Anal. Appl. **8** (1974), 54–66.

[119] S. Novo, C. Núñez and R. Obaya, *Almost automorphic and almost periodic dynamics for quasi monotone nonautonomous functional differential equations*, J. Dynam. Differential Equations **17** (2005), 689–719.

[120] S. Novo, C. Núñez and R. Obaya, *Ergodic properties and rotation number for linear Hamiltonian systems*, J. Differential Equations **148** (1998), 148–185.

[121] S. Novo and R. Obaya, *An ergodic and topological approach to almost periodic bidimensional linear systems*, Cont. Math., Vol. 215, Amer. Math. Soc., Providence, RI (1998), 299–322.

[122] S. Novo and R. Obaya, *An ergodic classification of bidimensional linear systems*, J. Dynam. Differential Equations **8** (1996), 373–406.

[123] S. Novo and R. Obaya, *On the dynamical behavior of an almost periodic linear system with an ergodic two-sheet studied by R. Johnson*, Israel J. Math. **105** (1998), 235–249.

[124] S. Novo, R. Obaya and A. Sans, *Attractor minimal sets for cooperative and strongly convex delay differential equations*, J. Differential Equations **208** (2005), 86–123.

[125] S. Novo, R. Obaya and A. Sans, *Exponential stability in non-autonomous delayed equations with applications to neural networks*, Proc. Roy. Soc. London Ser. A **2061** (2005), 2767–2783.

[126] R. Ortega and M. Tarallo, *Almost periodic linear differential equations with non-separated solutions*, J. Func. Anal. **237** (2006), 402–426.

[127] R. Ortega and M. Tarallo, *Almost periodic upper and lower solutions*, J. Differential Equations **193** (2003), 343–358.

[128] R. Ortega and M. Tarallo, *Masserás theorem for quasi-periodic differential equations*, Topol. Math. Nonlin. Anal. **19** (2002), 39–61.

[129] V. Oseledets, *A multiplicative ergodic theorem, Lyapunov characteristic number for dynamical systems*, Trans. Moscow Math. Soc. **19** (1968), 197–231.

[130] K. Palmer, *A perturbation theorem for exponential dichotomies*, Proc. Roy. Soc. Edinburgh Sect. A **106** (1987), 25–37.

[131] R.R. Phelps, *Lectures on Choquet's Theorem*, American Book Co., New York (1966).

[132] J. Poeschl and E. Trubowitz, *Inverse Spectral Theory*, Pure and Applied Mathematics, Vol. 120, Academic Press (1987).

[133] J. Puig, *A nonperturbative Eliasson's reducibility theorem*, Nonlinearity **19** (2006), 355–376.

[134] J. Puig, *Cantor spectrum for the almost Mathieu operator*, Commun. Math. Phys. **244** (2004), 297–309.

[135] J. Puig and C. Simó, *Analytic families of reducible linear quasi-periodic differential equations*, Ergodic Theory Dynam. Systems **26** (2006), 481–524.

[136] W. Rudin, *Real and Complex Analysis*, McGraw-Hill (1970).

[137] R.J. Sacker and G.R. Sell, *A spectral theory for linear differential systems*, J. Differential Equations **27** (1978), 320–358.

[138] R.J. Sacker and G.R. Sell, *Existence of dichotomies and invariant splittings for linear differential systems I*, J. Differential Equations **15** (1974), 429–458.

[139] R.J. Sacker and G.R. Sell, *Lifting properties in skew-product flows with applications to differential equations*, Mem. Amer. Math. Soc., Vol. 190, Amer. Math. Soc., Providence, RI (1977).

[140] S. Schwartzman, *Asymptotic cycles*, Ann. of Math. **66** (2) (1957), 270–284.

[141] G. Segal and G. Wilson, *Loop groups and equations of KdV type*, Publ. IHES **61** (1985), 5–65.

[142] J. Selgrade, *Isolated invariant sets for flows on vector bundles*, Trans. Amer. Math. Soc. 203 (1975), 350–390.

[143] G.R. Sell, *Lectures on Topological Dynamics and Differential Equations*, Van Nostrand–Reinhold, London (1971).

[144] W. Shen and Y. Yi, *Almost Automorphic and Almost Periodic Dynamics in Skew-Product Semiflows*, Mem. Amer. Math. Soc., Vol. 647, Amer. Math. Soc., Providence, RI (1998).

[145] W. Shen and Y. Yi, *On minimal sets of scalar parabolic equations with skew-product structures*, Trans. Amer. Math. Soc. **347** (1995), 4413–4431.

[146] C.L. Siegel, *Topics in Complex Function Theory*, Vols. 1,2,3, Wiley-Interscience (1969).

[147] B. Simon, *Kotani theory for one-dimensional stochastic Jacobi matrices*, Commun. Math. Phys. **89** (1983), 297–309.

[148] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2,* Amer. Math. Soc. Colloq. Publ., Vol. 54, Providence, RI (2005).

[149] T.C. Titchmarsh, *Eigenfunction Expansion Associated with Second Order Differential Equations*, Oxford (1946).

[150] W. Veech, *Almost automorphic functions on groups*, Amer. J. Math. **87** (1965), 719–751.

[151] W. Veech, *Topological dynamics*, Bull. Amer. Math. Soc. **83** (1977), 775–830.

[152] R. Vinograd, *A problem suggested by N.P. Erugin*, Differ. Uravn. **11** (1975), 632–638.

[153] H. Weinberger, *A First Course on Partial Differential Equations*, Dover Publications (1965).

[154] Y. Yi, *On almost automorphic oscillations*, Difference and Differential Equations, Proc 7th ICDEA, Fields Int. Comm., Vol. 42, Amer. Math. Soc., Providence, RI (2004), 75–99.

[155] G. Zames, *Feedback and optimal sensitivity: model reference transformation, multiplicative seminorms, and approximate inverses*, IEEE Trans. Autom. Control, AC-26, 301–320, 1981.

[156] L. Zampogni, *On algebro-geometric solutions of the Camassa–Holm hierarchy*, J. Adv. Nonlin. Stud., 7 (2007), 845–880.

[157] L. Zampogni, *On the Inverse Sturm–Liouville problem and the Camassa–Holm equation*, Ph.D. Thesis, Universitá di Firenze (2006).

[158] V. Zhikov and B. Levitan, *Favard Theory*, Russ. Math. Surveys **32** (1977), 129–180.

CHAPTER 3

Complex Differential Equations

Ilpo Laine
Department of Mathematics, University of Joensuu, Finland
E-mail: *ilpo.laine@joensuu.fi*

Contents

Preface . . . 271
1. Local theory of complex differential equations . . . 271
    1.1. Elementary background from several complex variables . . . 271
    1.2. Basic local existence and uniqueness results . . . 273
    1.3. Some remarks on the behavior of local solutions . . . 276
    1.4. Local behavior of solutions of $w' = P(z, w)/Q(z, w)$ . . . 279
2. Linear differential equations in the complex plane . . . 284
    2.1. Homogeneous linear differential equations with polynomial coefficients . . . 287
    2.2. Second order homogeneous linear differential equations with entire coefficients . . . 292
    2.3. Higher order homogeneous linear differential equations . . . 296
    2.4. Homogeneous linear differential equations with meromorphic coefficients . . . 301
    2.5. Non-homogeneous linear differential equations . . . 304
3. Linear differential equations in the unit disc . . . 307
    3.1. Classical results on zeros of solutions . . . 307
    3.2. Complex function spaces in the unit disc . . . 312
    3.3. Growth estimates for homogeneous linear differential equations . . . 315
    3.4. Equations with fast growing solutions . . . 317
    3.5. Equations with slowly growing solutions . . . 319
4. Non-linear differential equations in a complex domain . . . 322
    4.1. Introductory remarks . . . 322
    4.2. Malmquist type theorems . . . 324
    4.3. Riccati differential equation . . . 330
    4.4. Schwarzian differential equations . . . 332
    4.5. Painlevé differential equations . . . 334
    4.6. Briot–Bouquet differential equations . . . 342
    4.7. Algebraic differential equations . . . 343
    4.8. Algebraic differential equations and differential fields . . . 348
5. Algebroid solutions of complex differential equations . . . 350
    5.1. Introduction to algebroid functions . . . 350
    5.2. Malmquist type theorems for algebroid functions . . . 353

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 4
Edited by F. Battelli and M. Fečkan
© 2008 Elsevier B.V. All rights reserved

- 5.3. Algebroid solutions of algebraic differential equations . . . 354
- 5.4. Algebroid solutions of linear differential equations . . . 355

Acknowledgement . . . 356

References . . . 356

Preface

This presentation is mainly devoted to considering growth and value distribution of meromorphic solutions of complex differential equations, both in the complex plane as well as in the unit disc. Our main emphasis is in results and considerations not to be found in the key references of the present field, namely [98] and [107] below. Therefore, results achieved during the last fifteen years or so play a special role, hence in particular recent investigations concerning linear differential equations in the unit disc. Conjectures and open problems as well as comments about defectively understood specific areas have been scattered in the text, hopefully prompting additional interest.

The book reference list [63,74,77,99,141] and [189] includes items known to us that cover, at least partially, our present point of view. Of course, these references include a lot of material not to be found in this presentation, due to lack of space. Additional book references of more specialized type can be found in the reference lists of specific chapters below.

After a short introduction into the local theory of complex differential equations, we first collect essentials about growth and value distribution of solutions of linear differential equations in the complex plane, followed by a survey of the very recent investigations about solutions of linear differential equations in the unit disc. The next chapter, non-linear differential equations in a complex domain, takes about one third of the whole presentation, including Malmquist type theorems, Riccati, Schwarzian, Painlevé and Briot–Bouquet differential equations, algebraic differential equations in general and their connections to differential fields. The final chapter offers a short presentation of algebroid solutions of complex differential equations, a field that definitely needs more studies.

1. Local theory of complex differential equations

In this section, we recall basic results of the local theory of complex differential equations. These results serve as a device to the existence of local solutions to a given differential equation in the complex plane, usually under given initial condition(s). In addition to the local existence, the global existence of solutions in a domain $G \subset \mathbf{C}$ may often be deduced. The solutions to be looked for are typically analytic or meromorphic or some slight extensions of these such as algebroid functions. For the convenience of the reader, we have included some proofs, not immediately available, in this section.

1.1. *Elementary background from several complex variables*

Let $\Omega \subset \mathbb{C}^n$ be an open set, equipped with the topology obtained from the standard Euclidean topology by identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$. For a point $(z_1, \dots, z_n) \in \Omega$, let $x_1, \dots, x_{2n}$ denote its corresponding real coordinates, with the proviso that $z_j = x_{2j-1} + ix_{2j}$ for $j = 1, \dots, n$. Consider now a function $f : \Omega \to \mathbb{C}$ continuously differentiable with respect to $x_1, \dots, x_{2n}$. Defining the complex partial derivatives by

$$\frac{\partial f}{\partial z_j} := \frac{1}{2}\left(\frac{\partial f}{\partial x_{2j-1}} - \mathrm{i}\frac{\partial f}{\partial x_{2j}}\right), \tag{1.1.1}$$

$$\frac{\partial f}{\partial \bar{z}_j} := \frac{1}{2}\left(\frac{\partial f}{\partial x_{2j-1}} + \mathrm{i}\frac{\partial f}{\partial x_{2j}}\right), \tag{1.1.2}$$

the total differential

$$\mathrm{d}f := \sum_{j=1}^{2n} \frac{\partial f}{\partial x_j}\,\mathrm{d}x_j \tag{1.1.3}$$

may be written in the form

$$\mathrm{d}f = \partial f + \bar{\partial} f := \sum_{j=1}^{n} \frac{\partial f}{\partial z_j}\,\mathrm{d}z_j + \sum_{j=1}^{n} \frac{\partial f}{\partial \bar{z}_j}\,\mathrm{d}\bar{z}_j. \tag{1.1.4}$$

Recall now

DEFINITION 1.1.1. A continuously differentiable function $f : \Omega \to \mathbb{C}$ is *analytic* (*holomorphic*) in Ω, if $\bar{\partial} f = 0$ in Ω.

REMARK. If $n = 1$ and we denote $f = u + \mathrm{i}v$, $z = x + \mathrm{i}y$, Definition 1.1.1 reduces to the familiar Cauchy–Riemann equations in the classical sense.

In the local theory, the following several variables counterpart to the classical Taylor expansion is frequently needed. Using the notion of a *polydisc*

$$B(\beta, r) := \prod_{j=1}^{n} B(\beta_j, r_j), \tag{1.1.5}$$

where $\beta_j \in \mathbb{C}$, $r_j > 0$ for $j = 1, \ldots, n$ and $B(\beta_j, r_j) \subset \mathbb{C}$ is the open disc of radius r_j, centered at β_j, we prove the following theorem, see [85], p. 27:

THEOREM 1.1.2. *Suppose* $f : B(\beta, r) \to \mathbb{C}$ *is analytic. Then* f *may be uniquely represented as an absolutely converging power series*

$$f(z_1, \ldots, z_n) = \sum_{k_1,\ldots,k_n=0}^{\infty} a_{k_1,\ldots,k_n}(z_1 - \beta_1)^{k_1} \cdots (z_n - \beta_n)^{k_n}, \tag{1.1.6}$$

converging uniformly in compact subsets of $B(\beta, r)$. *If* $|f(z_1, \ldots, z_n)| \leqslant M$ *in* $\prod_{j=1}^{n} B(\beta_j, \rho_j)$, *where* $0 < \rho_j < r_j$, $j = 1, \ldots, n$, *then*

$$|a_{k_1,\ldots,k_n}| \leqslant \frac{M}{r_1^{k_1} \cdots r_n^{k_n}}. \tag{1.1.7}$$

By Theorem 1.1.2, it is not too difficult to prove the following theorem (due to Hartogs), which is of frequent (implicit) use in practical considerations of complex differential equations. For the proof of this theorem, see [85], pp. 28–29.

THEOREM 1.1.3. *If $f:\Omega\to\mathbb{C}$ is analytic in each of the complex variables $z_1,\dots,z_n$ separately, when all other variables are given arbitrary fixed values, then f is analytic in the sense of Definition* 1.1.1.

1.2. *Basic local existence and uniqueness results*

The classical existence result for solutions of first order differential equations reads as follows:

THEOREM 1.2.1. *Suppose $f:\Omega\to\mathbb{C}$ in a domain $\Omega\subset\mathbb{C}^2$ is analytic and $(z_0,w_0)\in\Omega$. Then there exists $R>0$ such that the initial value problem*

$$w'(z)=f\bigl(z,w(z)\bigr),\quad w(z_0)=w_0, \tag{1.2.1}$$

admits a unique analytic solution $w(z)$ in $B(z_0,R)$ such that $(z,w(z))\in\Omega$ for all $z\in B(z_0,R)$.

There are several classical ways to prove this theorem. We recall here three of the most standard of these proofs: (1) the method of successive approximations, (2) the Banach fixed point method and (3) the Cauchy majorant method. See [74,77] for more details.

The previous theorem may easily be extended to the corresponding existence and uniqueness result for solutions of systems of first order differential equations:

THEOREM 1.2.2. *Suppose that $f_j:\Omega\to\mathbb{C}$, $j=1,\dots,n$, are analytic in a domain $\Omega\subset\mathbb{C}^{n+1}$ and that $(z_0,w_{1,0},\dots,w_{n,0})\in\Omega$. Suppose also that for some $r>0$,*

$$\overline{B}(z_0,w_{1,0},\dots,w_{n,0};r):=\overline{B}(z_0,r)\times\prod_{j=1}^{n}\overline{B}(w_{j,0},r)\subset\Omega \tag{1.2.2}$$

and that $|f_j(z,w_1,\dots,w_n)|\leqslant M$, $j=1,\dots,n$, for all $(z,w_1,\dots,w_n)\in\overline{B}(z_0,w_{1,0},\dots,w_{n,0};r)$. Then the initial value problem

$$w_j'(z)=f_j\bigl(z,w_1(z),\dots,w_n(z)\bigr),\quad w_j(z_0)=w_{j,0},\quad j=1,\dots,n, \tag{1.2.3}$$

admits a unique analytic solution $(w_1(z),\dots,w_n(z))$ in some neighborhood U of z_0 such that

$$B\left(z_0,r\left(1-\exp\left(-\frac{1}{(n+1)M}\right)\right)\right)\subset U. \tag{1.2.4}$$

REMARK. The previous theorem has been formulated for a proof by the Cauchy majorant method. An interested reader may find a complete proof in [74], pp. 19–22.

A slight modification of the majorant method may be applied to prove that local solutions of a system of linear differential equations remain analytic as long as all coefficients are analytic, see [74], pp. 23–25:

THEOREM 1.2.3. *Let* $b_{jk}(z)$, $a_j(z)$, $j, k = 1, \dots, n$, *be analytic functions in* $B(z_0, r)$. *Then the initial value problem*

$$\begin{cases} w_j' = \displaystyle\sum_{k=1}^{n} b_{jk}(z) w_k + a_j(z), \quad j = 1, \dots, n, \\ w_1(z_0) = w_{1,0}, \dots, w_n(z_0) = w_{n,0} \end{cases} \tag{1.2.5}$$

possesses a unique solution $(w_1(z), \dots, w_n(z))$ *analytic in* $B(z_0, r)$.

Theorem 1.2.3 has two immediate corollaries of frequent use in practical considerations. To formulate these consequences, let $p_1(z), \dots, p_n(z), a(z)$ be analytic in $B(z_0, r)$ and consider the initial value problem

$$\begin{cases} w^{(n)} + p_1(z) w^{(n-1)} + \cdots + p_n(z) w = a(z), \\ w(z_0) = w_{0,0}, w'(z_0) = w_{1,0}, \dots, w^{(n-1)}(z_0) = w_{n-1,0}. \end{cases} \tag{1.2.6}$$

Clearly, Eq. (1.2.6) may be represented as the system of linear differential equations,

$$\begin{cases} w_j' = w_{j+1}, \quad j = 1, \dots, n-1, \\ w_n' = -p_n(z) w_1 - \cdots - p_1(z) w_n + a(z). \end{cases} \tag{1.2.7}$$

Therefore, an immediate consequence of Theorem 1.2.3 is

COROLLARY 1.2.4. *The initial value problem* (1.2.6) *possesses a unique solution* $w(z)$ *analytic in* $B(z_0, r)$.

In particular, this implies the following important global consequence:

COROLLARY 1.2.5. *If the coefficients of a linear differential equation* (1.2.6) *are entire functions, then all solutions of* (1.2.6) *are entire functions as well.*

In the case of non-linear differential equations, the conclusion corresponding to Corollary 1.2.4, i.e. proving that all local solutions remain meromorphic as long as the coefficients are analytic, is a complicated problem in general. Standard cases where global existence of meromorphic solutions can be proved, are the Riccati differential equation and the Painlevé differential equations. Concerning the Riccati differential equation, we get

THEOREM 1.2.6. *Suppose that the coefficients* a_0, a_1, a_2 *of the Riccati differential equation*

$$w' = a_0(z) + a_1(z)w + a_2(z)w^2 \tag{1.2.8}$$

are analytic in $B(z_0, r)$. *Then all local solutions of* (1.2.8) *in a simply connected domain* $\Omega \subset B(z_0, r)$ *admit a meromorphic continuation into* $B(z_0, r)$.

The basic idea to prove this theorem is to fix three distinct complex numbers c_1, c_2, c_3, and consider the following pair of differential equations:

$$\begin{cases} u' = a_0(z)v + a_1(z)u, \\ v' = -a_2(z)u. \end{cases} \tag{1.2.9}$$

By Theorem 1.2.3, (1.2.9) possesses a unique solution (u_j, v_j), $j = 1, 2, 3$, analytic in $B(z_0, r)$ and satisfying the initial conditions

$$u_j(z_0) = c_j, \quad v_j(z_0) = 1, \quad j = 1, 2, 3. \tag{1.2.10}$$

By elementary computation, $w_j = u_j/v_j$, $j = 1, 2, 3$, are three distinct solutions of (1.2.8), meromorphic in $B(z_0, r)$. To complete the proof, it suffices to recall that an arbitrary local solution w of a Riccati differential equation may be represented in terms of these three solutions, or actually in terms of their restrictions to the domain to be considered, see [9], Proposition 2.1. By the standard uniqueness theorem of meromorphic functions, w possesses the required meromorphic continuation.

COROLLARY 1.2.7. *Suppose the coefficients of the Riccati differential equation* (1.2.8) *are entire functions. Then all local solutions of* (1.2.8) *admit a meromorphic continuation into* $\mathbb{C}$.

Concerning the Painlevé differential equations, see Section 4.5.

Cauchy majorant method may easily be applied to prove the classical fact that local solutions depend analytically on initial values and parameters. For simplicity, we restrict ourselves to formulate the result for one parameter only. However, it is no difficulty indeed to permit several (finitely many) parameters in the assertion.

THEOREM 1.2.8. *Consider a parameter-depending initial value problem*

$$w' = f(z, w, \lambda), \quad w(z_0, \lambda_0) = w_0, \tag{1.2.11}$$

where f *is analytic in a domain* $\Omega \subset \mathbb{C}^3$, *containing the point* (z_0, w_0, λ_0). *Then* (1.2.11) *admits a unique analytic solution* $w(z, \lambda)$ *in a domain* $\Delta \subset \mathbb{C}^2$ *such that* $(z, w(z), \lambda) \in \Omega$ *for all* $(z, \lambda) \in \Delta$. *In particular, the dependence on the parameter* λ *is analytic in a neighborhood of* λ_0.

For the proof, see [74], pp. 26–27.

1.3. *Some remarks on the behavior of local solutions*

This section is devoted to considering some complements for the preceding basic existence and uniqueness results. We also consider analytic continuation of local solutions along curves on the complex plane and, secondly, possibilities of obtaining some information of solutions around such points in $\mathbb{C}^2$, where $f(z, w)$ has a singularity. Concerning the proofs in this section, see [22].

THEOREM 1.3.1. *Suppose* $f : \Omega \to \mathbb{C}$ *is analytic in a domain* $\Omega \subset \mathbb{C}^2$ *and* $(z_0, w_0) \in \Omega$. *Let further* U *be a neighborhood of* z_0. *Then the differential equation*

$$w' = f(z, w) \tag{1.3.1}$$

cannot possess a solution $w : U \setminus \{z_0\} \to \mathbb{C}$ *such that* w *would be analytic in* $U \setminus \{z_0\}$ *with an essential singularity at* z_0.

PROOF. Fix first $\overline{B}(z_0, w_0; r) \subset \Omega$ such that $\overline{B}(z_0, r) \subset U$, and assume that $|f(z, w)| \leqslant M$ in $\overline{B}(z_0, w_0; r)$. Assume then, contrary to the assertion, that w is analytic in $U \setminus \{z_0\}$, with an essential singularity at z_0. By the Casorati–Weierstraß theorem, we may construct a sequence $(z_j)_{j\in\mathbb{N}}$ of distinct points in $B(z_0, r)$ converging to z_0 such that $(w(z_j))_{j\in\mathbb{N}}$ is a sequence of distinct points in $B(w_0, r)$ converging to w_0. Moreover, we may assume that $z_j \neq z_0$, $w(z_j) = w_j \neq w_0$, and that $(z_j, w_j) \in \overline{B}(z_0, w_0; \frac{r}{2})$. By the local existence theorem, the initial value problem

$$w'(z) = f\bigl(z, w(z)\bigr), \quad w(z_j) = w_j \tag{1.3.2}$$

has a unique analytic solution in some neighborhood U_j of z_j such that $B(z_j, \frac{r}{2}(1 - \exp(-\frac{1}{2M})) \subset U_j$. Therefore, $w(z)$ equals with this local solution determined by the existence theorem. If $|z_j - z_0| < \frac{r}{2}(1 - \exp(-\frac{1}{2M}))$, we conclude that $w(z)$ must be analytic at z_0, a contradiction. □

Actually, the above theorem is a corollary of the following theorem, originally due to Painlevé:

THEOREM 1.3.2. *Suppose that* $f : \Omega \to \mathbb{C}$ *is analytic in a domain* $\Omega \subset \mathbb{C}^2$ *and that* $(z_0, w_0) \in \Omega$. *Let* $w(z; z_0, w_0)$ *be the unique analytic solution of the initial value problem*

$$w' = f(z, w), \quad w(z_0) = w_0. \tag{1.3.3}$$

Moreover, let $\gamma : [0, 1] \to \Omega$ *be a Jordan curve such that* $\gamma(0) = z_0$, $\gamma(1) = a \neq z_0$, *and assume that* $w(z; z_0, w_0)$ *permits an analytic continuation along the curve* γ *such that* $(\gamma(t), w(\gamma(t); z_0, w_0)) \in \Omega$ *for all* $t \in [0, 1)$. *Finally, suppose that for a sequence* $(z_j)_{j\in\mathbb{N}}$ *on the curve* γ *such that* $z_j \to a$, *we have* $\lim_{j\to\infty} w(z_j; z_0, w_0) = \alpha \in \mathbb{C}$ *with* $(a, \alpha) \in \Omega$. *Then* $w(z; z_0, w_0)$ *permits an analytic continuation along the curve* γ *over the end-point* a *and we have* $w(z; z_0, w_0) = w(z; a, \alpha)$ *locally in some neighborhood of* a.

PROOF. We now fix $r > 0$ such that $\overline{B}(a, \alpha; 2r) \subset \Omega$, assuming that $|f(z, w)| \leqslant M$ in $\overline{B}(a, \alpha; 2r)$. We also choose, as we may by our assumptions, $(\zeta, \omega) \in \Omega$ such that

$$\zeta \in \gamma,\ \zeta \neq a, \quad |\zeta - a| < R := r\left(1 - \exp\left(-\frac{1}{2M}\right)\right) \tag{1.3.4}$$

and that for $\omega := w(\zeta; z_0, w_0)$ we have $|\omega - \alpha| < r$. Clearly, we then have $w(z; z_0, w_0) = w(z; \zeta, \omega)$ in some neighborhood of ζ by the basic existence and uniqueness theorem and by our assumption on the analytic continuation of $w(z; z_0, w_0)$. Since $\overline{B}(\zeta, \omega; r) \subset \overline{B}(a, \alpha; 2r)$ by the triangle inequality, the local solution $w(z; \zeta, \omega)$ must be analytic at least in $B(\zeta, R)$. Since $w(z_j; z_0, w_0) \to \alpha$ as $j \to \infty$, we must have $w(a; \zeta, \omega) = \alpha$. By the uniqueness of local solutions, $w(z; a, \alpha) = w(z; \zeta, \omega)$, implying the second assertion. □

In order to continue our considerations on the analytic continuations of local solutions, we turn our attention to the poles of $f(z, w)$. We say that (z_0, w_0) is a *pole* of $f(z, w)$, provided that $1/f(z, w)$ is analytic in a neighborhood of (z_0, w_0) and that $1/f(z_0, w_0) = 0$. As an example, $f(z, w) := 1/(z + w)$ has a pole at every point $(z_0, -z_0)$, $z_0 \in \mathbb{C}$.

PROPOSITION 1.3.3. *Suppose that* $f(z, w)$ *has a pole at* $(z_0, w_0) \in \mathbb{C}^2$ *and that* $1/f(z_0, w)$ *does not vanish identically. Then the differential equation* $w' = f(z, w)$ *admits, locally around* z_0*, a multivalent solution* $\widetilde{w}(z)$ *such that* $\widetilde{w}(z_0) = w_0$.

PROOF. Consider the initial value problem

$$z'(w) = \frac{1}{f(z(w), w)}, \quad z(w_0) = z_0. \tag{1.3.5}$$

By Theorem 1.2.1, this initial value problem admits a unique analytic solution $z(w)$ locally around z_0. Since $1/f(z, w)$ does not vanish identically, $z(w)$ cannot be a constant. Moreover,

$$z'(w_0) = \frac{1}{f(z(w_0), w_0)} = \frac{1}{f(z_0, w_0)} = 0. \tag{1.3.6}$$

Therefore, the Taylor expansion of $z(w)$ around w_0 takes the form

$$z(w) = z_0 + c_k(w - w_0)^k + \cdots, \quad c_k \neq 0,\ k > 1. \tag{1.3.7}$$

Clearly, by (1.3.7), $z(w)$ admits, locally around z_0, a multivalent inverse function $w(z)$ satisfying the assertion. □

REMARK. The solution $\widetilde{w}$, $\widetilde{w}(z_0) = w_0$ constructed in Proposition 1.3.3 is unique in the following sense: If $\widetilde{w}(z) \to w_0$ as $z \to z_0$ along a Jordan curve γ, and if $z \in \gamma$ is sufficiently close to z_0, then $\widetilde{w}(z)$ must locally coincide around z with a branch of the multivalued solution of $w' = f(z, w)$ in Proposition 1.3.3. Indeed, since $\widetilde{w}'(z) = f(z, \widetilde{w}(z))$

along the curve γ, and (z_0, w_0) is a pole of $f(z, w)$, then $\widetilde{w}'(z) \neq 0$ for any $z \in \gamma$ sufficiently close to z_0. Therefore, $\widetilde{w}(z)$ admits an analytic local inverse around z, satisfying the differential equation in (1.3.5). Clearly, this inverse function may be continued along the curve γ, approaching to z_0. By Theorem 1.3.2, the inverse function must coincide with the $z(w)$ constructed in the proof of Proposition 1.3.3.

REMARK. Before proceeding, we introduce a couple of standard notations. In the case of Proposition 1.3.3, we say that the multi-valued function $w(z)$ has an *algebraic branch point* at z_0. Similarly, we say that z_0 is an *algebraic pole* of w at z_0, if w is multivalent around z_0, $t(z_0) := 1/w(z_0) = 0$, and if the local inverse $z(t)$ of $t(z) := 1/w(z)$ around 0 admits a local expansion of type

$$z(t) = z_0 + c_k t^k + \cdots, \quad c_k \neq 0,\ k > 1. \tag{1.3.8}$$

To complete this section, we consider the equation (1.3.1) in the special case where $f(z, w)$ is either a polynomial or a rational function in w.

THEOREM 1.3.4. *Considering*

$$w' = P(z, w) = \sum_{j=0}^{n} p_j(z) w^j, \quad n \geqslant 2, \tag{1.3.9}$$

assume that the coefficients $p_j(z)$ are analytic in a domain $\Omega \subset \mathbb{C}$. If $a \in \Omega$ and $p_n(a) \neq 0$, for $n > 2$, and if $w(z)$ is a solution of (1.3.9) *analytic on a Jordan curve $\gamma : [0, 1) \to \Omega$ from z_0 to a, then either* (i) *$w(z)$ permits a meromorphic continuation over a, or* (ii) *$w(z)$ has an algebraic pole at a.*

PROOF. Suppose first that $w(z)$ remains bounded on γ as $z \to a$. Then there is a sequence (z_j) on γ converging to a such that $w(z_j)$ converges to some $\alpha \neq \infty$ as $z_j \to a$. By Theorem 1.3.2, $w(z)$ permits an analytic continuation over a. Otherwise, we may find a sequence (z_j) on γ such that $w(z_j) \to \infty$ as $z_j \to a$. By the transformation $v := 1/w$ we obtain

$$v' = Q(z, v) = -\frac{1}{v^{n-2}}\big(p_0(z)v^n + \cdots + p_n(z)\big). \tag{1.3.10}$$

If now $n = 2$, then $Q(z, v)$ is analytic around $(a, 0)$. By the first part of the proof, $v(z)$ has an analytic continuation over a, hence w has a meromorphic continuation over a with a pole at a. Finally, if $n > 2$, then $Q(z, v)$ has a pole at $(a, 0)$ and $1/Q(a, v)$ does not vanish identically. By Proposition 1.3.3, the case (ii) follows. □

THEOREM 1.3.5. *Consider now*

$$w' = \frac{P(z, w)}{Q(z, w)}, \tag{1.3.11}$$

where $P(z,w)$, $Q(z,w)$ *are polynomials in* w *with analytic coefficients in a domain* $\Omega \subset \mathbb{C}$. *Assume that* $a \in \Omega$, *that* $\deg_w Q \geqslant 1$ *and that* $Q(a,w)$ *does not vanish identically. Moreover, we assume that* $P(a,w)$ *and* $Q(a,w)$ *have no common factors. If now* $w(z)$ *is a solution of* (1.3.11) *analytic on a Jordan curve* $\gamma : [0,1) \to \mathbb{C}$ *from* z_0 *to* a *in* Ω, *then either* (i) $w(z)$ *permits a meromorphic continuation over* a, *or* (ii) $w(z)$ *has an algebraic branch point at* a, *or* (iii) $w(z)$ *has an algebraic pole at* a.

PROOF. Since $Q(a,w)$ does not vanish identically, it has finitely many zeros as a polynomial of w, say $w_1,\dots,w_k$. By the non-existence of common factors, $P(a,w_j) \neq 0$ for all $j=1,\dots,k$. Given $\varepsilon > 0$, consider the set

$$H(\varepsilon) := \bigl\{ w \mid |w| \leqslant 1/\varepsilon, |w-w_j| \geqslant \varepsilon, j=1,\dots,k \bigr\}, \tag{1.3.12}$$

and fix ε_0 sufficiently small. Assume first that for every ε, $0<\varepsilon<\varepsilon_0$, there exists some $t_\varepsilon \in [0,1)$ such that $w(\gamma(t)) \in \mathbb{C} \setminus H(\varepsilon)$ for all $t \in [t_\varepsilon, 1)$. By continuity of w, we may conclude that only one of the inequalities

$$\bigl|w\bigl(\gamma(t)\bigr) - w_j\bigr| < \varepsilon, \quad j=1,\dots,k, \quad \bigl|w\bigl(\gamma(t)\bigr)\bigr| > 1/\varepsilon \tag{1.3.13}$$

holds for all $t \in [t_\varepsilon, 1)$, provided ε_0 is small enough. Therefore, we must have either $w(z) \to \infty$ or $w(z) \to w_j$ for some $j=1,\dots,k$ as $z \to a$ along the curve γ. If $w(z) \to w_j$, we have $Q(a,w_j)=0$, $P(a,w_j) \neq 0$, while $Q(a,w)$ does not vanish identically. Therefore, by Proposition 1.3.3, we have the case (ii) at hand. If then $w(z) \to \infty$, we substitute $w = 1/v$, resulting in

$$v' = -\frac{P_1(z,v)}{Q_1(z,v)} = v^2 \frac{P(z,1/v)}{Q(z,1/v)}, \tag{1.3.14}$$

and we have two subcases to consider. If $Q_1(a,0) \neq 0$, then $v(z)$ permits an analytic continuation over a, and $v(a) = 0$. Therefore, w has a pole at a, permitting a meromorphic continuation over a. If, on the other hand, $Q_1(a,0)=0$, then clearly $P_1(a,0) \neq 0$. Moreover, $Q_1(a,v)$ cannot vanish identically, since $Q(a,w)$ does not vanish identically. By Proposition 1.3.3 again, v has an algebraic branch point at a with $v(a)=0$. But this implies that w has an algebraic pole at a, and we have the case (iii).

It remains to consider the case that for some sequence (t_j) converging to 1 as $j \to \infty$, $w(\gamma(t_j)) \in H(\varepsilon)$. Since $H(\varepsilon)$ is a compact set, we may assume, by taking a subsequence, if needed, that $w(\gamma(t_j))$ converges to some point $w_0 \neq w_1,\dots,w_k,\infty$. Therefore, $P(z,w)/Q(z,w)$ is analytic at (a,w_0). By Theorem 1.3.2, w permits an analytic continuation over the point a. □

1.4. *Local behavior of solutions of* $w' = P(z,w)/Q(z,w)$

In this section, we consider the local behavior of solutions of the differential equation

$$w' = \frac{P(z,w)}{Q(z,w)}, \tag{1.4.1}$$

where $P(z,w)$, $Q(z,w)$ both are analytic in both variables. If the initial value $w(z_0)=w_0$ is to be satisfied, and if $Q(z_0,w_0)\neq 0$, resp. $Q(z_0,w_0)=0$ and $P(z_0,w_0)\neq 0$, we may apply Theorem 1.3.2, resp. Proposition 1.3.3, provided that $Q(z_0,w)/P(z_0,w)$ does not vanish identically. If then $P(z_0,w_0)=Q(z_0,w_0)=0$, the situation becomes more complicated. We may assume that $(z_0,w_0)=(0,0)$, hence we have to consider the differential equation (1.4.1) under the initial value $w(z_0)=w_0$ in the case of

$$\begin{aligned} P(z,w) &= \alpha z+\beta w+\sum_{j+k\geqslant 2} p_{jk}z^j w^k,\\ Q(z,w) &= \gamma z+\delta w+\sum_{j+k\geqslant 2} q_{jk}z^j w^k. \end{aligned} \tag{1.4.2}$$

The basic idea to handle this initial value problem is to normalize the equation by a suitable linear transformation. Indeed, it appears that the linear terms in (1.4.2) are essential for the behavior of solutions of (1.4.1) around $z=0$. In this section, we restrict ourselves to considering two special cases only, leaving aside more general considerations. Concerning the role of non-linear terms, it should perhaps be mentioned, however, that they may have a substantial influence in some cases. As an example, the non-vanishing solutions of

$$w'=\frac{-w+zw^2}{z} \tag{1.4.3}$$

may be expressed, locally around a point $z\neq 0$ close to the origin, in the form

$$w(z)=\frac{1}{z(C-\log z)},\quad C\in\mathbb{C}, \tag{1.4.4}$$

while for

$$w'=-\frac{w}{z} \tag{1.4.5}$$

we have the solutions

$$w(z)=\frac{C}{z},\quad C\in\mathbb{C}. \tag{1.4.6}$$

1.4.1. *The linear case* This section is devoted to considering the non-singular linear case

$$w'=\frac{\alpha w+\beta z}{\gamma w+\delta z},\quad \alpha,\beta,\gamma,\delta\in\mathbb{C},\ \alpha\delta-\beta\gamma\neq 0. \tag{1.4.7}$$

It appears that (1.4.7) is, essentially, equivalent to the following pair of linear differential equations:

$$\begin{cases} w'=\alpha w+\beta z,\\ z'=\gamma w+\delta z. \end{cases} \tag{1.4.8}$$

Indeed, let $w(z)$ be a meromorphic solution of (1.4.8), satisfying the initial condition $w(z_0) = w_0 \neq \infty$. Then, obviously, $\gamma w(z) + \delta z$ does not vanish identically. Fixing now $t_0 \in \mathbb{C}$ arbitrarily, we may solve, at least locally, $z(t)$ from the linear differential equation

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \gamma w(z) + \delta z, \quad z(t_0) = z_0. \tag{1.4.9}$$

But then, considering $w = w(t) := w(z(t))$, we have

$$w(t_0) = w\big(z(t_0)\big) = w(z_0) = w_0 \tag{1.4.10}$$

and

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\mathrm{d}w}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\alpha w(z) + \beta z}{\gamma w(z) + \delta z}\big(\gamma w(z) + \delta z\big) = \alpha w(z) + \beta z. \tag{1.4.11}$$

Conversely, let $(w(t), z(t))$ be a meromorphic solution of (1.4.9), satisfying the finite initial conditions $w(t_0) = w_0$, $z(t_0) = z_0$ for an arbitrarily given t_0. However, we have to assume that $(z_0, w_0) \neq (0, 0)$, since $z_0 = w_0 = 0$ would result in $w(t) \equiv 0$, $z(t) \equiv 0$ by Theorem 1.2.1. Since $\alpha\delta - \beta\gamma \neq 0$, we conclude from (1.4.9) that $|w'(t_0)|^2 + |z'(t_0)|^2 > 0$. Assuming that $z'(t_0) \neq 0$, we infer that the inverse function $t(z)$ is locally defined around z_0. But then

$$\widetilde{w}(z) := w\big(t(z)\big) \tag{1.4.12}$$

satisfies

$$\widetilde{w}(z_0) = w\big(t(z_0)\big) = w(t_0) = w_0 \tag{1.4.13}$$

and

$$\widetilde{w}'(z) = w'\big(t(z)\big) = \frac{\mathrm{d}w}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}z} = \frac{\alpha w + \beta}{\gamma w + \delta}. \tag{1.4.14}$$

Therefore, it means no essential restriction to consider the pair (1.4.8) of differential equations. Denoting

$$\Phi = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{1.4.15}$$

we may represent (1.4.9) in vector notation as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} w \\ z \end{pmatrix} = \Phi \begin{pmatrix} w \\ z \end{pmatrix}. \tag{1.4.16}$$

Let now Γ be a 2×2-matrix of complex numbers such that $\det \Gamma \neq 0$, to apply the linear transformation

$$\begin{pmatrix} w \\ z \end{pmatrix} = \Gamma \begin{pmatrix} \omega \\ \zeta \end{pmatrix} \tag{1.4.17}$$

to (1.4.16). This results in

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \omega \\ \zeta \end{pmatrix} = \Gamma^{-1}\Phi\Gamma \begin{pmatrix} \omega \\ \zeta \end{pmatrix}. \tag{1.4.18}$$

By an elementary fact from linear algebra, we may choose Γ to satisfy

$$\Gamma^{-1}\Phi\Gamma = \begin{pmatrix} \lambda_1 & \mu \\ 0 & \lambda_2 \end{pmatrix}, \tag{1.4.19}$$

where $\mu = 0$, if $\lambda_1 \neq \lambda_2$. Therefore, two possibilities essentially cover the linear case:

(1) If $\lambda_1 = \lambda_2$, then (1.4.18) takes the form

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = \lambda_1\omega + \mu\zeta, \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \lambda_1\zeta, \tag{1.4.20}$$

and so

$$\frac{\mathrm{d}\omega}{\mathrm{d}\zeta} = \frac{\omega + \alpha\zeta}{\zeta}, \quad \alpha \neq 0. \tag{1.4.21}$$

Now, any local solution of (1.4.21) at a point $\zeta \neq 0$ near the origin takes the form

$$\omega(\zeta) = \zeta(\alpha \log \zeta + C), \quad C \in \mathbb{C}, \tag{1.4.22}$$

where we may take the branch of the logarithm to satisfy $\log 1 = 0$. If now γ is a Jordan curve ending at $\zeta = 0$ with the representation $\zeta = \rho \exp(\mathrm{i}\varphi(\rho))$, then $\zeta \log \zeta = (\rho \cos \varphi + \mathrm{i} \sin \varphi)(\log \rho + i\varphi(\rho))$, from which we compute that

$$|\zeta \log \zeta|^2 = (\rho \log \rho)^2 + \bigl(\rho\varphi(\rho)\bigr)^2. \tag{1.4.23}$$

Combining now (1.4.22) and (1.4.23), we observe that $\lim_{\zeta\to 0} \omega(\zeta) = 0$ if and only if $\lim_{\rho\to 0} \rho\varphi(\rho) = 0$.

(2) If $\lambda_1 \neq \lambda_2$, then (1.4.18) reduces to

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = \lambda_1\omega, \qquad \frac{\mathrm{d}\zeta}{\mathrm{d}t} = \lambda_2\zeta, \tag{1.4.24}$$

hence

$$\frac{\mathrm{d}\omega}{\mathrm{d}\zeta} = \beta\frac{\omega}{\zeta}, \quad \beta \neq 0. \tag{1.4.25}$$

In this case, we get

$$\omega(\zeta) = C\zeta^{\beta} = C\mathrm{e}^{\beta \log \zeta}, \quad C \in \mathbb{C}. \tag{1.4.26}$$

Let again γ be a curve ending at $\zeta = 0$, represented by $\zeta = \rho \exp(\mathrm{i}\varphi(\rho))$. Denoting $\beta = \beta_1 + \mathrm{i}\beta_2$, where β_1, β_2 are real numbers, we obtain

$$\mathrm{e}^{\beta \log \zeta} = \mathrm{e}^{\beta \log(\rho \exp(\mathrm{i}\varphi))} = \mathrm{e}^{(\beta_1 + \mathrm{i}\beta_2)(\log \rho + \mathrm{i}\varphi(\rho))}, \tag{1.4.27}$$

and therefore

$$|\mathrm{e}^{\beta \log \zeta}| = \mathrm{e}^{\beta_1 \log \rho - \beta_2 \varphi(\zeta)}. \tag{1.4.28}$$

To consider the behavior of $\omega(\zeta)$ as $\zeta \to 0$, let us first assume that $\beta_2 = 0$. Then we have $\lim_{\zeta \to 0} \omega(\zeta) = 0$, if $\beta < 0$, and $\lim_{\zeta \to 0} \omega(\zeta) = \infty$, if $\beta > 0$. Supposing now that $\beta_2 \neq 0$, then we have $\lim_{\zeta \to 0} \omega(\zeta) = 0$, provided that $\beta_1 > 0$. On the other hand, if $\beta_1 \leqslant 0$, we then may always find a curve γ such that $\omega(\zeta)$ doesn't converge to 0 as $\zeta \to 0$ along γ.

1.4.2. *A non-linear case: a glimpse at the Briot–Bouquet theory* As an example of a non-linear case, we first consider

$$zw' = \lambda w + p_{10} z + \sum_{j,k=1}^{\infty} p_{jk} z^j w^k, \quad \lambda \in \mathbb{C} \tag{1.4.29}$$

with constant coefficients. This is the most simple case of what are usually called as Briot–Bouquet equations.

THEOREM 1.4.1. *If* $\lambda \in \mathbb{C} \setminus \mathbb{N}$, *then* (1.4.29) *admits a local analytic solution in a neighborhood of* $z = 0$. *The solution satisfying the initial condition* $w(0) = 0$ *is unique.*

PROOF. For a proof based on the Cauchy majorant method, see [74], pp. 55–56. □

This theorem is a special case of more general systems of differential equations having a singular point of regular type at $z = 0$. We may write such a system in the form

$$zy_j' = f_j(z, y_1, \ldots, y_n), \quad j = 1, \ldots, n, \tag{1.4.30}$$

where the right-hand sides $f_j(z, y_1, \ldots, y_n) = f_j(z, \mathbf{y})$ are analytic around $(z, \mathbf{y}) = (0, \mathbf{0})$, and assume that the initial conditions $f_j(0, \mathbf{0}) = 0$ are satisfied for $j = 1, \ldots, n$. Systems of this type are called Briot–Bouquet systems. Of course, we may write (1.4.30) in vectorial form as

$$zY' = F(z, Y), \tag{1.4.31}$$

where $Y = {}^t(\mathbf{y}) \in \mathbb{C}^n$ and $F(z,Y) = {}^t(f_1(z,Y),\dots,f_n(z,Y))$. Denoting $K := (k_1, \dots, k_n) \in (\mathbb{N} \cup \{0\})^n$, $|K| := k_1 + \cdots + k_n$, $A_{jK} := {}^t(a_{jK}^{(1)}, \dots, a_{jK}^{(n)}) \in \mathbb{C}^n$, $||Y|| := \max_j |y_j|$ and $Y^K := y_1^{k_1} \cdots y_n^{k_n}$, the power series expansion

$$F(z,Y) = \sum_{j+|K| \geqslant 1} A_{jK} z^j Y^K \tag{1.4.32}$$

converges in some domain around the origin, say $|z| < r_0$, $||Y|| < R_0$. Then we obtain the following

THEOREM 1.4.2. *If none of the eigenvalues of the matrix*

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(0,\mathbf{0}) & \cdots & \frac{\partial f_1}{\partial y_n}(0,\mathbf{0}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial y_1}(0,\mathbf{0}) & \cdots & \frac{\partial f_n}{\partial y_n}(0,\mathbf{0}) \end{pmatrix}$$

is a natural number, then Eq. (1.4.31) *admits a unique solution of the form*

$$Y(z) = \sum_{j=1}^{\infty} C_j z^j, \quad C_j \in \mathbb{C}^n, \tag{1.4.33}$$

analytic in a neighborhood of the origin.

PROOF. See [97], pp. 261–263. □

2. Linear differential equations in the complex plane

Despite of the extensive literature on linear differential equations in a real domain as well as in abstract spaces such as Banach spaces etc., the global theory of linear differential equations in the complex plane has a surprisingly short research history. Of course, global theory may be understood in a variety of different ways. In this connection, it will be meant to studying properties of solutions defined in the whole complex plane. Moreover, we assume that the solutions under consideration are meromorphic (or algebroid) in the plane. In this sense, the birth of Nevanlinna theory in 1920's was a decisive step towards such global investigations. Indeed, Nevanlinna theory has been the key method in this field since 1929 when F. Nevanlinna considered [136] the differential equation $f'' + A(z)f = 0$ in the case of a polynomial coefficient $A(z)$. A more continuous activity had to wait until 1950's, started by Wittich, see e.g. [189] and Frei [44]. In late 1960's and in 1970's, most of the research made in this field was due to S. Bank and former students of Wittich (J. Nikolaus, G. Frank, E. Mues). The papers [14,15] by S. Bank and the present author in 1982 prompted extensive activity in the field of linear differential equations, having been continued ever since. In this section, we concentrate ourselves to presenting recent developments in the field during the last fifteen years, as there are two relevant sources [98], [107] covering previous investigations. Also, the book [76] by E. Hille contains a lot of classical material about complex differential equations.

For the convenience of the reader, we recall here some of the key notations and basic results of the Nevanlinna theory, due to its importance to understand a lot of the material presented below.

Given a non-constant meromorphic function, denote by $n(r, f)$ the number of poles of f, resp. by $n(r, \frac{1}{f-a}) = n(r, a, f)$ the number of a-points of f in the disc $|z| \leqslant r$, each such point being counted according to its multiplicity. The respective *counting functions* are now obtained by logarithmic integration:

$$N\left(r, \frac{1}{f-a}\right) := \int_0^r \frac{n(t, a, f) - n(0, a, f)}{t}\,\mathrm{d}t + n(0, a, f)\log r,$$

resp.

$$N(r, f) := \int_0^r \frac{n(t, f) - n(0, f)}{t}\,\mathrm{d}t + n(0, f)\log r.$$

The corresponding *proximity functions* will be defined as the following integrals over the boundary $|z| = r$:

$$m\left(r, \frac{1}{f-a}\right) := \frac{1}{2\pi}\int_0^{2\pi} \log^+\left|\frac{1}{f(re^{\mathrm{i}\theta}) - a}\right| \mathrm{d}\theta$$

and

$$m(r, f) := \frac{1}{2\pi}\int_0^{2\pi} \log^+\left|f(re^{\mathrm{i}\theta})\right| \mathrm{d}\theta,$$

where $\log^+ a := \max(0, \log a)$ for $a \geqslant 0$. The *characteristic function* of f is nothing but the sum of the proximity function and the counting function:

$$T(r, f) := m(r, f) + N(r, f).$$

Using this notion, the *order* of f is defined as

$$\rho(f) := \limsup_{r\to\infty} \frac{\log T(r, f)}{\log r}.$$

Observe that this definition extends the classical order notion for entire functions, where $\log M(r, f)$ is used instead of $T(r, f)$; these two notions are not identical, but close enough to result in the same order value (in the complex plane, but not necessarily in the unit disc, say).

The celebrated *first main theorem* in Nevanlinna theory now tells an almost complete balance between the number of a-points and the affinity of f towards the value $a \in \mathbb{C}$:

THEOREM 2.0.3. *For a non-constant meromorphic function f and for an arbitrary complex value a,*

$$T\left(r, \frac{1}{f-a}\right) = T(r, f) + \mathrm{O}(1).$$

Here the bounded error term slightly depends on the value a.

The characteristic function may be used to separate rational functions from the non-rational ones: A meromorphic function f is rational if and only if $T(r, f) = \mathrm{O}(\log r)$.

Although the first main theorem is one of the great achievements in function theory during the last century, the really deep component in the Nevanlinna theory is the *lemma of the logarithmic derivative*:

THEOREM 2.0.4. *Let f be a non-constant meromorphic function, and k be any natural number. Then*

$$m\left(r, \frac{f^{(k)}}{f}\right) = S(r, f),$$

where $S(r, f) = \mathrm{o}(T(r, f))$ outside of a possible exceptional set of finite linear measure as $r \to \infty$. If f is of finite order, then

$$m\left(r, \frac{f^{(k)}}{f}\right) = \mathrm{O}(\log r)$$

without an exceptional set.

The logarithmic derivative lemma is the key device to prove the *second main theorem*, which is a far-reaching refinement of the great Picard theorem. We give here a simplified version only:

THEOREM 2.0.5. *Let f be a non-constant meromorphic function, let $q \geqslant 2$ and let $a_1, \dots, a_q$ be distinct complex numbers. Then*

$$m(r, f) + \sum_{j=1}^{q} m\left(r, \frac{1}{f-a_j}\right) \leqslant 2T(r, f) + S(r, f).$$

Defining now the *deficiency* for $a \in \widehat{\mathbb{C}}$ by

$$\delta(a, f) := \liminf_{r\to\infty} \frac{m(r, 1/(f-a))}{T(r, f)},$$

the second main theorem immediately implies the fact that $\delta(a, f) = 0$ except for at most countably many exceptional values, and that

$$\sum_{a \in \widehat{\mathbf{C}}} \delta(a, f) \leqslant 2.$$

In addition to the deficiency, a *ramification index* $\vartheta(a, f)$ will also be applied in what follows. This is used to obtain some information of multiple values of f. To this end, we just count a-points, resp. poles, of f in such a way that its multiplicity will be reduced by one, using notation $N_1(r, \frac{1}{f-a})$, resp. $N_1(r, f)$, for the corresponding integrated counting functions. We then define

$$\vartheta(a, f) := \liminf_{r \to \infty} \frac{N_1(r, 1/(f-a))}{T(r, f)}, \qquad \vartheta(\infty, f) := \liminf_{r \to \infty} \frac{N_1(r, f)}{T(r, f)}.$$

Later on, at beginning of Section 4 below, we recall a few more important background results from Nevanlinna theory, frequently applied in considering complex differential equations.

For more details, several standard textbooks may be found. In our present list of references, [98] and [107] may be useful.

2.1. *Homogeneous linear differential equations with polynomial coefficients*

In studying global properties of solutions of linear differential equations in the complex plane, the non-trivial model case is

$$f'' + A(z) f = 0. \tag{2.1.1}$$

In this section, we assume that the coefficient $A(z) = c_n z^n + \cdots + c_0$ is a non-constant polynomial of degree n, $c_n \neq 0$. Then it is a classical result that certain symmetrically spaced radii from the origin essentially describe the zeros of non-trivial solutions of (2.1.1), see e.g. Gundersen [52], pp. 279–280, and [76], Chapter 7.4:

PROPOSITION 2.1.1. *Consider a differential equation* (2.1.1) *with a non-constant polynomial coefficient* $A(z)$ *of degree* n. *Denote* $\alpha = \frac{n+2}{2}$. *For* $0 < \delta < \frac{\pi}{2\alpha}$ *and* $j = 0, \ldots, n+1$, *consider the sectors*

$$S_j(\delta) := \left\{ z;\ \frac{j\pi - \arg c_n}{\alpha} + \delta \leqslant \arg z \leqslant \frac{(j+2)\pi - \arg c_n}{\alpha} - \delta \right\}. \tag{2.1.2}$$

If now a non-trivial solution of (2.1.1) *has infinitely many zeros in a sector* $S_j(\delta)$, *then for any* $\varepsilon > 0$, *all but finitely many of these zeros lie in the sector*

$$V_j(\varepsilon) := \left\{ z;\ \left| \arg z - \frac{2(j+1)\pi - \arg c_n}{n+2} \right| < \varepsilon \right\}. \tag{2.1.3}$$

Moreover, the non-integrated counting function for the zeros of f in the sector $V_j(\varepsilon)$ is

$$n_j(r,1/f)=\bigl(1+\mathrm{o}(1)\bigr)\frac{\sqrt{c_n}}{n\alpha}r^{\alpha}. \tag{2.1.4}$$

This proposition enables us to obtain fairly good estimates for the growth and zero distribution of non-trivial solutions f of (2.1.1). By the Wiman–Valiron theory, see e.g. [98], all non-trivial solutions are of order $\rho(f)=\alpha:=\frac{n+2}{2}$. To describe their growth and value distribution, we denote by $p(f)$ the number of sectors $V_j(\varepsilon)$ with the property that for some $\varepsilon>0$, the non-trivial solution f of (2.1.1) has only finitely many zeros in $V_j(\varepsilon)$. Clearly, $0\leqslant p(f)\leqslant n+2$. We now obtain, see [52], Theorems 5 and 6, the following asymptotic equalities:

THEOREM 2.1.2. *Given a non-trivial solution f of* (2.1.1), *and using the notations described above, $p(f)$ is an even number and*

$$n(r,1/f)=\bigl(1+\mathrm{o}(1)\bigr)\frac{2\alpha-p(f)}{\pi\alpha}\sqrt{|c_n|}\,r^{\alpha}, \tag{2.1.5}$$

$$N(r,1/f)=\bigl(1+\mathrm{o}(1)\bigr)\frac{2\alpha-p(f)}{\pi\alpha^2}\sqrt{|c_n|}\,r^{\alpha}, \tag{2.1.6}$$

$$T(r,f)=\bigl(1+\mathrm{o}(1)\bigr)\frac{4\alpha-p(f)}{2\pi\alpha^2}\sqrt{|c_n|}\,r^{\alpha}, \tag{2.1.7}$$

$$\log M(r,f)=\bigl(1+\mathrm{o}(1)\bigr)\frac{\sqrt{|c_n|}}{\alpha}\,r^{\alpha}. \tag{2.1.8}$$

Moreover, we have for the deficiency of zeros of f;

$$\delta(0,f)=\Delta(0,f)=\frac{p(f)}{4\alpha-p(f)}. \tag{2.1.9}$$

REMARK. Equation (2.1.1) with $\deg A=n$ possesses $m\geqslant 2$ pairwise linearly independent solutions $f_1,\dots,f_m$ such that $p(f_j)\geqslant 2$ for $j=1,\dots,m$ and $\sum_{j=1}^{m}p(f_j)=2(n+2)$. Moreover, any non-trivial solution f of (2.1.1) that is not a constant multiple of some f_j, $j=1,\dots,m$, has $p(f)=0$. Hence, most of the non-trivial solutions f have $p(f)=0$. See [52], pp. 281–283, for more details. Finally, observe that the maximum modulus of f is independent of $p(f)$.

Because of the symmetric spacing of the critical rays described in Proposition 2.1.1, and the fact that $p(f)$ is even, it is not surprising that the number of non-real zeros of a non-trivial solution f is maximal in most cases, see [52]:

THEOREM 2.1.3. *Given two linearly independent solutions f_1,f_2 of* (2.1.1), *at least one of f_1,f_2 has the property that its sequence of non-real zeros has exponent of convergence equal to* $(n+2)/2$. *Moreover, if the coefficient polynomial $A(z)$ in* (2.1.1) *is of degree $n=2+4k$ for a non-negative integer k, then any non-trivial solution f either has finitely* *many zeros only, or the exponent of convergence of the non-real zero-sequence of* f *is* $(n+2)/2$.

Combining [52], Theorem 3 and Lemma 5, see also [147], Theorem 1.1, the following conclusion appears:

THEOREM 2.1.4. *If Eq.* (2.1.1) *admits a non-trivial solution* f *with infinitely many real zeros, then* $A(z)$ *is a real polynomial* (*i.e. real on the real axis*), *and* f *is a constant multiple of a real solution of* (2.1.1).

By the preceding result, it makes sense to call special attention to the case when $A(z)$ is a real polynomial. In this direction, Eremenko and Merenkov [41] proved the following interesting results recently:

THEOREM 2.1.5. *For each* $n \geqslant 1$, *there exists Eq.* (2.1.1) *with a real polynomial* $A(z)$ *of degree* n *and a solution* f *of* (2.1.1) *such that all zeros of* f *are real. Concerning different values of* n,

(a) *if* $n = 0 \pmod 4$, *then the zero-set of* f *is either finite or it is unbounded from above and below,*
(b) *if* $n = 2 \pmod 4$, *then* f *has finitely many zeros only and*
(c) *if* n *is odd, then* f *has infinitely many zeros all of which lie on a ray of the real axis.*

The proof of this result in [41] makes use of the graph theory (Speiser graphs).

EXAMPLE. (See [52].) The classical Airy differential equation

$$f'' - zf = 0$$

admits a solution whose zeros are all real and negative, namely the standard Airy function $Ai(z)$, compare Theorem 2.1.5(c). Clearly, $p(Ai) = 2$. Therefore, by Theorem 2.1.2,

$$N(r, 0, Ai) = \bigl(1 + \mathrm{o}(1)\bigr)\frac{4}{9\pi}r^{3/2}, \qquad T(r, Ai) = \bigl(1 + \mathrm{o}(1)\bigr)\frac{8}{9\pi}r^{3/2}$$

and

$$\log M\bigl(r, Ai\bigr) = \bigl(1 + \mathrm{o}(1)\bigr)\frac{2}{3}r^{3/2}.$$

For an analysis of a generalized Airy differential equation

$$f'' - z^n f = 0,$$

see [54].

REMARK. Proposition 2.1.1 depicts what is called the *finite oscillation property*, see [7], p. 266: There exist finitely many rays $\arg z = \theta_j$, $j = 1, \dots, m$, such that for any $\varepsilon > 0$, all but finitely many zeros of any non-trivial solution of a linear differential equation

$$f^{(n)} + p_{n-1}(z) f^{(n-1)} + \cdots + p_0(z) f = 0 \tag{2.1.10}$$

lie in the union of the sectors $|\arg z - \theta_j| < \varepsilon$, $j = 1, \dots, m$. On the other hand, we say that a linear differential equation possesses the *global oscillation property*, if there is a non-trivial solution f such that for any ray $\arg z = \theta$ and any $\varepsilon > 0$, f has infinitely many zeros in the sector $|\arg z - \theta| < \varepsilon$. By Proposition 2.1.1, the finite oscillation property holds for the basic second order differential equation (2.1.1) with a polynomial coefficient $A(z)$. However, the finite oscillation property does not hold in general in the higher order case (2.1.10). For instance, equation

$$f^{(n)} + z^2 f'' + z f' + f = 0, \quad n > 2,$$

possesses the global oscillation property, see [6], p. 37. In fact, given any Eq. (2.1.10) with polynomial coefficients, it can be shown that it has either the finite oscillation property, or the global oscillation property, see [8], Theorem 1.

Concerning homogeneous linear differential equations of the form (2.1.10) with $n \geqslant 2$ and with polynomial coefficients such that $p_0(z)$ does not vanish identically, it is well known that all solutions f of Eq. (2.1.10) are entire functions of finite rational order. Moreover, denoting $d_j := \deg p_j$, if p_j does not vanish, and $d_j = -\infty$ otherwise, we have that

$$\rho(f) \leqslant \lambda := 1 + \max_{0 \leqslant j \leqslant n-1} \frac{d_j}{n-j} \tag{2.1.11}$$

for all solutions of (2.1.10). These are classical results usually proved by the Wiman–Valiron theory and, to determine the possible rational orders, the Newton–Puiseux diagram, see [98]. However, the list of possible rational orders may be obtained by simple arithmetic with the degrees of polynomial coefficients of (2.1.10), as shown by Gundersen, Steinbart and Wang in [56]. To recall this reasoning, we first determine a strictly decreasing finite sequence $s_1, \dots, s_p$ of non-negative integers by defining

$$s_0 = n, \tag{2.1.12}$$

$$s_1 := \min\left\{ j; \frac{d_j}{n-j} = \max_{0 \leqslant m \leqslant n-1} \frac{d_m}{n-m} \right\} \leqslant n-1 \tag{2.1.13}$$

and then, inductively,

$$s_{j+1} := \min\left\{ k; \frac{d_k - d_{s_j}}{s_j - m} = \max_{0 \leqslant m \leqslant s_j} \frac{d_m - d_{s_j}}{s_j - m} > -1 \right\}. \tag{2.1.14}$$

Clearly, this process terminates after a finite number $p \leqslant n$ of steps. We then define, for $j = 1, \dots, p$,

$$\alpha_j := 1 + \frac{d_{s_j} - d_{s_{j-1}}}{s_{j-1} - s_j} > 0, \tag{2.1.15}$$

where, of course, $d_{s_0} = 0$. We are now ready to state

THEOREM 2.1.6. *Concerning the growth of solutions of* (2.1.10), *the following assertions hold*:

(a) *Equation* (2.1.10) *possesses at most* s_p *linearly independent polynomial solutions*.

(b) *If* f *is a transcendental solution, then* $\rho(f) = \alpha_j$ *for some* j, $1 \leqslant j \leqslant p$. *Moreover, if* $s_1 \geqslant 1$ *and* $p \geqslant 2$, *then*

$$\alpha_1 > \alpha_2 > \cdots > \alpha_p \geqslant \frac{1}{s_1 - s_p} \geqslant \frac{1}{s_1} \geqslant \frac{1}{n-1},$$

and if $s_1 = 0$, *then all non-trivial solutions* f *are of order* $\rho(f) = 1 + d_0/n$.

(c) *For any* $j = 1, \dots, p$, *there exist at most* s_j *linearly independent solutions* f *of order* $\rho(f) < \alpha_j$. *Also, there always exists a solution* f *of order* $\rho(f) = \lambda = 1 + \max_{0 \leqslant j \leqslant n-1} \frac{d_j}{n-j}$.

(d) *If* $\{f_1, f_2, \dots, f_n\}$ *is any fundamental set of solutions of* (2.1.10), *then*

$$\sum_{j=1}^{n} \rho(f_j) \geqslant n + d_0.$$

For a collection of examples showing the sharpness of the preceding theorem, see [56], pp. 1240–1246.

The frequency of zeros of solutions of (2.1.10) with polynomial coefficients has been somewhat studied since 1950's; we refer here to a number of papers by Pöschl, Wittich, Frank, Brüggemann and Steinmetz. However, the recent papers by Wang [186] and Steinbart [168] offer the most complete understanding of this topics until now. To describe these recent results, we first recall the real numbers $s_0, \dots, s_p$ and $\alpha_1, \dots, \alpha_p$ defined by the degrees of the polynomial coefficients of (2.1.10) as described above. Observe that we then have

$$d_{s_j} - d_{s_{j-1}} = (\alpha_j - 1)(s_{j-1} - s_j).$$

Let now A_j be the leading coefficient of $p_j = A_j z^{d_j} + \cdots$, and define

$$A^{\star}_{k,m} := \begin{cases} A_k & \text{if } d_k = d_{s_{m-1}} + (\alpha_m - 1)(s_{m-1} - k), \\ 0 & \text{otherwise} \end{cases} \tag{2.1.16}$$

and

$$A^{\sharp}_{k} := \begin{cases} A_k & \text{if } \frac{d_k}{n-k} = \max_{0 \leqslant j \leqslant n-1} \frac{d_j}{n-j}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.17}$$

We then define the following polynomials corresponding to (2.1.16) and (2.1.17) as

$$H_m(t) := \sum_{k=0}^{s_m-1} A^{\star}_{k,m} t^k, \quad 1 \leqslant m \leqslant p, \tag{2.1.18}$$

and

$$H^{\sharp}(t) := t^n + \sum_{k=0}^{n-1} A^{\sharp}_k t^k. \tag{2.1.19}$$

We may now express the key results in [186] and [168] as follows:

THEOREM 2.1.7. *Concerning the frequency of zeros of solutions of* (2.1.10), *the following assertions hold*:

(a) *If f is a non-trivial solution such that $\rho(f)$ is not an integer, then f has infinitely many zeros and $\rho(f) = \lambda(f)$.*

(b) *If f is a transcendental solution of order $\rho(f) = \alpha_m \in \mathbb{N}$ for some m, $1 \leqslant m \leqslant p$, and if all non-zero roots of equation $H_m(t) = 0$ are simple, then either $\rho(f) = \lambda(f)$, or f has finitely many zeros only.*

(c) *Suppose at least one of the coefficients of* (2.1.10) *is non-constant. Then the following properties are equivalent*:

  (1) *Equation* (2.1.10) *admits a fundamental solution set $\{f_1, \ldots, f_n\}$ such that $\lambda(f_j) < \rho(f_j)$, $j = 1, \ldots, n$.*

  (2) *Equation* (2.1.10) *admits a fundamental solution set $\{f_1, \ldots, f_n\}$ such that*

$$\max_{1\leqslant j\leqslant n} \lambda(f_j) < \lambda = 1 + \max_{0\leqslant j\leqslant n-1} \frac{d_j}{n-j}.$$

  (3) *The polynomial $H^{\sharp}(t)$ takes the form $H(t) = (t+b)^n$ for some non-zero constant b.*

For some examples illuminating the preceding theorem, see [168], pp. 368–370.

REMARK. It seems quite likely that multiple roots of $H_m(t) = 0$, resp. of $H^{\sharp}(t) = 0$, correspond to subsets of the fundamental solution set $\{f_1, \ldots, f_n\}$ such that the solutions f in these subsets satisfy $\lambda(f) < \rho(f)$. As far as we know, there is no explicit proof for this claim yet. See, however, Theorem 2.5.5 below.

2.2. *Second order homogeneous linear differential equations with entire coefficients*

Looking at the non-trivial model case (2.1.1) with a transcendental entire coefficient $A(z)$, it immediately follows from the logarithmic derivative lemma that $\rho(f) = \infty$. Therefore, the key problem here is to consider the frequency of zeros of solutions of (2.1.1). This topic has been under active research since 1982, see [15].

The key idea to consider the zero frequency of solutions of (2.1.1) with $A(z)$ transcendental entire, is to make use of the product E of two linearly independent solutions f_1, f_2. By the classical Abel identity, we know that the Wronskian $W(f_1, f_2)$ is a non-zero constant. Therefore, we may normalize it to be equal to 1. Then it is routine to check that

$$4A = \left(\frac{E'}{E}\right)^2 - \frac{1}{E^2} - 2\frac{E''}{E}. \tag{2.2.20}$$

Since the zero-sets of f_1 and f_2 are distinct, we have for the exponent of convergence of the zeros of E that

$$\lambda(E) = \max\bigl(\lambda(f_1), \lambda(f_2)\bigr).$$

Then the basic results from [15], complemented with an improvement due to Shen [154] and Rossi [146] for the case $\rho(A) = 1/2$, may be written as

THEOREM 2.2.1. *Concerning the frequency of zeros of solutions of* (2.1.1) *with* $A(z)$ *transcendental entire, the following assertions hold*:

(a) *Suppose* $\rho(A) < \infty$ *but is not a natural number. If* $\rho(A) \leqslant 1/2$, *then* $\lambda(E) = \infty$; *if* $\rho(A) < 1$, *then* $\lambda(E) > 1$; *and if* $\rho(A) \geqslant 1$, *then* $\lambda(E) \geqslant \rho(A)$.

(b) *Suppose next that* $\lambda(A) < \rho(A)$ (*which implies by the Hadamard factorization that* $\rho(A)$ *is a natural number or infinite*). *Then the exponent of convergence* $\lambda(f)$ *of the zero-sequence of an arbitrary non-trivial solution* f *satisfies* $\lambda(f) \geqslant \rho(A)$.

(c) *If* $\bar{\lambda}(A) < \rho(A)$, *then* $\lambda(E) \geqslant \rho(A)$.

(d) *There exist two linearly independent zero-free solutions* f_1, f_2 *of* (2.1.1) *in the complex plane if and only if* $A(z)$ *is of the form*

$$-4A(z) = h'(z)^2 + \phi'(z)^2 - 2\phi''(z)$$

where ϕ *a non-constant entire function and* h *a primitive of* $\exp(\phi)$.

(e) *If* $\lambda(E) = \max(\lambda(f_1), \lambda(f_2)) < \infty$, *then* $\lambda(f) = \infty$ *for all solutions* f *of* (2.1.1) *not being of the form* αf_1 *or* αf_2 *for some* $\alpha \in \mathbb{C}$.

PROOF. As a typical example of applying the product E of two linearly independent solutions of (2.1.1), we prove the last assertion of (a). If $\lambda(E) < \rho(A)$, then (2.2.20) easily implies that $\rho(E) \leqslant \rho(A)$. Since $\rho(A)$ is not a natural number, the Hadamard product representation implies that $\lambda(E) = \rho(A)$. ☐

EXAMPLE. As an example of the case when $\rho(A)$ is a natural number, we collect here a few facts related to equation

$$f'' + \bigl(e^{P(z)} + Q(z)\bigr) f = 0, \tag{2.2.21}$$

where P is polynomial of degree $\deg P = \lambda > 0$ and Q is an entire function of order $\rho(Q) < \lambda$. Then the following properties are valid:

(1) If (2.2.21) admits a non-trivial solution f such that $\lambda(f) < \lambda$, then f has no zeros, $Q = -\frac{1}{16}(P')^2 + \frac{1}{4}P''$ is a polynomial and there is another zero-free solution of (2.2.21), linearly independent of f.

(2) If $Q(z) \equiv 0$, then $\lambda(f) = \infty$ for all non-trivial solutions f of (2.2.21).

(3) In the special case of $f'' + (e^z - K)f = 0$, $K \in \mathbb{C}$, a non-trivial solution f with $\lambda(f) < \infty$ exists if and only if $K = q^2/16$ for an odd integer q. In this case, the equation admits two linearly independent solutions such that for their product E, $\lambda(E) = 0$, if $|q| = 1$ and $\lambda(E) = 1$ otherwise.

Theorem 2.2.1, together with a collection of examples to be found in the literature, prompted the following conjecture, known as the Bank–Laine conjecture (BL-conjecture for short), implicitly contained in [15]:

CONJECTURE. *Let* $A(z)$ *be transcendental entire of finite order of growth* ρ. *If* f_1, f_2 *are two linearly independent solutions of* (2.1.1) *such that* $\max(\lambda(f_1), \lambda(f_2) < \infty$, *then* ρ *is a natural number.*

In its full generality, the conjecture remains open. Due to the wide interest this problem has implied, a few remarks follow. Recalling (2.2.20), it is not difficult to see that all zeros of f_1, f_2 and $E := f_1 f_2$ are simple, and $E'(z_0) = \pm 1$ at all zeros z_0 of E, provided the Wronskian $W(f_1, f_2)$ has been normalized to $W(f_1 f_2) = 1$. This condition is known as the BL-condition and any entire function E satisfying the BL-condition is called a BL-function. To resolve the BL-conjecture, it is sufficient to show that whenever E is a BL-function of finite order, then the corresponding coefficient function $A(z)$ of equation (2.1.1), determined by (2.2.20), is of an integer order of growth. A number of examples of BL-functions fall in two classes of functions: (1) Entire functions of type $P(z)e^{Q(z)}$, where P is a polynomial with simple zeros only, and Q is a (non-constant) polynomial and (2) entire functions of type $P(e^{\alpha z})e^{\beta z}$, where P is a polynomial with simple zeros only, and α, β are constants. However, not all known BL-functions of finite order belong in these two function classes as shown by Langley, see [112,113] and [114]. It has also been shown recently by Drasin and Langley, see [35], that there exists a Bank–Laine function E, associated with a transcendental coefficient function $A(z)$ determined by (2.2.20), such that for any given positive integer n, $n \leqslant \rho(E) \leqslant \lambda(E) < \infty$. For more details about the BL-conjecture, see [109].

The BL-conjecture has been very much in the background to direct research concerning Eq. (2.1.1) during the last quarter of century. A special situation having attracted a lot of attention is the case when $A(z)$ is a periodic function. The key problem here has been to explain when such an equation admits a non-trivial solution f with a few zeros in the sense of $\lambda(f) < \infty$. The first paper in this line was [16], to our knowledge. Recently, Chiang [29] improved the results in [16]. To recall the key results in these papers, suppose that $A(z)$ is of period $2\pi\mathrm{i}$. Then we may write

$$A(z)=B(\mathrm{e}^z), \qquad B(\zeta)=\sum_{j=k}^{l}K_j\zeta^j=g(\zeta)+h(1/\zeta),$$

where $K_kK_l\neq 0$, $l\geqslant k$, and $g(\zeta), h(1/\zeta)$ are entire functions. Making the transformation $\zeta\mapsto 1/\zeta$, if needed, we may assume that $l\geqslant 0$.

THEOREM 2.2.2. (See [29], Proposition 1.) *Given a differential equation* (2.1.1), *where* $A(z)=B(\mathrm{e}^z)$ *is a rational function of* e^z, *write* $B(\zeta)=\sum_{j=k}^{l}K_j\zeta^j$, $K_lK_k\neq 0$, *and assume that* $\lambda(f)<\infty$ *for a non-trivial solution* f. *Then the following assertions hold*:

(1) *If* l *is odd and positive, then* $k=0$ *and* f *has the following representation*:

$$f(z)=\psi(\mathrm{e}^{z/2})\exp\left(\sum_{j=0}^{l}d_j\mathrm{e}^{jz/2}+dz\right),$$

where d, d_j *are constants and* $\psi(\zeta)$ *is a polynomial with simple zeros only. Moreover,* $d_j=0$ *for all* $j>0$ *even.*

(2) *If* l *is even and positive, then* k *is even as well, and*

$$f(z)=\psi(\mathrm{e}^{z})\exp\left(\sum_{j=-k/2}^{l/2}d_j\mathrm{e}^{jz}+dz\right),$$

where again d, d_j *are constants and* $\psi(\zeta)$ *is a polynomial with simple zeros only.*

By the presentations for $f(z)$ in Theorem 2.2.2, it follows that f is zero-free if and only if $\lambda(f)=0$. In this special situation, more detailed representations for $f(z)$ may be obtained, see [29], Theorems 1, 2 and 3.

Continuing to the more general case when $A(z)$ is a non-constant periodic entire function of period $2\pi\mathrm{i}$ and transcendental in e^z, we obtain [16]:

THEOREM 2.2.3. *Let* $A(z)$ *be a non-constant periodic entire function such that* $A(z)=B(\mathrm{e}^z)$, *where* $B(\zeta)$ *is analytic in* $\mathbb{C}\setminus\{0\}$, *and assume that* f *is a non-trivial solutions of* (2.1.1) *such that* $\lambda(f)<\infty$. *Then* f *has the representation*

$$f(z)=\mathrm{e}^{dz}H(\mathrm{e}^{z/q})\exp\bigl(g(\mathrm{e}^{z/q})\bigr),$$

where (1) d *is a complex number,* (2) $q=1$ *or* $q=2$ *depending whether* $f(z)$ *and* $f(z+2\pi\mathrm{i})$ *are linearly dependent or not,* (3) $H(\zeta)$ *and* $g(\zeta)$ *are analytic in* $\mathbb{C}\setminus\{0\}$, *and* (4) $g(\zeta)$ *has at most a pole at* $\zeta=\infty$, *resp.* $\zeta=0$, *if and only if* $B(\zeta)$ *has at most a pole at* $\zeta=\infty$, *resp.* $\zeta=0$. *Moreover, letting* ρ_∞, *resp.* ρ_0, *denote the local order at* ∞, *resp. at* 0, *we have* $\rho_0(H)=\rho_\infty(H)=0$, $\rho_0(g)=\rho_0(B)$ *and* $\rho_\infty(g)=\rho_\infty(B)$.

For additional results concerning the case $A(z) = B(\mathrm{e}^z)$, where $A(z)$ is transcendental in e^z, see [30].

2.3. *Higher order homogeneous linear differential equations*

We now proceed to considering higher order homogeneous linear differential equations

$$L(f) := f^{(n)} + a_{n-1}(z) f^{(n-1)} + \cdots + a_0(z) f = 0 \tag{2.3.22}$$

with entire coefficients. Since (2.3.22) may easily be reduced to

$$f^{(n)} + a_{n-2}(z) f^{(n-1)} + \cdots + a_0(z) f = 0, \tag{2.3.23}$$

a substantial part below will be formulated for (2.3.23). Contrary to the normalized second-order case of (2.1.1), it may now appear, even in the second order case, that two solutions may be of substantially different growth. The first classical result in this direction is due to Frei in [44]:

THEOREM 2.3.1. *Let a_j be the last transcendental function in the coefficient sequence $a_0, \ldots, a_{n-1}$ of* (2.3.22). *Then Eq.* (2.3.22) *possesses at most j linearly independent solutions of finite order.*

REMARK. Combining Theorem 2.3.1 with (2.1.11) we conclude that all solutions of (2.3.22) are of finite order of growth if and only if all coefficients of (2.3.22) are polynomials, see Wittich [190].

REMARK. In the special case of $n = 2$, Theorem 2.3.1 implies that all non-trivial solutions are of infinite order, provided a_0 is transcendental and a_1 rational. More generally, the same conclusion follows in a variety of situations, where a_0 is dominant. A nice example of this type has been proved by Kwon in [106]:

THEOREM 2.3.2. *Given two constants $\alpha, \beta > 0$, and a set E in the real axis of strictly positive upper density, suppose that the entire functions a_1, a_0 satisfy the growth conditions*

$$\left|a_1(z)\right| \leqslant \exp\bigl(\mathrm{o}(1)|z|^\beta\bigr), \qquad \left|a_0(z)\right| \geqslant \exp\bigl(\bigl(1 + \mathrm{o}(1)\bigr)\alpha|z|^\beta\bigr)$$

as $|z| \to \infty$, $|z| \in E$. Then all non-trivial solutions of

$$f'' + a_1(z) f' + a_0(z) f = 0$$

are of infinite order.

For another result of this type, see Theorem 2.3.6 below.

Making use of the notion of an iterated order,

$$\rho_j(z) := \limsup_{r\to\infty} \frac{\log_j T(r, f)}{\log r}, \tag{2.3.24}$$

where $\log_1(r) := \log r$, $\log_j(r) := \log\log_{j-1}(r)$, we may perform a more detailed analysis of the growth of solutions of (2.3.22), see [102]. To this end, we define the *finiteness degree of growth* $i(f)$ of a meromorphic function f as $i(f) = 0$ for rational functions, $i(f) := \min\{j \in \mathbb{N}; \rho_j(f) < \infty\}$, and $i(f) = \infty$ otherwise. Concerning the linear differential operator $L(f)$ with entire coefficients, we define

$$\begin{aligned}
&\delta(L) := \max\bigl\{i(f); L(f) = 0\bigr\},\\
&\gamma_j(L) := \max\bigl\{\rho_j(f); L(f) = 0\bigr\}, \quad j \in \mathbb{N},\\
&p(L) := \max\bigl\{i(a_j); j = 0, \dots, n-1\bigr\}.
\end{aligned}$$

Provided $0 < p := p(L) < \infty$, we also define

$$\kappa(L) := \max\bigl\{\rho_p(a_j); j = 0, \dots, n-1\bigr\}.$$

Using these notations, we have

THEOREM 2.3.3. *If* $0 < p < \infty$, *then* $\delta(L) = p + 1$ *and* $\gamma_{p+1}(L) = \kappa(L)$. *Moreover, if* a_j *be the last coefficient in the sequence* $a_0, \dots, a_{n-1}$ *such that* $i(a_j) = p$, *then* (2.3.22) *possesses at most* j *linearly independent solutions* f *such that* $i(f) \leqslant p$.

PROOF. The key idea of the proof makes use of the standard order reduction procedure of linear differential equations, combined with the generalized logarithmic derivative lemma, see [102], Lemma 1.3: Given a meromorphic function f such that $i(f) = p \geqslant 1$ and $\rho_p(f) = \rho$, then for any $\varepsilon > 0$ and any natural number k, $m(r, f^{(k)}/f) = O(\exp_{p-2} r^{\rho+\varepsilon})$ outside of a possible exceptional set of finite linear measure. □

In the general case of (2.3.22) with $n \geqslant 2$, the preceding theorems leave open the problem of determining the situations when all solutions of (2.3.22) are, say, of infinite order. Such a situation typically appears, if one of the coefficients is dominating in the sense of growth, see [110]:

THEOREM 2.3.4. *Let* $a_0(z), \dots, a_{n-1}(z)$ *be entire functions such that for some integer* s, $1 \leqslant s \leqslant n-1$, *we have* $\rho_p(a_j) < \rho_p(a_s) \leqslant \infty$ *for all* $j \neq s$. *Then every transcendental solution* f *of* (2.3.22) *satisfies* $\rho_p(f) \geqslant \rho_p(a_s)$.

However, one may obtain similar results if one of the coefficients is dominant just almost everywhere only, in certain sense. Recently, a number of papers of such type have been written. Prototypes of such results may be found in [53]:

THEOREM 2.3.5. *Let* $a_0(z) \not\equiv 0$, $a_1(z)$ *be entire functions such that for some real constants* $\alpha > 0$, $\beta > 0$ *and* $\theta_1 < \theta_2$ *we have*

$$|a_1(z)| \geqslant \exp\bigl((1 + \mathrm{o}(1))\alpha |z|^\beta\bigr)$$

and

$$|a_0(z)| \leqslant \exp\bigl(\mathrm{o}(1)|z|^\beta\bigr)$$

as $z \to \infty$ *in the sector* $S(0) := \{\theta = \arg z; \theta_1 \leqslant \theta \leqslant \theta_2\}$. *Given* $\varepsilon > 0$ *small enough, consider the sector* $S(\varepsilon) := \{\theta = \arg z; \theta_1 + \varepsilon \leqslant \theta \leqslant \theta_2 - \varepsilon\}$. *If* f *is a non-trivial solution of*

$$f'' + a_1(z) f' + a_0(z) f = 0$$

of finite order, then the following conditions hold:

(i) *There exists a constant* $b \neq 0$ *such that* $f(z) \to b$ *as* $z \to \infty$ *in* $S(\varepsilon)$. *More precisely,*

$$|f(z) - b| \leqslant \exp\bigl(-(1 + \mathrm{o}(1))\alpha |z|^\beta\bigr).$$

(ii) *For each integer* $k \geqslant 1$,

$$|f^{(k)}(z)| \leqslant \exp\bigl(-(1 + \mathrm{o}(1))\alpha |z|^\beta\bigr)$$

as $z \to \infty$ *in* $S(\varepsilon)$.

This result, combined with the classical Phragmén–Lindelöf principle, applies now to prove

THEOREM 2.3.6. *Let* $a_0(z) \not\equiv 0$, $a_1(z)$ *be entire functions, and let* $\alpha > 0$ *and* $\beta > 0$ *be constants such that* $\rho(a_0) < \beta$. *Suppose that for any given* $\varepsilon > 0$, *there exist two finite collections of real numbers that satisfy* $\phi_1 < \theta_1 < \phi_2 < \theta_2 < \cdots < \phi_n < \theta_n < \phi_{n+1} = \phi_1 + 2\pi$, *and* $\sum_{j=1}^{n} (\phi_{j+1} - \phi_j) < \varepsilon$, *such that*

$$|a_1(z)| \geqslant \exp\bigl((1 + \mathrm{o}(1))\alpha |z|^\beta\bigr)$$

as $z \to \infty$ *in each of the sectors* $\phi_j \leqslant \arg z \leqslant \theta_j$, $j = 1, \ldots, n$. *Then all non-trivial solutions* f *of*

$$f'' + a_1(z) f' + a_0(z) f = 0$$

are of infinite order of growth.

Applying a certain type of nested induction process, see [110], the preceding theorems may be extended to (2.3.22), in the form that the dominating coefficient may be any one of the intermediate coefficients:

THEOREM 2.3.7. *Let* $\theta_1 < \theta_2$ *be given to fix the sector* $S(0)$ *as in Theorem* 2.3.5, *and let* $\delta > 0$ *be any real number such that* $n\delta < 1$. *Suppose that* $a_0(z) \not\equiv 0, a_1(z), \ldots, a_{n-1}(z)$ *are entire functions such that for real constants* $\alpha > 0$, $\beta > 0$, *we have for some* s, $1 \leqslant s \leqslant n-1$, *that*

$$\left|a_s(z)\right| \geqslant \exp\bigl((1+\delta)\alpha|z|^{\beta}\bigr),$$

$$\left|a_j(z)\right| \leqslant \exp\bigl(\delta\alpha|z|^{\beta}\bigr)$$

for all $j = 0, \ldots, s-1, s+1, \ldots, n-1$ *whenever* $|z|$ *is large enough in the sector* $S(0)$. *Given* $\varepsilon > 0$ *small enough, if* f *is a transcendental solution of finite order* ρ *of Eq.* (2.3.22), *then the following conditions hold*:

(i) *There exists* j, $0 \leqslant j \leqslant s-1$, *and a constant* $b_j \neq 0$ *such that* $f^{(j)} \to b_j$ *as* $z \to \infty$ *in the sector* $S(\varepsilon)$. *More precisely,*

$$\left|f^{(j)}(z) - b_j\right| \leqslant \exp\bigl(-(1-n\delta)\alpha|z|^{\beta}\bigr)$$

in $S(\varepsilon)$, *provided* $|z|$ *is large enough.*

(ii) *For each integer* $m \geqslant j+1$,

$$\left|f^{(m)}(z)\right| \leqslant \exp\bigl(-(1-n\delta)\alpha|z|^{\beta}\bigr)$$

in $S(3\varepsilon)$ *for all* $|z|$ *large enough.*

THEOREM 2.3.8. *Let* $a_0(z) \not\equiv 0, a_1(z), \ldots, a_{n-1}(z)$ *be entire functions, let* $\alpha > 0$, $\beta > 0$ *be given constants, let* $\delta > 0$ *be a real number such that* $n\delta < 1$ *and let* s *be an integer such that* $1 \leqslant s \leqslant n-1$. *Suppose that* $\rho(a_s) < \beta$ *for all* $j \neq s$ *and that for any given* $\varepsilon > 0$, *there exist two finite collections of real numbers that satisfy* $\phi_1 < \theta_1 < \phi_2 < \theta_2 < \cdots < \phi_m < \theta_m < \phi_{m+1} = \phi_1 + 2\pi$ *such that* $\sum_{j=1}^{m}(\phi_{j+1} - \phi_j) < \varepsilon$ *and*

$$\left|A_s(z)\right| \geqslant \exp\bigl((1+\delta)\alpha|z|^{\beta}\bigr)$$

as $z \to \infty$ *in each of the sectors* $\phi_j \leqslant \arg z \leqslant \theta_j$, $j = 1, \ldots, m$. *Then every transcendental solution* f *of* (2.3.22) *is of infinite order.*

Another result in the same line is due to Yang in [193], as a generalization of the corresponding theorem when $n = 2$:

THEOREM 2.3.9. *Let* E *be a set in the positive real axis of strictly positive upper density, and suppose that for some constants* $0 \leqslant \beta < \alpha$ *and* $\mu > 0$, *we have*

$$\left|a_0(z)\right| \geqslant \exp\bigl(\alpha|z|^{\mu}\bigr)$$

and

$$\left|a_j(z)\right| \leqslant \exp\bigl(\beta|z|^{\mu}\bigr),$$

for $j = 1, \dots, n-1$ *as* $|z| \to \infty$, $|z| \in E$. *Then all non-trivial solutions of* (2.3.22) *are of infinite order and, in fact,* $\rho_2(f) \geqslant \mu$.

Proceeding now to the zero distribution of solutions, the present research status is much more scattered than it is in the special second order case (2.1.1), even in the reduced case of (2.3.23). The product $E := f_1 \cdots f_n$ still remains important, but no counterpart of the identity (2.2.20) is available in this more general situation. However, Bank and Langley were able to prove in [18] several results corresponding to previous ones in the case of (2.1.1):

THEOREM 2.3.10. *Suppose that* $n \geqslant 3$ *in Eq.* (2.3.23) *and that the coefficients* $a_0(z), \dots, a_{n-2}(z)$ *are entire functions of finite order. Moreover, suppose that* (2.3.23) *admits a solution base* $f_1, \dots, f_n$ *each with* $\lambda(f_j) < \infty$. *Then their product* E *is of finite order as well.*

THEOREM 2.3.11. *Suppose that* $n \geqslant 3$ *in Eq.* (2.3.23) *and that the coefficients* $a_0(z), \dots, a_{n-2}(z)$ *are entire functions. Moreover, assume that* (i) a_0 *is transcendental of order* $\rho(a_0) < 1/2$ *and* (ii) $\rho(a_j) < \rho(a_0)$ *for* $j = 1, \dots, n-2$, *if* $\rho(a_0) > 0$, *while* $a_1, \dots, a_{n-2}$ *are polynomials, if* $\rho(a_0) = 0$. *Then Eq.* (2.3.23) *cannot admit two linearly independent solutions* f_1, f_2 *each with* $\lambda(f_j) < \infty$.

THEOREM 2.3.12. *Suppose that* $n \geqslant 3$ *in Eq.* (2.3.23) *and that the coefficients* $a_0(z), \dots, a_{n-2}(z)$ *are entire functions of finite order such that* $\rho(a_j) < \rho := \rho(a_0)$ *for all* $j = 1, \dots, n-2$. *Moreover, suppose that Eq.* (2.3.23) *admits a solution base* $f_1, \dots, f_n$ *each with* $\lambda(f_j) < \infty$. *Then* ρ *is a natural number and* $\rho(E) = \rho$.

THEOREM 2.3.13. *Suppose that* $n \geqslant 3$ *in Eq.* (2.3.23) *and that the coefficients* $a_0(z), \dots, a_{n-2}(z)$ *are entire functions. Moreover, assume that* (i) a_0 *is transcendental of order* $\rho(a_0) < 1/2$ *and* (ii) $\rho(a_j) < \rho(a_0)$ *for* $j = 1, \dots, n-2$, *if* $\rho(a_0) > 0$, *while* $a_1, \dots, a_{n-2}$ *are polynomials, if* $\rho(a_0) = 0$. *Then, given two linearly independent solutions* f_1, f_2 *of* (2.3.23), *we have that the quotient* f_1/f_2 *is of infinite order.*

A natural question is now whether somewhat similar results could be obtained, provided the dominant coefficient is some of the $a_1, \dots, a_{n-2}$ instead of a_0. Most important results of this type are due to Langley [115] and to Hellerstein, Miles and Rossi [72]. We first give a growth result from [72]:

THEOREM 2.3.14. *Considering linear differential equation* (2.3.22), *suppose that for some* k, $0 \leqslant k \leqslant n-1$, $\rho(a_j) < \rho(a_k) \leqslant 1/2$ *for all* $j \neq k$. *Then every solution of* (2.3.22) *is either a polynomial or an entire function of infinite order.*

Concerning zero distribution of a solution base of the reduced equation (2.3.23), corresponding to Theorem 2.2.1(a), we have the following result, see [115]:

THEOREM 2.3.15. *In the higher order case* $n \geqslant 3$, *suppose that all coefficients* a_j, $j = 0, \dots, n-2$, *of* (2.3.23) *are of order* $\rho(a_j) < \frac{1}{2(n-1)}$. *Then the linear differential* *equation* (2.3.23) *cannot have linearly independent solution* $f_1, \dots, f_n$ *each satisfying* $\lambda(f_j) < \infty$.

For additional results of similar type, see [111].

2.4. *Homogeneous linear differential equations with meromorphic coefficients*

If a linear differential equation (2.3.22) has meromorphic coefficients $a_j(z)$, $j = 1, \dots, n-1$, of which at least one is non-entire, then the existence of a meromorphic solution, resp. of a meromorphic solution base, in the whole complex plane becomes the first problem to be treated. Indeed, algebraic and logarithmic branch points may appear for such linear differential equations more as a rule than as an exception. As a simple classical example, consider a linear differential equations of type

$$f'' + \frac{h(z)}{(z - z_0)^2} f = 0$$

with $h(z)$ analytic in an open disc $B(z_0, R)$. Supposing that the indicial equation

$$\rho(\rho - 1) + h(z_0) = 0$$

has its roots ρ_1, ρ_2 such that $\rho_1 - \rho_2$ is a non-zero integer, then the equation admits two linearly independent solutions

$$f_1(z) = (z - z_0)^{\rho_1} \sum_{j=0}^{\infty} a_j (z - z_0)^j, \quad a_0 \neq 0,$$

$$f_2(z) = k f_1(z) \log(z - z_0) + (z - z_0)^{\rho_2} \sum_{j=0}^{\infty} b_j (z - z_0)^j,$$

where $k = 0$ or $k = 1$, in some slit disc $D := B(z_0, r) \setminus \{z_0 + t; 0 \leqslant t < r\}$, $r \leqslant R$. On the other hand, any meromorphic function f may appear as a global solution by just defining $A(z) := -f''(z)/f(z)$. Therefore, restricting ourselves to the most simple case of (2.1.1) with a meromorphic coefficient $A(z)$, a natural problem is to ask, when Eq. (2.1.1) admits two linearly independent meromorphic solutions in the whole complex plane, i.e. to characterizing when all solutions of (2.1.1) are globally meromorphic. To this end, the following result may be found, see [107], pp. 120–124:

THEOREM 2.4.1. *A linear differential equation* (2.1.1) *with* $A(z)$ *meromorphic in the complex plane admits two meromorphic solutions in the complex plane if and only if at all poles* z_0 *of* $A(z)$, *the Laurent expansion of* $A(z)$ *at* z_0 *is of the form*

$$A(z) = \frac{b_0}{(z - z_0)^2} + \frac{b_1}{z - z_0} + b_2 + \cdots,$$

where

$$4b_0 = 1 - m^2$$

for an odd integer $m \geqslant 3$, *and where the determinant*

$$D(z_0) = \begin{vmatrix} 1-m & 0 & \cdots & 0 & b_1 \\ b_1 & 4-2m & & & b_2 \\ b_2 & b_1 & \ddots & (0) & b_3 \\ \vdots & \vdots & & & \vdots \\ b_{m-2} & b_{m-3} & \cdots & (m-1)^2-(m-1)m & b_{m-1} \\ b_{m-1} & b_{m-2} & \cdots & b_1 & b_m \end{vmatrix} \tag{2.4.25}$$

vanishes.

In such a situation, when all solutions are meromorphic in the plane, a number of results are known to describe the distribution of zeros and poles of solutions, see [17]. It remains mostly open, however, to express these results in terms of the coefficient $A(z)$ only. To give a simple example, let f_1, f_2 be a meromorphic solution base. Then it is easy to see, from $(f_2/f_1)' = 1/f_1^2$, that all solutions have the same order of growth. On the other hand, contrary to the case of $A(z)$ entire, it may happen that all solutions are of finite order. A simple example may be provided by $f'' - (2/\cos^2 z)f = 0$ with $f_1(z) = \tan z$, $f_2(z) = 1 + z\tan z$. However, it remains open to characterize, in terms of $A(z)$ only, when the solutions are of finite order. Similarly, it is not known, in terms of $A(z)$ again, for which $A(z)$, there exist two linearly independent solutions f_1, f_2, each being zero-free in the whole complex plane.

Concerning the case when all solutions of (2.1.1) with a meromorphic coefficient $A(z)$ are of finite order, we may look at the Lamé equation

$$f'' - \big(q(q+1)\wp(z) + B\big)f = 0, \tag{2.4.26}$$

where $\wp(z)$ is the Weierstraß $\wp$-function, q a natural number and B a complex constant. In this case of (2.4.26), all solutions are meromorphic and if non-trivial, they are of order $\rho(f) = 2$. Moreover, we have $\lambda(f - a) = 2$ for all $a \in \mathbb{C}$, see [157], pp. 594–595. More generally, the same conclusions hold for all non-trivial solutions of (2.1.1), provided the coefficient $A(z)$ is a doubly periodic meromorphic function such that its Laurent expansion $A(z) = (z - z_0)^2 \sum_{j=0}^{\infty} b_j (z - z_0)^j$ satisfies around every pole z_0 of $A(z)$ the following conditions:

(1) $b_0 = -q(z_0)(q(z_0) + 1)$ for an integer $q(z_0)$,

$$(2)\ D(a) = \begin{vmatrix} \mu_1 & 0 & \cdots & & 0 & b_1 \\ b_1 & \mu_2 & \ddots & (0) & \vdots & b_2 \\ b_2 & b_1 & \ddots & \ddots & \vdots & b_3 \\ \vdots & \vdots & \ddots & \ddots & 0 & \vdots \\ b_{2q(a)-1} & b_{2q(a)-2} & \cdots & b_1 & \mu_{2q(a)} & b_{2q(a)} \\ b_{2q(a)} & b_{2q(a)-1} & \cdots & b_2 & b_1 & b_{2q(a)+1} \end{vmatrix} = 0,$$

where $\mu_j = j^2 - (2q(z_0)+1)j$, $1 \leqslant j \leqslant 2q(z_0)$. This more general case may be regarded as a generalized Lamé equation.

In the case, when the coefficient $A(z)$ in (2.1.1) is periodic of period ω, but not doubly periodic, and all solutions are supposed to be meromorphic, we may apply, under mild conditions, the Floquet theory to conclude that there exists a solution base, the members of which are of the type $f_j(z) = R_j(z)\exp(\beta_j z)$, where β_j is a complex constant, and $R_j(z)$ is rational in $\exp(2\pi i z/\omega)$, see [187]. For a number of results in the periodic case, see [160].

For a few examples of results describing the distribution of zeros and poles of meromorphic solutions in terms of $A(z)$ only, see [17]:

THEOREM 2.4.2. *Let $A(z)$ be a transcendental meromorphic function of order ρ such that all solutions of* (2.1.1) *are meromorphic in the complex plane. Then we have*

(a) *If $0 < \rho < \infty$, and if f_1, f_2 are linearly independent, then*

$$\max\bigl(\bar{\lambda}(f_1), \bar{\lambda}(f_2), \lambda(1/f_1)\bigr) \geqslant \rho.$$

(b) *If $\rho = 0$, then at least one of the following three sets is an infinite set of points: the zero-set of f_1, the zero-set of f_2, the pole-set of f_1.*

(c) *If $\rho > 0$ and if $\bar{\lambda}(A) < \rho$, then $\max(\bar{\lambda}(f), \bar{\lambda}(1/f)) \geqslant \rho$ for all non-trivial solutions f.*

(d) *If $\bar{\lambda}(f_1) < \infty$ and $\bar{\lambda}(f_2) < \infty$ for two linearly independent solutions f_1, f_2, then $\max(\bar{\lambda}(f), \bar{\lambda}(1/f)) = \infty$, unless all solutions are of finite order.*

Observe that these assertions represent a meromorphic counterpart of Theorem 2.2.1.

Of course, more refined information about the zeros of solutions of (2.3.22) follow if one is restricting the subset of $\mathbb{C}$ to be considered. An interesting paper of this type is due to Y. Il'yashenko and S. Yakovenko [89]. They consider a finite interval in $K \subset \mathbb{R}$, and its sufficiently small neighborhood U such that the coefficients $a_j(z)$ of (2.3.22) are analytic in U. Provided $|a_j(z)| \leqslant A$ in U for all $j = 0, \ldots, n-1$, then all non-trivial solutions that are real on K have at most $\beta(A+n)$ distinct zeros on K, where the constant β depends on the geometry of the pair (K, U) only. A more general result, removing the assumption that a solutions is real on K, was due to D. Novikov and S. Yakovenko in [138]:

THEOREM 2.4.3. *Let $K \subset U$ be a compact set in an open set $U \subset \mathbb{C}$ and suppose that K, U both have simply connected interiors and piecewise smooth boundaries. More-*

over, suppose that K *equals to the closure of its interior. Then all non-trivial solutions* f *of* (2.3.22) *with coefficients analytic in a neighborhood of* $\bar{U}$ *have at most* $\beta(A + n\log n)$ *distinct zeros in* K, *where now* $A := \max_{j=0,\dots,n-1}\max_{z\in U}|a_j(z)|$.

For a number of related generalizations, see [138].

2.5. *Non-homogeneous linear differential equations*

In this section, we describe key results concerning the growth and oscillation of solutions of non-homogeneous linear differential equations

$$f^{(n)} + a_{n-1}(z)f^{(n-1)} + \cdots + a_0(z)f = H(z) \tag{2.5.27}$$

with entire coefficients. It is well-known that all solutions of (2.5.27) are entire functions, see [107], Proposition 8.1. A typical phenomenon here is that the growth of solutions is essentially determined by the coefficients in the homogeneous part of (2.5.27), provided the finiteness degree of growth of $H(z)$ satisfies that $i(H) \leqslant 1 + \max_{j=0,\dots,n-1} i(a_j)$. As for the situation, when the coefficients $a_j(z)$, $j = 0, \dots, n-1$ are polynomials, while $H(z)$ is of finite order of growth, see two papers [55,57] by Gundersen, Steinbart and Wang. In what follows, we partially apply the same notations (2.1.12)–(2.1.15) as used with respect to the corresponding homogeneous equation (2.1.10) above. While looking at the results below, it also makes sense to compare with the corresponding results related to the homogeneous case (2.1.10) with polynomial coefficients. Recall that all solutions of the corresponding homogeneous linear differential equation are of growth

$$\rho(f) \leqslant \lambda := 1 + \max_{0\leqslant j\leqslant n-1} \frac{d_j}{n-j}$$

at most, while there always exist solutions of growth

$$\rho(f) = 1 + \max_{0\leqslant j\leqslant n-1} \frac{d_j}{n-j} = \alpha_1.$$

In the present non-homogeneous situation, elementary order considerations show that we always have $\rho(f) \geqslant \rho(H)$ for all solutions of (2.5.27). More precisely, we obtain

THEOREM 2.5.1. *All solutions* f *of* (2.5.27) *satisfy either* $\rho(f) = \rho(H)$, *or* $\rho(f) = \alpha_j$ *for some* $1 \leqslant j \leqslant p$, *provided* $\alpha_j > \rho(H)$, *and there always exists a solution* f *such that* $\rho(f) = \max(\rho(H), 1 + \alpha_1)$. *Moreover, if* $\rho(H)$ *is an integer, and if* $\rho(f) > \rho(H)$, *then* $\rho(f) \geqslant \rho(H) + 1/n$.

The last inequality here is sharp as shown by Example 4.1 in [57], and the assumption that $\rho(H)$ is an integer cannot be deleted, see Example 4.2 in [57].

Perhaps more interesting than the growth only is to see the influence of the non-homogeneity part H to the zero-distribution of f:

THEOREM 2.5.2. *For all solutions f of* (2.5.27), *we have*

$$\rho(f) - \lambda(f) \leqslant \rho(H) - \lambda(H).$$

Therefore, we always have $\lambda(f) \geqslant \lambda(H)$ and $\lambda(f) = \rho(f)$ as soon as $\lambda(H) = \rho(H)$.

THEOREM 2.5.3. *If $\lambda(H)$ is an integer, and if f is a solution of* (2.5.27) *such that $\lambda(f) > \lambda(H)$, then*

$$\lambda(f) \geqslant \lambda(H) + 1/n. \tag{2.5.28}$$

Again, observe that the inequality (2.5.28) is sharp, and the assumption that $\lambda(H)$ is an integer cannot be deleted, see examples in [57].

THEOREM 2.5.4. *If $\lambda(H) < \rho(H)$, and if f is a solution of* (2.5.27) *such that $\lambda(f) < \rho(f)$, then*

$$\rho(f) - \lambda(f) \geqslant \min\bigl(1/n, \rho(H) - \lambda(H)\bigr). \tag{2.5.29}$$

In particular, if $\rho(H) - \lambda(H) \leqslant 1/n$, and $\lambda(f) < \rho(f)$, then

$$\rho(f) - \lambda(f) = \rho(H) - \lambda(H). \tag{2.5.30}$$

Examples in [57] again show the sharpness of Theorem 2.5.4.

To obtain a counterpart of Theorem 2.1.7 in the non-homogeneous case (2.5.27), where $\lambda(H) < \rho(H) < \infty$, we need to write $H(z) = h(z)\mathrm{e}^{Q(z)}$, where $Q(z) = \frac{1}{\beta} b z^{\beta} + \cdots$ is a polynomial of degree β and h is an entire function such that $\lambda(h) = \rho(h) < \beta$. Let next A_j be the leading coefficient of $a_j(z) = A_j z^{d_j}$ in (2.5.27), and set $a_n(z) \equiv 1$. Setting then

$$\tau := \max_{0 \leqslant j \leqslant n} \bigl\{ d_j + j(\beta - 1) \bigr\},$$

we may define the constants $A_j^{\star}$ for $j = 0, \ldots, n$ as

$$A_j^{\star} := \begin{cases} A_k & \text{if } d_j + j(\beta - 1) = \tau, \\ 0 & \text{otherwise} \end{cases} \tag{2.5.31}$$

and a polynomial $H^{\star}(t)$ as

$$H^{\star}(t) := \sum_{j=0}^{n} A_j^{\star} t^j.$$

Using these notations, we obtain the following conclusions, see [55]:

THEOREM 2.5.5. *If b is a zero of multiplicity $m \geqslant 0$ of the polynomial $H^{\star}(t)$. Then Eq.* (2.5.27) *admits at most $m + 1$ linearly independent solutions f such that $\lambda(f) < \rho(f)$.*

In particular, each solution f *of* (2.5.27) *satisfies* $\lambda(f) < \rho(f)$ *if and only if* $H^{\star}(t) = (t-b)^n$.

Observe that this theorem also covers the case, when $H^{\star}(b) \neq 0$ showing that in this case their is at most one solution f such that $\lambda(f) < \rho(f)$.

THEOREM 2.5.6. *If the non-homogeneity part* H *in Eq.* (2.5.27) *is of infinite order, then all solutions* f *of* (2.5.27) *satisfy* $\lambda(f) = \rho(f) = \infty$ *with at most one exceptional solution* f_0.

Returning back to the finite vs. global oscillation properties treated above in the case of homogeneous linear differential equations (2.1.10) with polynomial coefficients, the corresponding behavior in the non-homogeneous case appears to be more complicated, see [7], Theorem 1.1:

THEOREM 2.5.7. *Suppose that all coefficients in Eq.* (2.5.27) *are polynomials with* $H(z) \not\equiv 0$. *Then there exist finitely many real numbers* s_j, $j = 1, \dots, m+1$, *with* $0 = s_0 < s_1 < \cdots < s_m < s_{m+1} = 2\pi$ *such that for each* k, $0 \leqslant k \leqslant m$, *one of the following properties hold*: (1) *for any* $\theta \in (s_k, s_{k+1})$ *and for any* $\varepsilon > 0$, *there is a solution of* (2.5.27) *having infinitely many zeros in* $|\arg z - \theta| < \varepsilon$ *or* (2) *for any* $\varepsilon > 0$, *any solution of* (2.5.27) *admits at most finitely many solutions in the sector* $s_k + \varepsilon \leqslant \arg z \leqslant s_{k+1} - \varepsilon$.

Therefore, in the non-homogeneous case, Eq. (2.5.27) has either the finite oscillation property, the global oscillation property, or a mixture of these two properties. An example of the last situation is

$$f''' + (z^2+3)f'' + (2z^2+z+3)f' + (z^2+z+2)f = H(z),$$

which has the global oscillation property in the left half plane, while four rays $\theta = -\pi/2, -\pi/6, \pi/6, \pi/2$ determine the finite oscillation property of this equation in the right half plane, see [7], pp. 290–291, for details.

We finally remark that there are very few results concerning the growth and zero distribution of solutions of (2.5.27), with coefficients being entire functions so that at least one of the coefficients in the homogeneous part is transcendental. We restrict ourselves here to giving just a couple of typical results, see [27] and [49]:

THEOREM 2.5.8. *Suppose that the coefficients in* (2.5.27) *are entire functions of finite order so that* $\rho(a_j) < \rho(a_0)$, $j = 1, \dots, n-1$. *Then all solutions* f *of* (2.5.27) *satisfy* $\bar{\lambda}(f) = \lambda(f) = \rho(f) = \infty$ *with at most one exceptional solution* f_0 *of finite order. If such an exceptional solution* f_0 *exists, then* $\rho(f_0) \leqslant \max\{\rho(a_0), \rho(H), \bar{\lambda}(f_0)\} < \infty$.

THEOREM 2.5.9. *Suppose that the coefficients in* (2.5.27) *are entire functions of finite order so that at least one of* $a_j(z)$, $j = 0, \dots, n-1$ *is transcendental. Then at least one of the solutions of* (2.5.27) *satisfies* $\bar{\lambda}(f) = \lambda(f) = \rho(f) = \infty$.

See also [24] for a result concerning the iterated order of solutions.

3. Linear differential equations in the unit disc

Meromorphic solutions of linear differential equations in the unit disc behave to some extent similarly as to the case of the complex plane, provided the solutions are fast growing in the sense of the Nevanlinna theory, i.e. that $T(r, f)/\log(1/(1-r))$ is unbounded. This case will be considered in Section 3.4. The complementary case of slowly growing solutions, which becomes trivial in the plane case, needs extensive consideration in the unit disc. In fact, analytic functions in the unit disc typically appear to be of slow growth when they belong to certain complex function spaces such as Hardy spaces, Bloch space, Q_p-spaces for some p etc. This gives rise to the general problem of investigating connections between the function spaces where the solutions, resp. the coefficients belong to. This topic has attracted systematic studies only very recently. Quite likely, new observations and improvements to the presentation in this chapter are to be expected in the near future. Another key aspect in investigating linear differential equations in the unit disc is the distribution of zeros of solutions. Historically, this topic preceded growth considerations, and combining these classical results on zero distribution with the recent growth studies is an open task of research today.

3.1. *Classical results on zeros of solutions*

Certain results about the zeros of solutions of linear differential equations in a disc of the complex plane are classical, although perhaps not so well-known, see [76]. To give an example of such classical theorems, consider the linear differential equation

$$f'' + A(z)f = 0, \tag{3.1.1}$$

where $A(z)$ is analytic:

THEOREM 3.1.1. (See [76], p. 580.) *Suppose $A(z)$ is analytic in a disc $|z - z_0| < R$ and bounded there by $|A(z)| \leqslant M$. Let f be a solution of $f'' + A(z)f = 0$ in this disc. Then we have*:

(i) *If f satisfies the initial conditions $f(z_0) = 0$, $f'(z_0) = 1$, then f has no zeros in the punctured disc $0 < |z - z_0| < \min(R, \pi M^{-1/2})$.*

(ii) *If f satisfies the initial conditions $f(z_0) = 1$, $f'(z_0) = 0$, then f has no zeros in the disc $|z - z_0| < \min(R, \frac{1}{2}\pi M^{-1/2})$.*

This type of classical results, connected with the closely related problem of the univalence of solutions of (3.1.1) prompted quite a lot of activity in 1950's. Unfortunately, this activity gradually faded out, being perhaps again reviving. The basic notions to be considered in this connection are related to the oscillatory nature of solutions. For the sake of normalization, we shall consider the situation in the unit disc $\mathbb{D}$ only. Of course, by a simple rescaling, any finite disc in the complex plane could be treated as well.

The differential equation

$$f'' + A(z)f = 0, \tag{3.1.2}$$

where $A(z)$ is analytic in $\mathbb{D}$, is said to be *oscillatory*, if at least one of its non-trivial solutions has infinitely many zeros in the unit disc $\mathbb{D}$. Otherwise, Eq. (3.1.2) is called *non-oscillatory*. Two more notions applied in this connection are *disconjugate*, meaning that every non-trivial solution of (3.1.2) has at most one zero in $\mathbb{D}$, and *Blaschke-oscillatory* meaning that the zero sequence (z_j) of every non-trivial solution of (3.1.2) is a Blaschke sequence, i.e. that $\sum_{j=1}^{\infty}(1-|z_j|)$ converges.

The research activity in early 1950's was very much due to the short but important paper [132] by Nehari, see also [134,135]. In particular, he proved the following theorem, which connects the notion of univalence to solutions of (3.1.2) in the unit disc:

THEOREM 3.1.2. *Given two linearly independent solutions f_1, f_2 of Eq.* (3.1.2), *their quotient g takes a complex value α n times in $\mathbb{D}$ if and only if there exists a non-trivial solutions of* (3.1.2) *that has n zeros in $\mathbb{D}$. In particular, g is univalent if and only if Eq.* (3.1.2) *is disconjugate.*

Actually, the proof of this theorem is quite elementary: If $g = f_1/f_2$ takes a certain value $\alpha \in \mathbb{C}$ at some point z, then $f_1 - \alpha f_2$ vanishes at z.

Making use of Theorem 3.1.2, Nehari [132] proved the following non-oscillation theorem:

THEOREM 3.1.3. *If*

$$\left|A(z)\right| \leqslant \frac{1}{(1-|z|^2)^2}$$

in the unit disc $\mathbb{D}$, then Eq. (3.1.2) *is disconjugate.*

The same conclusion holds, if $|A(z)| \leqslant \frac{2}{1-|z|^2}$, see Pokornyi [142]. Observe that Theorem 3.1.3 is in some sense close to the best possible, as shown by Schwarz [152]:

THEOREM 3.1.4. *If the inequality in Theorem* 3.1.3 *holds close to the boundary of $\mathbb{D}$, i.e. in $r_0 \leqslant |z| < 1$ for some $r_0 > 0$, then equation* (3.1.2) *is non-oscillatory. Moreover, for each $\gamma > 0$, there exists $A(z)$ such that*

$$|A(z)| \leqslant \frac{1+4\gamma^2}{(1-|z|^2)^2}$$

and that Eq. (3.1.2) *becomes oscillatory.*

For the proof of the first assertion, see [152], pp. 160–162. On the other hand, taking

$$A(z) = \frac{1+4\gamma^2}{(1-z^2)^2},$$

the corresponding Eq. (3.1.2) has an analytic solution

$$f(z) = (z^2 - 1)^{1/2} \sin\left(\gamma \log \frac{1+z}{1-z} - C\right)$$

in the unit disc, which has infinitely many zeros in $\mathbb{D}$, provided C is real (and provided that the branches of square root and logarithm have been chosen suitably).

A substantial number of oscillation results for (3.1.2) in the unit disc come out by integration of $|A(z)|$. Denoting by $\mathrm{d}\sigma$ the standard area element in the complex plane, the following conclusions may be found in [124]:

THEOREM 3.1.5. *Equation* (3.1.2) *is disconjugate, resp. non-oscillatory, if* $\int_{\mathbb{D}} |A(z)|\,\mathrm{d}\sigma \leqslant \pi$, *resp.* $< \infty$.

PROOF. The first assertion is an immediate consequence of Theorem 3.1.3, combined with the standard estimate

$$\pi\left(1 - |\zeta|^2\right)^2 \left|A(\zeta)\right| \leqslant \int_{\mathbb{D}} \left|A(z)\right| \mathrm{d}\sigma, \quad |\zeta| < 1$$

for analytic functions in the unit disc.

The second assertion essentially follows from the invariance of area integrals of analytic functions under Möbius transformations of the independent variable, see [124], pp. 982–986. □

A corresponding pair of oscillation results follows by computing the boundary integral of $|A(z)|$ instead of the area integral. Observing that $I(\rho, A) := \int_0^{2\pi} |A(\rho e^{i\theta})|\,\mathrm{d}\theta$ is nondecreasing, we may define

$$I(1, A) := \lim_{\rho \to 1} I(\rho, A).$$

A counterpart to Theorem 3.1.5 is now

THEOREM 3.1.6. *Equation* (3.1.2) *is disconjugate, resp. non-oscillatory, if* $I(1, A) \leqslant 4\pi$, *resp.* $I(1, A) < \infty$.

PROOF. The first assertion follows by combining the disconjugacy criterion due to Pokornyi, see the remark after Theorem 3.1.3, with the inequality

$$2\pi\left(1 - |\zeta|^2\right)\left|A(\zeta)\right| \leqslant I(1, A)$$

due to Nehari, see [133], p. 127. On the other hand, the second assertion is an immediate consequence of the corresponding non-oscillation criterion in Theorem 3.1.5. □

For additional, related results, see [34].

An immediate problem open for further investigations is the sharpness of the bounds in disconjugacy criteria π and 4π in Theorems 3.1.5 and 3.1.6, respectively. The example

$f'' + \frac{\pi^2}{4} f = 0$ due to London [124] shows that the bounds are definitely not sharp. Indeed, we have in this case $\int_{\mathbb{D}} |A(z)|\, d\sigma = \frac{\pi^3}{4} \approx 2.467\pi$ and $I(1, A) = \frac{\pi^3}{2} \approx 4.935\pi$. On the other hand, the boundary integral value is not too far from the best possible bound as shown by the following example due to Heittokangas [67]: Considering

$$f'' + \frac{10\zeta^2 - z^2}{4\zeta^4} f = 0, \quad |\zeta| < 1,$$

it is not difficult to see that this equation has a solution with two zeros at $\pm\zeta$, and so the equation is not disconjugate. On the other hand, by numerical computation, we see that the area integral $\int_{\mathbb{D}} |A(z)|\, d\sigma(z) \approx 5\pi/(2\zeta^2)$, while $I(1, A) \approx 5.013\pi$ as $|\zeta|$ is conveniently chosen close to one. Therefore, it is natural to conjecture that the possible constant might be $5\pi/2$ and 5π, respectively.

It is trivial to observe that disconjugate and non-oscillatory equations (3.1.2) are Blaschke-oscillatory. However, characterizing Blaschke-oscillatory equations (3.1.2) remains defectively investigated by now. The key results into this direction read as follows, see Pommerenke [143] and Heittokangas [65,66]:

THEOREM 3.1.7. *If* $A(z)$ *is analytic in* $\mathbb{D}$ *and either* $\int_{\mathbb{D}} |A(z)|^{1/2}\, d\sigma(z) < \infty$ *or* $\int_{\mathbb{D}} |A(z)|(1 - |z|)\, d\sigma(z) < \infty$, *then Eq.* (3.1.2) *is Blaschke-oscillatory.*

THEOREM 3.1.8. *If* $A(z)$ *is analytic in* $\mathbb{D}$ *and Eq.* (3.1.2) *is Blaschke-oscillatory, then* $\int_{\mathbb{D}} |A(z)|^{\alpha}\, d\sigma(z) < \infty$ *for each* $\alpha \in (0, 1/2)$.

The first of these results is an easy consequence of [143], Theorem 5, while the second one can be proved by using the unit disc version of the second main theorem of Nevanlinna theory, see [66], p. 126. Observe that the two integral conditions in Theorem 3.1.7 are independent of each other, as one may see by [70], Section 4. Also observe that there exist Blaschke-oscillatory equations of the form (3.1.2) for which neither one of the conditions in Theorem 3.1.7 is not satisfied, as one may easily see by looking at the equation

$$f'' - \frac{4}{(1 - z)^4} f = 0,$$

see [66], pp. 121–122. Concerning the next result, see [65], pp. 55–56:

THEOREM 3.1.9. *Suppose the coefficient* $A(z)$ *in* (3.1.2) *belongs to the Hardy space* H^p. *If* $1/4 \leqslant p \leqslant \infty$, *then* (3.1.2) *is Blaschke-oscillatory. Moreover, if* $1 \leqslant p \leqslant \infty$, *then* (3.1.2) *is non-oscillatory.*

This is sharp in the following sense: Eq. (3.1.2) with $A(z) := 4/(1 - z)^4$ is not Blaschke-oscillatory, while $A \in H^p$ for all $p < 1/4$, but $A \notin H^{1/4}$, see [65], p. 56.

A natural question about zeros of solutions of (3.1.2) is their geometric distribution inside of the unit disc. This problem has been addressed by Schwarz [152], Beesack [21] and Herold [75] by considering the hyperbolic non-Euclidean distance between zeros, while

Schwarz [152] and Beesack [20] also considered Euclidean distance between zeros. To describe these results, recall first that the hyperbolic non-Euclidean distance between two points z_1, z_2 in the unit disc will be defined as

$$[z_1, z_2] := \int_\gamma \frac{\mathrm{d}\zeta}{1-|\zeta|^2},$$

where the path of integration γ is the circle arc joining z_1 and z_2, orthogonal to the unit circle. Schwarz now proved the following theorem:

THEOREM 3.1.10. *Let f be a non-trivial solution of* (3.1.2).

(a) *If*

$$|A(z)| \leqslant \frac{1+a^2}{(1-|z|^2)^2}$$

for some $a > 0$ and if z_1, z_2 are two zeros of f in $\mathbb{D}$, then

$$[z_1, z_2] > \log \frac{\sqrt{1+a^2}+1}{\sqrt{1+a^2}-1}.$$

(b) *On the other hand, if for any two zeros z_1, z_2 in the unit disc $\mathbb{D}$,*

$$[z_1, z_2] \geqslant \log \frac{\sqrt{1+a^2}+1}{\sqrt{1+a^2}-1}, \quad a > 0,$$

then

$$|A(z)| < \frac{3(1+a^2)}{(1-|z|^2)^2}.$$

Herold [75] improved the first part of the preceding theorem by proving

THEOREM 3.1.11. *Let f be a non-trivial solution of* (3.1.2) *and suppose that*

$$|A(z)| \leqslant \frac{1+a^2}{(1-|z|^2)^2}.$$

Then for any two zeros z_1, z_2 of f in the unit disc $\mathbb{D}$,

$$[z_1, z_2] \geqslant \frac{\pi}{a}.$$

Theorem 3.1.11 is in fact best possible, as shown by the following example which goes back to Schwarz, see [152], p. 171, and [75], p. 638:

The function in the unit disc $\mathbb{D}$ defined as

$$f(z) := \sqrt{1-z^2}\sin\left(\frac{\alpha}{2}\log\frac{1+z}{1-z}\right),$$

where $\log 1 = 0$, is a solution of

$$f'' + \frac{1+\alpha^2}{(1-z^2)^2} f = 0.$$

The zeros z_n of f in the unit disc are determined by

$$\log\frac{1+z_n}{1-z_n} = n\frac{2\pi}{\alpha}, \quad n \in \mathbb{Z}.$$

By an elementary computation,

$$[z_n, z_{n+1}] = \int_{z_n}^{z_{n+1}} \frac{\mathrm{d}t}{1-t^2} = \Big/_{z_n}^{z_{n+1}} \frac{1}{2}\log\frac{1+t}{1-t} = \frac{\pi}{\alpha}.$$

3.2. *Complex function spaces in the unit disc*

Describing the growth and zero distribution of solutions of linear differential equations in the unit disc essentially relies on characterizing the coefficients of the equation as well as of its solutions in terms of complex functions spaces in the unit disc. Indeed, we may consider meromorphic functions f in the unit disc in two classes: Defining

$$D(f) := \limsup_{r\to 1^-} \frac{T(r,f)}{\log(1/(1-r))},$$

we say that f is of *finite degree*, if $D(f) < \infty$, while if $D(f) = \infty$, we say that f is of *infinite degree*. Of course, we may also say that f is of degree β, if $D(f) = \beta$. Observe that the finiteness of degree of f could also be defined for analytic functions as

$$D_M(f) := \limsup_{r\to 1^-} \frac{\log M(r,f)}{\log(1/(1-r))}$$

instead of $D(f)$. Clearly, $D(f) \leqslant D_M(f)$. However, $D_M(f)$ is not equivalent to $D(f)$ in defining the finiteness of degree, as shown by $f(z) := \mathrm{e}^{1/(1-z)}$. In this case, $D_M(f) = \infty$ and $D(f) < \infty$ by simple computation.

It is important to observe that the unit disc version of Nevanlinna theory remains a powerful tool while considering fast growing solutions of complex differential equations in the unit disc, meaning that the solution under consideration is of infinite degree and is growing essentially faster than the coefficients of the equation. On the other hand, solutions of slow growth need different tools, as in this situation the growth of a solution is comparable with the growth of coefficients. Due to recent studies, the theory of complex functions spaces takes the role of a key device to attack the situation.

For the convenience of the reader, we shortly recall basic definitions of the complex analytic spaces needed in what follows.

Perhaps the most classical linear space of analytic functions in the unit disc is the *Hardy space* H^p of analytic functions f satisfying the inequality

$$\sup_{0\leqslant r<1}\left(\frac{1}{2\pi}\int_0^{2\pi}\left|f(re^{i\theta})\right|^p d\theta\right)^{1/p}<\infty$$

for $0<p<\infty$, while for $p=\infty$, the corresponding Hardy space H^∞ is just the space of analytic functions bounded in the unit disc. For basic properties of Hardy spaces, see [36]. To define a natural extension of Hardy spaces H^p, $0<p<\infty$, we consider the p-characteristic of an analytic function f defined as

$$m_p(r,f):=\left(\frac{1}{2\pi}\int_0^{2\pi}\left(\log^+\left|f(re^{i\theta})\right|\right)^p d\theta\right)^{1/p}\tag{3.2.3}$$

for $0\leqslant r<1$. Now, the *generalized Nevanlinna class* N^p consists of all analytic functions f in the unit disc for which $\sup_{0\leqslant r<1}m_p(r,f)<\infty$. Observe that N^p is actually a linear space, and for each p, $\bigcup_0^\infty H^p\subset N^p$. The classical Nevanlinna class then is nothing but $N:=N^1$.

We next define the *weighted Hardy spaces* H_q^p, resp. H_q^∞ for $0\leqslant q<\infty$ as analytic functions f in the unit disc satisfying

$$\sup_{0\leqslant r<1}(1-r^2)^q\left(\frac{1}{2\pi}\int_0^{2\pi}\left|f(re^{i\theta})\right|^p d\theta\right)^{1/p}<\infty,$$

resp.

$$\sup_{z\in\mathbb{D}}\left(1-|z|^2\right)^q\left|f(z)\right|<\infty.$$

Clearly, $H^p=H_0^p$ and $H^\infty=H_0^\infty$. Finally, we say that an analytic function $f\in H_{q,0}^\infty$, if

$$\lim_{|z|\to1^-}\left(1-|z|^2\right)^q\left|f(z)\right|=0,$$

and correspondingly for $f\in H_{q,0}^p$. A closely related analytic function space is the *Korenblum space*

$$\mathcal{H}:=\bigcup_{q\in[0,\infty)}H_q^\infty.$$

In some sense, Bergman, resp. weighted Bergman, spaces may be understood as the area integral counterparts of Hardy spaces. Given $0<p<\infty$ and $-1<\alpha<\infty$, the *weighted*

Bergman space A^p_α is defined as the family of analytic functions f in the unit disc $\mathbb{D}$ for which the area integral over the unit disc satisfies

$$\left(\int_{\mathbb{D}} \left|f(z)\right|^p \left(1-|z|^2\right)^\alpha \,\mathrm{d}\sigma\right)^{1/p} < \infty.$$

For $\alpha = 0$, we obtain the classical *Bergman space* A^p. Observe that $A^p \subset H^\infty_{2/p}$ and $H^\infty_p \subset A^{\varepsilon/p}$ for each $\varepsilon \in (0,1)$, see [64], p. 11.

Applying the derivative of an analytic function f in the unit disc, a collection of further complex function spaces follow. The classical starting point in this direction is the space $\mathcal{B}$ of *Bloch functions* defined as those functions satisfying

$$\sup_{z\in\mathbb{D}} \left(1-|z|^2\right)\left|f'(z)\right| < \infty.$$

Slightly more generally, we may consider the *α-Bloch space* $\mathcal{B}^\alpha$ defined by

$$\sup_{z\in\mathbb{D}} \left(1-|z|^2\right)^\alpha \left|f'(z)\right| < \infty$$

for $0 < \alpha < \infty$. A connection to weighted Hardy spaces appears by the equality $\mathcal{B}^\alpha = H^\infty_{\alpha-1}$ for $1 < \alpha < \infty$, see [196], Proposition 7.

There are now two standard types of corresponding area integral counterparts of Bloch spaces. The more classical of them is the notion of *Dirichlet spaces*, resp. *weighted Dirichlet spaces*. For $-1 < q < \infty$, an analytic function f in the unit disc is said to be in the weighted Dirichlet space $\mathcal{D}_q$, provided

$$\int_{\mathbb{D}} \left|f'(z)\right|^2 \left(1-|z|^2\right)^q \,\mathrm{d}\sigma < \infty.$$

Of course, $\mathcal{D} := D_0$ is the classical Dirichlet space.

We also need related notion of Dirichlet type spaces $\mathcal{D}^p$ consisting of analytic functions f in the unit disc such that

$$\int_{\mathbb{D}} \left|f'(z)\right|^p \left(1-|z|^2\right)^{p-1} \,\mathrm{d}\sigma < \infty.$$

Of course, this means that $f' \in A^p_{p-1}$.

More recent is the notion of Q_p-spaces, introduced by Aulaskari and Lappan [1] in 1994. These spaces are defined for $0 < p < \infty$ as analytic functions f in the unit disc that satisfy

$$\sup_{a\in\mathbb{D}} \int_{\mathbb{D}} \left|f'(z)\right|^2 \left(g(z,a)\right)^p \,\mathrm{d}\sigma(z) < \infty,$$

where $g(z,a) = \log((1-\bar{a}z)/(z-a))$ is the usual Green function of the unit disc with a logarithmic singularity at $a \in \mathbb{D}$. As shown by Aulaskari and Lappan in [1], Q_p equals to the Bloch space $\mathcal{B}$, provided $p > 1$. Moreover, $Q_1 = BMOA$, where *BMOA* is the space of analytic functions of *bounded mean oscillation*. It is also important to keep in mind the strict nesting property $Q_p \subset Q_q \subset BMOA \subset \mathcal{B}$ whenever $0 < p < q < 1$.

Finally, replacing the derivative with the spherical derivative, we obtain the definition of *α-normal functions* instead of *α-Bloch functions*: A meromorphic function f in the unit disc is said to belong to the class of *α-normal functions* $\mathcal{N}^\alpha$, provided

$$\sup_{z\in\mathbb{D}}\bigl(1-|z|^2\bigr)^\alpha f^\sharp(z) < \infty.$$

Of course, $\mathcal{N}^1$ is nothing but the classical class $\mathcal{N}$ of *normal meromorphic functions* in the unit disc. Finally, we say that f is strongly normal, if

$$\lim_{|z|\to 1^-}\bigl(1-|z|^2\bigr) f^\sharp(z) = 0.$$

3.3. *Growth estimates for homogeneous linear differential equations*

In this section, we recall some recent results, which have appeared to be extremely powerful while considering the growth of solutions of homogeneous linear differential equations

$$L(f) := f^{(n)} + a_{n-1}(z) f^{(n-1)} + \cdots + a_0(z) f = 0 \tag{3.3.4}$$

in the unit disc, with analytic coefficients. The first two results [68] below make use of the classical Gronwall lemma to obtain a growth estimate for solutions of (3.3.4). Surprisingly, these estimates are, qualitatively, best possible.

THEOREM 3.3.1. *Let f be a solution of* (3.3.4) *in a disc D_R of radius R, centered at the origin.*

(a) *If $0 < R \leqslant 1$, then there exist a constant $C_1 = C_1(k) > 0$, depending on the initial conditions at the origin, and a constant $C_2 = C_2(k) > 0$, such that*

$$\bigl|f(r\mathrm{e}^{\mathrm{i}\theta})\bigr| \leqslant C_1 \exp\left(C_2 \sum_{j=0}^{n-1}\sum_{k=0}^{j}\int_0^r \bigl|a_j^{(k)}(s\mathrm{e}^{\mathrm{i}\theta})\bigr|(R-s)^{n-j+k-1}\,\mathrm{d}s\right) \tag{3.3.5}$$

for all $\theta \in [0, 2\pi)$ and $r \in [0, R)$.

(b) *If $1 < R \leqslant \infty$, then there exist a constant $C_1 = C_1(k) > 0$, depending on the initial conditions at the origin, and a constant $C_2 = C_2(k) > 0$, such that*

$$\bigl|f(r\mathrm{e}^{\mathrm{i}\theta})\bigr| \leqslant C_1 r^{n-1} \exp\left(C_2 \sum_{j=0}^{n-1}\sum_{k=0}^{j}\int_0^r \bigl|a_j^{(k)}(s\mathrm{e}^{\mathrm{i}\theta})\bigr| s^{n-j+k-1}\,\mathrm{d}s\right) \tag{3.3.6}$$

for all $\theta \in [0, 2\pi)$ and $r \in (1, R)$.

Combining the p-characteristic of an analytic function with the preceding theorem it is a straightforward conclusion to obtain

COROLLARY 3.3.2. *Let f be a solution of* (3.3.4) *in a disc D_R of radius R, centered at the origin, and let $1 \leqslant p < \infty$ be given.*

(a) *If $0 < R \leqslant 1$, then there exist a constant $C_1 = C_1(k) > 0$, depending on the initial conditions at the origin, and a constant $C_2 = C_2(k) > 0$, both depending on p, such that*

$$m_p(r, f)^p \leqslant C_1 + C_2 \sum_{j=0}^{n-1} \sum_{k=0}^{j} \int_0^{2\pi} \int_0^r \left|a_j^{(k)}(s\mathrm{e}^{\mathrm{i}\theta})\right|(R-s)^{p(n-j+k-1)}\,\mathrm{d}s\,\mathrm{d}\theta \tag{3.3.7}$$

for all $r \in [0, R)$.

(b) *If $1 < R \leqslant \infty$, then there exist a constant $C_1 = C_1(k) > 0$, depending on f, and a constant $C_2 = C_2(k) > 0$, both depending on p, such that*

$$m_p(r, f)^p \leqslant C_1(\log r^{n-1})^p + C_2 \sum_{j=0}^{n-1} \sum_{k=0}^{j} \int_0^{2\pi} \int_0^r \left|a_j^{(k)}(s\mathrm{e}^{\mathrm{i}\theta})\right| s^{p(n-j+k-1)}\,\mathrm{d}s\,\mathrm{d}\theta \tag{3.3.8}$$

for all $r \in (1, R)$.

REMARK. Of course, the right-hand side integrals in (3.3.7) and (3.3.8) may be replaced by area integrals over the disc of radius r, if needed. The same remark also applies to (3.3.10).

Closely related estimates follow by applying a slight variant, see [68], of a comparison theorem due to Herold ([73], Satz 1):

THEOREM 3.3.3. *Let f be a solution of* (3.3.4) *in a disc D_R of radius R, where $0 < R \leqslant \infty$, centered at the origin, let b be the number of non-zero elements in the coefficient sequence $a_0, \dots, a_{n-1}$. Given $\theta \in [0, 2\pi)$ and $\varepsilon > 0$, if $z_\theta = \sigma \mathrm{e}^{\mathrm{i}\theta} \in D_R$ is such that $a_j(z_\theta) \neq 0$ for some of the coefficients, then we have for all r such that $\sigma < r < R$,*

$$\left|f(r\mathrm{e}^{\mathrm{i}\theta})\right| \leqslant C \exp\left(b \int_\sigma^r \max_{0 \leqslant j \leqslant n-1} \left|a_j(t\mathrm{e}^{\mathrm{i}\theta})\right|^{1/(n-j)}\,\mathrm{d}t\right), \tag{3.3.9}$$

where $C > 0$ is a constant depending on b and $a_j(z_\theta)$, $f^{(j)}(z_\theta)$, $j = 0, \dots, n-1$. Actually,

$$C \leqslant (1+\varepsilon) \max_{j=0,\dots,n-1} \left(\frac{|f^{(j)}(z_\theta)|}{b^j \max_{k=0,\dots,n-1} |a_k(z_\theta)|^{j(n-k)}} \right).$$

Similarly as in the Gronwall estimate case, the following corollary is an easy consequence:

COROLLARY 3.3.4. *Let f be a solution of* (3.3.4) *in a disc D_R of radius R, where $0 < R \leqslant \infty$, centered at the origin, and let $1 \leqslant p < \infty$. Then, for all $0 \leqslant r < R$, we have*

$$m_p(r, f)^p \leqslant C\left(1 + \sum_{j=0}^{n-1} \int_0^{2\pi}\int_0^r \left|a_j(se^{i\theta})\right|^{p/(n-j)} \,\mathrm{d}s\,\mathrm{d}\theta\right), \tag{3.3.10}$$

where $C = C(n) > 0$ is a constant depending on p, and on the initial values of f at a point z_θ, where at least one of the coefficients a_j takes a non-zero value.

We also recall certain pointwise estimates for the generalized meromorphic functions in the unit disc, corresponding to similar estimates due to Gundersen in the plane. For a more detailed exposition, see [33]. The first estimate here, see [69], Lemma B, is a variant of the original version in [33]:

THEOREM 3.3.5. *Given two integers k, j such that $k > j \geqslant 0$, $\varepsilon > 0$ and $\beta \in (0,1)$, if f is meromorphic in the unit disc $\mathbb{D}$ and $f^{(j)}$ does not vanish identically, then for all z such that $r = |z| \notin E$, where E is of finite logarithmic measure, i.e. $\int_E (1-r)^{-1}\,\mathrm{d}r < \infty$, we have*

$$\left|\frac{f^{(k)}(z)}{f^{(j)}(z)}\right| \leqslant \left(\left(\frac{1}{1-r}\right)^{2+\varepsilon} \max\left(\log\frac{1}{1-r}, T\big(s(r), f\big)\right)\right)^{n-j}, \tag{3.3.11}$$

where $s(r) := 1 - \beta(1-r)$. If f is of finite order $\rho(f)$ in the unit disc, then

$$\left|\frac{f^{(k)}(z)}{f^{(j)}(z)}\right| \leqslant \left(\frac{1}{1-r}\right)^{(n-j)(\rho(f)+2+\varepsilon)}, \quad r \notin E, \tag{3.3.12}$$

while if f is of finite iterated order $\rho_k(f)$ for some $k \geqslant 2$, then

$$\left|\frac{f^{(k)}(z)}{f^{(j)}(z)}\right| \leqslant \exp_{k-1}\left(\left(\frac{1}{1-r}\right)^{\rho(f)+\varepsilon}\right), \quad r \notin E. \tag{3.3.13}$$

Finally, we mention that the notion of disconjugacy treated relative to Eq. (3.1.1) above may be extended to the higher order case (3.3.4), see [116]: Equation (3.3.4) is said to be desconjugate in the unit disc $\mathbb{D}$, if all non-trivial solutions have at most $n-1$ zeros in $\mathbb{D}$. For disconjugacy results in this setting, see [58,101,116] and [117].

3.4. *Equations with fast growing solutions*

The key notion in these considerations, essentially started by Heittokangas in his thesis [64] in 2000, is the order, resp. iterated order, of an analytic, resp. meromorphic, function

f in the unit disc $\mathbb{D}$. Defining for analytic functions

$$\rho_j(f) := \limsup_{r\to 1^-} \frac{\log_j^+ T(r,f)}{-\log(1-r)}$$

and

$$\rho_{M,j}(f) := \limsup_{r\to 1^-} \frac{\log_{j+1}^+ M(r,f)}{-\log(1-r)},$$

it is immediate to observe that $\rho(f) \leqslant \rho_M(f) \leqslant \rho(f) + 1$ for the usual order, while $\rho_j(f) = \rho_{M,j}(f)$ for $j \geqslant 2$. Therefore, unless otherwise specified, we use $\rho_j(f)$ to describe the order in what follows in this section. A starting point for order considerations of solutions of homogeneous linear differential equations (3.3.4) in the unit disc with analytic coefficients is the following counterpart of the Frei theorem, Theorem 2.3.1, in the complex plane:

THEOREM 3.4.1. *Let a_j be the last coefficient in the sequence $a_0, \ldots, a_{n-1}$ not belonging to the Korenblum space $\mathcal{H}$. Then Eq.* (3.3.4) *possesses at most j linearly independent solutions of finite order of growth in $\mathbb{D}$.*

This theorem, due to Heittokangas, see [64], Theorem 6.3, has been recently generalized by using the notion of iterated order, see [25] and [69]. To this end, we introduce the unit disc counterpart of the *finiteness degree* $i(f)$ of f, and the related notions, see Section 2.3. So, $i(f) := 0$ for meromorphic functions of finite degree and $i(f) := \min\{s \in \mathbb{N};\ \rho_s(f) < \infty\}$ for meromorphic functions of infinite degree. Of course, this means that $i(f) = \infty$, if $\rho_s(f) = \infty$ for all $s \in \mathbb{N}$. Of course, we may use $\rho_M(f)$ as well, equivalently, to determine $i(f)$ for analytic functions in the unit disc. Exactly as in Section 2.3, we then define in this unit disc setting $\delta(L), \gamma_j(L), p(L)$ and $\kappa(L)$. Also, we may use $\gamma_{M,j}(L)$ and $\kappa_M(L)$, if needed, observing that $\gamma_\delta(L) = \gamma_{M,\delta}(L)$, if $\delta > 1$, and $\gamma_1(L) \leqslant \gamma_{M,1}(L) \leqslant \gamma_1(L) + 1$, if $\delta = 1$. Similarly, $\kappa(L) = \kappa_M(L)$, if $p(L) > 1$ and $\kappa(L) \leqslant \kappa_M(L) \leqslant \kappa(L) + 1$, if $p = 1$.

THEOREM 3.4.2. (See [25], Theorem 2.1.) *For Eq.* (3.3.4), *the following conditions always hold*:

(i) $\delta(L) \leqslant 1 + p(L)$,
(ii) *if* $1 < p < \infty$, *then* $\delta(L) = 1 + p(L)$ *and* $\kappa_M(L) = \gamma_{M,p+1}(L) = \gamma_{p+1} = \kappa(L)$,
(iii) *if* $p = 0$ *and* $m := \max_j D_M(a_j)$, *then* $\gamma_1(L) \leqslant \gamma_{M,1}(L) \leqslant 1 + m$.

There are a number of recent results, see [25,69] and [104], that are improvements of the preceding theorem:

THEOREM 3.4.3. (See [69], Theorem 1.1, [25], Theorem 2.2.) *Given $k \in \mathbb{N}$ and $\alpha \geqslant 0$, all solutions of* (3.3.4) *in $\mathbb{D}$ satisfy $\rho_{M,k+1}(f) \leqslant \alpha$ if and only if $\rho_{M,k}(a_j) \leqslant \alpha$ for all $j = 0, \ldots, n-1$. Moreover, if q, $0 \leqslant q \leqslant n-1$ is the largest index such that $\rho_{M,k}(a_q) =$*

$\max_{0\leqslant j\leqslant n-1}\rho_{M,k}(a_j)$, *then each solution base of* (3.3.4) *possesses at least* $n-q$ *linearly independent solutions* f *such that* $\rho_{M,k+1}(f)=\rho_{M,k}(a_q)$.

This result is sharp in the sense that examples exist of a solution base containing exactly $n-q$ solutions of maximal growth, see [69], Example 1.3. On the other hand, examples of (3.3.4) exist such that all non-trivial solutions are of maximal growth, see [69], Example 1.4.

THEOREM 3.4.4. (See [25], Theorem 2.3, [69], Theorem 1.2.) *Suppose* $0<p<\infty$ *and* $i(a_0)=p$. *If* $\max_{1\leqslant j\leqslant n-1} i(a_j)<p$ *or if* $\max_{1\leqslant j\leqslant n-1}\rho_{M,p}(a_j)<\rho_{M,p}(a_0)$, *then for all solutions* f *of* (3.3.4), $i(f)=p+1$ *and* $\rho_{M,p+1}(f)=\rho_{M,p}(a_0)$.

The preceding theorems may be slightly refined by introducing the notion of *iterated type*: Given $k\in\mathbb{N}$, and an analytic function f in the unit disc $\mathbb{D}$ such that $0<\rho_{M,k}(f)<\infty$, we define the iterated k-type as

$$\tau_{M,k}(f):=\limsup_{r\to 1^-}(1-r)^{\rho_{M,k}(f)}\log_k^+ M(r,f).$$

THEOREM 3.4.5. (See [69], Theorem 1.5.) *Given* $k\in\mathbb{N}$, *suppose that*

$$\max_{1\leqslant j\leqslant n-1}\rho_{M,k}(a_j)\leqslant\rho_{M,k}(a_0)$$

and that

$$\sum_{\rho_{M,k}(a_j)=\rho_{M,k}(a_0)}\tau_{M,k}(a_j)<\tau_{M,k}(a_0),$$

then all non-trivial solutions f *of* (3.3.4) *satisfy* $\rho_{M,k+1}(f)=\rho_{M,k}(a_0)$.

REMARK. Concerning Theorems 3.4.4 and 3.4.5, it remains open, whether a similar conclusion holds, if a dominating coefficient a_j exists, but $j\neq 0$.

In addition to the growth considerations of fast growing solutions of (3.3.4), published results about the oscillation of such solutions in the unit disc are practically non-existent. The only exception we know is a forth-coming paper [26] by Cao and Yi, devoted to considering zero distribution of solutions of $f''+A(z)f=0$, where the coefficient $A(z)$ is of infinite degree $D(f)=\infty$. In this case it is well known that $\rho_2(f)=\rho_M(A)$, and oscillation results parallel to what are known from the complex plane case immediately follow by applying the unit disc version of the Nevanlinna theory, see [26], Theorem 2.2 through Theorem 2.4. See also [28].

3.5. *Equations with slowly growing solutions*

As pointed out above, Nevanlinna theory becomes ineffective as soon as a solution f of a differential equation, say of (3.3.4), is slowly growing in the rough sense that the growth of a solution is not essentially faster than the growth of coefficients. Indeed, it may happen that the growth of solution is slower than that of the coefficients, as shown by the equation $f^{(n)} = (1-z)^{-\alpha} f = 0$, the solutions of which are bounded in the unit disc, provided $\alpha < n$. In the case of slowly growing solutions, complex function spaces appear to be the correct tool. The key questions regarding the growth may be formulated as follows, see [70], Section 1:

(1) Providing the coefficients a_j, $j = 0, \dots, n-1$ in (3.3.4) belong to a certain analytic function space (possibly depending on j), determine the function space(s) where solutions of (3.3.4) belong to.

(2) Providing that solutions f of (3.3.4) belong to a certain analytic function space, determine the function space(s) where the coefficients a_j belong to.

In what follows, the first problem here is called the *direct growth problem*, while the second is termed as the *inverse growth problem*. Concerning these two problems, the key results, see [70], may be expressed as follows:

THEOREM 3.5.1. *Let f be a non-trivial solution of* (3.3.4) *with analytic coefficients a_j, $j = 0, \dots, n-1$, in the unit disc $\mathbb{D}$. Then we conclude that*

(1) *$f \in N^p$, provided $1 \leqslant p < \infty$ and $a_j \in A^{p/(n-j)}$ for all $j = 0, \dots, n-1$, or $a_j \in A^p_{p(n-j-1)}$ for all $j = 0, \dots, n-1$,*

(2) *$f \in N$, provided $-1 < \alpha < 0$ and $a_j \in H^{1/(n-j)}_{(\alpha+1)(n-j)}$ for all $j = 0, \dots, n-1$,*

(3) *f is of finite degree $D(f)$, provided $a_j \in H^{1/(n-j)}_{n-j}$ for all $j = 0, \dots, n-1$,*

(4) *f is of finite order $\rho(f) \leqslant \alpha$, provided $0 < \alpha < \infty$ and $a_j \in A^{1/(n-j)}_{\alpha}$ for all $j = 0, \dots, n-1$, or $a_j \in A^1_{n-j-1+\alpha}$ for all $j = 0, \dots, n-1$,*

(5) *f is of finite order $\rho(f) \leqslant \alpha$, provided $\alpha > 0$ and $a_j \in H^{1/(n-j)}_{(\alpha+1)(n-j)}$ for all $j = 0, \dots, n-1$.*

Similarly, for the inverse problem, we obtain

THEOREM 3.5.2. *Let f be a non-trivial solution of* (3.3.4) *with analytic coefficients a_j, $j = 0, \dots, n-1$, in the unit disc $\mathbb{D}$. Then we conclude that*

(1) *The coefficients $a_j \in \bigcap_{0<p<1/(n-j)} A^p$ for all $j = 0, \dots, n-1$, provided all non-trivial solutions $f \in N$,*

(2) *the coefficients $a_j \in \bigcap_{0<p<1/(n-j)} H^p_{1/p}$ for all $j = 0, \dots, n-1$, provided all non-trivial solutions f are of finite degree $D(f)$,*

(3) *the coefficients $a_j \in \bigcap_{0<p<1/(n-j)} A^p_{\alpha}$ for all $j = 0, \dots, n-1$, provided all non-trivial solutions f are of finite order $\rho(f) \leqslant \alpha$,*

(4) *the coefficients $a_j \in \bigcap_{0<p<1/(n-j)} H^p_{(\alpha+1)/p}$ for all $j = 0, \dots, n-1$, provided $0 \leqslant \alpha < \infty$ and all non-trivial solutions f are of finite order $\rho(f) \leqslant \alpha$.*

For a quick, not quite complete, overview of the preceding two theorems, see the table in [70], Section 1. See also [70], Section 4, for a number of inclusions between the function spaces applied here.

Refining the reasoning used to prove Theorems 3.5.1(4) and 3.5.2(3), one may indeed even prove the following characterization, see [105]:

THEOREM 3.5.3. *All solutions f of the linear differential equations* (3.3.4) *satisfy $\rho(f) \leqslant \alpha$ if and only if the coefficients $a_j \in \bigcap_{0<p<1/(n-j)} A_\alpha^p$ for all $j = 0, \dots, n-1$.*

A slight improvement has been proved in [105] as follows:

THEOREM 3.5.4. *Denoting by $\mathbb{A}_\alpha^p$ the space of analytic function f in the unit disc $\mathbb{D}$ such that* $\alpha = \inf\{t \geqslant 0;\ f \in \bigcap_{0<s<p} A_t^s\}$, *we conclude that all non-trivial solutions f of* (3.3.4) *satisfy*

$$\min\left\{\alpha_0, \min_{j=1,\dots,n-1}\left\{\alpha_j + \frac{k(\alpha_0 - \alpha_j)}{j}\right\}\right\} \leqslant \rho(f) \leqslant \max_{j=0,\dots,n-1} \alpha_j,$$

provided the coefficients $a_j \in \mathbb{A}_{\alpha_j}^{1/(n-j)}$ for all $j = 0, \dots, n-1$.

Moreover, making use of the standard order reduction procedure, and using the same notations as in the preceding theorem, we conclude

THEOREM 3.5.5. *If $q \in \{0, \dots, n-1\}$ is the smallest index such that $\alpha_q = \max_{j=0,\dots,n-1} \alpha_j$, then each solution base of* (3.3.4) *possesses at least $n-q$ linearly independent solutions f of order $\rho(f) = \alpha_q$.*

Replies to the direct problem under the requirement that all solutions of (3.3.4) belong to, say H_p^∞, are not included in the preceding theorems. This type of recent investigations may be found in [71]. Main results in this important paper can be collected as follows:

THEOREM 3.5.6. *Given $0 < \delta < 1$, for every $p > 0$ there exists a constant $\alpha = \alpha(p, n) > 0$ such that if the coefficients a_j of* (3.3.4) *satisfy*

$$\sup_{|z| \geqslant \delta} \left|a_j(z)\right| \left(1 - |z|^2\right)^{n-j} \leqslant \alpha \tag{3.5.14}$$

for $j = 0, \dots, n-1$, then all solutions f of (3.3.4) *are in the weighted Hardy space H_p^∞. In particular, the conclusion remains true, if the supremum in* (3.5.14) *is taken over the whole unit disc $\mathbb{D}$.*

Proving the next theorem, see also [71], one has to make use of the Carleson measure, and the related notion of a Carleson box

$$S(I) := \left\{z \in \mathbb{D};\ \frac{z}{|z|} \in I, 1 - |I| \leqslant |z|\right\}$$

attached to a subarc of the unit disc boundary $\partial\mathbb{D}$ such that the arc length $|I| \leqslant 1$.

THEOREM 3.5.7. *Given* $0 < \delta < 1$, *all solutions* f *of* (3.3.4) *belong to the space* $\mathcal{D}^p \cap H_p^\infty$, *provided one of the following conditions hold*:

(1) *For every* $0 < p \leqslant 2$, *there exists a constant* $\alpha = \alpha(p, n)$ *such that the coefficients* a_j *satisfy the condition*

$$\sup_{\zeta \geqslant \delta} \int_{\mathbb{D}} \left|a_j(z)\right|^p \left(1 - |z|^2\right)^{p(n-j)-1} \frac{1 - |\zeta|^2}{|1 - \bar{\zeta} z|^2} \, \mathrm{d}\sigma_z \leqslant \alpha$$

for all $j = 0, \ldots, n-1$.

(2) *For every* $0 < p \leqslant 2$, *there exists a constant* $\beta = \beta(p, n)$ *such that the coefficients* a_j *satisfy the condition*

$$\sup_{0<|I|\leqslant \delta} \int_{S(I)} \left|a_j^k(z)\right|^p \left(1 - |z|^2\right)^{p(n+k-j)-1} \mathrm{d}\sigma_z \leqslant \beta$$

for all $j = 0, \ldots, n-1$ *and all* $k = 0, \ldots, n-1$.

For additional results of similar type as the two theorems above, see [144]. We also refer to [43] for results concerning the growth of $A(z)$, provided f is an α-normal solution of $f^{(k)} + A(z) f = 0$.

4. Non-linear differential equations in a complex domain

4.1. *Introductory remarks*

We first recall three key results, repeatedly used in studying non-linear algebraic differential equations in the complex plane, namely those called as the Clunie, Mohon'ko and Valiron–Mohon'ko lemmas. In this section, we mostly refer to [107], where complete proofs and references to the original works can be found.

We start with the Clunie lemma, see [107], Lemma 2.4.2:

THEOREM 4.1.1. *Let* f *be a meromorphic solution of*

$$f^n P(z, f) = Q(z, f),$$

where $P(z, f), Q(z, f)$ *are polynomials in* f *and its derivatives with meromorphic coefficients, say* $a_\lambda, \lambda \in I$, *such that* $m(r, a_\lambda) = S(r, f)$ *for all* $\lambda \in I$. *If the total degree of* $Q(z, f)$ *in* f *and its derivatives is* $\leqslant n$, *then*

$$m\big(r, P(z, f)\big) = S(r, f).$$

REMARK. Observe that there are numerous variants of this theorem in the literature, see e.g. [107], Section 2.4. Perhaps the most general version may be found in a forth-coming paper [192] by Yang and Ye. This version reads as follows:

THEOREM 4.1.2. *Let f be a meromorphic solution of*

$$S(z,f)P(z,f)=Q(z,f),$$

where $S(z,f)$ is a polynomials in f, and $P(z,f)$, $Q(z,f)$ are differential polynomials in f, all three with small coefficients $a(z)$ in the sense of $m(r,a)=S(r,f)$. If the degree of $S(z,f)$ is greater or equal to the total degree of $Q(z,f)$ in f and its derivatives, then $m(r,P(z,f))=S(r,f)$.

We next proceed to the Mohon'ko lemma, see [107], Proposition 9.2.3:

THEOREM 4.1.3. *Let f be a meromorphic solution of an algebraic differential equation*

$$P(z,f,f',\dots,f^{(n)})=0, \tag{4.1.1}$$

where P is a polynomial in $f,f',\dots,f^{(n)}$ with meromorphic coefficients small with respect to f. If a complex constant c does not satisfy Eq. (4.1.1), *then*

$$m\left(r,\frac{1}{f-c}\right)=S(r,f).$$

In this case as well, there are several variants in the literature. We recall here the following version, due in this generality to Ishizaki and Yanagihara [95]:

THEOREM 4.1.4. *Let f be a meromorphic solution of an algebraic differential equation* (4.1.1), *with meromorphic coefficients small with respect to f. Moreover, let $Q(z,f)$ be an irreducible polynomial in f with meromorphic coefficients small with respect to f. If the function elements determined by equation $Q(z,f)=0$ do not satisfy Eq.* (4.1.1), *then we have for $Q=Q(z,f(z))$ that $m(r,1/Q)=S(r,f)$.*

REMARK. The Ishizaki–Yanagihara variant of the Mohon'ko lemma has not been used very often in the theory of complex differential equations. However, we believe the importance of this version has been underestimated.

To close this section, we recall the Valiron–Mohon'ko theorem, see [107], Theorem 2.2.5:

THEOREM 4.1.5. *Let f be a meromorphic function, and let $R(z,f)=P(z,f)/Q(z,f)$ be an irreducible rational function in f with meromorphic coefficients small with respect to f. Then the characteristic function of $R(z,f(z))$ satisfies*

$$T\bigl(r,R(z,f)\bigr)=dT(r,f)+S(r,f),$$

where $d:=\max(\deg_f P(z,f),\deg_f Q(z,f))$.

4.2. *Malmquist type theorems*

Malmquist type theorems essentially contain necessary existence conditions for meromorphic solutions of algebraic differential equations that are fast growing with respect to the coefficients of the equation. The original version, due to Malmquist [125] in 1913, reads as follows:

THEOREM 4.2.1. *If a differential equation of the form*

$$f' = R(z, f),$$

where $R(z, f)$ is rational in both arguments and irreducible in f, admits a transcendental meromorphic solution, then the equation reduces to a Riccati differential equation

$$f' = a_0(z) + a_1(z)f + a_2(z)f^2$$

with rational coefficients.

Proving this theorem with the aid of the Nevanlinna theory is almost trivial, just making use of the Valiron–Mohon'ko theorem and comparing the characteristic functions of both sides of the equation. Yosida was the first one to offer this type of proof in 1933. Introducing the notion of a *small function* $\alpha(z)$ with respect to another function f by the requirement $T(r, \alpha) = S(r, f)$, where $S(r, f)$ stands for a quantity $o(T(r, f))$ as $r \to \infty$ outside of a possible exceptional set of finite linear measure, a slight improvement easily follows by the same reasoning:

THEOREM 4.2.2. *If a differential equation of the form*

$$(f')^n = R(z, f), \tag{4.2.2}$$

where $R(z, f)$ is rational and irreducible in f with meromorphic coefficients, admits a meromorphic solution f such that all coefficients of $R(z, f)$ are small with respect to f, then (4.2.2) *reduces into*

$$(f')^n = \sum_{j=0}^{2n} \alpha_j(z) f^j, \tag{4.2.3}$$

where at least one of the coefficients $\alpha_j(z)$ does not vanish identically.

In fact, the reduced equation (4.2.3) may still be reduced further. This reduction was first made by Steinmetz [169] in the basic case of rational coefficients, then partially extended to the case of meromorphic coefficients by von Rieth [185], and finally completed in [62]. In the general case (with non-rational coefficients), the proof in [62] makes use of a slight extension of the Hayman–Miles theorem estimating $T(r, f)$ in terms of $T(r, f')$ in a set of positive lower logarithmic density, see [60] and [62], Lemma D:

THEOREM 4.2.3. *If a differential equation of the form* (4.2.2), *where* $R(z, f)$ *is rational and irreducible in* f *with meromorphic coefficients, admits a meromorphic solution* f *such that all coefficients of* $R(z, f)$ *are small with respect to* f, *then* (4.2.2) *reduces by a suitable non-singular Möbius transformation* $f = (\alpha y + \beta)/(\gamma y + \delta)$ *into one of the following equation types* (*or an integer power of one of these*):

$$y' = a(z) + b(z)y + c(z)y^2, \tag{4.2.4}$$

$$(y')^2 = a(z)\big(y - b(z)\big)^2(y - \tau_1)(y - \tau_2), \tag{4.2.5}$$

$$(y')^2 = a(z)(y - \tau_1)(y - \tau_2)(y - \tau_3)(y - \tau_4), \tag{4.2.6}$$

$$(y')^3 = a(z)(y - \tau_1)^2(y - \tau_2)^2(y - \tau_3)^2, \tag{4.2.7}$$

$$(y')^4 = a(z)(y - \tau_1)^2(y - \tau_2)^3(y - \tau_3)^3, \tag{4.2.8}$$

$$(y')^6 = a(z)(y - \tau_1)^3(y - \tau_2)^4(y - \tau_3)^5, \tag{4.2.9}$$

where $\tau_1, \tau_2, \tau_3, \tau_4$ *are complex constants, and the coefficients* $a(z), b(z), c(z)$ *are small meromorphic functions. Moreover,* $a(z)$ *does not vanish identically in* (4.2.5)–(4.2.9).

As pointed out by Steinmetz in [169], examples of all six types in Theorem 4.2.3 with a meromorphic solution such that the coefficients are small may easily be constructed.

Somewhat similar considerations concerning the following two differential equations

$$(f')^n = q(z)P(f)P_0(f')(f - z)^m \tag{4.2.10}$$

and

$$(f')^n = q(z)\mathrm{e}^{P_1(z)}P(f)(f - z)^m, \tag{4.2.11}$$

where q is a rational function, P, P_0, P_1 are polynomials, $P_0 \neq 0$ and m, n are natural numbers, have been made by Ishizaki and Wang in [94]. We remark that these differential equations have a certain interest in the field of complex dynamics, as shown by Bergweiler, Terglane and Wang. As in the case of Malmquist–Yosida equations, existence of transcendental meromorphic solutions implies that the equations reduce to certain special forms:

THEOREM 4.2.4. *If Eq.* (4.2.10) *admits a transcendental meromorphic solution, then* (4.2.10) *reduces to one of the following special forms*:

$$f' = q(z)(f - z),$$

$$f' = q(z)(f - z)^2,$$

$$f' = q(z)(f - \tau_1)(f - z),$$

$$(f')^2 = q(z)(f - \tau_1)(f - z)^2,$$

$$(f')^2 = q(z)(f - \tau_1)(f - \tau_2)(f - z)^2,$$

where τ_1, τ_2 *are distinct complex constants.*

THEOREM 4.2.5. *If Eq.* (4.2.11) *admits a transcendental meromorphic solution, then* (4.2.11) *reduces to one of the following special forms*:

$$\begin{aligned}
&f' = q(z)\mathrm{e}^{P_1(z)}(f - z),\\
&f' = q(z)\mathrm{e}^{P_1(z)}(f - z)^2,\\
&f' = q(z)\mathrm{e}^{P_1(z)}(f - \tau_1)(f - z),\\
&(f')^2 = q(z)\mathrm{e}^{P_1(z)}(f - \tau_1)(f - z)^2,\\
&(f')^2 = q(z)\mathrm{e}^{P_1(z)}(f - \tau_1)(f - \tau_2)(f - z)^2,
\end{aligned}$$

where τ_1, τ_2 *are distinct complex constants.*

Concerning first order differential equations in general, first recall a classical theorem due to L. Fuchs:

THEOREM 4.2.6. *If the solutions of a differential equation*

$$P(z, f, f') := \sum_{j=0}^{n} A_j(z, f)(f')^j = 0 \tag{4.2.12}$$

where $P(z, f, f')$ *is an irreducible polynomial in* f, f' *with analytic coefficients, is free of movable singularities other than poles, then*

$$\deg_f A_j(z, f) \leqslant 2(n - j).$$

In particular, $A_0(z, f)$ *is independent of* f.

We remark in this connection that the notion of *movable singularities* has had wide use in the theory of complex differential equations, although it is not easy to find a rigorous definition in the literature. We recall the following one due to Bieberbach in [23]: Considering a complex differential equation (D), let S_D be the set of those points z_0 where at least one of the solutions of (D) has a singularity other than a pole. Then (D) is *free of movable singularities*, if the interior of S_D is empty. Isolated points in S_D are called *fixed singularities* of (D).

Somewhat surprisingly, the assertion as in Theorem 4.2.6 results while extending Theorem 4.2.2 to equations of type (4.2.12). This extension is called to be of the second Malmquist theorem type, see [126] for the original paper and [38], Theorem 6. To find a rigorous proof of this result, see [38], Sections 2–4. In a form corresponding to the previous versions of the (first) Malmquist theorem, we may write

THEOREM 4.2.7. *If a differential equation of the form* (4.2.12), *where the left-hand side* $P(z, f, f')$ *is an irreducible polynomial in* f, f' *with meromorphic coefficients, admits a meromorphic solution* f *such that all coefficients are small with respect to* f, *then* $\deg_f A_j(z, f) \leqslant 2(n - j)$. *In particular,* $A_0(z, f)$ *is independent of* f.

As pointed out by A. Eremenko in [38], the preceding way of considering a solution f of a differential equation such that all coefficients are small with respect to f, may easily be expressed in a slightly more general form. To this end, let $\varphi : [0, \infty) \to (0, \infty)$ satisfy $\log r = \mathrm{O}(\varphi(r))$ as $r \to \infty$, and denote by $\mathcal{M}_\varphi$ the family of all meromorphic functions that satisfy $T(r, f) = \mathrm{O}(\varphi(r))$ as $r \to \infty$ outside of a possible exceptional set E, depending on f, of finite linear measure. As one may easily see, $\mathcal{M}_\varphi$ is an algebraically closed field. Supposing now that all coefficients of $P(z, f, f')$ are in $\mathcal{M}_\varphi$, we say that a solution f of Eq. (4.2.12) is *admissible*, provided $f \notin \mathcal{M}_\varphi$. We may now reformulate Theorem 4.2.7 as follows, see [38], Theorem 6, and [86], Section 4:

THEOREM 4.2.8. *If a differential equation of the form* (4.2.12), *where the left-hand side $P(z, f, f')$ now is an irreducible polynomial in f, f' with meromorphic coefficients in $\mathcal{M}_\varphi$, admits a meromorphic solution $f \notin \mathcal{M}_\varphi$, then*

$$\deg_f A_j(z, f) \leqslant 2(n - j).$$

In particular, $A_0(z, f)$ is independent of f.

It remains now to consider such issues as the growth, value distribution, factorization etc. of the special types of differential equations described above with fast growing meromorphic solutions. As an example of such results we recall the following theorem due to S. Bank and R. Kaufman in [11]:

THEOREM 4.2.9. *Let f be a transcendental meromorphic solutions of the Malmquist–Yosida equation* (4.2.2) *with polynomial coefficients. Then the order $\rho(f)$ is either zero, or a positive multiple of* $1/2$, *or a positive multiple of* $1/3$. *Conversely, any such number appears as the order of growth of a transcendental meromorphic solution in the plane of such Eq.* (4.2.2).

Observe that the assertion of Theorem 4.2.9 remains valid for all first-order algebraic differential equations with polynomial coefficients, as shown by Eremenko in [40].

Concerning Malmquist type theorems in the second order case, it would be natural to ask for a complete list of differential equations of type

$$f'' = R(z, f, f'), \tag{4.2.13}$$

where $R(z, f, f')$ is rational in f, f' with meromorphic coefficients, such that all the coefficients (4.2.13) are small with respect to some solution f that does not satisfy first order differential equations of type (4.2.12), save for a possible Möbius transformation type change of variable in f. The requirement not to satisfy a first order algebraic differential equations is quite obvious, since almost arbitrarily complicated second order differential equations may be derived from (4.2.12) by differentiation. It has been conjectured that the existence of a fast growing solution implies that (4.2.13) reduces to

$$f'' = L(z, f)(f')^2 + M(z, f)f' + N(z, f), \tag{4.2.14}$$

where $L(z, f), M(z, f), N(z, f)$ are rational in f with small meromorphic coefficients, save for a Möbius transformation in f. In this full generality, the problem remains open, and we are aware of just one result that could be called as a second order Malmquist type theorem:

THEOREM 4.2.10. *If Eq.* (4.2.13), *where* $R(z, f, f')$ *is rational in all of its arguments, possesses a meromorphic solution* f *of infinite order, then* (4.2.13) *reduces into* (4.2.14).

Observe that the proof in [120] is quite different to the usual Malmquist theorem reasoning based on Nevanlinna theory; this time the well known Zalcman lemma and some pointwise derivative estimates due to Barsegian will be applied. Observe as well that such a solution f cannot satisfy a first-order algebraic differential equation by the classical Gol'dberg theorem [50], see Theorem 4.7.1.

Except for the preceding Theorem 4.2.10, all results below related to the second order Malmquist type topic are restricted to considering equation (4.2.14) under the assumption that all coefficients are small with respect to a meromorphic solution f. We recall two results of this type:

THEOREM 4.2.11. (See [172], Theorem 2.) *Writing* (4.2.14) *in the form*

$$P_3(z, f)f'' = P_2(z, f)(f')^2 + P_1(z, f)f' + P_0(z, f),$$

where P_0, P_1, P_2, P_3 *are relatively prime polynomials with rational coefficients, suppose that* f *is a transcendental meromorphic solution. Then*

$$\deg_f P_3 \leqslant 4, \quad \deg_f(fP_2 - 2P_3) \leqslant 3, \quad \deg_f P_1 \leqslant 4, \quad \deg_f P_0 \leqslant 6.$$

These bounds are best possible as shown by [172], Example 1.

THEOREM 4.2.12. (See [121], Theorem 1.) *Suppose that* $L(z, f), M(z, f), N(z, f)$ *in Eq.* (4.2.14) *are rational in both arguments, and suppose that* f *is a transcendental meromorphic solution of* (4.2.14). *Moreover, we assume that*

$$L(z, f) = \frac{a_n(z)f^n + \cdots + a_s(z)f^s}{b_m(z)f^m + \cdots + b_r(z)f^r},$$

where either $m - n < 1$ *or* $r - s > 1$. *Then* $\rho(f) < \infty$.

Concerning this last result, the simple example $f(z) = \exp(e^z)$ solving $ff'' - (f')^2 - ff' = 0$ shows that the assumptions $m - n < 1, r - s > 1$ cannot be deleted simultaneously.

In the autonomous case, Liao and Ye [122] were able to extract the possible types with fast growing solutions as follows:

THEOREM 4.2.13. *If the differential equation* (4.2.14) *with constant coefficients possesses a transcendental meromorphic solution* f *of order* $\rho(f) > 2$, *then the equation reduces into one of the following nine special cases*:

$$f'' = \sum_{j=1}^{4} \frac{1}{f-a_j}\left(\frac{1}{2}(f')^2 + C_j\right) + b_0 f' + c_2 f^2 + c_1 f + c_0, \tag{4.2.15}$$

$$\begin{aligned} f'' = {} & \frac{1}{f-a_1}\big((f')^2 + B_1 f' + C_1\big) + \frac{1}{f-a_2}\left(\frac{1}{2}(f')^2 + C_2\right) \\ & + \frac{1}{f-a_3}\left(\frac{1}{2}(f')^2 + C_3\right) + b_0 f' + c_2 f^2 + c_1 f + c_0, \end{aligned} \tag{4.2.16}$$

$$\begin{aligned} f'' = {} & \frac{1}{f-a_1}\left(\frac{1}{2}(f')^2 + C_1\right) + \frac{1}{f-a_2}\left(\frac{1}{2}(f')^2 + C_2\right) \\ & + \frac{1}{f-a_3}\left(\frac{1}{2}(f')^2 + C_3\right) + (b_1 f + b_0) f' \\ & + c_3 f^3 + c_2 f^2 + c_1 f + c_0, \end{aligned} \tag{4.2.17}$$

$$\begin{aligned} f'' = {} & \frac{1}{f-a_1}\left(\frac{2}{3}(f')^2 + B_1 f' + C_1\right) + \frac{1}{f-a_2}\left(\frac{2}{3}(f')^2 + B_2 f' + C_2\right) \\ & + \frac{1}{f-a_3}\left(\frac{2}{3}(f')^2 + B_3 f' + C_3\right) + b_0 f' + c_2 f^2 + c_1 f + c_0, \end{aligned} \tag{4.2.18}$$

$$\begin{aligned} f'' = {} & \frac{1}{f-a_1}\left(\frac{1}{2}(f')^2 + C_1\right) + \frac{1}{f-a_2}\left(\frac{3}{4}(f')^2 + B_2 f' + C_2\right) \\ & + \frac{1}{f-a_3}\left(\frac{3}{4}(f')^2 + B_3 f' + C_3\right) + b_0 f' + c_2 f^2 + c_1 f + c_0, \end{aligned} \tag{4.2.19}$$

$$\begin{aligned} f'' = {} & \frac{1}{f-a_1}\left(\frac{1}{2}(f')^2 + C_1\right) + \frac{1}{f-a_2}\left(\frac{2}{3}(f')^2 + B_2 f' + C_2\right) \\ & + \frac{1}{f-a_3}\left(\frac{5}{6}(f')^2 + B_3 f' + C_3\right) + b_0 f' + c_2 f^2 + c_1 f + c_0, \end{aligned} \tag{4.2.20}$$

$$\begin{aligned} f'' = {} & \frac{1}{f-a_1}\big(A_1 (f')^2 + B_1 f' + C_1\big) + \frac{1}{f-a_2}\big(A_2 (f')^2 + B_2 f' + C_2\big) \\ & + (b_1 f + b_0) f' + c_3 f^3 + c_2 f^2 + c_1 f + c_0, \end{aligned} \tag{4.2.21}$$

$$\begin{aligned} f'' = {} & \frac{1}{f-a_1}\big(A_1 (f')^2 + B_1 f' + C_1\big) + (b_1 f + b_0) f' \\ & + c_3 f^3 + c_2 f^2 + c_1 f + c_0, \end{aligned} \tag{4.2.22}$$

$$f'' = (b_1 f + b_0) f' + c_3 f^3 + c_2 f^2 + c_1 f + c_0, \tag{4.2.23}$$

where a_j *for* $j = 1, \dots, 4$, B_j, C_j *for* $j = 1, \dots, 4$, b_1, b_0 *and* c_j $j = 1, \dots, 3$ *are complex constants, and where either* $A_1 = A_2 = 1/2$ *or* $A_1 = (k_1 - 1)/k_1$, $A_2 = (k_2 - 1)/k_2$ *and* $(k_1 - 1)/k_1 + (k_2 - 1)/k_2 = (k+1)/k$ *for some integers* $k_1 \geqslant 2, k_2 \geqslant 2$ *and* $k \geqslant 3$.

The proof of Theorem 4.2.13 again relies, among other tools, to the Zalcman lemma. As shown by concrete examples, see [122], all nine cases in Theorem 4.2.13 may actually appear. The non-autonomous case of (4.2.14) remains open. Although it is quite likely that a result similar to Theorem 4.2.13 would remain true for fast growing solutions, some modification of the proof in [122] is definitely needed.

In addition to Theorems 4.2.11 and 4.2.13, special cases of (4.2.14) have been treated in a number of papers: The case when $L(z, f) = M(z, f) \equiv 0$ has been due to Wittich in [188], the case $M(z, f) = N(z, f) \equiv 0$ by von Rieth [185] and the case $L(z, f) \equiv 0$ by Steinmetz [171]. Ishizaki gives in [92] a fairly detailed treatment of the case $L(z, f) \equiv 0$ as well as of the general case, provided $L(z, f) \equiv 1$. See also [91].

Finally, we mention that a number of Malmquist type theorems have been proved for complex differential equations of higher order. As a typical example of such theorems, we recall

THEOREM 4.2.14. (See [46], Satz 5, [169], p. 21.) *If a differential equation of the form*

$$P(z, f, f', \dots, f^{(n)}) := \sum_{\lambda \in I} c_\lambda(z) f^{\lambda_0} (f')^{\lambda_1} \cdots (f^{(n)})^{\lambda_n} = R(z, f) \tag{4.2.24}$$

where $P(z, f, f', \dots, f^{(n)})$ *is an irreducible polynomial in* $f, f', \dots, f^{(n)}$ *with meromorphic coefficients, and* $R(z, f)$ *is rational and irreducible in* f *with meromorphic coefficients, admits a meromorphic solution* f *such that all coefficients of* P *and* R *are small with respect to* f, *then* $R(z, f)$ *is a polynomial in* f *and*

$$\deg_f R(z, f) \leqslant \Delta := \max_{\lambda \in I} (\lambda_0 + \lambda_1 + \cdots + \lambda_n).$$

4.3. *Riccati differential equation*

As the most simple non-linear differential equation, the Riccati differential equation

$$f' = a(z) + b(z) f + c(z) f^2 \tag{4.3.25}$$

deserves a separate treatment, not the least due to its important role in optimal control (usually in the frame of matrix-valued complex differential equation). First recall (Corollary 1.2.7) that all solutions of (4.3.25) are meromorphic functions whenever the coefficients $a(z), b(z), c(z)$ are entire. The situation becomes more complicated, if the coefficients (or at least one of them) is non-entire meromorphic. Since (4.3.25) may be transformed to

$$f' = A(z) + f^2 \tag{4.3.26}$$

by a simple transformation, which preserves the meromorphic nature of the coefficients, we mostly refer to (4.3.26) in what follows.

THEOREM 4.3.1. *If the Riccati differential equation* (4.3.26) *admits three distinct meromorphic solutions in the complex plane, then all solutions of* (4.3.26) *are meromorphic, and they constitute a one-parameter family* $\mathfrak{F} := (f_C)_{C\in\mathbb{C}}$, *with* $f_0 \notin \mathfrak{F}$ *for one of the three original solutions. Two distinct meromorphic solutions are sufficient to imply the same conclusion, if for one of them, say* f_0, $2f_0$ *has integer residues at all of its poles.*

Concerning the number of distinct meromorphic solutions of (4.3.26) deduced from the properties of the meromorphic coefficient $A(z)$, the following conclusions can be made, see [9] and [107], Section 9.1:

THEOREM 4.3.2.

(1) *If* $A(z)$ *is entire, or if all poles* ζ *of* $A(z)$ *are double such that the coefficients of their Laurent expansions*

$$A(z) = \sum_{j=-2}^{\infty} b_{j+2}(z-\zeta)^j$$

satisfy that (i) $4b_0 = 1 - m^2$ *for a natural number* $m \geqslant 2$ *and that* (ii) *the determinant condition* (2.4.25) *holds, then all solutions of* (4.3.26) *are meromorphic.*

(2) *If* $A(z)$ *has a simple pole, then* (4.3.26) *has at most one meromorphic solution.*

(3) *If* $A(z)$ *has a double pole such that the coefficients of the Laurent expansion at this pole does not satisfy the conditions in* (1), *then* (4.3.26) *has at most two meromorphic solutions. If, however,* $4b_0 = 1$, *then at most one meromorphic solution exists.*

(4) *If* $A(z)$ *has a multiple pole of odd multiplicity, then* (4.3.26) *has no meromorphic solutions. If* $A(z)$ *has multiple poles only, all with even multiplicity, then two meromorphic solutions may exist.*

As an example of Theorem 4.3.2(1), see [108], we may consider equation $f'' - \frac{2}{\cos^2 z} f = 0$ which has two linearly independent meromorphic solutions $f_1(z) = \tan z$ and $f_2(z) = 1 + z\tan z$. By a simple computation, we see that the conditions given in Theorem 4.3.2(1) are satisfied. For a collection of examples covering the remaining cases, see [9].

It is well known that all meromorphic solutions of (4.3.25) are of finite order whenever $a(z), b(z), c(z)$ are rational functions. As for the case of non-rational coefficients, very few papers only may be found. We may refer to [9] for the normalized equation (4.3.26) with $A(z)$ entire, and to [166] for (4.3.26) with $A(z)$ doubly periodic meromorphic. Concerning basic observations for the value distribution of meromorphic solutions of (4.3.25), resp. of (4.3.26), see [107] and [9].

What also remains more or less open, is to consider Riccati differential equation in the unit disc. We mention here a result due to Yamashita [191]:

THEOREM 4.3.3. *Assume that in Eq.* (4.3.25) *the coefficients belong to the following weighted Hardy spaces*: $a(z) \in H_2^\infty, b(z) \in H_1^\infty, c(z) \in H_0^\infty$. *If* $\inf_{z\in\mathbb{D}} |c(z)| > 0$, *then each holomorphic solution* f *of* (4.3.25) *may be expressed in the form* $f = Cy''/y'$, *where* $C > 0$ *is a constant and* y *is a non-constant function analytic and univalent in* $\mathbb{D}$.

Quite likely, some more results could be easily obtained by recalling the close connection between (4.3.26) and the linear differential equation $u'' + A(z)u = 0$, and making use of the results described in Section 3.

4.4. *Schwarzian differential equations*

A Schwarzian differential equation means a complex differential equation of the form

$$\big(S(f)(z)\big)^p = R(z, f) = \frac{P(z, f)}{Q(z, f)}, \tag{4.4.27}$$

where p is a natural number, $R(z, f)$ is an irreducible rational function in f with meromorphic coefficients, and where

$$S(f)(z) := \left(\frac{f''}{f'}\right)' - \frac{1}{2}\left(\frac{f''}{f'}\right)^2 = \frac{f'''}{f'} - \frac{3}{2}\left(\frac{f''}{f'}\right)^2 \tag{4.4.28}$$

is the standard Schwarzian derivative of a meromorphic function f. We also denote by d the degree of $R(z, f)$ with respect to f.

Observe first that Schwarzian differential equations are closely related with Riccati differential equations. In fact, denote by $\mathcal{R}$ the class of meromorphic functions f that satisfy a Riccati differential equation (4.3.25), $f' = a(z) + b(z)f + c(z)f^2$, such that the coefficients $a(z), b(z), c(z)$ are small with respect to f. Moreover, denote by $\mathcal{R}_k$ the subclass of functions f in $\mathcal{R}$ that have k Picard exceptional values. Then we have the following result, see [92], Theorem 3.2.13:

THEOREM 4.4.1. *If* $f \in \mathcal{R}_k$ *with* $k = 1, 2$, *then* f *also satisfies a Schwarzian differential equation* (4.4.27) *with* $p = 1$ *and* $d = 4 - 2k$. *If* $f \in \mathcal{R}_0$, *then* f *also satisfies a Schwarzian differential equation* (4.4.27) $p = 1$ *and* $d = 4$ *or* $d = 2$ *depending on whether* $D(z) = b(z)^2 - 4a(z)c(z) \equiv 0$, *or not.*

Not surprisingly, the Schwarzian differential equation (4.4.27) has similarities with the Malmquist–Yosida differential equation (4.2.2). To this end, compare the following theorem from [90] with Theorem 4.2.3:

THEOREM 4.4.2. *Suppose that the autonomous Schwarzian differential equation*

$$S(f)(z) = R(f) \tag{4.4.29}$$

admits a transcendental meromorphic solution f. *Then for a non-singular Möbius transformation* $u := (\alpha f + \beta)/(\gamma f + \delta)$ *with complex coefficients, Eq.* (4.4.29) *reduces into one of the following forms*:

$$\begin{aligned}
S(u)(z) &= c\frac{(u-\sigma_1)(u-\sigma_2)(u-\sigma_3)(u-\sigma_4)}{(u-\tau_1)(u-\tau_2)(u-\tau_3)(u-\tau_4)},\\
S(u)^3(z) &= c\frac{(u-\sigma_1)^3(u-\sigma_2)^3}{(u-\tau_1)^3(u-\tau_2)^2(u-\tau_3)},\\
S(u)^3(z) &= c\frac{(u-\sigma_1)^3(u-\sigma_2)^3}{(u-\tau_1)^2(u-\tau_2)^2(u-\tau_3)^2},\\
S(u)^2(z) &= c\frac{(u-\sigma_1)^2(u-\sigma_2)^2}{(u-\tau_1)^2(u-\tau_2)(u-\tau_3)},\\
S(u)(z) &= c\frac{(u-\sigma_1)(u-\sigma_2)}{(u-\tau_1)(u-\tau_2)},\\
S(u)(z) &= c,
\end{aligned}$$

where $c \in \mathbb{C}$, *and* $\tau_1, \tau_2, \tau_3, \tau_4, \sigma_1, \sigma_2, \sigma_3, \sigma_4$ *are complex constants so that* $\tau_1, \tau_2, \tau_3, \tau_4$ *are mutually distinct.*

As for the corresponding non-autonomous case, Ishizaki proved in [93] the following corresponding result:

THEOREM 4.4.3. *Suppose that the non-autonomous Schwarzian differential equation* (4.4.27) *with* $p = 1$ *admits a transcendental meromorphic solution* f *such that all meromorphic coefficients of* $R(z, f)$ *are small with respect to* f. *Then either the solution* f *satisfies a Riccati differential equation with small meromorphic coefficients, or* f *satisfies a first order algebraic differential equation*

$$(f')^2 + B(z, f)f' + A(z, f) = 0 \tag{4.4.30}$$

where $A(z, f), B(z, f)$ *are polynomials in* f *with small meromorphic coefficients, or* (4.4.27) *reduces to one of the following two forms*:

$$\begin{aligned}
S(u)(z) &= \frac{P(z,u)}{(f+b(z))^2},\\
S(u)(z) &= c(z),
\end{aligned}$$

where $b(z), c(z)$ *are small meromorphic functions.*

REMARK. In the case of (4.4.30), we may apply a transformation $f = \tau + 1/w$, $\tau \in \mathbb{C}$, to obtain a first order differential equation of the same form (4.4.30) such that $\deg_w B(z, w) \leqslant 1$, $\deg_w A(z, w) \leqslant 3$.

For corresponding results concerning a more complicated Schwarzian type differential equation, see [87].

The growth of solutions of Schwarzian differential equations has been treated by Liao and Ye recently, see [123]:

THEOREM 4.4.4. *Suppose f is a transcendental meromorphic solution of a Schwarzian differential equation* (4.4.27) *with polynomial coefficients. Then the following conclusions hold*:

(1) *If f has a Picard exceptional value, then $\rho(f)$ is finite.*
(2) *If f has two completely ramified values, then $\rho(f) \leqslant 2$.*
(3) *If the Schwarzian equation under consideration is autonomous, then $\rho(f) \leqslant 2$.*
(4) *If the right-hand side of* (4.4.27) *is independent of f, hence a rational function $R(z) = P(z)/Q(z)$, then f is of finite order and*

$$\rho(f) = (m - n + 2p)/(2p) > 0,$$

where $m = \deg P$, $n = \deg Q$. In particular, if $R(z)$ reduces to a constant, then $\rho(f) = 1$.

Finally, we recall the following result concerning the deficiencies, see [90]:

THEOREM 4.4.5. *Suppose f is a transcendental meromorphic solution of a Schwarzian differential equation* (4.4.27) *with polynomial coefficients, and let $\alpha_1, \dots, \alpha_s$ be distinct complex numbers. Then*

$$d + 2p \sum_{j=1}^{s} \delta(\alpha_j, f) \leqslant 4p,$$

where $d = \deg_f R(z, f)$.

4.5. *Painlevé differential equations*

The following problem has been originally proposed by Picard: Given $R(z, f, f')$ rational in its arguments, what are the second order ordinary differential equations in the complex plane of the form $f'' = R(z, f, f')$ having the property that the singularities other than poles of any solution of this equation depend on the equation in question only and not on the constants of integration? Painlevé (1895) and Gambier (1910) proved that there are fifty canonical equations of this form having the property proposed by Picard. This property is known as the *Painlevé property*, and the differential equations possessing this property are called equations of Painlevé type. Among the fifty equations obtained by Painlevé and Gambier, the following six equations, known as the *Painlevé equations,* are the most interesting ones:

$$f'' = 6f^2 + z, \tag{4.5.31}$$

$$f'' = 2f^3 + zf + \alpha, \tag{4.5.32}$$

$$f'' = \frac{(f')^2}{f} - \frac{1}{z}f' + \frac{1}{z}(\alpha f^2 + \beta) + \gamma f^3 + \frac{\delta}{f}, \tag{4.5.33}$$

$$f'' = \frac{(f')^2}{2f} + \frac{3}{2}f^3 + 4zf^2 + 2(z^2 - \alpha)f + \frac{\beta}{f}, \tag{4.5.34}$$

$$\begin{aligned} f'' = {} & \frac{3f-1}{2f(f-1)}(f')^2 - \frac{1}{z}f' + \frac{1}{z^2}(f-1)^2\left(\alpha f + \frac{\beta}{f}\right) \\ & + \frac{\gamma f}{z} + \frac{\delta f(f+1)}{f-1}, \end{aligned} \tag{4.5.35}$$

$$\begin{aligned} f'' = {} & \frac{1}{2}\left(\frac{1}{f} + \frac{1}{f-1} + \frac{1}{f-z}\right)(f')^2 - \left(\frac{1}{z} + \frac{1}{z-1} + \frac{1}{f-z}\right)f' \\ & + \frac{f(f-1)(f-z)}{z^2(z-1)^2}\left(\alpha + \frac{\beta z}{f^2} + \frac{\gamma(z-1)}{(f-1)^2} + \frac{\delta z(z-1)}{(f-z)^2}\right), \end{aligned} \tag{4.5.36}$$

where $\alpha, \beta, \gamma, \delta$ are arbitrary complex constants. In fact, solutions of eleven other equations of the Painlevé type may be expressed in terms of solutions of the preceding six Painlevé equations, while the thirty-three remaining Painlevé type equations are solvable in terms of solutions of linear differential equations, or in terms of elliptic functions.

The Painlevé equations form a vast field nowadays, in particular due to the importance of their solutions as a kind of non-linear special functions. The field has remarkably extended during the last four decades or so, very much falling in the borderline area of mathematical physics. As for some of the several key references of this field, we may refer to [42,96,137] and [139], as examples of a number of excellent references. In this connection, we restrict ourselves to considering solutions of Painlevé differential equations from the point of view of complex analysis, with main emphasis in the value distribution of solutions of Painlevé differential equations. In particular, we omit results related to asymptotic representations, integrability, Riemann–Hilbert problems, isomonodromic deformations and inverse scattering theory. Moreover, we do not go in detail into Bäcklund transformations.

The Painlevé property proposes that solutions of the six Painlevé equations are close to being meromorphic functions in the complex plane. It is easy to find non-meromorphic solutions to the third (4.5.33), fifth (4.5.35) and sixth (4.5.36) Painlevé equations. However, for (4.5.33) and (4.5.35) the only possible non-polar singularities are at $z = 0$ and $z = \infty$, and hence the transformation $z = \exp t$ might result in an equation with meromorphic solutions only, while for (4.5.36) with three possible fixed singularities at $z = 0, 1, \infty$, such a transformation does not work. Indeed, applying for (4.5.33) a slightly different transformation, we obtain the modified forms of (4.5.33) and (4.5.35):

$$f'' = \frac{(f')^2}{f} + \alpha f^2 + \gamma f^3 + \beta e^z + \frac{\delta e^{2z}}{f} \tag{4.5.37}$$

and

$$\begin{aligned} f'' = {} & \left(\frac{1}{2f} + \frac{1}{f-1}\right)(f')^2 + (f-1)^2\left(\alpha f + \frac{\beta}{f}\right) + \gamma e^z f \\ & + \frac{\delta e^{2z} f(f+1)}{f-1}. \end{aligned} \tag{4.5.38}$$

Making use of these modified forms, we may express the meromorphic nature of the solutions of Painlevé equations as follows:

THEOREM 4.5.1. *All local solutions of first* (4.5.31), *second* (4.5.32), *modified third* (4.5.37), *fourth* (4.5.34) *and modified fifth* (4.5.38) *Painlevé equations can be analytically continued to single-valued meromorphic functions in the complex plane.*

THEOREM 4.5.2. *Given* $z_0 \in \mathbb{C} \setminus \{0, 1\}$ *and a local solution* f *of the sixth Painlevé equation* (4.5.36) *in a neighborhood of* z_0, f *can be analytically continued without restriction to a possible multi-valued meromorphic function in* $\mathbb{C} \setminus \{0, 1\}$.

The original idea of proving these two theorems goes back to Painlevé in the early years of 1900’s. The first paper containing a rigorous function theoretic proof for these two theorems, as far as we know, is due to Hukuhara [88]. Unfortunately, this paper was never published, and so it remained quite unknown. Recently, several papers appeared to offer such proofs, see [80–83,140,161] and [173].

After the meromorphic nature of the Painlevé functions has been settled, natural questions about their growth and value distribution appears. Concerning the growth, we have the following results:

THEOREM 4.5.3. *All solutions* f *of the first Painlevé equation* (4.5.31) *are of order* $\rho(f) = 5/2$.

The inequality $\rho(f) \geqslant 5/2$ goes back to Mues and Redheffer [130]. The inverse inequality is due to Steinmetz [174] and Shimomura [163], see also [162]. The proofs are quite different: Steinmetz makes use of a clever rescaling of the equation, while Shimomura carries through a careful analysis of solutions f in the neighborhoods of its poles, making use of the fact that f is the derivative of a meromorphic function.

THEOREM 4.5.4. *The solutions* f *of the second Painlevé equation* (4.5.32), *resp. the fourth Painlevé equation* (4.5.34), *are either rational functions, or transcendental meromorphic of order* $\rho(f)$ *satisfying* $3/2 \leqslant \rho(f) \leqslant 3$, *resp.* $2 \leqslant \rho(f) \leqslant 4$.

This theorem was proved almost simultaneously in three papers with different methods, see [84,162] and [175].

Observe that the rational solutions for Painlevé differential equations may be completely characterized and constructed. Indeed, (4.5.31) admits no rational solutions, while for (4.5.32), there exists exactly one rational solution if and only if the parameter $\alpha \in \mathbb{Z}$. This solution vanishes identically, when $\alpha = 0$, from which the rational may be computed, at least in principle, for $\alpha \neq 0$, by the standard Bäcklund transformations, see [51], p. 100. Concerning the fourth Painlevé equation (4.5.34), rational solutions exist if and only if the parameters α, β satisfy that $\alpha = n_1$ and either $\beta = -\frac{2}{9}(6n_2 - 3n_1 + \varepsilon)^2$, or $\beta = -2(1 + 2n_3 - n_1)^2$, where n_1, n_2, n_3 are integers, and $\varepsilon = \pm 1$, see [131] and [195]. As for solutions f of order $\rho(f) = 3/2$ for (4.5.32), a family of such solutions, called Airy solutions, are characterized by the fact that they simultaneously satisfy a first order algebraic differential equation, see [51], Theorem 21.1. Concerning solutions f of order $\rho(f) = 2$ for (4.5.34), several such families of solutions are known, see [51], Chapter 6. Concerning the rational solutions of (4.5.33), (4.5.35) and (4.5.36), the case of (4.5.35) has been solved by Kitaev, Law and McLeod, see [103], while for (4.5.33) and (4.5.36), we refer to [195] by Yuan and Li. Of course, rational solutions of (4.5.33) and (4.5.35) result in special solutions f of (4.5.37) and (4.5.38) of growth $T(r, f) = O(r^K)$ for some $K > 0$.

Proceeding to the growth of modified third and fifth Painlevé equations, Shimomura proved in [165] that these are of finite hyper-order:

THEOREM 4.5.5. *All solutions f of* (4.5.37), *resp. of* (4.5.38), *satisfy $T(r, f) = O(\exp(\Lambda r))$ for some strictly positive real number Λ independent of f.*

The best possible value of Λ remains unknown. However, a result in [19], Theorem 2, indicates that $\Lambda \leqslant 2$ might be true. Moreover, a similar bound might be true for solutions f of the sixth Painlevé equation as well in the sense that around an essential singularity of f, the local hyper-order of f might be $\leqslant 2$.

We now proceed to considering the value distribution of Painlevé functions. The key quantities to be considered in this connection are their defect and ramification indices, including constant and slowly moving targets. We remark that most results of this type presently known are, quite likely, not best possible. Therefore, we mention a number of open questions in this connection. Most of the results referred to below are due to Shimomura, [155,158,156,159].

THEOREM 4.5.6. *Let f be a solution of* (4.5.31), *and let α be a meromorphic function small with respect to f. Then the defect of f, resp. the ramification index of f, with respect to α satisfies the inequality $\delta(\alpha, f) \leqslant 1/2$, resp. $\vartheta(\alpha, f) \leqslant 5/12$. If $\alpha(z) \equiv a \in \mathbb{C}$, then $\vartheta(a, f) \leqslant 1/6$. Finally, $\delta(\infty, f) = 0$.*

REMARK. Theorem 4.5.6 is originally due to Schubart and Wittich [151]. In the proof due to Shimomura [158], the key ingredients are the classical Clunie and Mohon'ko lemmas, see Theorems 4.1.1 and 4.1.3. The proof in [158] immediately implies that $\delta(\alpha, f) = 0$, if $\alpha(z)$ does not satisfy the same Painlevé equation (4.5.31). As the existence of such a pair

$\alpha(z)$, $f(z)$ seems quite unlikely, we conjecture that actually $\delta(\alpha, f) = 0$. Also, it is not known whether the upper bound 1/6 for the ramification index is best possible.

THEOREM 4.5.7. *Let f be a solution of* (4.5.32), *and let α be a meromorphic function small with respect to f. Then the defect of f with respect to α satisfies the inequality $\delta(\alpha, f) \leqslant 1/2$.*

REMARK. As for the second Painlevé equation (4.5.32), results related to the ramification indices $\vartheta(\alpha, f)$ of their solutions seem to be missing in the literature.

Concerning the modified third Painlevé differential equation (4.5.37), we first restrict our considerations to its fast growing solutions f meaning that $S(r) \equiv r$ is small with respect to f, see [155]. Secondly, we assume that the parameters satisfy the conditions $(\alpha, \gamma) \neq (0, 0)$ and $(\beta, \delta) \neq (0, 0)$.

THEOREM 4.5.8. *Under the conditions just described, $\delta(a, f) = 0$ for each extended complex number a. If the parameter $\delta \neq 0$, then $\vartheta(a, f) \leqslant 1/4$ for each complex number $a \neq 0$, while if $\delta = 0$ and $\beta \neq 0$, then $\vartheta(a, f) \leqslant 1/6$.*

REMARK. If $(\alpha, \gamma) = (0, 0)$ or $(\beta, \delta) = (0, 0)$, then (4.5.37) reduces to a Riccati differential equation with constant coefficients as shown in [81], pp. 326–327. In this case, defects and ramification indices may easily be computed.

Defect indices for solutions of the fourth Painlevé equation (4.5.34) become more complicated. We first consider the case of constant targets:

THEOREM 4.5.9. (See [170].) *Let f be a transcendental solution of* (4.5.34), *and let $a \in \mathbb{C} \cup \{\infty\}$ be a constant target value. We then have*:

(1) *If $a = \infty$, then $\delta(a, f) = 0$.*
(2) *If $(\beta, a) \in \mathbb{C} \times \mathbb{C} \setminus (0, 0)$, then $\delta(a, f) = 0$.*
(3) *If $(\beta, a) = (0, 0)$, and f satisfies one of the two Riccati differential equations $f' = \mp(f^2 + 2zf)$, then $\delta(a, f) = 1$.*
(4) *If $(\beta, a) = (0, 0)$, and f does not satisfy the two Riccati differential equations $f' = \mp(f^2 + 2zf)$, then $\delta(a, f) \leqslant 1/2$.*

PROOF. The case (1) is an immediate consequence of the Clunie lemma. As for the case (2), we refer to [170]. For the remaining cases (3) and (4), one needs to solve the Riccati differential equations $f' = \mp(f^2 + 2zf)$. Since $f = 0$ is a solution trivially, and for each $c \in \mathbb{C}$,

$$f_c^{\pm}(z) = \exp(\mp z^2)\left(c \pm \int_0^z \exp(\mp t^2)\,\mathrm{d}t\right)^{-1}$$

also solves the corresponding Riccati equation, we conclude by [107], Proposition 9.1.1, that the complete solution family is given by $\{0\} \cup \{f_c^{\pm};\, c \in \mathbb{C}\}$. Since the functions $f_c^{\pm}$ are entire, the conclusion of (3) immediately follows. The remaining case (4) is again referred to [170]. □

Observe that the solutions $f_c^{\pm}$ above are of order $\rho(f_c^{\pm}) = 2$.

Proceeding to the case of slowly moving targets, first recall the expression

$$H_\nu(z) = (-1)^\nu \exp(z^2)\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^\nu \exp(-z^2)$$

for the Hermite polynomials with $\nu \in \mathbb{N} \cup \{0\}$. For these values of ν, the Hermite–Weber differential equation $u'' + (-\frac{t^2}{4} + \nu + \frac{1}{2})u = 0$ admits two linearly independent solutions $D_\nu(t)$, $D_{-\nu-1}(\mathrm{i}t)$, for which we have

$$D_\nu(\sqrt{2}\,z) = 2^{-\nu/2}\exp(-z^2/2)H_\nu(z).$$

Let us first consider the special parameter values $\alpha = -(n + \frac{1}{2} \pm \frac{1}{2})$, $\beta = -2(n + \frac{1}{2} \mp \frac{1}{2})^2$ in (4.5.34), where $n \in \mathbb{Z}$ can be given arbitrarily. Then equation (4.5.34) possesses a one-parameter family $f_{n,c}^{\mp}$, $c \in \mathbb{C} \cup \{\infty\}$, of solutions, expressible in the form

$$f_{n,c}^{\mp}(z) = -z \mp \frac{\eta_{n,c}'(z)}{\eta_{n,c}(z)},$$

where

$$\eta_{n,c}(z) = c_1 D_n(\sqrt{2}\,z) + c_2 D_{-n-1}(\sqrt{2}\,\mathrm{i}z), \quad c = c_1/c_2.$$

These solutions are of order $\rho(f_{n,c}^{\mp}) = 2$, except for $g_n^{\mp} := f_{n,\infty}^{\mp}$, provided $n \in \mathbb{N} \cup \{0\}$ and for $g_n^{\mp} := f_{n,0}^{\mp}$, provided $n \in -\mathbb{N}$. These exceptional cases are rational solutions. For more details, as well as for proofs of the next theorem, see [159]. Again, the proofs strongly rely on using the Clunie and Mohon'ko lemmas.

THEOREM 4.5.10. *Let f be an arbitrary transcendental solution of* (4.5.34), *and let g be a meromorphic function small with respect to f. Then we have*:

(1) *If $(f, g) = (f_{n,c}^{\mp}, g_n^{\mp})$ is one of the special pairs determined above, then $\delta(g, f) = 1$.*

(2) *Otherwise, $\delta(g, f) \leqslant 1/2$, if $\beta \neq 0$, and $\delta(g, f) \leqslant 3/4$, if $\beta = 0$.*

REMARK. It is not known, whether the inequalities in Theorem 4.5.10(2) are best possible. Moreover, possible results concerning the ramification indices remain open.

The value distribution of (4.5.38) has been treated by Shimomura in [156]. First observe that if the parameters γ, δ vanish, then (4.5.38) easily reduces to a Riccati differential equation, see [82], pp. 134–135. Similarly as in the case of (4.5.37), we have to restrict ourselves to fast growing solutions f of (4.5.38), again meaning that $S(r) \equiv r$ is small with respect to f. We first have to extract two special families of solutions: (1) Provided

$\alpha = 0$ and $-4\beta\delta + (\gamma \pm (-2\delta)^{1/2})^2 = 0$, then (4.5.38) possesses a one-parameter family of solutions

$$f^{\pm}_{\gamma,\delta,c}(z) := \exp\bigl(\kappa_{\pm} z \mp (-2\delta)^{1/2} \mathrm{e}^{z}\bigr)\left(c - \kappa_{\pm} \int_0^z \exp\bigl(-\kappa_{\pm} s \pm (-2\delta)^{1/2}\mathrm{e}^{s}\bigr)\,\mathrm{d}s\right),$$

where $\kappa_{\pm} = 1 \pm \gamma(-2\delta)^{-1/2}$ and $c \in \mathbb{C}$, denoted by $\mathcal{V}$ below. (2) Provided $\beta = 0$ and $-4\alpha\delta + (-\gamma \pm (-2\delta)^{1/2})^2 = 0$, then (4.5.38) possesses a one-parameter family of solutions

$$g_{\gamma,\delta,c}(z) := 1/f_{-\gamma,\delta,c}(z),$$

denoted by $\mathcal{W}$ below.

THEOREM 4.5.11. *Let f be a fast growing solution of* (4.5.38), *where its parameters satisfy* $(\gamma, \delta) \neq (0, 0)$, *and let a be an extended complex number. Then* $\delta(a, f) = 0$ *except of the following special cases*:

(1) *If $f \in \mathcal{V}$, then $\delta(\infty, f) = 1$ and if $f \in \mathcal{W}$, then $\delta(0, f) = 1$.*
(2) *If $\alpha = 0$ and $f \notin \mathcal{V}$, then $\delta(\infty, f) \leqslant 1/2$ and if $\beta = 0$ and $f \notin \mathcal{W}$, then $\delta(0, f) \leqslant 1/2$.*
(3) *If $\alpha + \beta = 0$, $\gamma = 0$ and $\delta \neq 0$, then $\delta(-1, f) \leqslant 1/2$.*

THEOREM 4.5.12. *Let f be a fast growing solution of* (4.5.38), *where its parameters satisfy* $(\gamma, \delta) \neq (0, 0)$, *and let $a \in \mathbb{C} \setminus \{0, 1\}$. Then we have the following inequalities*:

(1) *If $\delta \neq 0$, then $\vartheta(a, f) \leqslant 1/4$.*
(2) *If $\delta = 0$ and $\gamma \neq 0$, then $\vartheta(a, f) \leqslant 1/6$.*
(3) *If $\alpha + \beta = 0$, $\gamma = 0$ and $\delta \neq 0$, then $\vartheta(-1, f) = 0$.*

Moreover,

(4) *If $\alpha \neq 0$, resp. $\beta \neq 0$, resp. $\delta \neq 0$, then $\vartheta(\infty, f) = 0$, resp. $\vartheta(0, f) = 0$, $\vartheta(1, f) = 0$. Finally, if $\delta = 0$, then $\vartheta(1, f) = 1/2$.*

REMARK. As one may immediately see, ramification indices remain open in several special cases. It is also not known, whether the inequalities in the preceding two theorems are best possible.

Concerning finally (4.5.36), the most general of the Painlevé equations, it would be natural to study the defects and the ramification indices of its solutions locally in a neighborhood of their essential singularities. However, this has not been done, so far we know.

Finally, it is worth to be mentioned that for meromorphic Painlevé functions, the asymptotic inequality in the second main theorem of Nevanlinna theory becomes an asymptotic equality, see [51], Section 11.

We then proceed to considering, shortly, higher order Painlevé equations, i.e. certain hierarchies of equations, each starting from a given Painlevé equation. Equations in these hierarchies all possess the Painlevé property. However, it remains open to what extent the solutions of these differential equations indeed are meromorphic functions. We conjecture that the solutions to equations in the hierarchies built on the first, second and fourth

Painlevé equations are meromorphic, but there is no proof for this by now. Also, considering the growth and value distribution of meromorphic solutions of higher order Painlevé equations is still very much an open area of research. The first and the second Painlevé hierarchies only have been somewhat treated presently, see below.

To define the *first Painlevé hierarchy*, let $D = \mathrm{d}/\mathrm{d}z$ be the usual differential operator, and D^{-1} stand for its inverse. Defining $d^1(f) := -4f$, we inductively determine

$$d^{n+1}(f) := D^{-1}\big((D^3 - 8fD - 4f')d^n(f)\big).$$

Then we may consider the equations in the first Painlevé hierarchy defined as

$$d^{n+1}(f) + 4z = 0. \tag{4.5.39}$$

It is immediate to observe that we get the first Painlevé equation (4.5.31) from the case $n = 1$, while for $n = 2$ we obtain

$$f^{(4)} = 20ff'' + 10(f')^2 - 40f^3 + z,$$

and so on. Value distribution of meromorphic solutions of equation in the first Painlevé hierarchy (4.5.39) has been treated by Shimomura [164], He [61] and Li [118]. Some of the main results in these papers are as follows:

THEOREM 4.5.13. *Given $n \in \mathbb{N}$, let f be a meromorphic solution of* (4.5.39). *Then the following assertions follow*:

(1) *The solution f is transcendental meromorphic with double poles only and of order $\rho(f) \geqslant \frac{2n+3}{n+1}$.*
(2) *For each extended complex number a, $\delta(a, f) = 0$.*
(3) *The second main theorem for f takes the form of an asymptotic equality*

$$m(r, f) + \sum_{j=1}^{q} m\left(r, \frac{1}{f - a_j}\right) + N_1(r, f) + N\left(r, \frac{1}{f'}\right) = 2T(r, f) + S(r, f).$$

REMARK. It is a natural conjecture that $\rho(f) = \frac{2n+3}{n+1}$. Unfortunately, this remains open.

As for the *second Painlevé hierarchy*, we first define the operator

$$S[f] := 4f^2 + 4f'D^{-1}f - D^2$$

and then construct the algebraic differential equation

$$D^{-1}S^{m-1}[f]f' + zf + \alpha = 0, \tag{4.5.40}$$

where $\alpha \in \mathbb{C}$ is a parameter, for $m \in \mathbb{N} \setminus \{1\}$. This differential equation is of order $2m - 2$ and for $m = 2$ we obtain the second Painlevé equation (4.5.32), for $m = 3$ we get

$$f^{(4)} = 10f^2f'' + 10f(f')^2 - 6f^5 - zf - \alpha,$$

and so on. Now, the growth and value distribution of meromorphic solutions of equations in the second Painlevé hierarchy (4.5.40) has been investigated in two papers by Li and He [119,118]. Some of their results read as follows:

THEOREM 4.5.14. *Given* $m \in \mathbb{N} \setminus \{1\}$, *let* f *be a transcendental meromorphic solution of* (4.5.40). *Then the following assertions hold*:

(1) *All poles of* f *are simple and* $\delta(\infty, f) = 0$.
(2) *If* $\alpha \neq 0$, *then* $\delta(a, f) = 0$ *for each* $a \in \mathbb{C}$.
(3) *If* $\alpha = 0$, *then* $\delta(a, f) = 0$ *for each* $a \in \mathbb{C} \setminus \{0\}$ *and* $\delta(0, f) \leqslant 1/2$.

REMARKS.

(1) There exist no results concerning the growth of meromorphic solutions of (4.5.40), although it is natural again to conjecture that these functions are of finite order.
(2) It would be interesting to investigate the residues of the poles of f for the parameter values $\alpha = \varepsilon/2$, $\varepsilon^2 = 1$. This has been done in the case of $m = 3$, see [119], Theorem 2.4. Compare also with the corresponding results for (4.5.32), i.e. in the case of $m = 2$, see [51], Section 21.
(3) In the second Painlevé hierarchy as well, transcendental meromorphic solutions again have an asymptotic equality in the second main theorem in the form

$$m\left(r, \frac{1}{f}\right) + \sum_{j=1}^{q} m\left(r, \frac{1}{f - a_j}\right) + N_1(r, f) + N\left(r, \frac{1}{f'}\right) = 2T(r, f) + S(r, f),$$

where each $a_j \neq 0$, see [118], Theorem 3.1.

4.6. *Briot–Bouquet differential equations*

By Briot–Bouquet differential equations we understand autonomous differential equations in the complex plane of the form $P(f, f^{(k)}) = 0$, where P is a polynomial in both variables, having constant coefficients, of course. Investigations concerning Briot–Bouquet differential equations go back to a classical paper by Briot and Bouquet in 1856, where they claim that all meromorphic solutions of $P(f, f') = 0$ belong to one of the following classes of functions: (1) rational functions, (2) rational functions of e^{az} for some complex constant $a \neq 0$, (3) elliptic functions. Denoting this collection of meromorphic functions by $\mathfrak{W}$, one can show [79] that any meromorphic function $f \in \mathfrak{W}$ satisfies a Briot–Bouquet equation, see also [78]. In the reverse direction, this remains an open

CONJECTURE. *Any meromorphic solution in the complex plane of a Briot–Bouquet differential equation is in the collection* $\mathfrak{W}$ *of meromorphic functions.*

Concerning the existence of meromorphic functions for Briot–Bouquet differential equations, the following fact is classical: Denote by $S = S(P)$ the compact Riemann surface determined by the algebraic curve $P(u, v) = 0$. A necessary condition for the existence of a non-constant meromorphic solution f to $P(f, f^{(k)}) = 0$ is that the genus $g(S)$ of the Riemann surface S is either $g(S) = 0$ or $g(S) = 1$. The following partial result to the conjecture above are known:

THEOREM 4.6.1. *Let f be a meromorphic solution of a Briot–Bouquet differential equation $P(f, f^{(k)}) = 0$ of the first or second order, i.e. if $k = 1$ or $k = 2$. Then $f \in \mathfrak{W}$.*

REMARK. There are several presentations of this result in the classical literature of the 19th century. However, these proofs hardly can be accepted of being rigorous according to the present standards. For a modern proof, see [12], Theorems 1 and 5.

THEOREM 4.6.2. (See [37].) *Suppose f is a non-constant meromorphic solution of a Briot–Bouquet differential equation $P(f, f^{(k)}) = 0$, and let g be the genus of the Riemann surface determined by the algebraic curve $P(u, v) = 0$. Then the following assertions hold*:

(1) *If $g = 1$, then f is an elliptic function.*
(2) *If $g = 0$, k is an odd natural number and f is entire, then $f \in \mathfrak{W}$.*

REMARK. Observe that there is an extensive literature in the unit disc about the differential equation

$$f + \frac{zf'}{\beta f + \gamma} = h(z),$$

where β, γ are complex constants and $h(z)$ is an analytic function. This differential equation is said to be a Briot–Bouquet differential equation as well. We omit these investigations in this connection, as their main connection is to the subordination theory, and not so much to the theory of complex differential equations.

4.7. *Algebraic differential equations*

Algebraic differential equations

$$\Omega(z, f, f', \dots, f^{(n)}) = \sum_{\lambda \in I} a_\lambda(z) f^{\lambda_0} (f')^{\lambda_1} \cdots (f^{(n)})^{\lambda_n} = 0, \tag{4.7.41}$$

where Ω is a polynomial in the variables $f, f', \dots, f^{(n)}$ with meromorphic coefficients, and I is a finite index set, form an extensive field of research. However, open problems abound, and the overall impression is still somehow scattered. It seems to us that obtaining some coherence in the whole area needs to combine tools from various fields such as complex analysis and differential field theory. To get a quick overview of a number of open research problems in the field of algebraic differential equations, we refer to the problem collections [148–150] composed by Rubel.

While considering Eq. (4.7.41) below, we need some degree and weight notations as follows: The *degree* $|\lambda|$ of a single term in (4.7.41) will now be defined as

$$|\lambda| := \lambda_0 + \cdots + \lambda_n,$$

and its *weight* $\|\lambda\|$ as

$$\|\lambda\| := \lambda_0 + 2\lambda_1 + \cdots + (n+1)\lambda_n.$$

Of course, the degree and the weight of the whole Eq. (4.7.41) will then be defined as

$$|\Omega| := \max_{\lambda \in I} |\lambda|, \qquad \|\Omega\| := \max_{\lambda \in I} \|\lambda\|.$$

One of the key open problems in the field of algebraic differential equations is the following

CONJECTURE. *All meromorphic solutions of algebraic differential equations* (4.7.41) *with polynomial coefficients satisfy the growth inequality*

$$T(r, f) \leqslant a \exp_{n-1}(br^c),$$

where $\exp_1(z) \equiv e^z$, $\exp_{j+1}(z) := \exp(\exp_j(z))$, *and where* a, b, c *are positive real constants.*

This conjecture has been verified for the first order algebraic differential equations as shown by the classical result due to Gol'dberg [50]:

THEOREM 4.7.1. *All meromorphic solutions* f *of first order algebraic differential equations* $\Omega(z, f, f') = 0$ *with polynomial coefficients are of finite order of growth.*

REMARKS.

(1) As mentioned above, see the remark after Theorem 4.2.9, the order of transcendental meromorphic solutions is either zero, or a multiple of $1/2$ or of $1/3$, see [40].

(2) If f is an entire solution of $\Omega(z, f, f') = 0$ with polynomial coefficients, then $\rho(f) > 1/2$. This is an improvement, due to Strelitz [177], of a classical result by Pólya who proved that $\rho(f) > 0$. However, a meromorphic solution may well be of order zero, see an example in [107], pp. 234–236, due to Bank and Kaufman. We remark here that the claim $\rho(f) > 0$ remains true for entire solutions f of second order algebraic differential equations $\Omega(z, f, f', f'') = 0$ as shown by Zimogljad [197], while for algebraic differential equations of order $\geqslant 3$, examples of entire solutions of zero order have been constructed by Valiron.

Proceeding to first order algebraic differential equations with meromorphic coefficients, observe first that it is not possible to estimate the growth of meromorphic solutions in terms of the growth of the coefficients only, uniformly for all such equations, not even in the linear case. This follows of a somewhat surprising arbitrary growth theorem due to Bank in [2]:

THEOREM 4.7.2. *Given any strictly positive real function* Φ *on the positive real axis, there exists a meromorphic function* h *such that* h'/h *is of finite order and* $T(r_n, h) > \Phi(r_n)$ *for a sequence tending to* $+\infty$.

Some of the results describing quantities needed for such a uniform estimate are presented below. Such a uniform estimate in terms of the coefficients only may be found for entire solutions, see [107], Theorem 11.10:

THEOREM 4.7.3. *Let* $\Omega(z, f, f') = 0$ *be a first order algebraic differential equation with entire coefficients of finite order* $\leqslant \rho < \infty$. *Then all entire solutions* f *satisfy*

$$T(r, f) = \mathrm{O}(\exp r^{\rho+\varepsilon})$$

for any $\varepsilon > 0$. *Hence, the hyper-order of* f *satisfies* $\rho_2(f) \leqslant \rho$.

As for second order algebraic differential equations $\Omega(z, f, f', f'') = 0$, we first remark that the growth conjecture above has been verified for entire solutions in the homogeneous case by Steinmetz, see [176]:

THEOREM 4.7.4. *Suppose an algebraic differential equation*

$$\Omega(f, f', f'') = 0$$

is homogeneous in f, f', f'' *with polynomial coefficients. Then each meromorphic solution of such an equation takes the form* $f(z) = \frac{g_1(z)}{g_2(z)} \exp(g_3(z))$, *where* g_1, g_2, g_3 *are entire functions of finite order.*

However, the non-homogeneous case remains open in general, even for entire solutions.

Concerning algebraic differential equations of arbitrary order, there seems to be very few systematic investigations about this topic. As for differential equations of a suitable form, reasoning of Malmquist or Clunie type may sometimes help. Recalling Theorem 4.2.14, if we are considering algebraic differential equations of the form

$$\Omega(z, f, f', \ldots, f^{(n)}) = R(z, f) \tag{4.7.42}$$

with meromorphic coefficients, all of them small with respect to f, and if $\deg_f R(z, f) > \|\Omega\|$, then all possible meromorphic solutions have their growth comparable, in some sense, with the growth of the coefficients. Indeed, denoting by $\Phi(r)$ the maximum of $\log r$ and the maximum of the characteristic functions of the meromorphic coefficients of Eq. (4.7.42), then we have $T(r, f) = \mathrm{O}(\Phi(r))$, at least on a set of r-values of infinite linear measure. However, it is quite likely that such an estimate actually holds nearly everywhere. However, this has been proved in the polynomial case of $R(z, f)$ only, see [178], as far as we know.

Three natural questions concerning meromorphic solutions of algebraic differential equations (4.7.41) appear: (1) to obtain growth estimates for solutions, (2) to work out criteria under which meromorphic solutions are of finite order of growth and (3) to work criteria under which meromorphic solutions exist and if so, produce the number of meromorphic solutions, resp. the dimension of this family. The last question remains untouched here; all studies we know in this direction are restricted to first order algebraic differential equations of a special form such as $f' = P(z, f)$ with polynomial coefficients.

Concerning the growth estimates, Theorem 4.7.2 above implies that no uniform growth estimates exist in terms of the coefficients only. It appears that information about the poles and distinct zeros are needed to obtain a uniform estimate. To this end, we divide Eq. (4.7.41) in its homogeneous parts as

$$\Omega(z, f, f', \ldots, f^{(n)}) = \sum_{q=0}^{|\Omega|} \Omega_q(z, f, f', \ldots, f^{(n)}) = 0,$$

denoting by $A_q(z)$ the sum of all coefficients in the homogeneous part Ω_q. As before, we define

$$\Phi(r) := \max_{\lambda \in I}\bigl(\log r, T(r, a_\lambda)\bigr)$$

over all coefficients of P.

THEOREM 4.7.5. (See [13].) *Let f be a meromorphic solution of* (4.7.41). *If f does not satisfy some of the homogeneous equations $\Omega_q(z, f, f', \ldots, f^{(n)}) = 0$ of* (4.7.41), *then for any $\sigma > 1$, we have, for some $K > 0$,*

$$T(r, f) \leqslant K G(\sigma r)$$

for all r sufficiently large, where

$$G(r) = \overline{N}\left(r, \frac{1}{f}\right) + \overline{N}(r, f) + \Phi(r).$$

On the other hand, if f satisfies all of the homogeneous equations $\Omega_q = 0$ of (4.7.41), *while for some q, $0 \leqslant q \leqslant |\Omega|$, we have $A_q(z) \not\equiv 0$, then*

$$T(r, f) = \mathrm{O}\bigl(rN(\sigma r, f) + r^2 \exp\bigl(\mathrm{d}G(\sigma r)\log\bigl(rG(\sigma r)\bigr)\bigr)\bigr).$$

REMARKS.

(1) In principle, this enables us to determine all quantities needed to obtain a uniform growth estimate for meromorphic solutions. Indeed, if A_q vanishes in the last part of the preceding theorem, we may continue by looking at

$$\sum_{q=0}^{|\Omega|-1} \Omega_q(z, f, f', \ldots, f^{(n)}) = 0.$$

Inductively, if needed, we may proceed this way down to a first order algebraic differential equation, which always can be treated by Theorem 4.7.1.

(2) In the special case of linear differential equations, the original result has been substantially improved recently by Chiang and Hayman, see [32], p. 456. In particular, if the meromorphic solution f has a few poles only in the sense that $\delta(\infty, f) > 0$, then a uniform growth estimate follows in terms of the coefficients only.

Concerning criteria to deduce existence of finite order meromorphic solutions, only a few results are known for non-linear algebraic differential equations. We recall here the following interesting result due to Hayman, see [59]:

THEOREM 4.7.6. *Consider an algebraic differential equation* (4.7.41) *with polynomial coefficients, and denote by* Λ *the collection of terms that have maximal weight among all terms of maximal degree* $|\Omega|$. *Suppose that*

$$\sum_{\lambda\in\Lambda} a_\lambda(z) \not\equiv 0.$$

Then all entire solutions f *of* (4.7.41) *are of finite order*

$$\rho(f) \leqslant \max\bigl(2|\Omega|, 1 + |\Omega|\bigr).$$

Observe that the assertion fails for meromorphic solutions as shown by equation

$$f^2 f'' - 2f(f')^2 - f^2 f' + f'' - f' = 0,$$

where $\sum_{\lambda\in\Lambda} a_\lambda(z) = -1$, having a meromorphic solution $f(z) = \tan \mathrm{e}^z$ of infinite order.

A simple example not covered by Theorem 4.7.6 is equation

$$ff'' - (f')^2 = a_2(z)f'' + a_1(z)f' + a_0(z)f + b(z) \tag{4.7.43}$$

with rational coefficients. Hayman conjectured [59] that all entire solutions are of finite order. This has been proved in the case of constant coefficients by Chiang and Halburd [31], and considerations in [19] indicate that the conjecture might be true.

The growth of meromorphic solutions of algebraic differential equations in the unit disc still remains very much an open field of research, and some of the results obtained are too complicated to be repeated here, see e.g. [167]. Concerning first order algebraic differential equations, a unit disc counterpart to Theorem 4.7.1 has been proved by Bank in [3]:

THEOREM 4.7.7. *Consider a first order algebraic differential equation*

$$\Omega(z, f, f') = \sum_{\lambda\in I} a_\lambda(z) f^{\lambda_0} (f')^{\lambda_1} = 0$$

with analytic coefficients of finite order in the unit disc. Denote $d := \max_{\lambda\in I}\rho_M(a_\lambda)$. *Then all analytic solutions* f *of* $\Omega(z, f, f') = 0$ *in the unit disc are of finite order. More precisely,*

$$M(r, f) \leqslant \exp\bigl(\exp\bigl((1-r)^{-(d+4)}\bigr)\bigr)$$

for all $r < 1$ *close enough to* 1.

A natural problem would be to explore connections between the function spaces the solutions, resp. the coefficients of an algebraic differential equation belong to, in the same spirit as was presented for linear differential equations in Chapter 3. There are a couple of results only of this type. We quote here

THEOREM 4.7.8. *Given* $1 \leqslant \alpha < \infty$, *let* f *be a meromorphic solution of*

$$(f')^n = \sum_{j=1}^{m} P_j(z, f) D_j[f],$$

where

$$P_j(z, f) \sum_{k=0}^{m_j} a_{kj}(z) f^k, \qquad D_j[f] = (f')^{j_1} \cdots (f^{(s)})^{j_s}$$

with analytic coefficients a_{kj} *in the unit disc so that*

$$n > \nu := \max_{1\leqslant j\leqslant m} (j_1 + \cdots + s j_s).$$

If

$$\sup_{z\in\mathbb{D}}\bigl(1-|z|\bigr)^{\alpha(n-\nu)} \max_{1\leqslant j\leqslant m} \sum_{k=0}^{m_j}\bigl|a_{kj}(z)\bigr| < \infty,$$

then $f \in \mathcal{N}^\alpha$, *i.e.* f *is* α*-normal.*

Respectively, if

$$\lim_{|z|\to 1^-}\bigl(1-|z|\bigr)^{\alpha(n-\nu)} \max_{1\leqslant j\leqslant m} \sum_{k=0}^{m_j}\bigl|a_{kj}(z)\bigr| = 0,$$

then f *is strongly normal.*

4.8. *Algebraic differential equations and differential fields*

The theory of differential fields is an extensive topic which apparently has natural connections to investigations in the area of algebraic differential equations. However, these two areas have been developed rather separately of each other, and their intersection clearly is much more narrow than it could be.

To fix our notations, we restrict ourselves to considering differential fields $\mathcal{L}$ that are subfields of the field of meromorphic functions. Moreover, we steadily assume that $\mathbb{C} \subset \mathcal{L}$. Of course, results to be applied from the theory of differential fields primarily are those dealing with fields of characteristic zero. In what follows, we are considering meromorphic functions (and their subfields) in the complex plane, although some of the results apply to the unit disc case as well. Meromorphic functions that satisfy an algebraic differential equation with coefficients in $\mathcal{L}$ may be called *differentially algebraic over* $\mathcal{L}$, being denoted as $\mathcal{A}(\mathcal{L})$. It is easy to see that $\mathcal{A}(\mathcal{L})$ is a differential field. But $\mathcal{A}(\mathcal{L})$ is also closed in the sense that whenever $f \in \mathcal{A}(\mathcal{A}(\mathcal{L}))$, then $f \in \mathcal{A}(\mathcal{L})$. A meromorphic function $f \notin \mathcal{A}(\mathcal{L})$ is called *differentially transcendental over* $\mathcal{L}$. Being differential algebraic has the elementary nesting property: If $\mathcal{L} \subset \mathcal{L}_1 \subset \mathcal{L}_2$ are subfields of meromorphic functions, and if $\mathcal{L}_1 \subset \mathcal{A}(\mathcal{L})$ and $\mathcal{L}_2 \subset \mathcal{A}(\mathcal{L}_1)$, then $\mathcal{L}_2 \subset \mathcal{A}(\mathcal{L})$. Finally, we say that the differential field $\mathcal{A}(\mathbb{C})$ is the field of *differentially elementary* meromorphic functions.

Perhaps the most typical papers on the borderline of algebraic differential equations and differential fields are those proving a differentially transcendental nature of a given meromorphic function (or a certain family of meromorphic functions). This type of investigations go back to Hölder who proved (in 1887) that the gamma-function $\Gamma \notin \mathbb{C}(z)$, hence is not differentially elementary. The Hölder theorem has been improved by Bank and Kaufman in [10]. For the convenience, we denote by T_1 the translation $z \mapsto z+1$.

THEOREM 4.8.1. *Suppose* $\mathcal{L}$ *is a subfield of meromorphic functions, satisfying the following three properties*: (i) $\mathbb{C}(z) \subset \mathcal{L}$, (ii) *if* $f \in \mathcal{L}$, *then* $f \circ T_1 \in \mathcal{L}$, (iii) *if* $f \in \mathcal{L}$ *and* $f \circ T_1 - f \in \mathbb{C}(z)$, *then* $f \in \mathbb{C}(z)$. *Then the gamma-function* Γ *is differentially transcendental over* $\mathcal{L}$, *hence over* $\mathcal{A}(\mathcal{L})$ *as well. In particular,* Γ *is differentially transcendental over the field of meromorphic functions* f *that satisfy* $T(r, f) = o(r)$.

For a simple proof of the following result, see [127]:

THEOREM 4.8.2. *The Riemann zeta-function* ζ *is not differentially elementary.*

A natural extension of the notion of algebraic independence may be applied to define *differential independence* of meromorphic functions. In particular, if f, g are two meromorphic functions, they are said to differentially independent (over $\mathbb{C}$), if either f is differentially transcendental over $\mathbb{C}(g)$, or g is differentially transcendental over $\mathbb{C}(f)$.

THEOREM 4.8.3. (*See [127].*) *The following pairs of meromorphic functions are differentially independent over* $\mathbb{C}$:

(1) $\Gamma(\sin(2\pi z))$ *and* $\Gamma(z)$;

(2) $\zeta(\sin(2\pi z))$ *and* $\Gamma(z)$.

REMARK. A natural conjecture would be that $\zeta(z)$ and $\Gamma(z)$ are differentially independent over $\mathbb{C}$, not the least due to Theorem 4.8.3(2). However, this remains an open problem.

Another group of results in the borderline area of algebraic differential equations and differential fields are related to the following classical lemma, see [107], Lemma 14.3.1, originally due to Siegel. We remark that this lemma is part of the bridge between complex differential equations and transcendental number theory, again a borderline area, which apparently needs more exploration.

THEOREM 4.8.4. *Consider a second-order linear differential equation*

$$f'' + a(z)f' + b(z)f = 0 \tag{4.8.44}$$

over a subfield $\mathcal{L}$ of meromorphic functions. Suppose (4.8.44) *admits a solution f_0 that is* (*algebraically*) *transcendental over $\mathcal{L}$, but satisfies a first-order algebraic differential equation over $\mathcal{L}$. Then there exists another non-trivial solution f of* (4.8.44) *such that f'/f is algebraic over $\mathcal{L}$.*

Example of a consequence of the Siegel lemma is concerned with the basic linear differential equation $f'' + A(z)f = 0$, where $A(z)$ is a non-constant polynomial of degree n. Denoting by $\mathcal{L}_{(n+2)/2}$ the differential field of meromorphic functions of order $< (n+2)/2$, we get

THEOREM 4.8.5. *If n is odd, then no non-trivial solution of $f'' + A(z)f = 0$ with a polynomial $A(z)$ of degree n can satisfy a first-order algebraic differential equations over $\mathcal{L}_{(n+2)/2}$. On the other hand, if n is even, then at least one of any two linearly independent solutions of $f'' + A(z)f = 0$ cannot satisfy a first-order algebraic differential equation over $\mathcal{L}_{(n+2)/2}$.*

For further results of similar type, see [4,5].

5. Algebroid solutions of complex differential equations

5.1. *Introduction to algebroid functions*

Algebroid functions in the complex plane are formed by collections of analytic function elements that satisfy an irreducible algebraic equation

$$A_n(z)f^n + A_{n-1}(z)f^{n-1} + \cdots + A_1(z)f + A_0(z) = 0 \tag{5.1.1}$$

with meromorphic coefficients. This means that the collection of algebroid functions in the complex plane form an extension field of the field of meromorphic functions. Moreover, an algebroid function f as defined above is n-valued, hence f may be understood as a one-valued meromorphic function on an n-sheeted Riemann surface. Typical examples of elementary algebroid functions are nth roots of meromorphic functions. In many aspects, algebroid functions behave similarly as to meromorphic functions. However, there are fundamental differences such as, for example, the phenomenon that the derivative f' of an algebroid function may have a pole at a point where the function f itself takes a finite value. A simple example is $\sqrt{z}$ whose derivative $1/(2\sqrt{z})$ has a pole at the origin. Of course, the branch points of f also add an aspect which has no direct counterpart in the realm of meromorphic functions.

Nevanlinna theory of meromorphic functions has a natural extension to algebroid functions, soon developed in early 1930's, most notably by Valiron [184], Selberg [153] and Ullrich [183]. Since Nevanlinna theory is a powerful tool to investigating algebroid solutions of complex differential equations, we shortly describe some of its similarities and deviations as compared to the original meromorphic version.

Let f be a non-constant ν-valued algebroid function, that has at z_0 an a-point, resp. a pole, with local expansions about z_0 of the form

$$f(z) = a + b_\tau (z - z_0)^{\tau/\lambda} + \cdots, \quad \tau \geqslant 1, \tag{5.1.2}$$

resp.

$$f(z) = b_{-\tau} (z - z_0)^{-\tau/\lambda} + \cdots, \quad \tau \geqslant 1, \tag{5.1.3}$$

where $\lambda \leqslant \nu$. The *non-integrated counting function* $n(r, 1/(f-a))$ for the a-points, resp. $n(r, f)$ for the poles in the disc $|z| \leqslant r$, is the sum of the τ's in the preceding expansions, over all a-points, resp. over all poles, in $|z| \leqslant r$. The *integrated counting function* is now defined as in the usual meromorphic theory, taking however into account the fact that we are now operating on an ν-sheeted surface S_f over the disc $|z| \leqslant r$, i.e.,

$$N\left(r, \frac{1}{f-a}\right) := \frac{1}{\nu} \int_0^r \frac{n(t,a) - n(0,a)}{t}\, \mathrm{d}t + \frac{1}{\nu} n(0,a) \log r,$$

resp.

$$N(r, f) := \frac{1}{\nu} \int_0^r \frac{n(t,f) - n(0,f)}{t}\, \mathrm{d}t + \frac{1}{\nu} n(0,f) \log r.$$

Similarly, we define the *proximity function* as

$$m\left(r, \frac{1}{f-a}\right) := \frac{1}{\nu} \sum_{j=1}^{\nu} \frac{1}{2\pi} \int_0^{2\pi} \log^+ \frac{1}{|f_j(r\mathrm{e}^{\mathrm{i}\theta}) - a|}\, \mathrm{d}\theta,$$

resp.

$$m(r, f) := \frac{1}{\nu} \sum_{j=1}^{\nu} \frac{1}{2\pi} \int_0^{2\pi} \log^+ \left| f_j(r\mathrm{e}^{\mathrm{i}\theta}) \right| \mathrm{d}\theta,$$

where f_j stands for the values of f on the jth sheet of the surface S_f. The *characteristic function* $T(r, f) := m(r, f) + N(r, f)$ is now defined as usual, and the first main theorem takes its usual form $T(r, 1/(f-a)) = T(r, f) + \mathrm{O}(1)$. Moreover, an algebroid function f is algebraic if and only if $T(r, f) = \mathrm{O}(\log r)$, and the lemma of the logarithmic derivative takes its usual form $m(r, f^{(k)}/f) = S(r, f)$, $k \in \mathbb{N}$, in general and $m(r, f^{(k)}/f) = \mathrm{O}(\log r)$ for an algebroid function of finite order.

The key deviation from the usual meromorphic theory comes out of the *counting function of branch points* of f, which may be determined from the expansions (5.1.2), (5.1.3) as

$$n_3(r, f) := \sum_{|z| \leqslant r,\, a \in \widehat{\mathbb{C}}} (\lambda - 1)$$

first in the non-integrated form, and then the integrated counting function of branch points $N_3(r, f)$ follows exactly as above for $N(r, f)$, say. We also make use of the corresponding function $N_{3,\infty}(r, f)$ for branched poles above the disc $|z| \leqslant r$. The essential consequences may now be expressed as

PROPOSITION 5.1.1. *Let f be a ν-valued algebroid function. Then*
(1) $N_3(r, f) \leqslant 2(\nu - 1)T(r, f) + \mathrm{O}(1)$
and
(2) $T(r, f') \leqslant 2T(r, f) + N_3(r, f) + S(r, f) \leqslant 2\nu T(r, f) + S(r, f)$.

For more details and proofs concerning the basic Nevanlinna theory of algebroid functions, see [184,153,183] and [100].

We next give the algebroid versions of the Clunie, Mohon'ko and Valiron–Mohon'ko lemmas, see Section 4.1 for the corresponding meromorphic versions:

THEOREM 5.1.2. *Let f be an algebroid solution of*

$$f^n P(z, f) = Q(z, f),$$

where $P(z, f)$, $Q(z, f)$ are polynomials in f and its derivatives with algebroid coefficients, say a_λ, $\lambda \in I$, such that $m(r, a_\lambda) = S(r, f)$ for all $\lambda \in I$. If the total degree of $Q(z, f)$ in f and its derivatives is $\leqslant n$, then

$$m\bigl(r, P(z, f)\bigr) = S(r, f).$$

REMARK. The standard proof of the meromorphic version applies here. One just needs to replace the notions from meromorphic value distribution theory with the corresponding algebroid ones. The key point is that the logarithmic derivative lemma applies in the same form as in the meromorphic case.

THEOREM 5.1.3. (See [128].) *Let f be a meromorphic solution of an algebraic differential equation*

$$P(z, f, f', \ldots, f^{(n)}) = 0, \tag{5.1.4}$$

where P is a polynomial in $f, f', \dots, f^{(n)}$ with meromorphic coefficients small with respect to f. If a complex constant c does not satisfy Eq. (5.1.4), *then*

$$m\left(r, \frac{1}{f-c}\right) = S(r, f).$$

THEOREM 5.1.4. (See [128,129].) *Let f be an algebroid function, and let $R(z, f) = P(z, f)/Q(z, f)$ be an irreducible rational function in f with algebroid coefficients small with respect to f. Then the characteristic function of $R(z, f(z))$ satisfies*

$$T\big(r, R(z, f)\big) = dT(r, f) + S(r, f),$$

where $d := \max(\deg_f P(z, f), \deg_f Q(z, f))$.

5.2. *Malmquist type theorems for algebroid functions*

We start by giving the algebroid counterpart of Theorem 4.2.2:

THEOREM 5.2.1. *If a differential equation of the form*

$$(f')^n = R(z, f), \tag{5.2.5}$$

where $R(z, f) = P(z, f)/Q(z, f)$ is rational and irreducible in f with meromorphic coefficients, admits a ν-valued algebroid solution f such that all coefficients of $R(z, f)$ are small with respect to f, then

$$q := \deg Q \leqslant 2n(\nu - 1), \qquad p := \deg P \leqslant q + 2n \leqslant 2n\nu.$$

REMARKS.

(1) This result was originally proved by Yosida [194]. He also proved that all algebroid solutions of (5.2.5) with small meromorphic coefficients are of finite order of growth, see [194], Theorem 7.

(2) As a deviation from the meromorphic case, observe that $R(z, f)$ does not necessarily reduce to a polynomial in f as shown by an example in [45]: The equation $f^2 + f \tan z + z = 0$ defines a two-valued algebroid function that satisfies

$$f' = \frac{z^2 + f + (1 + 2z)f^2 + f^4}{z - f^2}.$$

However, if the solution has not too much of the branch points, then $R(z, f)$ reduces to polynomial:

THEOREM 5.2.2. (See [100].) *Under the same assumptions as in Theorem* 5.2.1, *if*

$$\xi(f) := \limsup_{r\to\infty} \frac{N_3(r, f)}{T(r, f)} < \frac{1}{n},$$

then $R(z, f)$ reduces to a polynomial in f of degree $\leqslant 2n$.

REMARKS.

(1) The assumption that $\xi(f) < 1/n$ cannot be improved, as shown in [145], Beispiel 7.1: Consider $f(z) := (e^z - 1)^{-1}$, and define algebroid functions w, v by the algebraic equation $w^2 - f(z) = 0$ and $v := (w-1)^{-1}$. Then v satisfies

$$v' = -\frac{(v+1)(2v^2+2v+1)}{2v},$$

while $\xi(v) = 1$.

(2) In the case of Theorem 5.2.2, the reduction may still be continued further in the algebroid case similarly as in the meromorphic case, see [100], Theorem 3.10 and [145], Satz 7.3.

(3) Concerning some introductory remarks of the value distribution of algebroid solutions of (5.2.5) with small meromorphic coefficients, see [145], Section 8.

(4) The Malmquist type theorems above for algebroid solutions correspond to Theorem 4.2.2 for meromorphic solutions. Concerning a counterpart of Theorem 4.2.7 for the algebroid case, see [39], [86], Satz 4.5 and [128], Theorem 5.

5.3. *Algebroid solutions of algebraic differential equations*

Investigations concerning algebroid solutions of algebraic differential equations

$$\Omega(z, f, f', \ldots, f^{(n)}) = 0, \tag{5.3.6}$$

where Ω is a polynomial in the variables $f, f', \ldots, f^{(n)}$ with meromorphic coefficients, are quite few, although this could also be an extensive field of research.

In considering an expression such as

$$\Omega(z, f, f', \ldots, f^{(n)}) = \sum_{\lambda \in I} a_\lambda(z) f^{\lambda_0} \cdots (f^{(n)})^{\lambda_n},$$

we apply the same notations concerning degree and weight as in Section 4.7, i.e.

$$|\Omega| := \max_{\lambda \in I} |\lambda|, \qquad \|\Omega\| := \max_{\lambda \in I} \|\lambda\|.$$

In addition, we need a couple of notations for certain modified versions of the weight:

$$|\Omega|_0 := \max_{\lambda \in I} \sum_{j=1}^{n} j\lambda_j, \qquad |\Omega|_1 := \max_{\lambda \in I} \sum_{j=1}^{n} (2j-1)\lambda_j.$$

Similarly as in the meromorphic case, see Theorem 4.2.14, Theorem 5.2.1 may be extended to algebraic differential equations of the form

$$\Omega(z, f, f', \dots, f^{(n)}) = R(z, f) = \frac{P(z, f)}{Q(z, f)}, \tag{5.3.7}$$

where $P(z, f)$, $Q(z, f)$ are polynomials in f with meromorphic coefficients, and $\Omega(z, f, f', \dots, f^{(n)})$ is a differential polynomial in f with meromorphic coefficients. Making use of the notations above for degree, weight etc. of Ω, and denoting $p := \deg P$, $q := \deg Q$, we get the following

THEOREM 5.3.1. (See [181].) *Suppose f is a ν-valued algebroid solution of Eq.* (5.3.7) *with small meromorphic coefficients. Then we have*

$$\max\bigl(p, q + \|\Omega\|\bigr) \leqslant \|\Omega\| + |\Omega|_1 \xi(f),$$
$$p \leqslant \min\bigl(q + |\Omega| + |\Omega|_0\bigl(1 - \Theta(\infty, f) + \xi_\infty(f)\bigr), \|\Omega\| + |\Omega|_1 \xi(f)\bigr),$$

where $\xi_\infty(f) := \limsup_{r\to\infty} \frac{N_{3,\infty}(r, f)}{T(r, f)}$.

REMARK.

(1) For several closely related variants of Theorem 5.3.1, see [48] and [47].

(2) As in the model case of Theorem 5.2.1, the reduced form may have a non-trivial denominator, contrary to the meromorphic case, see Theorem 4.2.14. However, if $\xi(f) = 0$, then the denominator becomes trivial, and we have

COROLLARY 5.3.2. *If* $\xi(f) = 0$, *then*

$$q = 0, \qquad p \leqslant \min\bigl(\|\Omega\|, |\Omega| + |\Omega|_0\bigl(1 - \Theta(\infty, f)\bigr)\bigr).$$

REMARK. Several of the items in the list of references below contain some additional results concerning algebroid solutions of algebraic differential equations. In particular, we refer to [39] and [86], Section 4. See also [179,180].

5.4. *Algebroid solutions of linear differential equations*

The only paper we know investigating algebroid solutions of linear differential equations of type

$$f'' + A(z) f = 0,$$

where $A(z)$ is meromorphic, is due to Katajamäki, see [100], Section 5. We quote here a couple of his results, THEOREM 5.4.1. *If the quotient of two local solutions of the equation* $f'' + A(z)f = 0$ *admits in a disc, where* $A(z)$ *is holomorphic, a* ν*-valued,* $\nu \geqslant 2$*, algebroid continuation into the complex plane, then at each pole* z_0 *of* $A(z)$ *either* (1) *the Laurent expansion of* $A(z)$ *about* z_0 *has the form*

$$A(z) = \frac{1 - m^2}{4(z - z_0)^2} + \cdots ,$$

where $m = p/q$ *is rational but not an integer,* p, q *have no non-trivial common factors, and* $q|\nu$*, or* (2) m *is an integer* $\geqslant 2$ *and the determinant condition* $D(z_0) = 0$ *in* (2.4.25) *holds.*

This theorem now has the following consequence:

THEOREM 5.4.2. *If equation* $f'' + A(z)f = 0$ *admits a* ν*-valued,* $\nu \geqslant 2$*, algebroid solution in the complex plane, then the equation has no non-trivial meromorphic solutions.*

THEOREM 5.4.3. *If equation* $f'' + A(z)f = 0$ *admits a* 2*-valued algebroid solution* f *in the complex plane, then* f *must be a square-root of a meromorphic function, i.e.* f^2 *is meromorphic.*

As for algebroid solutions of linear differential equations of higher order, see Tohge [182].

Acknowledgement

My best gratitude is due to all of my friends and colleagues who have had a role in the preparation of this survey, either directly or indirectly. Special thanks are due to my junior colleagues, Janne Heittokangas and Risto Korhonen. Of course, the final survey reflects my personal opinion of the field, and all possible defects, mistakes and misunderstandings are on my responsibility. In any case, I hope that this survey might serve as a completion to previous key references in the field, assisting beginners to penetrate in the wealth of material available.

References

[1] R. Aulaskari and P. Lappan, *Criteria for an analytic function to be Bloch and a harmonic or meromorphic function to be normal*, Complex Analysis and its Applications, Longman Sci. Tech., Harlow (1994), 136–146.

[2] S. Bank, *A note on algebraic differential equations whose coefficients are entire functions of finite order*, Ann. Scuola Norm. Sup. Pisa **26** (1972), 291–297.

[3] S. Bank, *On solutions of algebraic differential equations whose coefficients are analytic functions in the unit disk*, Ann. Mat. Pura Appl. **92** (1972), 323–335.

[4] S. Bank, *On certain properties of solutions of second-order linear differential equations*, Monatsh. Math. **94** (1982), 179–200.

[5] S. Bank, *On the value distribution theory for entire solutions of second-order linear differential equations*, Proc. London Math. Soc. **50** (1985), 505–534.

[6] S. Bank, *A note on the location of complex zeros of solutions of linear differential equations*, Bull. Amer. Math. Soc. **18** (1988), 35–38.

[7] S. Bank, *On the oscillation of solutions of non-homogeneous linear differential equations*, Analysis **10** (1990), 265–293.

[8] S. Bank, *On the complex zeros of solutions of linear differential equations*, Ann. Mat. Pura Appl. **161** (1992), 83–112.

[9] S. Bank, G. Gundersen and I. Laine, *Meromorphic solutions of the Riccati differential equation*, Ann. Acad. Sci. Fenn. A I Math. **6** (1981), 369–398.

[10] S. Bank and R. Kaufman, *An extension of Hölder's theorem concerning the Gamma function*, Funkcial. Ekvac. **19** (1976), 53–63.

[11] S. Bank and R. Kaufman, *On the growth of meromorphic solutions of the differential equation* $(y')^m = R(z, y)$, Acta Math. **144** (1980), 223–248.

[12] S. Bank and R. Kaufman, *On Briot–Bouquet differential equations and a question of Einar Hille*, Math. Z. **177** (1981), 549–559.

[13] S. Bank and I. Laine, *On the growth of meromorphic solutions of linear and algebraic differential equations*, Math. Scand. **40** (1977), 119–126.

[14] S. Bank and I. Laine, *On the oscillation theory of* $f'' + Af = 0$ *where* A *is entire*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), 95–98.

[15] S. Bank and I. Laine, *On the oscillation theory of* $f'' + Af = 0$ *where* A *is entire*, Trans. Amer. Math. Soc. **273** (1982), 351–363.

[16] S. Bank and I. Laine, *Representations of solutions of periodic second order linear differential equations*, J. Reine Angew. Math. **344** (1983), 1–21.

[17] S. Bank and I. Laine, *On the zeros of meromorphic solutions of second order liner differential equations*, Comment. Math. Helv. **58** (1983), 656–677.

[18] S. Bank and J. Langley, *Oscillation theory for higher order linear differential equations with entire coefficients*, Complex Variables Theory Appl. **16** (1991), 163–175.

[19] G. Barsegian, I. Laine and D. Lê, *On topological behavior of solutions of some algebraic differential equations*, Complex Var. Elliptic Equ., in press.

[20] P. Beesack, *Nonoscillation and disconjugacy in the complex plane*, Trans. Amer. Math. Soc. **81** (1956), 212–248.

[21] P. Beesack and B. Schwarz, *On the zeros of solutions of second-order linear differential equations*, Canad. J. Math. **8** (1956), 504–515.

[22] L. Bieberbach, *Theorie der gewöhnlichen Differentialgleichungen auf funktionentheoretischer Grundlage dargestellt*, Springer-Verlag, Berlin–New York (1965).

[23] L. Bieberbach, *Theorie der gewöhnlichen Differentialgleichungen auf funktionentheoretischer Grundlage dargestellt*, Springer-Verlag, Berlin–New York (1965).

[24] T. Cao, *Complex oscillation of entire solutions of higher order linear differential equations*, Electron. J. Differential Equations **81** (2006), 1–8.

[25] T. Cao and H.-X. Yi, *The growth of solutions of differential equations with coefficients of iterated order in the unit disc*, J. Math. Anal. Appl. **319** (2006), 278–294.

[26] T. Cao and H.-X. Yi, *On the complex oscillation theory of* $f'' + A(z)f = 0$, *where* $A(z)$ *is analytic in the unit disc*, in press.

[27] Z. Chen and S. Gao, *The complex oscillation theory of certain nonhomogeneous linear differential equations with transcendental entire coefficients*, J. Math. Anal. Appl. **179** (1993), 403–416.

[28] Z. Chen and K. Shon, *The growth of solutions of differential equations with coefficients of small growth in the disc*, J. Math. Anal. Appl. **297** (2004), 285–304.

[29] Y.-M. Chiang, *On the zero-free solutions of linear periodic differential equations in the complex plane*, Results Math. **38** (2000), 213–225.

[30] Y.-M. Chiang and S.-A. Gao, *On a problem in complex oscillation theory of periodic second order linear differential equations and some related perturbation results*, Ann. Acad. Sci. Fenn. Math. **27** (2002), 273–290.

[31] Y.-M. Chiang and R. Halburd, *On the meromorphic solutions of an equation of Hayman*, J. Math. Anal. Appl. **281** (2003), 663–677.

[32] Y.-M. Chiang and W. Hayman, *Estimates on the growth of meromorphic solutions of linear differential equations*, Comment. Math. Helv. **79** (2004), 451–470.

[33] I. Chyzhykov, G. Gundersen and J. Heittokangas, *Linear differential equations and logarithmic derivative estimates*, Proc. London Math. Soc. **86** (2003), 735–754.

[34] J. Cima and J. Pfaltzgraff, *Oscillatory behavior of* $u''(z) + h(z)u(z) = 0$, J. Analyse Math. **25** (1972), 311–322.

[35] D. Drasin and J. Langley, *Bank–Laine functions via quasi-conformal surgery*, Transcendental Dynamics and Complex Analysis, Cambridge Univ. Press, in press.

[36] P. Duren, *Theory of* H^p *Spaces*, Academic Press, New York–London (1970).

[37] A. Eremenko, *Meromorphic solutions of equations of Briot–Bouquet type*, Teor. Funktsiĭ Funktsional. Anal. i Prilozhen. **38** (1982), 48–56.

[38] A. Eremenko, *Meromorphic solutions of algebraic differential equations*, Uspekhi Mat. Nauk **37:4** (1982), 53–82.

[39] A. Eremenko, *Meromorphic solutions of algebraic differential equations*, Russian Math. Surveys **37:4** (1982), 61–95.

[40] A. Eremenko, *Meromorphic solutions of first-order algebraic differential equations*, Funktsional. Anal. i Prilozhen. **18** (1984), 78–79.

[41] A. Eremenko and S. Merenkov, *Nevanlinna functions with real zeros*, Illinois J. Math. **49** (2005), 1093–1110.

[42] A. Fokas, A. Its, A. Kapaev and V. Novokshenov, *Painlevé Transcendents. The Riemann–Hilbert Approach*, Amer. Math. Soc., Providence, RI (2006).

[43] K. Fowler and L. Sons, *Interaction between coefficient conditions and solution conditions of differential equations in the unit disk*, Int. J. Math. Math. Sci.

[44] M. Frei, *Über die Lösungen linearer Differentialgleichungen mit ganzen Funktionen als Koeffizienten*, Comment. Math. Helv. **35** (1961), 201–222.

[45] F. Gackstatter and I. Laine, *Zur Theorie der gewöhnlichen Differentialgleichungen im Komplexen*, Ann. Polon. Math. **38** (1980), 259–287.

[46] F. Gackstatter and I. Laine, *Zur Theorie der gewöhnlichen Differentialgleichungen im Komplexen*, Ann. Polon. Math. **58** (1980), 259–287.

[47] L. Gao, *On some generalized higher-order algebraic differential equations with admissible algebroid solutions*, Indian J. Math. **43** (2001), 163–175.

[48] L. Gao, *Some results on admissible algebroid solutions of complex differential equations*, Indian J. Pure Appl. Math. **32** (2001), 1041–1050.

[49] S. Gao, S. Wang and Z. Chen, *A note on the complex oscillation of nonhomogeneous linear differential equations with transcendental entire coefficients*, Ann. Differential Equations **11** (1995), 176–182.

[50] Gol'dberg, A., *On single-valued solutions of first-order differential equations*, Ukrain. Mat. Zh. **8** (1956), 254–261 (in Russian)

[51] V. Gromak, I. Laine and S. Shimomura, *Painlevé Differential Equations in the Complex Plane*, W. de Gruyter, Berlin–New York (2002).

[52] G. Gundersen, *On the real zeros of solutions of* $f'' + A(z)f = 0$ *where* $A(z)$ *is entire*, Ann. Acad. Sci. Fenn A I Math. **11** (1986), 275–294.

[53] G. Gundersen, *Finite order solutions of second order linear differential equations*, Trans. Amer. Math. Soc. **305** (1988), 415–429.

[54] G. Gundersen and E. Steinbart, *A generalization of the Airy integral for* $f'' - z^n f = 0$, Trans. Amer. Math. Soc. **337** (1993), 737–755.

[55] G. Gundersen, E. Steinbart and S. Wang, *Solutions of nonhomogeneous linear differential equations with exceptionally few zeros*, Ann. Acad. Sci. Fenn. Math. **23** (1998), 429–452.

[56] G. Gundersen, E. Steinbart and S. Wang, *The possible orders of solutions of linear differential equations with polynomial coefficients*, Trans. Amer. Math. Soc. **350** (1998), 1225–1247.

[57] G. Gundersen, E. Steinbart and S. Wang, *Growth and oscillation theory of non-homogeneous linear differential equations*, Proc. Edinburgh Math. Soc. **43** (2000), 343–359.

[58] R. Hadass, *On the zeros of the solutions of the differential equation* $y^{(n)}(z) + p(z)y(z) = 0$, Pacific J. Math. **31** (1969), 33–46.

[59] W. Hayman, *The growth of solutions of algebraic differential equations*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **7** (1996), 67–73.

[60] W. Hayman and J. Miles, *On the growth of a meromorphic function and its derivatives*, Complex Variables Theory Appl. **12** (1989), 245–260.

[61] Y. He, *Value distribution of the higher order analogues of the first Painlevé equation*, Value Distribution Theory and Related Topics, Kluwer Acad. Publ., Boston (2004), 209–217.

[62] Y. He and I. Laine, *The Hayman–Miles theorem and the differential equation* $(y')^n = R(z, y)$, Analysis **10** (1990), 387–396.

[63] Y. He and X. Xiao, *Algebroid Functions and Ordinary Differential Equations*, Beijing (1988) (in Chinese).

[64] J. Heittokangas, *On complex differential equations in the unit disc*, Ann. Acad. Sci. Fenn. Math. Diss. **122** (2000), 1–54.

[65] J. Heittokangas, *Solutions of* $f'' + A(z)f = 0$ *in the unit disc having Blaschke sequences as the zeros*, Comput. Methods Funct. Theory **5** (2005), 49–63.

[66] J. Heittokangas, *Blaschke-oscillatory equations of the form* $f'' + A(z)f = 0$, J. Math. Anal. Appl. **318** (2006), 120–133.

[67] J. Heittokangas, private communication.

[68] J. Heittokangas, R. Korhonen and J. Rättyä, *Growth estimates for analytic solutions of complex linear differential equations*, Ann. Acad. Sci. Fenn. Math. **29** (2004), 233–246.

[69] J. Heittokangas, R. Korhonen and J. Rättyä, *Fast growing solutions of linear differential equations in the unit disc*, Results Math. **49** (2006), 265–278.

[70] J. Heittokangas, R. Korhonen and J. Rättyä, *Linear differential equations with solutions in weighted Bergman and Hardy spaces*, Trans. Amer. Math. Soc. **360** (2008), 1005–1034.

[71] J. Heittokangas, R. Korhonen and J. Rättyä, *Linear differential equations with solutions in Dirichlet type subspace of the Hardy space*, Nagoya Math. J. **187** (2007), 91–113.

[72] S. Hellerstein, J. Miles and J. Rossi, *On the growth of solutions of certain linear differential equations*, Ann. Acad. Sci. Fenn. A I Math. **17** (1992), 343–365.

[73] H. Herold, *Ein Vergleichssatz für komplexe lineare Differentialgleichungen*, Math. Z. **126** (1972), 91–94.

[74] H. Herold, *Differentialgleichungen im Komplexen*, Vandenhoeck et Ruprecht, Göttingen (1975).

[75] H. Herold, *Nichteuklidischer Nullstellenabstand der Lösungen von* $w'' + p(z)w = 0$, Math. Ann. **287** (1990), 637–642.

[76] E. Hille, *Lectures on Ordinary Differential Equations*, Addison–Wesley, Reading, MA (1969).

[77] E. Hille, *Ordinary Differential Equations in the Complex Domain*, Wiley, New York–London–Sydney (1976).

[78] E. Hille, *Higher order Briot–Bouquet differential equations*, Ark. Math. **16** (1978), 271–286.

[79] E. Hille, *Remarks on Briot–Bouquet differential equations I*, Comment. Math. Special Issue **1** (1978), 119–132.

[80] A. Hinkkanen and I. Laine, *Solutions of the first and second Painlevé equations are meromorphic*, J. Analyse Math. **79** (1999), 345–377.

[81] A. Hinkkanen and I. Laine, *Solutions of a modified third Painlevé equation are meromorphic*, J. Analyse Math. **85** (2001), 323–337.

[82] A. Hinkkanen and I. Laine, *Solutions of a modified fifth Painlevé equation are meromorphic*, Report Univ. Jyväskylä **83** (2001), 133–146.

[83] A. Hinkkanen and I. Laine, *The meromorphic nature of the sixth Painlevé transcendents*, J. Analyse Math. **94** (2004), 319–342.

[84] A. Hinkkanen and I. Laine, *Growth results of Painlevé transcendents*, Math. Proc. Camb. Phil. Soc. **137** (2004), 645–655.

[85] L. Hörmander, *An Introduction to Complex Analysis in Several Variables*, D. Van Nostrand Co., Princeton–Toronto–London (1966).

[86] R. Hotzel, *Algebraische Differentialgleichungen mit zulässigen Lösungen: ein algebraischer Ansatz*, Diplomarbeit, Technische Hochscule Aachen (1994).

[87] R. Hotzel and G. Jank, *Algebraic Schwarzian differential equations*, Ann. Acad. Sci. Fenn. Math. **21** (1996), 353–366.

[88] M. Hukuhara, *Lecture notes on differential equations*, unpublished manuscript (in Esperanto).

[89] Y. Il'yashenko and S. Yakovenko, *Counting real zeros of analytic functions satisfying linear ordinary differential equations*, J. Differential Equations **126** (1996), 87–105.

[90] K. Ishizaki, *Admissible solutions of the Schwarzian differential equations*, J. Austral. Math. Soc. A **50**, (1991), 258–278.

[91] K. Ishizaki, *Admissible solutions of second order differential equations*, Tôhoku Math. J. **44** (1992), 305–325.

[92] K. Ishizaki, *Meromorphic solutions of complex differential equations*, Ph.D. thesis, Chiba University (1993).

[93] K. Ishizaki, *On the Schwarzian differential equation* $\{w, z\} = R(z, w)$, Kodai Math. J. **20** (1997), 67–78.

[94] K. Ishizaki and Y. Wang, *Non-linear differential equations with transcendental meromorphic solutions*, J. Aust. Math. Soc. **70** (2001), 88–118.

[95] K. Ishizaki and N. Yanagihara, *Admissible solutions of algebraic differential equations*, Funkcial. Ekvac. **38** (1995), 433–442.

[96] A. Its and V. Novokshenov, *The Isomonodromic Deformation Method in the Theory of Painlevé Equations*, Lecture Notes in Mathematics, Vol. 1191, Springer-Verlag, Berlin (1986).

[97] K. Iwasaki, H. Kimura, S. Shimomura and M. Yoshida, *From Gauss to Painlevé: A Modern Theory of Special Functions*, F. Vieweg & Sohn, Braunschweig (1991).

[98] G. Jank and L. Volkmann, *Einführung in die Theorie der ganzen und meromorphen Funktionen mit Anwendungen auf Differentialgleichungen*, Birkhäuser, Basel–Boston (1985).

[99] V. Jarnik, *Differential Equations in the Complex Domain*, Academia, Prague (1975) (in Czech)

[100] K. Katajamäki, *Algebroid solutions of binomial and linear differential equations*, Ann. Acad. Sci. Fenn. Ser. A I Math. Diss. **90** (1993), 48 p.

[101] W. Kim, *On the zeros of solutions of* $y^{(n)} + py = 0$, J. Math. Anal. Appl. **25** (1969), 189–208.

[102] L. Kinnunen, *Linear differential equations with solutions of finite iterated order*, Southeast Asian Math. Bull. **22** (1998), 385–405.

[103] A. Kitaev, C. Law and J. McLeod, *Rational solutions of the fifth Painlevé equation*, Differential Integral Equations **7** (1994), 967–1000.

[104] R. Korhonen and J. Rättyä, *Linear differential equations in the unit disc with analytic solutions of finite order*, Proc. Amer. Math. Soc. **135** (2007), 1355–1363.

[105] R. Korhonen and J. Rättyä, *Finite order solutions of linear differential equations in the unit disc*, in press.

[106] K.H. Kwon, *On the growth of entire functions satisfying second order linear differential equations*, Bull. Korean Math. Soc. **33** (1996), 487–496.

[107] I. Laine, *Nevanlinna Theory and Complex Differential Equations*, W. de Gruyter, Berlin–New York (1993).

[108] I. Laine and T. Sorvali, *Local solutions of* $w'' + A(z)w = 0$ *and branched polymorphic functions*, Results Math. **10** (1986), 107–129.

[109] I. Laine and K. Tohge, *The Bank–Laine conjecture – a survey*, Some Topics in Value Distribution Theory and Defferentiability in Complex and p-Adic Analysis, Science Press, Beijing, in press.

[110] I. Laine and R. Yang, *Finite order solutions of complex linear differential equations*, Electron. J. Differential Equations **65** (2004), 1–8.

[111] J. Langley, *Some oscillation theorems for higher order linear differential equations with entire coefficients of small growth*, Results Math. **20** (1991), 517–529.

[112] J. Langley, *Quasiconformal modifications and Bank–Laine functions*, Arch. Math. **71** (1998), 233–239.

[113] J. Langley, *Bank–Laine functions with sparse zeros*, Proc. Amer. Math. Soc. **129** (2000), 1969–1978.

[114] J. Langley, *Composite Bank–Laine functions and a question of Rubel*, Trans. Amer. Math. Soc. **354** (2001), 1177–1191.

[115] J. Langley, *Linear differential equations with entire coefficients of small growth*, Arch. Math. **78** (2002), 291–296.

[116] M. Lavie, *The Schwarzian derivative and disconjugacy of* n*th order linear differential equations*, Canad. Math. J. **21** (1969), 235–249.

[117] M. Lavie, *Disconjugacy of linear differential equations in the complex domain*, Pacific J. Math. **32** (1970), 435–457.

[118] Y. Li, *Some properties of the solutions of higher analogue of the Painlevé equation*, Acta Math. Appl. Sinica **22** (2006), 59–64.

[119] Y. Li and Y. He, *On analytic properties of higher analogues of the second Painlevé equation*, J. Math. Phys. **43** (2002), 1106–1115.

[120] L. Liao, W. Su and C.-C. Yang, *A Malmquist–Yosida type of theorem for the second-order algebraic differential equations*, J. Differential Equations **187** (2003), 63–71.

[121] L. Liao and Yang, C-C., *On the growth of meromorphic and entire solutions of second-order algebraic differential equations*, Ann. Mat. Pura Appl. **179** (2001), 149–158.

[122] L. Liao and Z. Ye, *A class of second order differential equations*, Israel J. Math. **146** (2005), 281–301.

[123] L. Liao and Z. Ye, *On the growth of meromorphic solutions of the Schwarzian differential equations*, J. Math. Anal. Appl. **309** (2005), 91–102.

[124] D. London, *On the zeros of solutions of* $w'' + p(z)w(z) = 0$, Pacific J. Math. **12** (1962), 979–991.

[125] J. Malmquist, *Sur les fonctions á un nombre fini des branches définies par les équations différentielles du premier ordre*, Acta Math. **36** (1913), 297–343.

[126] J. Malmquist, *Sur les fonctions á un nombre fini des branches satisfaisant á une équation différentielle du premier ordre*, Acta Math. **42** (1920), 317–325.

[127] L. Markus, *Differential independence of* Γ *and* ζ, J. Dynam. Differential Equations **19** (2007), 133–154.

[128] A. Mohon'ko, *Estimates of Nevanlinna characteristics of algebroid functions and their applications to differential equations*, Sib. Math. J. **23** (1982), 80–88.

[129] A. Mohon'ko, *Nevanlinna characteristics of a composition of rational and algebroid functions*, Ukrain. Math. J. **34** (1982), 319–325.

[130] E. Mues and R. Redheffer, *On the growth of logarithmic derivatives*, J. London Math. Soc. **8** (1974), 412–425.

[131] Y. Murata, *Rational solutions of the second and fourth Painlevé equations*, Funkcial. Ekvac. **28** (1985), 1–32.

[132] Z. Nehari, *The Schwarzian derivative and schlicht functions*, Bull. Amer. Math. Soc. **55** (1949), 545–551.

[133] Z. Nehari, *Conformal Mapping*, McGraw Hill Book Co., New York–Toronto–London (1952).

[134] Z. Nehari, *On the zeros of solutions of second order linear differential equations*, Amer. J. Math. **76** (1954), 689–697.

[135] Z. Nehari, *Univalent-functions and linear differential equations*, Lectures on Functions of a Complex Variable, Ann Arbor (1955), 49–60.

[136] F. Nevanlinna, *Über eine Klasse meromorpher Funktionen*, Septième Congrès Math. Scand., Oslo 1930, 81–83.

[137] M. Noumi, *Painlevé Equations through Symmetry*, Amer. Math. Soc., Providence, RI (2004).

[138] D. Novikov and S. Yakovenko, *A complex analogue of the Rolle theorem and polynomial envelopes of irreducible differential equations in the complex domain*, J. London Math. Soc. **56** (1997), 305–319.

[139] K. Okamoto, *Studies on the Painlevé equations*; I: Ann. Mat. Pura Appl. **146** (1987), 337–381; II: Japan. J. Math. **13** (1987), 47–76; III: Math. Ann. **275** (1986), 221–255; IV: Funkcial. Ekvac. **30** (1987), 305–332.

[140] K. Okamoto and K. Takano, *The proof of the Painlevé property by M. Hukuhara*, Funkcial. Ekvac. **44** (2001), 201–217.

[141] V. Petrenko, *Growth of Meromorphic Functions*, Viščа Škola, Kharkov (1978) (in Russian)

[142] V. Pokornyi, *On some sufficient conditions for univalence*, Dokl. Akad. Nauk SSSR (N.S.) **79** (1951), 743–746.

[143] C. Pommerenke, *On the mean growth of the solutions of complex linear differential equations in the disk*, Complex Variables Theory Appl. **1** (1982), 23–38.

[144] J. Rättyä, *Linear differential equations with solutions in the Hardy space*, Complex Var. Elliptic Equ. **52** (2007), 785–795.

[145] J. v. Rieth, *Untersuchungen gewisser Klassen gewöhnlicher Differentialgleichungen erster und zweiter Ordnung im Komplexen*, Ph.D. thesis, Aachen (1986).

[146] J. Rossi, *Second order differential equations with transcendental coefficients*, Proc. Amer. Math. Soc. **97** (1986), 61–66.

[147] J. Rossi and S. Wang, *The radial oscillation of solutions to ODE's in the complex domain*, Proc. Edinburgh Math. Soc. **39** (1996), 473–483.

[148] L. Rubel, *Some research problems about algebraic differential equations*, Trans. Amer. Math. Soc. **280** (1983), 43–52.

[149] L. Rubel, *A survey of transcendentally transcendental functions*, Amer. Math. Monthly **96** (1989), 777–788.

[150] L. Rubel, *Some research problems about algebraic differential equations II*, Illinois J. Math. **36** (1992), 659–680.

[151] H. Schubart and H. Wittich, *Über die Lösungen der beiden ersten Painlevéschen Differentialgleichungen*, Math. Z. **66** (1957), 364–370.

[152] B. Schwarz, *Complex nonoscillation theorems and criteria of univalence*, Trans. Amer. Math. Soc. **80** (1955), 159–186.

[153] H. Selberg, *Über die Wertverteilung der algebroiden Funktionen*, Math. Z. **31** (1930), 709–728.

[154] L.-C. Shen, *Proof of a conjecture of Bank and laine regarding product of two linearly independent solutions of* $y'' + Ay = 0$, Proc. Amer. Math. Soc. **100**, (1987), 301–308.

[155] S. Shimomura, *Value distribution of Painlevé transcendents of the third kind*, Complex Variables Theory Appl. **40** (1999), 51–62.

[156] S. Shimomura, *Value distribution of Painlevé transcendents of the fifth kind*, Results Math. **38** (2000), 348–361.

[157] S. Shimomura, *On meromorphic solutions of a linear differential equation with doubly periodic coefficients*, Illinois J. Math. **44** (2000), 593–601.

[158] S. Shimomura, *Value distribution of Painlevé transcendents of the first and the second kind*, J. Analyse Math. **82** (2000), 333–346.

[159] S. Shimomura, *On deficiencies of small functions for Painlevé transcendents of the fourth kind*, Ann. Acad. Sci. Fenn. Math. **27** (2002), 109–120.

[160] S. Shimomura, *Oscillation results for n-th order linear differential equations with meromorphic periodic coefficients*, Nagoya Math. J. **166** (2002), 55–82.

[161] S. Shimomura, *Proofs of the Painlevé property for all Painlevé equations*, Japan. J. Math. **29** (2003), 159–180.

[162] S. Shimomura, *Lower estimates for the growth of Painlevé transcendents*, Funkcial. Ekvac. **46** (2003), 287–295.

[163] S. Shimomura, *Growth of the first, the second and the fourth Painlevé transcendents*, Math. Proc. Camb. Phil. Soc. **134** (2003), 259–269.

[164] S. Shimomura, *Poles and* α*-points of meromorphic solutions of the first Painlevé hierarchy*, Publ. RIMS **40** (2004), 471–485.

[165] S. Shimomura, *Growth of modified Painlevé transcendents of the fifth and third kind*, Forum Math. **16** (2004), 231–247.

[166] S. Shimomura, *Meromorphic solutions of a Riccati differential equation with a doubly periodic coefficient*, J. Math. Anal. Appl. **304** (2005), 644–651.

[167] L. Sons, *Growth properties for solutions of differential equations in the disk*, Ann. Mat. Pura Appl. **162** (1992), 252–262.

[168] E. Steinbart, *A further note on a result of Bank, Frank and Laine*, Results Math. **42** (2002), 365–383.

[169] N. Steinmetz, *Eigenschaften eindeutiger Lösungen gewöhnlicher Differentialgleichungen im Komplexen*, Ph.D. thesis, Universität Karlsruhe (1978).

[170] N. Steinmetz, *Zur Wertverteilung der Lösungen der vierten Painlevéschen Differentialgleichung*, Math. Z. **181** (1982), 553–561.

[171] N. Steinmetz, *Ein Malmquistscher Satz für algebraische Differentialgleichungen zweiter Ordnung*, Results Math. **10** (1986), 152–167.

[172] N. Steinmetz, *Meromorphic solutions of second-order algebraic differential equations*, Complex Variables Theory Appl. **13** (1989), 75–83.

[173] N. Steinmetz, *On Painlevé's equations I, II and IV*, J. Analyse Math. **82** (2000), 363–377.

[174] N. Steinmetz, *Value distribution of the Painlevé transcendents*, Israel J. Math. **128** (2002), 29–52.

[175] N. Steinmetz, *Boutroux's method vs, re-scaling*, Port. Math. **61** (2004), 369–374.

[176] S. Steinmetz, *Über das Anwachsen der Lösungen homogenen algebraischer Differentialgleichungen zweiter Ordnung*, Manuscripta Math. **32** (1980), 303–308.

[177] S. Strelitz, *Three theorems on the growth of entire transcendental solutions of algebraic differential equations*, Canad. J. Math. **35** (1983), 1110–1128.

[178] N. Toda, *On the growth of meromorphic solutions of an algebraic differential equation*, Proc. Japan Acad. Ser. A Math. Sci. **60** (1984), 117–120.

[179] N. Toda, *On algebroid solutions of some algebraic differential equations in the complex plane*, Proc. Japan Acad. Ser. A Math. Sci. **65** (1989), 94–97.

[180] N. Toda, *On algebroid solutions of algebraic differential equations in the complex plane II*, J. Math. Soc. Japan **45** (1993), 705–717.

[181] N. Toda and M. Kato, *On some algebraic differential equations with admissible algebroid solutions*, Proc. Japan Acad. Ser. A Math. Sci. **61** (1985), 325–328.

[182] K. Tohge, *Logarithmic derivatives of meromorphic or algebroid solutions of some homogeneous linear differential equations*, Analysis **19** (1999), 273–297.

[183] E. Ullrich, *Über den Einfluß der Verzweigtheit einer Algebroide auf ihre Wertverteilung*, J. Reine Angew. Math. **167** (1931), 198–220.

[184] G. Valiron, *Sur la dérivée des fonctions algébroïdes*, Bull. Soc. Math. France **59** (1931), 17–39.

[185] J. von Rieth, *Untersuchungen gewisser Klassen Gewöhnlicher Differentialgleichungen erster und zweiter Ordnung im Komplexen*, Ph.D. thesis, Technische Hochschule Aachen (1986).

[186] S. Wang, *On the frequency of zeros of a fundamental solutions set of complex linear differential equations*, Kodai Math. J. **20** (1997), 143–155.

[187] R. Weikard, *Floquet theory for linear differential equations with meromorphic solutions*, Electron. J. Qual. Theory Differ. Equ. **8** (2000), 1–6.

[188] H. Wittich, *Eindeutige Lösungen der Differentialgleichungen* $w'' = P(z, w)$, Math. Ann. **125** (1953), 355–365.

[189] H. Wittich, *Neuere Untersuchungen über eindeutige analytische Funktionen*, Springer–Verlag, Berlin-Göttingen-Heidelberg (1955).

[190] H. Wittich, *Zur Theorie linearer Differentialgleichungen im Komplexen*, Ann. Acad. Sci. Fenn. Ser. **A I 379** (1966).

[191] S. Yamashita, *Schlicht holomorphic functions and the Riccati differential equation*, Math. Z. **157** (1977), 19–22.

[192] C.-C. Yang and Z. Ye, *Estimates of the proximate function of differential polynomials*, Proc. Japan Acad. Ser. A Math. Sci. **83** (2007), 50–55.

[193] L.Z. Yang, *The growth of linear differential equations and their applications*, Israel J. Math. **147** (2005), 359–370.

[194] K. Yosida, *On algebroid solutions of ordinary differential equations*, Japan. J. Math. **10** (1934), 199–208.

[195] W. Yuan and Y. Li, *Rational solutions of Painlevé equations*, Canad. J. Math. **54** (2002), 648–670.

[196] K. Zhu, *Bloch type spaces of analytic functions*, Rocky Mountain J. Math. **23** (1993), 1143–1177.

[197] V. Zimogljad, *The order of growth of entire transcendental solutions of second order algebraic differential equations*, Mat. Sb. (N.S.) **85(127)** (1971), 286–302.

This page intentionally left blank

CHAPTER 4

Transversal Periodic-to-Periodic Homoclinic Orbits

Kenneth J. Palmer*

Department of Mathematics, National Taiwan University, Taipei 106, Taiwan
E-mail: *palmer@math.ntu.edu.tw*

Contents

1. Introduction . . . 367
2. Trichotomies . . . 368
3. Hyperbolic periodic orbits and their stable and unstable manifolds . . . 381
    3.1. Poincaré map . . . 382
    3.2. Local stable and unstable manifolds . . . 386
    3.3. Global stable and unstable manifolds and asymptotic phase . . . 395
4. Homoclinic orbits . . . 402
    4.1. Poincaré map associated with a homoclinic orbit . . . 403
    4.2. Transversal periodic-to-periodic homoclinic orbits and hyperbolicity . . . 405
    4.3. Chaos near a transversal periodic-to-periodic homoclinic orbit . . . 407
5. Robustness of transversal periodic-to-periodic homoclinic orbits . . . 412
6. Finding transversal periodic-to-periodic homoclinic orbits through regular perturbation . . . 415
7. Finding transversal periodic-to-periodic homoclinic orbits through numerical shadowing . . . 425
References . . . 438

---

*Partially supported by NSC.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 4
Edited by F. Battelli and M. Fečkan
© 2008 Elsevier B.V. All rights reserved

This page intentionally left blank

1. Introduction

The possible presence of transversal homoclinic points and the complicated dynamics associated with them was noticed by Poincaré [21] while studying the restricted three-body problem. Later Birkhoff [3] proved that every transversal homoclinic point of a two-dimensional diffeomorphism is accumulated by periodic points. Then Smale's [25] theorem showed that the existence of a transversal homoclinic orbit implied the presence of a horseshoe, and hence complicated dynamics.

Note that Smale's result was proved for diffeomorphisms. A transversal homoclinic point is a point where the stable and unstable manifolds of a hyperbolic fixed point intersect transversally. The corresponding notion for a flow is a point where the stable and unstable manifolds of a hyperbolic periodic orbit intersect transversally. We call the orbit of such a point a transversal periodic-to-periodic homoclinic orbit, the subject of the present chapter. Sil'nikov [22] proved a result analogous to Smale's for transversal periodic-to-periodic homoclinic orbits. A detailed history of homoclinic orbits is given by Sil'nikov [23].

A homoclinic orbit of a diffeomorphism is transversal if and only if the invariant set consisting of it and the associated hyperbolic fixed point is hyperbolic (confer [19]). This, in turn, happens if and only if the variational equation along the homoclinic orbit has an exponential dichotomy. Analogously a periodic-to-periodic homoclinic orbit of a flow is transversal if and only if the invariant set consisting of it and the associated hyperbolic periodic orbit is hyperbolic. It turns out that happens if and only if the variational equation along the homoclinic orbit has what we call a trichotomy. Trichotomy is similar to an exponential dichotomy but with the difference that there is an additional one-dimensional subspace of bounded solutions. Trichotomies play an important role in our development of the theory and so Section 2 is devoted to their study.

In Section 3 we first define the Poincaré map associated with a periodic orbit of an autonomous system of ordinary differential equations. Then we construct the local and global stable and unstable manifolds of a hyperbolic periodic orbit, thereby introducing the concept of asymptotic phase. These, of course, are well-known results but we give detailed proofs because, for later application, we need the results in a special form. Note that our smoothness assumptions may not be optimal.

In Section 4 we study properties of periodic-to-periodic homoclinic orbits. First we define a Poincaré map associated with such an orbit similar to that given by Hale and Lin in [11]. Then we give equivalent conditions for transversality of the periodic-to-periodic homoclinic orbits including its equivalence with the transversality of the corresponding homoclinic orbit of the Poincaré map. Note that the possibility of doing this was suggested to me in a conversation with Daniel Stoffer in Santander airport. Then Smale's theorem enables to deduce the existence of chaotic behaviour in the neighbourhood of a transversal periodic-to-periodic homoclinic orbit. Note that in [19] a shadowing theorem for hyperbolic sets of flows was used to give a complete symbolic description of the dynamics in the neighbourhood of a transversal periodic-to-periodic homoclinic orbit.

In Section 5 we consider an autonomous system depending on a parameter where we assume the unperturbed system has a hyperbolic periodic orbit with associated transversal homoclinic orbit. We show that the perturbed system has a similar structure.

Even though, as we found in Section 5, transversal periodic-to-periodic homoclinic orbits persist under perturbation, it does not seem possible to write down an explicit system with an explicit transversal periodic-to-periodic homoclinic orbit. This is also the situation with homoclinic points of diffeomorphisms. In the latter case, Melnikov [15] finds diffeomorphisms with transversal homoclinic points as period maps of differential equations with time-periodic vector fields which are perturbations of autonomous systems with nontransversal homoclinic orbits. In Section 6 we prove an analogous theorem for periodic-to-periodic homoclinic orbits of flows. We begin with a system with a nontransversal homoclinic orbit; such systems can be easily written down explicitly. Then we show that a perturbed system has a transversal orbit nearby provided a certain Melnikov function has a simple zero. Note that the existence of transversal periodic-to-periodic homoclinic orbits in singularly perturbed systems has been shown by Szmolyan in [26]; see also [1]. Note also that in [7] Deng and Sakamoto show that periodic-to-periodic homoclinic orbits arise as the result of a Sil'nikov–Hopf bifurcation. However, these orbits may be nontransversal.

In Section 7, which is based on Coomes, Koçak and Palmer [4], it is shown how transversal periodic-to-periodic homoclinic orbits can be constructed via numerical shadowing. We begin with a numerically computed orbit which appears to be bi-asymptotic to an also numerically computed apparent periodic orbit. The main theorem gives computable criteria for verifying the existence of a true periodic orbit and an associated true transversal homoclinic orbit near the numerically computed orbits. Other authors, for example, Dieci and Rebaza [9] and Pampel [20], inspired by the work of Beyn [2], approximate a connecting orbit by the solution of a certain boundary value problem and they derive estimates for the error in the approximation. However, these works are carried out on the assumption that a true connecting orbit exists. Examples of transversal periodic-to-periodic homoclinic in the restricted three-body problem are studied in Kirchgraber and Stoffer [14].

2. Trichotomies

It turns out that a periodic-to-periodic homoclinic orbit is transversal if and only if its variational equation has a trichotomy in the sense defined below. So trichotomies are an important tool in our study and, in this section, we prove the results about trichotomies that we will need in the sequel.

DEFINITION 2.1. Let $A(t)$ be a continuous $n \times n$ matrix function. The linear system

$$\dot{x} = A(t)x \tag{1}$$

is said to have a *trichotomy* on an interval J if there exist supplementary projections P_0, P_+, P_- with rank $P_0 = 1$ and positive constants K, α such that

$$\begin{aligned} &\left|X(t)P_0X^{-1}(s)\right| \leqslant K \quad \text{for all } t, s,\\ &\left|X(t)P_+X^{-1}(s)\right| \leqslant K\mathrm{e}^{-\alpha(t-s)} \quad \text{for } t \geqslant s,\\ &\left|X(t)P_-X^{-1}(s)\right| \leqslant K\mathrm{e}^{-\alpha(s-t)} \quad \text{for } t \leqslant s \end{aligned} \tag{2}$$

where $X(t)$ is the fundamental matrix for (1) with $X(0) = I$.

Note that if $P_0 = 0$ we have an exponential dichotomy. Sometimes we refer to K as the “constant” and to α as the “exponent”.

Trichotomies have similar properties to exponential dichotomies. We describe these properties in the remarks below.

REMARK 2.2. *Characterization of projections*. If $J = [t_0, \infty)$, then it is clear that

$$\begin{aligned}
\mathcal{R}(P_+) &= \bigl\{\xi\colon \bigl|X(t)\xi\bigr| \to 0 \text{ as } t \to \infty\bigr\},\\
\mathcal{R}(P_0 + P_+) &= \Bigl\{\xi\colon \sup_{t \geqslant t_0}\bigl|X(t)\xi\bigr| < \infty\Bigr\}.
\end{aligned}$$

Now we show that $\mathcal{R}(P_0)$ can be any complementary subspace U_0 to $\mathcal{R}(P_+)$ in $\mathcal{R}(P_0 + P_+)$ and that $\mathcal{R}(P_-)$ can be any complementary subspace W_0 to $\mathcal{R}(P_0 + P_+)$ in $\mathbf{R}^n$. To see this, let Q_0, Q_+ and Q_- be the projections associated with the splitting

$$\mathbf{R}^n = U_0 \oplus \mathcal{R}(P_+) \oplus W_0.$$

Then for $t, s \geqslant t_0$, since $\mathcal{R}(P_+) = \mathcal{R}(Q_+)$ so that $P_+Q_+ = Q_+$ and $Q_+P_+ = P_+$,

$$\begin{aligned}
&\bigl|X(t)(Q_+ - P_+)X^{-1}(s)\bigr|\\
&\quad = \bigl|X(t)P_+(Q_+ - P_+)(I - P_+)X^{-1}(s)\bigr|\\
&\quad \leqslant \bigl|X(t)P_+X^{-1}(t_0)\bigr|\bigl|X(t_0)(Q_+ - P_+)X^{-1}(t_0)\bigr|\bigl|X(t_0)(P_0 + P_-)X^{-1}(s)\bigr|\\
&\quad \leqslant K\mathrm{e}^{-\alpha(t-t_0)}\bigl|X(t_0)(Q_+ - P_+)X^{-1}(t_0)\bigr|2K\\
&\quad = 2K^2\bigl|X(t_0)(Q_+ - P_+)X^{-1}(t_0)\bigr|\mathrm{e}^{-\alpha(t-t_0)}\\
&\quad = K_1\mathrm{e}^{-\alpha(t-t_0)}
\end{aligned}$$

and since $\mathcal{N}(P_-) = \mathcal{N}(Q_-)$ so that $P_-Q_- = P_-$ and $Q_-P_- = Q_-$,

$$\begin{aligned}
&\bigl|X(t)(Q_- - P_-)X^{-1}(s)\bigr|\\
&\quad = \bigl|X(t)(P_0 + P_+)(Q_- - P_-)P_-X^{-1}(s)\bigr|\\
&\quad \leqslant \bigl|X(t)(P_0 + P_+)X^{-1}(t_0)\bigr|\bigl|X(t_0)(Q_- - P_-)X^{-1}(t_0)\bigr|\bigl|X(t_0)P_-X^{-1}(s)\bigr|\\
&\quad \leqslant 2K\bigl|X(t_0)(Q_- - P_-)X^{-1}(t_0)\bigr|K\mathrm{e}^{-\alpha(s-t_0)}\\
&\quad = K_2\mathrm{e}^{-\alpha(s-t_0)}.
\end{aligned}$$

In particular this implies that if $t \geqslant t_0$,

$$\bigl|X(t)(Q_+ - P_+)X^{-1}(t)\bigr| \leqslant K_1, \qquad \bigl|X(t)(Q_- - P_-)X^{-1}(t)\bigr| \leqslant K_2.$$

Then going through the first argument again with t_0 replaced by s, with $s \geqslant t_0$, we find that for $t \geqslant s \geqslant t_0$

$$\left|X(t)(Q_+ - P_+)X^{-1}(s)\right| \leqslant K_3 \mathrm{e}^{-\alpha(t-s)},$$

where $K_3 = 2K^2K_1$, and going through the second argument again with t_0 replaced by t, with $t \geqslant t_0$, we find that for $s \geqslant t \geqslant t_0$

$$\left|X(t)(Q_- - P_-)X^{-1}(s)\right| \leqslant K_4 \mathrm{e}^{-\alpha(s-t)},$$

where $K_4 = 2K^2K_2$. Also we note that for $t, s \geqslant t_0$

$$\begin{aligned}\left|X(t)(Q_0 - P_0)X^{-1}(s)\right| &\leqslant \left|X(t)(Q_+ - P_+)X^{-1}(s)\right| \\ &\quad + \left|X(t)(Q_- - P_-)X^{-1}(s)\right| \\ &\leqslant K_1 + K_2.\end{aligned}$$

Then we see that for $t, s \geqslant t_0$

$$\begin{aligned}\left|X(t)Q_0X^{-1}(s)\right| &\leqslant \left|X(t)P_0X^{-1}(s)\right| + \left|X(t)(Q_0 - P_0)X^{-1}(s)\right| \\ &\leqslant K + K_1 + K_2,\end{aligned}$$

for $t \geqslant s \geqslant t_0$

$$\begin{aligned}\left|X(t)Q_+X^{-1}(s)\right| &\leqslant \left|X(t)P_+X^{-1}(s)\right| + \left|X(t)(Q_+ - P_+)X^{-1}(s)\right| \\ &\leqslant (K + K_3)\mathrm{e}^{-\alpha(t-s)}\end{aligned}$$

and for $s \geqslant t \geqslant t_0$

$$\begin{aligned}\left|X(t)Q_-X^{-1}(s)\right| &\leqslant \left|X(t)P_-X^{-1}(s)\right| + \left|X(t)(Q_- - P_-)X^{-1}(s)\right| \\ &\leqslant (K + K_4)\mathrm{e}^{-\alpha(s-t)}.\end{aligned}$$

If $J = (-\infty, t_0]$, then it is clear that

$$\begin{aligned}\mathcal{R}(P_-) &= \left\{\xi\colon \left|X(t)\xi\right| \to 0 \text{ as } t \to -\infty\right\}, \\ \mathcal{R}(P_0 + P_-) &= \left\{\xi\colon \sup_{t \leqslant t_0}\left|X(t)\xi\right| < \infty\right\},\end{aligned}$$

and we can similarly prove that $\mathcal{R}(P_0)$ can be any complementary subspace to $\mathcal{R}(P_-)$ in $\mathcal{R}(P_0 + P_-)$ and that $\mathcal{R}(P_+)$ can be any complementary subspace to $\mathcal{R}(P_0 + P_-)$ in $\mathbf{R}^n$.

If $J = \mathbf{R}$, it is clear that

$$\begin{aligned}\mathcal{R}(P_+) &= \left\{\xi\colon \left|X(t)\xi\right| \to 0 \text{ as } t \to \infty\right\}, \\ \mathcal{R}(P_-) &= \left\{\xi\colon \left|X(t)\xi\right| \to 0 \text{ as } t \to -\infty\right\},\end{aligned}$$

and

$$\mathcal{R}(P_0) = \Big\{\xi\colon \sup_{t\in\mathbf{R}}\big|X(t)\xi\big| < \infty\Big\}.$$

So, in this case, up to a scalar multiple, the linear system has a unique nontrivial bounded solution $x_0(t)$ where $x_0(0)$ spans the range of P_0 and, moreover, $\inf_{t\in\mathbf{R}} |x_0(t)| \geqslant K^{-1}|x_0(0)| > 0$.

REMARK 2.3. *If* $t_0 < t_1$, *trichotomy on* $[t_1, \infty)$ *implies trichotomy on* $[t_0, \infty)$. Let (1) have a trichotomy on $[t_1, \infty)$ as in Definition 2.1. For $\sigma = 0, +, -$, define

$$M_\sigma = \sup\big\{\mathrm{e}^{\alpha(t-s)}\big|X(t)P_\sigma X^{-1}(s)\big|\colon\ t_0 \leqslant t, s \leqslant t_1\big\}.$$

Then for $t_0 \leqslant t, s \leqslant t_1$ and $\sigma = 0, +, -$,

$$\big|X(t)P_\sigma X^{-1}(s)\big| \leqslant M_\sigma \mathrm{e}^{-\alpha(t-s)}.$$

Then, for example, if $t_0 \leqslant s \leqslant t_1 \leqslant t$,

$$\big|X(t)P_+X^{-1}(s)\big| \leqslant \big|X(t)P_+X^{-1}(t_1)\big|\big|X(t_1)P_+X^{-1}(s)\big| \leqslant M_+K\mathrm{e}^{-\alpha(t-s)}.$$

It follows that if $t_0 \leqslant s \leqslant t$,

$$\big|X(t)P_+X^{-1}(s)\big| \leqslant K_1\mathrm{e}^{-\alpha(t-s)},$$

where

$$K_1 = \max\{K, M_+K\}.$$

The inequalities for P_0 and P_- can be similarly derived.

Similarly, if $t_0 > t_1$, trichotomy on $(-\infty, t_1]$ implies trichotomy on $(-\infty, t_0]$.

REMARK 2.4. *Trichotomy on* $\mathbf{R}$ *from trichotomies on* $\mathbf{R}_+$ *and* $\mathbf{R}_-$. If (1) has a trichotomy on $\mathbf{R}$, then it certainly has trichotomies on $\mathbf{R}_+$ and $\mathbf{R}_-$ with projections of the same rank and, as we saw at the end of Remark 2.2, (1) has, up to a scalar multiple, a unique nontrivial bounded solution $x_0(t)$ which also satisfies $\inf_{t\in\mathbf{R}} |x_0(t)| > 0$.

The converse statement is also true. Suppose the projections on $[0, \infty)$ are P_0, P_+, P_- and the projections on $(-\infty, 0]$ are Q_0, Q_+, Q_- with K and α as the constants in both cases. Now if $\xi \in \mathcal{R}(P_+) \cap \mathcal{R}(Q_0 + Q_-)$, then $X(t)\xi$ is bounded on the whole line and hence is a multiple of $x_0(t)$. However, $X(t)\xi \to 0$ as $t \to \infty$ and so $\xi = 0$. Hence $\mathcal{R}(P_+)$ is complementary to $\mathcal{R}(Q_0 + Q_-)$ and so, by Remark 2.2, the Q_σ can be chosen so that $\mathcal{R}(Q_+) = \mathcal{R}(P_+)$. Similarly, we can assume $\mathcal{R}(P_-) = \mathcal{R}(Q_-)$. Next since $\inf_{t\in\mathbf{R}} |x_0(t)| > 0$, $\operatorname{span}\{x_0(0)\}$ is complementary to $\mathcal{R}(P_+)$ in $\mathcal{R}(P_0 + P_+)$ and also complementary to

$\mathcal{R}(Q_-)$ in $\mathcal{R}(Q_0 + Q_-)$. Hence we can assume $\mathcal{R}(P_0) = \mathcal{R}(Q_0) = \operatorname{span}\{x_0(0)\}$. Thus we can arrange that $P_\sigma = Q_\sigma$ for $\sigma = 0, +, -$. Then we know that

$$\left|X(t)P_+X^{-1}(s)\right| \leqslant K\mathrm{e}^{-\alpha(t-s)}$$

for $s \leqslant t \leqslant 0$ and $0 \leqslant s \leqslant t$. If $s \leqslant 0 \leqslant t$, then

$$\begin{aligned}\left|X(t)P_+X^{-1}(s)\right| &\leqslant \left|X(t)P_+X^{-1}(0)\right|\left|X(0)P_+X^{-1}(s)\right| \\ &\leqslant K\mathrm{e}^{-\alpha t}K\mathrm{e}^{-\alpha(0-s)} \\ &= K^2\mathrm{e}^{-\alpha(t-s)}.\end{aligned}$$

Using two more similar arguments, we see that (1) has a trichotomy on **R** with constants K^2 and α.

REMARK 2.5. *The case when $A(t)$ is bounded.* If $A(t)$ is bounded, we can weaken the inequalities in the definition of trichotomy to having for all ξ

$$\begin{aligned}&\left|X(t)P_0\xi\right| \leqslant K\left|X(s)P_0\xi\right| \quad \text{for all } t, s,\\ &\left|X(t)P_+\xi\right| \leqslant K\mathrm{e}^{-\alpha(t-s)}\left|X(s)P_+\xi\right| \quad \text{for } t \geqslant s,\\ &\left|X(t)P_-\xi\right| \leqslant K\mathrm{e}^{-\alpha(s-t)}\left|X(s)P_-\xi\right| \quad \text{for } t \leqslant s\end{aligned}$$

since these inequalities imply that for $t \geqslant s$

$$\frac{|X(t)P_+\xi||X(s)P_0\xi|}{|X(s)P_+\xi||X(t)P_0\xi|} \leqslant K^2\mathrm{e}^{-\alpha(t-s)}$$

and

$$\frac{|X(t)P_0\xi||X(s)P_-\xi|}{|X(s)P_0\xi||X(t)P_-\xi|} \leqslant K^2\mathrm{e}^{-\alpha(t-s)}.$$

This means the solution subspaces corresponding to P_+, P_0 and P_- are exponentially separated and so by Lemma 1 in [16] the norms of the projections $X(t)P_\sigma X^{-1}(t)$ are bounded by a constant M (say) and so, for example, if $s \leqslant t$ then for all ξ

$$\left|X(t)P_+X^{-1}(s)\xi\right| \leqslant K\mathrm{e}^{-\alpha(t-s)}\left|X(s)P_+X^{-1}(s)\xi\right| \leqslant KM\mathrm{e}^{-\alpha(t-s)}|\xi|$$

so that

$$\left|X(t)P_+X^{-1}(s)\right| \leqslant KM\mathrm{e}^{-\alpha(t-s)}.$$

REMARK 2.6. *The adjoint equation*. By taking adjoints in the inequalities (2) and exchanging s and t, we find that

$$\begin{aligned} &\left|Y(t)P_0^*Y^{-1}(s)\right| \leqslant K \quad \text{for all } t, s, \\ &\left|Y(t)P_-^*Y^{-1}(s)\right| \leqslant K\mathrm{e}^{-\alpha(t-s)} \quad \text{for } t \geqslant s, \\ &\left|Y(t)P_+^*X^{-1}(s)\right| \leqslant K\mathrm{e}^{-\alpha(s-t)} \quad \text{for } t \leqslant s, \end{aligned}$$

where $Y(t) = X^{-1}(t)^*$. So the adjoint system has a trichotomy on J with projections P_0^*, P_-^*, P_+^*.

REMARK 2.7. *Roughness theorem*. For this, for convenience, we restrict to the case where J is $\mathbf{R}$, $\mathbf{R}_+$ or $\mathbf{R}_-$. We show the following result.

Suppose (1) *has a trichotomy as in Definition* 2.1, *letting* $x_0(t)$ *be the solution with* $P_0x_0(0) = x_0(0)$ *and* $|x_0(0)| = 1$. *Let* $B(t)$ *be a continuous matrix function with* $|B(t)| \leqslant \delta$ *for all* t *in* J *such that*

$$\dot{x} = \left[A(t) + B(t)\right]x \tag{3}$$

has a solution $y_0(t)$ *with*

$$0 < \Delta \leqslant \left|y_0(t)\right| \leqslant L < \infty$$

for all t *in* J *and also*

$$\left|y_0(0) - x_0(0)\right| < K^{-1}$$

when J *is a half-line.*

Let $\varepsilon > 0$ *be given with* $\varepsilon < \alpha/2$. *Then if* δ *is sufficiently small,* (3) *has a trichotomy on* J *with constant depending only on* K, L, Δ (*and* $|y_0(0) - x_0(0)|$ *when* J *is a half-line*) *and exponent* $\alpha - \varepsilon$.

Note first that the shifted equation

$$\dot{x} = \left[A(t) + \alpha/2\right]x$$

has an exponential dichotomy on J with projection P_+ and constants $2K$, $\alpha/2$. It follows from the roughness theorem for dichotomies in [5, Lecture 5] that, if δ is sufficiently small, the perturbed system

$$\dot{x} = \left[A(t) + B(t) + \alpha/2\right]x \tag{4}$$

has an exponential dichotomy on J with both exponents $\alpha/2 - \varepsilon$ and constant K_1 depending on K and projection Q_+ having the same rank as P_+. In particular, if $Y(t)$ is the fundamental matrix of (3) with $Y(0) = I$, then

$$\left|Y(t)Q_+Y^{-1}(s)\right| \leqslant K_1\mathrm{e}^{-(\alpha-\varepsilon)(t-s)} \tag{5}$$

for $t \geqslant s$.

Next the shifted equation

$$\dot{x} = \left[A(t) - \alpha/2\right]x$$

has an exponential dichotomy on J with projection $P_0 + P_+$ and constants $2K$, $\alpha/2$. It follows from the roughness theorem for dichotomies that, if δ is sufficiently small, the perturbed system

$$\dot{x} = \left[A(t) + B(t) - \alpha/2\right]x$$

has an exponential dichotomy on J with projection $I - Q_-$ having the same rank as $P_0 + P_+$ with exponent $\alpha/2 - \varepsilon$ and constant K_1. In particular,

$$\left|Y(t)Q_-Y^{-1}(s)\right| \leqslant K_1\mathrm{e}^{-(\alpha-\varepsilon)(s-t)} \tag{6}$$

for $t \leqslant s$.

Now if $J = \mathbf{R}$ or $\mathbf{R}_+$,

$$\mathcal{N}(Q_-) = \left\{\xi\colon \sup_{t\geqslant 0} \mathrm{e}^{-\alpha t/2}\left|Y(t)\xi\right| < \infty\right\},$$
$$\mathcal{R}(Q_+) = \left\{\xi\colon \sup_{t\geqslant 0} \mathrm{e}^{\alpha t/2}\left|Y(t)\xi\right| < \infty\right\}$$

so that $\mathcal{R}(Q_+) \subset \mathcal{N}(Q_-)$ and $y_0(0) \in \mathcal{N}(Q_-)$. It follows that $Q_-Q_+ = 0$ and $Q_-y_0(0) = 0$. Similarly, if $J = \mathbf{R}$ or $\mathbf{R}_-$, $Q_+Q_- = 0$ and $Q_+y_0(0) = 0$.

Then, when $J = \mathbf{R}$, if we define $Q_0 = I - Q_+ - Q_-$, we see that Q_0, Q_+, Q_- are supplementary projections such that the range of Q_0 is spanned by $y_0(0)$. Now for all t and s

$$\left|y_0(t)\right| \leqslant L\Delta^{-1}\left|y_0(s)\right|$$

and so for all ξ

$$\left|Y(t)Q_0\xi\right| \leqslant L\Delta^{-1}\left|Y(s)Q_0\xi\right|.$$

Hence

$$\left|Y(t)Q_0Y^{-1}(s)\xi\right| \leqslant L\Delta^{-1}\left|Y(s)Q_0Y^{-1}(s)\xi\right|.$$

However,

$$\left|Y(s)Q_0Y^{-1}(s)\right| = \left|Y(s)(I - Q_+ - Q_-)Y^{-1}(s)\right| \leqslant 2K_1 + 1$$

so that

$$\left|Y(t)Q_0Y^{-1}(s)\right| \leqslant L(2K_1+1)\Delta^{-1}$$

for all t and s. This together with (5) and (6) shows that (3) has a trichotomy if δ is sufficiently small with constant $\max\{K_1, L(2K_1+1)\Delta^{-1}\}$ and exponent $\alpha - \varepsilon$.

Now suppose $J = \mathbf{R}_+$. Coppel chooses Q_+ so that

$$\mathcal{N}(Q_+) = \mathcal{N}(P_+) = \operatorname{span}\bigl\{x_0(0)\bigr\} \oplus \mathcal{R}(P_-)$$

and $|Q_+ - P_+| \leqslant N\delta$, where N depends only on K and α. Similarly, he chooses Q_- so that

$$\mathcal{R}(Q_-) = \mathcal{R}(P_-)$$

and $|Q_- - P_-| \leqslant N\delta$. Since $\mathcal{R}(Q_+) \subset \mathcal{N}(Q_-)$, $y_0(0) \in \mathcal{N}(Q_-)$ and clearly $y_0(0) \notin \mathcal{R}(Q_+)$ it follows by consideration of the dimensions that

$$\mathcal{N}(Q_-) = \operatorname{span}\bigl\{y_0(0)\bigr\} \oplus \mathcal{R}(Q_+).$$

Then

$$\mathbf{R}^n = \mathcal{N}(Q_-) \oplus \mathcal{R}(Q_-) = \operatorname{span}\bigl\{y_0(0)\bigr\} \oplus \mathcal{R}(Q_+) \oplus \mathcal{R}(Q_-). \tag{7}$$

Let $\tilde{Q}_+$ be the projection with range $\mathcal{R}(Q_+)$ and nullspace $\operatorname{span}\{y_0(0)\} \oplus \mathcal{R}(Q_-)$ and set $\tilde{Q}_0 = I - \tilde{Q}_+ - Q_-$. We see that $\tilde{Q}_0$, $\tilde{Q}_+$, Q_- are supplementary projections such that the range of $\tilde{Q}_0$ is spanned by $y_0(0)$.

By [5, p. 14], we know we still get a dichotomy for (4) with the altered projection $\tilde{Q}_+$ with exponent $\alpha/2 - \varepsilon$ but different constant depending on K, α and $|\tilde{Q}_+ - Q_+|$. However we need to ensure that the constant can be chosen as a fixed number depending only on the other constants provided δ is sufficiently small. To this end, we need to estimate $|\tilde{Q}_+ - Q_+|$. Using $P_0 + P_+ + P_- = I$, $\tilde{Q}_+Q_+ = Q_+$, $\tilde{Q}_+Q_- = 0$ and $Q_+P_- = 0$, we find that

$$\begin{aligned}
|\tilde{Q}_+ - Q_+| &= \bigl|(\tilde{Q}_+ - Q_+)P_0 + (\tilde{Q}_+ - Q_+)(P_+ + P_-)\bigr| \\
&= \bigl|(\tilde{Q}_+ - Q_+)P_0 + (\tilde{Q}_+ - Q_+)(P_+ + P_- - Q_+ - Q_-)\bigr| \\
&\leqslant \bigl|(\tilde{Q}_+ - Q_+)P_0\bigr| + 2N\delta|\tilde{Q}_+ - Q_+|.
\end{aligned}$$

Now if $P_0x = \lambda x_0(0)$ so that $|\lambda| \leqslant K|x|$,

$$\begin{aligned}
&\bigl|(\tilde{Q}_+ - Q_+)P_0x\bigr| \\
&\quad = |\lambda|\bigl|(\tilde{Q}_+ - Q_+)x_0(0)\bigr| \\
&\quad \leqslant K|x|\bigl|(\tilde{Q}_+ - Q_+)x_0(0)\bigr|
\end{aligned}$$

$$\begin{aligned}&= K|x|\big|\tilde{Q}_+\big(x_0(0)-y_0(0)\big)\big| \quad \text{since } Q_+x_0(0)=\tilde{Q}_+y_0(0)=0\\ &\leqslant K|x||\tilde{Q}_+|\big|x_0(0)-y_0(0)\big|\\ &\leqslant K|x|\big[|Q_+|+|\tilde{Q}_+-Q_+|\big]\big|x_0(0)-y_0(0)\big|.\end{aligned}$$

It follows that

$$\begin{aligned}&|\tilde{Q}_+-Q_+|\\ &\quad\leqslant K\big[|Q_+|+|\tilde{Q}_+-Q_+|\big]\big|x_0(0)-y_0(0)\big|+2N\delta|\tilde{Q}_+-Q_+|\\ &\quad\leqslant KK_1\big|x_0(0)-y_0(0)\big|+\big(K\big|x_0(0)-y_0(0)\big|+2N\delta\big)|\tilde{Q}_+-Q_+|.\end{aligned}$$

So, provided

$$\begin{aligned}&2N\delta<1-K\big|x_0(0)-y_0(0)\big|,\\ &|\tilde{Q}_+-Q_+|\leqslant\big(1-K\big|x_0(0)-y_0(0)\big|-2N\delta\big)^{-1}KK_1\big|x_0(0)-y_0(0)\big|.\end{aligned}$$

Then it follows from [5, p. 14] that (4) has an exponential dichotomy with the altered projection $\tilde{Q}_+$, with the same exponent $\alpha/2-\varepsilon$ and with constant depending only on K, K_1 and $|x_0(0)-y_0(0)|$, provided δ is sufficiently small. Then we finish the proof that (3) has a trichotomy on $\mathbf{R}_+$ as in the case $J=\mathbf{R}$.

The case $J=\mathbf{R}_-$ is similarly handled.

REMARK 2.8. *Smoothness of the projections.* We can show the smoothness of the projections in a trichotomy with respect to a parameter. Suppose (1) has a trichotomy as in the definition. Let $B(t,\mu)$ be a continuous matrix function where the continuity in μ is uniform in t and $B(t,0)=0$. Suppose also that the system

$$\dot{x}=\big[A(t)+B(t,\mu)\big]x \tag{8}$$

has for μ small a solution $y_0(t,\mu)$ with $0<\Delta\leqslant|y_0(t,\mu)|\leqslant L<\infty$ such that $y_0(0,\mu)$ is continuous. Then by Remark 2.7, if μ is sufficiently small, system (8) has a trichotomy on J. Now we suppose in addition that $B_\mu(t,\mu)$ exists and is bounded and continuous where the continuity in μ is uniform with respect to t and also that $y_0(0,\mu)$ is C^1 in μ.

Let the projections be $Q_0(\mu)=I-Q_+(\mu)-Q_-(\mu)$, $Q_+(\mu)$ and $Q_-(\mu)$ as found in Remark 2.7. Now it follows from theorems on exponential dichotomies (see Theorem 4 in [13], Proposition 2.3 in [18]) that $Q_+(\mu)$ and $Q_-(\mu)$ are C^1 in μ. This completes the proof when $J=\mathbf{R}$. Suppose now $J=\mathbf{R}_+$. Then we found in (7) that

$$\mathbf{R}^n=\operatorname{span}\big\{y_0(0,\mu)\big\}\oplus\mathcal{R}\big(Q_+(\mu)\big)\oplus\mathcal{R}\big(Q_-(\mu)\big) \tag{9}$$

and we let $\tilde{Q}_+(\mu)$ be the projection with range $\mathcal{R}(Q_+(\mu))$ and nullspace $\operatorname{span}\{y_0(0,\mu)\}\oplus\mathcal{R}(Q_-(\mu))$ and set $\tilde{Q}_0(\mu)=I-\tilde{Q}_+(\mu)-Q_-(\mu)$. Since the subspaces in the splitting (9) are C^1 in μ, it follows that the associated projections $\tilde{Q}_0(\mu)$, $\tilde{Q}_+(\mu)$ and $Q_-(\mu)$ are C^1 in μ also.

Next we develop a necessary and sufficient condition for a trichotomy in terms of exponential dichotomy of an associated difference equation.

PROPOSITION 2.9. *Suppose $A(t)$ is a continuous matrix function with $|A(t)| \leqslant M$ for all t in $\mathbf{R}$. Let $x_0(t)$ be a nonzero bounded solution of* (1) *with $0 < \Delta \leqslant |x_0(t)| \leqslant L < \infty$ for all t.*

Suppose $\{s_k\}_{k\in Z}$ is a sequence of times such that for all k, $0 < s_{\min} \leqslant s_{k+1} - s_k \leqslant s_{\max} < \infty$ and let $\{Q_k\}_{k\in Z}$ be a sequence of projections of rank $n-1$ *with $|Q_k|, |I - Q_k| \leqslant N$ such that the nullspace of Q_k is spanned by $x_0(s_k)$.*

Then (1) *has a trichotomy on* $\mathbf{R}$ *if and only if the difference equation*

$$u_{k+1} = Q_{k+1}X(s_{k+1})X^{-1}(s_k)u_k, \quad u_k \in \mathcal{R}(Q_k),\ k \in Z, \tag{10}$$

has an exponential dichotomy, where $X(t)$ is the fundamental matrix for (1) *with $X(0) = I$.*

PROOF. We write $A_k = X(s_{k+1})X^{-1}(s_k)$ and note that

$$Q_{k+1}A_kQ_k = Q_{k+1}A_k \tag{11}$$

since $A_kx_0(s_k) = x_0(s_{k+1})$.

Suppose first that (1) has a trichotomy and let $P_0(t)$, $P_+(t)$ and $P_-(t)$ be the associated projections, where $P_\sigma(t) = X(t)P_\sigma X^{-1}(t)$. Then, by Remark 2.2, the range of $P_0(t)$ is spanned by $x_0(t)$. Define the sequence

$$R_k = Q_kP_k^+,$$

where $P_k^+ = P_+(s_k)$. Note that $R_kP_k^+ = R_k$ and since $P_k^+x_0(s_k) = 0$ also $R_kQ_k = R_k$. Then

$$R_k^2 = R_kQ_kP_k^+ = R_kP_k^+ = R_k$$

so that R_k is a projection. Moreover, R_k is invariant with respect to (10) since

$$\begin{aligned} Q_{k+1}A_kR_k &= Q_{k+1}A_kQ_kP_k^+ \\ &= Q_{k+1}A_kP_k^+ \\ &= Q_{k+1}P_{k+1}^+A_k \\ &= R_{k+1}A_k \\ &= R_{k+1}Q_{k+1}A_k. \end{aligned}$$

Now the transition matrix for (10) is given by

$$\Phi(k,m) = \begin{cases} Q_kA_{k-1}\cdots A_m & \text{if } k > m, \\ I & \text{if } k = m, \\ Q_kA_k^{-1}\cdots A_{m-1}^{-1} & \text{if } k < m \end{cases}$$

since, for example, $\Phi(k,k-1)=Q_kA_{k-1}$ and if $k>m+1$, then

$$\Phi(k,m)=\Phi(k,k-1)\Phi(k-1,m)$$

so that if we assume $\Phi(k-1,m)=Q_{k-1}A_{k-2}\cdots A_m$ then, using (11),

$$\Phi(k,m)=Q_kA_{k-1}Q_{k-1}A_{k-2}\cdots A_m=Q_kA_{k-1}A_{k-2}\cdots A_m$$

so that the formula for $k>m$ follows by induction on k. Then if $k>m$

$$\begin{aligned}\left|\Phi(k,m)R_m\right|&=|Q_kA_{k-1}\cdots A_mQ_mP_m^+|\\&=\left|Q_kX(s_k)X^{-1}(s_m)P_+(s_m)\right|\\&\leqslant NK\mathrm{e}^{-\alpha(s_k-s_m)}\\&\leqslant NK\mathrm{e}^{-\alpha s_{\min}(k-m)}.\end{aligned}$$

Similarly, if $k<m$,

$$\left|\Phi(k,m)(I-R_m)\right|=\left|Q_kX(s_k)X^{-1}(s_m)P_-(s_m)\right|\leqslant NK\mathrm{e}^{-\alpha s_{\min}(m-k)}.$$

Finally, if $k=m$,

$$\left|\Phi(k,m)R_m\right|=|R_m|\leqslant NK.$$

Hence (10) has an exponential dichotomy with projections R_k and constants NK and $\mathrm{e}^{-\alpha s_{\min}}$.

Now, conversely, suppose (10) has an exponential dichotomy with projections $R_k:\mathcal{R}(Q_k)\mapsto\mathcal{R}(Q_k)$ and constants K, λ. Then a solution u_k of Eq. (10) with $u_0\in\mathcal{R}(R_0)$ satisfies

$$|u_k|\leqslant K\lambda^{k-m}|u_m|$$

for $m\leqslant k$. Define

$$\eta_k=-\sum_{\ell=k}^{\infty}\alpha_{\ell+1}(A_\ell u_\ell),$$

where the linear functional $\alpha_k:\mathbf{R}^n\mapsto\mathbf{R}$ is defined by

$$(I-Q_k)v=\alpha_k(v)x_0(s_k).$$

Note that

$$|\alpha_k|\leqslant N\Delta^{-1}$$

so that if $k \geqslant m$

$$|\eta_k| \leqslant \sum_{\ell=k}^{\infty} N\Delta^{-1}\mathrm{e}^{Ms_{\max}}|u_\ell| \leqslant \sum_{\ell=k}^{\infty} N\Delta^{-1}\mathrm{e}^{Ms_{\max}}K\lambda^{\ell-m}|u_m| \leqslant K_1\lambda^{k-m}|u_m|,$$

where

$$K_1 = NK\mathrm{e}^{Ms_{\max}}\Delta^{-1}(1-\lambda)^{-1}.$$

Then we define

$$v_k = \eta_k x_0(s_k) + u_k.$$

Note that for all k

$$\begin{aligned} v_{k+1} &= \eta_{k+1}x_0(s_{k+1}) + u_{k+1} \\ &= \big[\eta_k + \alpha_{k+1}(A_k u_k)\big]x_0(s_{k+1}) + Q_{k+1}A_k u_k \\ &= \eta_k x_0(s_{k+1}) + (I - Q_{k+1})A_k u_k + Q_{k+1}A_k u_k \\ &= \eta_k A_k x_0(s_k) + A_k u_k \\ &= X(s_{k+1})X^{-1}(s_k)v_k \end{aligned}$$

and that for $k \geqslant m$

$$|v_k| \leqslant K_1 L\lambda^{k-m}|u_m| + K\lambda^{k-m}|u_m| = (K_1 L + K)|u_m|\lambda^{k-m},$$

where we see that

$$|u_m| = |Q_m v_m| \leqslant N|v_m|$$

so that for $k \geqslant m$

$$|v_k| \leqslant K_2|v_m|\lambda^{k-m},$$

where

$$K_2 = (K_1 L + K)N.$$

Next note that

$$\eta_0 = -\sum_{k=0}^{\infty}\alpha_{k+1}(A_k u_k) = -\sum_{k=0}^{\infty}\alpha_{k+1}(A_k Q_k A_{k-1}\cdots A_0 u_0).$$

So what we have shown is that if we define the subspace

$$E^s = \{v_0 = \beta(u_0)x_0(s_0) + u_0\colon\ u_0 \in \mathcal{R}(R_0)\}$$

of $\mathbf{R}^n$, where

$$\beta(u_0) = -\sum_{k=0}^{\infty} \alpha_{k+1}(A_k Q_k A_{k-1} \cdots A_0 u_0),$$

then for $v_0 \in E^s$,

$$|X(s_k)X^{-1}(s_0)v_0| \leqslant K_2 \lambda^{k-m} |X(s_m)X^{-1}(s_0)v_0|$$

for $m \leqslant k$. Now if $s \leqslant t$, we can find $m \leqslant k$ such that

$$s_m \leqslant s < s_{m+1}, \qquad s_k \leqslant t < s_{k+1}.$$

Then

$$\begin{aligned}
|X(t)X^{-1}(s_0)v_0| &\leqslant \mathrm{e}^{Ms_{\max}} |X(s_k)X^{-1}(s_0)v_0| \\
&\leqslant \mathrm{e}^{Ms_{\max}} K_2 \lambda^{k-m} |X(s_m)X^{-1}(s_0)v_0| \\
&\leqslant \mathrm{e}^{2Ms_{\max}} K_2 \lambda^{k-m} |X(s)X^{-1}(s_0)v_0| \\
&\leqslant K_2 \mathrm{e}^{2Ms_{\max}} \lambda^{\frac{t-s}{s_{\max}}-1} |X(s)X^{-1}(s_0)v_0| \\
&= K_2 \mathrm{e}^{2Ms_{\max}} \lambda^{-1} \mathrm{e}^{\frac{\ln\lambda}{s_{\max}}(t-s)} |X(s)X^{-1}(s_0)v_0|.
\end{aligned}$$

So if $v_0 \in E^s$, then for $s \leqslant t$

$$|X(t)X^{-1}(s_0)v_0| \leqslant K_3 \mathrm{e}^{-\alpha(t-s)} |X(s)X^{-1}(s_0)v_0|, \tag{12}$$

where

$$K_3 = K_2 \mathrm{e}^{2Ms_{\max}}/\lambda, \qquad \alpha = -\ln\lambda/s_{\max}.$$

Similarly, there is a subspace

$$E^u = \{v_0 = \gamma(v_0)x_0(s_0) + u_0\colon\ u_0 \in \mathcal{N}(R_0)\}$$

of $\mathbf{R}^n$ such that if $v_0 \in E^u$,

$$|X(t)X^{-1}(s_0)v_0| \leqslant K_3 \mathrm{e}^{-\alpha(s-t)} |X(s)X^{-1}(s_0)v_0| \tag{13}$$

for $s \geqslant t$.

Next we show that $E^s \cap E^u = \{0\}$. Let v_0 be in this intersection. Then there exist $u_0 \in \mathcal{R}(R_0)$ and $\tilde{u}_0 \in \mathcal{N}(R_0)$ such that

$$v_0 = \beta(u_0)x_0(s_0) + u_0 = \gamma(\tilde{u}_0)x_0(s_0) + \tilde{u}_0.$$

Multiplying through by $I - Q_0$, we see that $\beta(u_0) = \gamma(\tilde{u}_0)$ which implies $u_0 = \tilde{u}_0$ and so both are zero. Thus $v_0 = 0$ and our assertion is proved.

Next note that $x_0(s_0)$ is not in $E^s \oplus E^u$, since if there exist $u_0 \in \mathcal{R}(R_0)$ and $\tilde{u}_0 \in \mathcal{N}(R_0)$ such that

$$x_0(s_0) = \beta(u_0)x_0(s_0) + u_0 + \gamma(\tilde{u}_0)x_0(s_0) + \tilde{u}_0,$$

then

$$x_0(s_0) = \bigl(\beta(u_0) + \gamma(\tilde{u}_0)\bigr)x_0(s_0) + u_0 + \tilde{u}_0.$$

Since $x_0(s_0)$ and $u_0 + \tilde{u}_0$ are linearly independent, it follows that $u_0 + \tilde{u}_0 = 0$ and hence $u_0 = \tilde{u}_0 = 0$ and so $x_0(s_0) = 0$, which is not true. So $x_0(s_0)$ is not in $E^s \oplus E^u$. Then since

$$\dim E^s + \dim E^u = \operatorname{rank} R_0 + (n - 1 - \operatorname{rank} R_0) = n - 1,$$

it follows that

$$\mathbf{R}^n = \operatorname{span}\bigl\{x_0(s_0)\bigr\} \oplus E^s \oplus E^u.$$

Let P_0, P_+, P_- be the projections associated with this splitting. Then it follows from (12) and (13) that for all ξ

$$\bigl|X(t)X^{-1}(s_0)P_+\xi\bigr| \leqslant K_3 e^{-\alpha(t-s)}\bigl|X(t)X^{-1}(s_0)P_+\xi\bigr| \quad (s \leqslant t)$$

and

$$\bigl|X(t)X^{-1}(t_0)P_-\xi\bigr| \leqslant K_3 e^{-\alpha(s-t)}\bigl|X(s)X^{-1}(s_0)P_+\xi\bigr| \quad (s \geqslant t).$$

Also since $|x_0(t)| \leqslant L\Delta^{-1}|x_0(s)|$ for all s, t, it follows for all t, s, ξ that

$$\bigl|X(t)X^{-1}(s_0)P_0\xi\bigr| \leqslant L\Delta^{-1}\bigl|X(s)X^{-1}(s_0)P_0\xi\bigr|.$$

Then, in view of Remark 2.5, it follows that (1) has a trichotomy on $\mathbf{R}$. This completes the proof of the proposition. □

3. Hyperbolic periodic orbits and their stable and unstable manifolds

In this section we study an autonomous system in $\mathbf{R}^n$

$$\dot{x} = f(x), \tag{14}$$

where f is C^1 and for all x

$$\left|f(x)\right| \leqslant M_0, \qquad \left|f'(x)\right| \leqslant M_1.$$

We denote by $\phi^t(x)$ the corresponding flow which, under these conditions, is defined for all t and x. Frequently we shall use the fact that

$$\left|D\phi^t(x)\right| \leqslant \mathrm{e}^{M_1|t|},$$

which follows from Gronwall's inequality, and also the fact that

$$D\phi^t(x)f(x) = f\bigl(\phi^t(x)\bigr).$$

We first define the Poincaré map associated with a periodic orbit of this system. Then we construct the local and global stable and unstable manifolds of a hyperbolic periodic orbit, thereby introducing the concept of asymptotic phase.

3.1. *Poincaré map*

First we define a more general Poincaré map. For this, we need the following lemma.

LEMMA 3.1. *Let $f(x)$ be a C^1 vectorfield in $\mathbf{R}^n$ with $|f(x)| \leqslant M_0$ and $|f'(x)| \leqslant M_1$. Let x be a point in $\mathbf{R}^n$, v a unit vector and T a real number such that $\langle f(y), v\rangle > 0$, where $y = \phi^T(x)$. Set*

$$\alpha = \frac{\langle f(y), v\rangle}{4M_0M_1}, \qquad \delta = \frac{\mathrm{e}^{-M_1T}\langle f(y), v\rangle}{4M_1}\min\left\{1, \frac{\langle f(y), v\rangle}{2M_0}\right\}.$$

Then if $|z - x| < \delta$, there exists a unique $\tau = \tau(z)$ such that

$$\bigl\langle \phi^\tau(z) - y, v\bigr\rangle = 0$$

and $|\tau - T| \leqslant \alpha$. Moreover,

$$\left|\tau(z) - T\right| \leqslant 2\bigl\langle f(y), v\bigr\rangle^{-1}\mathrm{e}^{M_1T}|z - x|, \tag{15}$$

$\tau(z)$ is a C^1 function with

$$\left|\tau'(z)\right| \leqslant 2\bigl\langle f(y), v\bigr\rangle^{-1}\mathrm{e}^{M_1(T+\alpha)}$$

and

$$\bigl\langle f\bigl(\phi^{\tau(z)}(z)\bigr), v\bigr\rangle \geqslant \frac{1}{2}\bigl\langle f(y), v\bigr\rangle.$$

PROOF. We need to solve the equation

$$g(\tau, z) = \langle \phi^\tau(z) - y, v \rangle = 0$$

for τ when $|z - x| < \delta$. Note that

$$g(T, x) = 0, \qquad g_\tau(T, x) = \langle f(y), v \rangle.$$

Then the equation holds if and only if

$$S\tau = \tau,$$

where

$$S\tau = T - g_\tau(T, x)^{-1}\eta(\tau, z),$$

with

$$\eta(\tau, z) = g(\tau, z) - g(T, x) - g_\tau(T, x)(\tau - T).$$

Note that if $|\tau_i - T| \leqslant \alpha$ and $|z - x| < \delta$, then

$$\begin{aligned} \bigl|\eta(\tau_1, z) - \eta(\tau_2, z)\bigr| &= \bigl|g(\tau_1, z) - g(\tau_2, z) - g_\tau(T, x)(\tau_1 - \tau_2)\bigr| \\ &= \left| \int_0^1 g_\tau\bigl(\theta\tau_1 + (1-\theta)\tau_2, z\bigr) - g_\tau(T, x)\, \mathrm{d}\theta (\tau_1 - \tau_2) \right|. \end{aligned}$$

Now if $|\tau - T| \leqslant \alpha$ and $|z - x| < \delta$,

$$\begin{aligned} \bigl|g_\tau(\tau, z) - g_\tau(T, x)\bigr| &\leqslant \bigl|f\bigl(\phi^\tau(z)\bigr) - f\bigl(\phi^T(x)\bigr)\bigr| \\ &\leqslant M_1\bigl|\phi^\tau(z) - \phi^T(x)\bigr| \\ &\leqslant M_1\bigl[\bigl|\phi^\tau(z) - \phi^T(z)\bigr| + \bigl|\phi^T(z) - \phi^T(x)\bigr|\bigr] \\ &\leqslant M_1\bigl[M_0|\tau - T| + \mathrm{e}^{M_1 T}|z - x|\bigr] \\ &\leqslant M_1\bigl[M_0\alpha + \mathrm{e}^{M_1 T}\delta\bigr]. \end{aligned} \tag{16}$$

It follows that if $|\tau_i - T| \leqslant \alpha$ and $|z - x| < \delta$, then

$$\bigl|\eta(\tau_1, z) - \eta(\tau_2, z)\bigr| \leqslant M_1[M_0\alpha + \mathrm{e}^{M_1 T}\delta]|\tau_1 - \tau_2|.$$

Next note that

$$\bigl|\eta(T, z)\bigr| = \bigl|g(T, z) - g(T, x)\bigr| \leqslant \bigl|\phi^T(z) - \phi^T(x)\bigr| \leqslant \mathrm{e}^{M_1 T}|z - x|.$$

Using these we see that if $|\tau - T| \leqslant \alpha$ and $|z - x| < \delta$, then

$$\begin{aligned}|S\tau - T| &\leqslant \langle f(y), v\rangle^{-1}\big[\big|\eta(\tau, z) - \eta(T, z)\big| + \big|\eta(T, z)\big|\big]\\ &\leqslant \langle f(y), v\rangle^{-1}\big[M_1[M_0\alpha + \mathrm{e}^{M_1 T}\delta]|\tau - T| + \mathrm{e}^{M_1 T}|z - x|\big]\end{aligned}$$

so that

$$|S\tau - T| \leqslant \frac{1}{2}|\tau - T| + \langle f(y), v\rangle^{-1}\mathrm{e}^{M_1 T}|z - x|. \tag{17}$$

In particular, this means that

$$|S\tau - T| \leqslant \frac{1}{2}\alpha + \langle f(y), v\rangle^{-1}\mathrm{e}^{M_1 T}\delta \leqslant \alpha.$$

Next, by similar reasoning, if $|\tau_i - T| \leqslant \alpha$ and $|z - x| < \delta$, then

$$|S\tau_1 - S\tau_2| \leqslant \langle f(y), v\rangle^{-1}M_1[M_0\alpha + \mathrm{e}^{M_1 T}\delta]|\tau_1 - \tau_2| \leqslant \frac{1}{2}|\tau_1 - \tau_2|.$$

It follows by the contraction mapping principle that if $|z - x| < \delta$, S has a unique fixed point τ in $|\tau - T| \leqslant \alpha$. The inequality (15) follows from (17) by putting $S\tau = \tau$. This completes the proof of the first part of the lemma.

That $\tau(x)$ is C^1 follows from standard results concerning smoothness of fixed point of a uniform contraction (see, for example, Lemma 3.8 below). Now

$$\tau'(z)h = -\frac{g_z(\tau(z), z)h}{g_\tau(\tau(z), z)} = -\frac{\langle D\phi^{\tau(z)}(z)h, v\rangle}{\langle f(\phi^{\tau(z)}(z)), v\rangle}. \tag{18}$$

To estimate this, we use (16) which implies that

$$\big|g_\tau\big(\tau(z), z\big) - g_\tau(T, x)\big| \leqslant M_1[M_0\alpha + \mathrm{e}^{M_1 T}\delta] \leqslant \frac{1}{2}\langle f(y), v\rangle.$$

Hence, since $g_\tau(T, x) = \langle f(y), v\rangle$,

$$g_\tau\big(\tau(z), z\big) \geqslant \frac{1}{2}\langle f(y), v\rangle.$$

Next

$$\big|\langle D\phi^{\tau(z)}(z)h, v\rangle\big| \leqslant \big|D\phi^{\tau(z)}(z)\big||h| \leqslant \mathrm{e}^{M_1\tau(z)}|h| \leqslant \mathrm{e}^{M_1(T+\alpha)}|h|.$$

Thus we obtain the inequality for $|\tau'(z)|$ and we have also shown in passing the last inequality in the statement of the lemma. Thus we complete the proof of the lemma. □

REMARK 3.2. We can write the τ found in the lemma as $\tau(z, x, v, T)$. We observe here that τ is a C^1 function of (z, x, v, T) on the open set of (z, x, v, T) for which $|z-x| < \delta = \delta(x, v, T)$. This follows from the smoothness of the fixed point of a uniform contraction (see, for example, Lemma 3.8 below).

We use this lemma to define a general Poincaré map. Consider a system (14) with a nonconstant solution $u(t)$. Let $t_1 < t_2$ be two times and let v_i $(i = 1, 2)$ be a unit vector such that $\langle f(u(t_i)), v_i\rangle > 0$. Let $W_i = \{v\colon \langle v, v_i\rangle = 0\}$ and set $C_i = u(t_i) + W_i$. Also set

$$\alpha = \bigl\langle f\bigl(u(t_2)\bigr), v_2\bigr\rangle / 4M_0M_1$$

and

$$\delta = \min\left\{\frac{\mathrm{e}^{-M_2(t_2-t_1)}\langle f(u(t_2)), v_2\rangle}{4M_1}\min\left\{1, \frac{\langle f(u(t_2)), v_2\rangle}{2M_0}\right\}, \frac{\langle f(u(t_1)), v_1\rangle}{M_1}\right\}.$$

Then, by Lemma 3.1, if $x \in C_1$ with $|x - u(t_1)| < \delta$, there exists a unique $\tau = \tau(x)$ in $|\tau - (t_2 - t_1)| < \alpha$ such that $\phi^\tau(x) \in C_2$. Moreover, $\tau(x)$ is a C^1 function, $\tau(u(t_1)) = t_2 - t_1$ and

$$\bigl|\tau(x) - (t_2 - t_1)\bigr| \leqslant 2\bigl\langle f\bigl(u(t_2)\bigr), v_2\bigr\rangle^{-1}\mathrm{e}^{M_1(t_2-t_1)}\bigl|x - u(t_1)\bigr|.$$

Then for $x \in C_1$ with $|x - u(t_1)| < \delta$ we define the Poincaré map

$$P(x) = \phi^{\tau(x)}(x).$$

Note that

$$P\bigl(u(t_1)\bigr) = u(t_2).$$

Also, as in (18), for $x \in C_1$ and $\xi \in W_1$,

$$\tau'(x)\xi = -\frac{\langle D\phi^{\tau(x)}(x)\xi, v_2\rangle}{\langle f(\phi^{\tau(x)}(x)), v_2\rangle}$$

so that

$$DP(x)\xi = f\bigl(P(x)\bigr)\tau'(x)\xi + D\phi^{\tau(x)}(x)\xi = Q_x D\phi^{\tau(x)}(x)\xi, \tag{19}$$

where Q_x is the projection on to W_2 along $f(P(x))$. Note that if $DP(x)\xi = 0$, then $D\phi^{\tau(x)}(x)\xi = \lambda f(P(x))$, where $\lambda = -\tau'(x)\xi$, and so

$$D\phi^{\tau(x)}(x)\bigl(-\lambda f(x) + \xi\bigr) = 0$$

from which it follows that $-\lambda f(x) + \xi = 0$. Now since $|x - u(t_1)| < \delta$,

$$\bigl\langle f(x), v_1\bigr\rangle > \bigl\langle f\bigl(u(t_1)\bigr), v_1\bigr\rangle - M_1\delta \geqslant 0$$

and so we must have $\lambda = 0$ and $\xi = 0$. So $DP(x)$ is injective and hence invertible. It follows that P is a diffeomorphism on $x \in C_1$, $|x - u(t_1)| < \delta$.

We use the notation $P(u(t_1), t_1, t_2, v_1, v_2)$ for the Poincaré map just defined.

DEFINITION 3.3. We say a T-periodic orbit $u_0(t)$ of (14) is *hyperbolic* if all but one of its Floquet multipliers lie off the unit circle.

PROPOSITION 3.4. *A T-periodic orbit $u_0(t)$ of* (14) *is hyperbolic if and only if $u_0(0)$ is hyperbolic as a fixed point of the Poincaré map*

$$P\big(u_0(0), 0, T, v, v\big), \quad v = \frac{u_0'(0)}{|u_0'(0)|}.$$

PROOF. Write $P = P(u_0(0), 0, T, v, v)$. First we note that

$$P\big(u_0(0)\big) = u_0(T) = u_0(0).$$

Next we know from (19) that

$$DP\big(u_0(0)\big) = QD\phi^T\big(u_0(0)\big),$$

where Q is the orthogonal projection onto the orthogonal complement W (the domain of $DP(u_0(0))$) of $u_0'(0) = f(u_0(0))$. Also we know that

$$D\phi^T\big(u_0(0)\big) f\big(u_0(0)\big) = f\big(u_0(0)\big).$$

We choose a basis of $\mathbf{R}^n$ with $f(u_0(0))$ as the first basis vector and with the other basis vectors in W. Then the matrix of $D\phi^T(u_0(0))$ with respect to this basis has the partitioned form

$$\begin{bmatrix} 1 & b \\ 0 & B \end{bmatrix},$$

where B is the matrix of $DP(u_0(0))$. Then the proposition follows immediately since the characteristic polynomial is $(1 - \lambda)\det(B - \lambda I)$. □

3.2. *Local stable and unstable manifolds*

In this subsection we study the local stable and unstable manifolds of a hyperbolic periodic orbit. For later use, we consider a system

$$\dot{x} = f(x, \mu) \tag{20}$$

depending on a parameter μ. We assume that when $\mu = 0$ the system has a hyperbolic T-periodic orbit $u_0(t)$. By standard theorems (see [10]) the perturbed system (20) has a nearby hyperbolic periodic orbit $u(t,\mu)$ such that $u(t,0) = u_0(t)$. This orbit will have period $T(\mu)$ with $T(0) = T$. Note that if $f(x,\mu)$ is C^r, then $u(t,\mu)$ and $T(\mu)$ are C^r also.

In the following proposition, we find the local stable manifold near $u(0,\mu)$ with asymptotic phase 0, that is, we find the solutions $x(t)$ with $x(0)$ near $u(0,\mu)$ such that $|x(t) - u(t,\mu)| \to 0$ as $t \to \infty$.

PROPOSITION 3.5. *Let the function $f(x,\mu)$ be C^{r+2} $(r \geqslant 0)$ and suppose*

$$\dot{x} = f(x,0)$$

has a hyperbolic T-periodic solution $u_0(t)$. Then

(i) *the system*

$$\dot{x} = f_x\bigl(u_0(t),0\bigr)x$$

has a trichotomy with projections P_0, P_+ and P_- and constants K, α;

(ii) *the perturbed system* (20) *has a hyperbolic $T(\mu)$-periodic solution $u(t,\mu)$ where $u(t,\mu)$ and $T(\mu)$ are C^{r+2} functions with $u(t,0) = u_0(t)$ and $T(0) = T$;*

(iii) *suppose $0 < \gamma < \alpha$. For $\Delta > 0$ sufficiently small, there are positive numbers μ_0, ξ_0 depending on Δ such that if $|\mu| < \mu_0$ and $\xi \in \mathcal{R}(P_+)$ satisfies $|\xi| < \xi_0$, there exists a unique solution $x(t) = x^+(t,\xi,\mu)$ of* (20) *such that*

$$\bigl|x(t) - u(t,\mu)\bigr|\mathrm{e}^{\gamma t} \leqslant \Delta \quad (t \geqslant 0), \qquad P_+\bigl[x(0) - u(0,\mu)\bigr] = \xi.$$

Moreover,

$$x^+(t,0,\mu) = u(t,\mu),$$

$x^+(t,\xi,\mu)$ is C^r in its arguments and $x^+_\xi(t,\xi,\mu)$ is the unique solution restricted to $\mathcal{R}(P_+)$ of

$$\dot{Z} = f_x\bigl(x^+(t,\xi,\mu),\mu\bigr)Z$$

such that $\sup_{t\geqslant 0} \mathrm{e}^{\gamma t}|Z(t)| < \infty$ and $P_+Z(0) = P_+$ with, in particular,

$$x^+_\xi(t,0,0) = X(t)P_+,$$

and there is a constant N such that when $r \geqslant 1$,

$$\bigl|x^+_\xi(t,\xi,\mu)\bigr| \leqslant N\mathrm{e}^{-\gamma t}, \quad \bigl|x^+_\mu(t,\xi,\mu) - u_\mu(t,\mu)\bigr| \leqslant N\mathrm{e}^{-\gamma t} \quad (t \geqslant 0)$$

and when $r \geqslant 2$,

$$\bigl|x^+_{\xi\xi}(t,\xi,\mu)\bigr| \leqslant N\mathrm{e}^{-\gamma t} \quad (t \geqslant 0).$$

PROOF. For the linear system

$$\dot{x} = f_x\bigl(u_0(t), 0\bigr)x, \tag{21}$$

1 is a Floquet multiplier of algebraic multiplicity one and the others lie off the unit circle. Now by Floquet theory there is a linear invertible periodic transformation $x = S(t)y$ taking (21) into an autonomous system $\dot{y} = Ay$ with A having zero as an eigenvalue of multiplicity one and the other eigenvalues having nonzero real parts. We define P_0, P_+, P_- to be the supplementary projections corresponding to the splitting of $\mathbf{R}^n$ into the generalized eigenspaces corresponding respectively to the zero eigenvalue, the eigenvalues with negative real parts and the eigenvalues with positive real parts. Note that the projections P_σ commute with A and there exist positive constants L, α (note that α can be taken as any positive number such that $|\mathrm{Re}(\lambda)| > \alpha$ for all eigenvalues λ of A with nonzero real part) such that

$$\begin{aligned}
&|\mathrm{e}^{tA}P_0| \leqslant L \quad \text{for all } t,\\
&|\mathrm{e}^{tA}P_+| \leqslant L\mathrm{e}^{-\alpha t} \quad \text{for } t \geqslant 0,\\
&|\mathrm{e}^{tA}P_-| \leqslant L\mathrm{e}^{\alpha t} \quad \text{for } t \leqslant 0.
\end{aligned}$$

Then $X(t) = S(t)\mathrm{e}^{tA}$ is a fundamental matrix for (21) and, since $X(t)P_\sigma X^{-1}(s) = S(t)\mathrm{e}^{(t-s)A}P_\sigma S^{-1}(s)$, we obtain the inequalities

$$\begin{aligned}
&\bigl|X(t)P_0X^{-1}(s)\bigr| \leqslant K \quad \text{for all } t,\\
&\bigl|X(t)P_+X^{-1}(s)\bigr| \leqslant K\mathrm{e}^{-\alpha(t-s)} \quad \text{for } t \geqslant s,\\
&\bigl|X(t)P_-X^{-1}(s)\bigr| \leqslant K\mathrm{e}^{-\alpha(s-t)} \quad \text{for } t \leqslant s,
\end{aligned}$$

where $K = LM^2$ with M an upper bound for $|S(t)|$ and $|S^{-1}(t)|$. Now if we replace P_σ by $X(0)P_\sigma X^{-1}(0)$, it follows that (21) has a trichotomy on $\mathbf{R}$ with these projections, which we rename as P_σ, and constants K and α. So (i) is proved.

As noted before the proposition, (ii) follows from standard theorems. Now if we change the time scale $\tau = \omega(\mu)t$, where $\omega(\mu) = T/T(\mu)$, then the new system

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \omega(\mu)^{-1}f(x,\mu)$$

has the periodic orbit $u(\omega(\mu)^{-1}\tau, \mu)$ which has period T for all μ. So we can assume without loss of generality that in the original system the period of $u(t,\mu)$ is T, independent of μ. The advantage of this is that now $\sup_{t\in\mathbf{R}}|u(t,\mu) - u_0(t)| \to 0$ as $\mu \to 0$.

To prove (iii), we look for solutions $x(t)$ of (20) such that $|x(t) - u(t,\mu)| \to 0$ at an exponential rate as $t \to \infty$. We write

$$x = u(t,\mu) + z.$$

Then

$$\dot{z} = f_x\bigl(u_0(t), 0\bigr)z + h(t, z, \mu), \tag{22}$$

where

$$\begin{aligned} h(t, z, \mu) = {} & \bigl[f_x\bigl(u(t, \mu), \mu\bigr) - f_x\bigl(u_0(t), 0\bigr)\bigr]z \\ & + f\bigl(u(t, \mu) + z, \mu\bigr) - f\bigl(u(t, \mu), \mu\bigr) - f_x\bigl(u(t, \mu), \mu\bigr)z. \end{aligned}$$

Note that $h(t, 0, \mu) = 0$ and

$$\bigl|h_z(t, z, \mu)\bigr| = \bigl|f_x\bigl(u(t, \mu) + z, \mu\bigr) - f_x\bigl(u_0(t), 0\bigr)\bigr| \leqslant \delta\bigl(|z|, |\mu|\bigr),$$

where $\delta(|z|, |\mu|)$ is a continuous function nondecreasing in both arguments which $\to 0$ as $(|z|, |\mu|) \to (0, 0)$.

Given small $\xi \in \mathcal{R}(P_+)$, we look for solutions $z(t)$ of (22) such that

$$P_+z(0) = \xi, \qquad \bigl|z(t)\bigr| \leqslant \Delta e^{-\gamma t} \quad (t \geqslant 0),$$

where Δ is chosen suitably small. First we prove the following lemma.

LEMMA 3.6. *Suppose the linear system*

$$\dot{x} = A(t)x \tag{23}$$

has a trichotomy on $[0, \infty)$ *with projections* P_0, P_+, P_- *and constants* K, α *and let* $0 < \gamma < \alpha$. *Suppose* $\xi \in \mathcal{R}(P_+)$ *and* $h(t)$ *is a continuous function with* $\|h\|_\gamma = \sup_{t \geqslant 0} |h(t)|e^{\gamma t} < \infty$. *Then*

$$x(t) = X(t)\xi + \int_0^t X(t)P_+X^{-1}(s)h(s)\,ds - \int_t^\infty X(t)(I - P_+)X^{-1}(s)h(s)\,ds$$

is the unique solution of

$$\dot{x} = A(t)x + h(t) \tag{24}$$

such that $\|x\|_\gamma < \infty$ *and* $P_+x(0) = \xi$. *Moreover*

$$\|x\|_\gamma \leqslant K|\xi| + K\bigl[(\alpha - \gamma)^{-1} + 2\gamma^{-1}\bigr]\|h\|_\gamma.$$

PROOF. To prove the uniqueness first, note that the difference $x(t)$ between any two such solutions would be a solution of (23) such that $x(t) \to 0$ as $t \to \infty$ and $P_+x(0) = 0$. Now from Remark 2.2 it follows that $x(0) \in \mathcal{R}(P_+)$. So $x(0) = P_+x(0) = 0$. This gives the uniqueness.

To prove the existence, note that

$$\begin{aligned}
&\int_0^t \bigl|X(t)P_+X^{-1}(s)h(s)\bigr|\,\mathrm{d}s + \int_t^\infty \bigl|X(t)(I-P_+)X^{-1}(s)h(s)\bigr|\,\mathrm{d}s\\
&\quad\leqslant \int_0^t K\mathrm{e}^{-\alpha(t-s)}\|h\|_\gamma \mathrm{e}^{-\gamma s}\,\mathrm{d}s + \int_t^\infty 2K\|h\|_\gamma \mathrm{e}^{-\gamma s}\,\mathrm{d}s\\
&\quad= K\bigl[(\alpha-\gamma)^{-1} + 2\gamma^{-1}\bigr]\|h\|_\gamma \mathrm{e}^{-\gamma t}.
\end{aligned}$$

Hence the integrals defining $x(t)$ converge and we find by differentiation that $x(t)$ is a solution of (24). Moreover, using the fact that $|X(t)P_+| \leqslant K\mathrm{e}^{-\gamma t}$ and the previous estimate, we find that $\|x\|_\gamma < \infty$ and that the inequality given in the statement of the lemma holds. □

From this lemma, we deduce another.

LEMMA 3.7. *Suppose the linear system* (23) *satisfies the conditions of Lemma* 3.6 *and let* $0 < \gamma < \alpha$. *Let* $h(t, x, \mu)$ *be a continuous function,* T*-periodic in* t*, satisfying*

$$h(t,0,\mu) = 0, \qquad \bigl|h(t,x_1,\mu) - h(t,x_2,\mu)\bigr| \leqslant L|x_1 - x_2|$$

for all $t \geqslant 0$, $|x_1| \leqslant \Delta$, $|x_2| \leqslant \Delta$, $|\mu| < \mu_0$. *Suppose that*

$$K(\gamma)L = K\bigl[(\alpha-\gamma)^{-1} + 2\gamma^{-1}\bigr]L < 1.$$

Then if $|\mu| < \mu_0$ *and* $\xi \in \mathcal{R}(P_+)$ *with*

$$K|\xi| < \bigl(1 - K(\gamma)L\bigr)\Delta, \tag{25}$$

the equation

$$\dot{x} = A(t)x + h(t,x,\mu)$$

has a unique solution $x(t) = x(t,\xi,\mu)$ *such that* $\|x\|_\gamma \leqslant \Delta$ *and* $P_+x(0) = \xi$. *Moreover*

$$x(t,0,\mu) = 0$$

and

$$\|x\|_\gamma \leqslant K\bigl(1 - K(\gamma)L\bigr)^{-1}|\xi|. \tag{26}$$

Also, if $h(t,x,\mu)$ *is* C^{r+2} $(r \geqslant 0)$ *in* (x,μ) *with all derivatives continuous in* (t,x,μ), *then* $x(t,\xi,\mu)$ *is* C^r *in* (x,μ) *and is continuous in* (t,x,μ) *together with all its derivatives. Moreover,* $x_\xi(t,\xi,\mu)$ *is the unique solution restricted to* $\mathcal{R}(P_+)$ *of*

$$\dot{Z} = \bigl[A(t) + h_x\bigl(t, x(t,\xi,\mu),\mu\bigr)\bigr]Z$$

such that $\|Z\|_\gamma < \infty$ *and* $P_+Z(0) = P_+$,

$$x_\xi(t, 0, 0) = X(t)P_+$$

and there exists a constant N *such that for* $t \geqslant 0$

$$\left|x_\xi(t, \xi, \mu)\right| \leqslant N\mathrm{e}^{-\gamma t}, \quad \left|x_\mu(t, \xi, \mu)\right| \leqslant N\mathrm{e}^{-\gamma t}, \quad \left|x_{\xi\xi}(t, \xi, \mu)\right| \leqslant N\mathrm{e}^{-\gamma t}$$

when these derivatives exist.

PROOF. Consider the Banach space E of continuous functions $x(t)$ on $[0, \infty)$ with $\|x\|_\gamma < \infty$, with $\|x\|_\gamma$ as norm. Let B be the ball of radius Δ and centre 0. Then B is a complete metric space. Next suppose $|\mu| < \mu_0$ and $\xi \in \mathcal{R}(P_+)$ satisfies (25). Then we define a mapping $T : B \mapsto B$ as follows: if $x \in B$, we let $T(x)(t)$ be the unique solution of

$$\dot{x} = A(t)x + h\bigl(t, x(t), \mu\bigr)$$

such that $P_+T(x)(0) = \xi$ and $\|T(x)\|_\gamma < \infty$. This exists by Lemma 3.6 and it also follows from that lemma that

$$\bigl\|T(x)\bigr\|_\gamma \leqslant K|\xi| + K(\gamma)L\|x\|_\gamma \leqslant K|\xi| + K(\gamma)L\Delta \leqslant \Delta. \tag{27}$$

Next if x_1 and x_2 are in B, it follows by similar reasoning that

$$\bigl\|T(x_1) - T(x_2)\bigr\|_\gamma \leqslant K(\gamma)L\|x_1 - x_2\|_\gamma. \tag{28}$$

Since $K(\gamma)L < 1$, T is a contraction and so has a unique fixed point $x(t)$. That $x(t) = 0$ when $\xi = 0$ follows by uniqueness and the inequality (26) follows from (27) by putting $Tx = x$.

To prove the smoothness, we write $T(x)$ as $T(x, \xi, \mu)$ to indicate the dependence on ξ and μ. Note that if $x \in B$

$$\begin{aligned} T(x, \xi, \mu)(t) = X(t)\xi &+ \int_0^t X(t)P_+X^{-1}(s)h\bigl(s, x(s), \mu\bigr)\,\mathrm{d}s \\ &- \int_t^\infty X(t)(I - P_+)X^{-1}(s)h\bigl(s, x(s), \mu\bigr)\,\mathrm{d}s. \end{aligned} \tag{29}$$

It is not hard to see that for $1 \leqslant m + p + q \leqslant r$, $p \geqslant 1$, the partial derivatives are given by the formulae

$$\bigl(T_\xi(x, \xi, \mu)\xi_1\bigr)(t) = X(t)\xi_1, \quad p = 1,\ m + q = 0,$$

$$T_{x^m\xi^p\mu^q}(x, \xi, \mu) = 0, \quad p = 1,\ m + q \geqslant 1 \text{ or } p \geqslant 2$$

and for $p = 0$, $1 \leqslant m + q \leqslant r$ by

$$\begin{aligned}&\big(T_{x^m\mu^q}(x,\xi,\mu)u_1\cdots u_m\big)(t)\\&\quad= \int_0^t X(t)P_+X^{-1}(s)h_{x^m\mu^q}\big(s,x(s),\mu\big)u_1(s)\cdots u_m(s)\,\mathrm{d}s\\&\qquad - \int_t^\infty X(t)(I-P_+)X^{-1}(s)h_{x^m\mu^q}\big(s,x(s),\mu\big)u_1(s)\cdots u_m(s)\,\mathrm{d}s,\end{aligned}$$

where we use the notation $T_{x^m\xi^p\mu^q}$ for $\frac{\partial^{m+p+q}T}{\partial x^m\partial\xi^p\partial\mu^q}$, $\xi_1 \in \mathcal{R}(P_+)$ and $u_1, \ldots, u_m$ are in E. (Note when $m = 0$ and $q = r$, we use the facts that $h(t, 0, \mu) = 0$ and the boundedness of the derivatives of h to verify that this formula yields the derivative.) These derivatives are also continuous functions of (x, ξ, μ) for $x \in B$, $\xi \in \mathcal{R}(P_+)$ satisfying (25) and μ in $|\mu| < \mu_0$. Thus $(x, \xi, \mu) \mapsto T(x, \xi, \mu)$ is a C^r map. Now we use the following well-known lemma:

LEMMA 3.8. *Let B be a closed ball in a Banach space E and let V be an open set in another Banach space Λ. Suppose $T : B \times V \mapsto B$ is a mapping such that for all x_1, $x_2 \in B$ and $\lambda \in V$*

$$\big\|T(x_1,\lambda) - T(x_2,\lambda)\big\| \leqslant \sigma\|x_1 - x_2\|$$

where $0 \leqslant \sigma < 1$. Then for each $\lambda \in V$ the equation

$$T(x,\lambda) = x$$

has a unique solution $x = x(\lambda)$ in B. Moreover, if T is C^r $(0 \leqslant r \leqslant \infty)$, then so also is $x(\lambda)$, and the derivative $Dx(\lambda)$ satisfies the equation

$$Dx(\lambda) = \frac{\partial T}{\partial x}\big(x(\lambda),\lambda\big)Dx(\lambda) + \frac{\partial T}{\partial\lambda}\big(x(\lambda),\lambda\big).$$

(*Note*: *T is C^r means that T is defined and C^r in an open set containing $B \times V$.*)

We apply this lemma with V as the set of $(\xi, \mu) \in \Lambda = \mathcal{R}(P_+) \times \mathbf{R}$ satisfying (25) and $|\mu| < \mu_0$ and with $T(x, \xi, \mu)$ as defined in (29). In view of (28), we can take $\sigma = K(\gamma)L$. This means $x(\xi, \mu)$ defined by $x(\xi, \mu)(t) = x(t, \xi, \mu)$, which maps the (ξ, μ) space into the Banach space E is C^r.

In particular, since $x(\xi, \mu)$ is continuous as a function into E, it follows that $x(t, \xi, \mu)$ is continuous in (ξ, μ) uniformly with respect to t. Then since $x(t, \xi, \mu)$ is continuous in t for each fixed (ξ, μ), it follows that $x(t, \xi, \mu)$ is continuous in (t, ξ, μ).

Next, if $r \geqslant 1$, we know that

$$\sup_{t\geqslant 0} \mathrm{e}^{\gamma t}\big|x(t,\xi+\xi_1,\mu) - x(t,\xi,\mu) - \big(x_\xi(\xi,\mu)\xi_1\big)(t)\big| = \mathrm{o}\big(|\xi_1|\big)$$

as $\xi_1 \to 0$, where as follows from Lemma 3.8,

$$x_\xi(\xi,\mu) = T_x\big(x(\xi,\mu),\xi,\mu\big)x_\xi(\xi,\mu) + T_\xi\big(x(\xi,\mu),\xi,\mu\big)$$

so that, if we write $(x_\xi(\xi,\mu)\xi_1)(t) = Z(t)\xi_1$, we see that

$$\begin{aligned} Z(t)\xi_1 = X(t)\xi_1 &+ \int_0^t X(t)P_+X^{-1}(s)h_x\big(s,x(s,\xi,\mu),\mu\big)Z(s)\xi_1\,\mathrm{d}s \\ &- \int_t^\infty X(t)(I-P_+)X^{-1}(s)h_x\big(s,x(s,\xi,\mu),\mu\big)Z(s)\xi_1\,\mathrm{d}s. \end{aligned}$$

This means that $x_\xi(t,\xi,\mu)$ exists and equals $Z(t)$ restricted to $\mathcal{R}(P_+)$, where $Z(t)$ is the unique solution of

$$\dot{Z} = \big[A(t) + h_x\big(t,x(t,\xi,\mu),\mu\big)\big]Z$$

such that $\|Z\|_\gamma < \infty$ and $P_+Z(0) = P_+$. Moreover, since $x_\xi(\xi,\mu)$ is a continuous function of (ξ,μ), it follows that $x_\xi(t,\xi,\mu)$ is a continuous function of (ξ,μ), uniformly with respect to t. Since $x_\xi(t,\xi,\mu)$ is continuous in t for each fixed (ξ,μ), it follows that $x_\xi(t,\xi,\mu)$ is continuous in (t,ξ,μ). Furthermore from the equation for $Z(t)$ above and Lemma 3.6, we have the estimate

$$\|Z\|_\gamma \leqslant K + K(\gamma)L\|Z\|_\gamma$$

so that for $t \geqslant 0$

$$\big|x_\xi(t,\xi,\mu)\big| \leqslant \big(1-K(\gamma)L\big)^{-1}K\mathrm{e}^{-\gamma t}.$$

Also, since $h(t,0,0) = 0$, we see that

$$x_\xi(t,0,0) = X(t)P_+.$$

Similarly, we show that $x_\mu(t,\xi,\mu)$ is the unique solution of

$$\dot{x} = \big[A(t) + h_x\big(t,x(t,\xi,\mu),\mu\big)\big]x + h_\mu\big(t,x(t,\xi,\mu),\mu\big)$$

such that $\|x\|_\gamma < \infty$ and $P_+x(0) = 0$ and, if $r \geqslant 2$, that $x_{\xi\xi}(t,\xi,\mu)$ is the unique solution restricted to $\mathcal{R}(P_+) \times \mathcal{R}(P_+)$ of

$$\dot{Z} = \big[A(t) + h_x\big(t,x(t,\xi,\mu),\mu\big)\big]Z + h_{xx}\big(t,x(t,\xi,\mu),\mu\big)x_\xi(t,\xi,\mu)x_\xi(t,\xi,\mu)$$

such that $\|Z\|_\gamma < \infty$ and $P_+Z(0) = 0$. Also we get the required estimates from Lemma 3.6. Thus we complete the proof of Lemma 3.7. □

We apply Lemma 3.7 to Eq. (22). Suppose Δ satisfies

$$K(\gamma)\delta(\Delta,0) = K\big[(\alpha-\gamma)^{-1} + 2\gamma^{-1}\big]\delta(\Delta,0) < 1.$$

Then we take μ_0 as the smallest μ such that

$$K(\gamma)\delta(\Delta,\mu) = 1$$

and set

$$\xi_0 = K^{-1}\big(1 - K(\gamma)L\big)\Delta.$$

Then, using Lemma 3.7, if $|\mu| < \mu_0$ and if $\xi \in \mathcal{R}(P_+)$ with $|\xi| < \xi_0$, we get a unique solution $z(t) = z(t,\xi,\mu)$ of (22) with $P_+z(0) = \xi$ and $\|z\|_\gamma \leqslant \Delta$. Then we define

$$x^+(t,\xi,\mu) = u(t,\mu) + z(t,\xi,\mu)$$

and we see that (iii) and hence Proposition 3.5 has been proved. □

Similarly we find the local unstable manifold near $u(0,\mu)$ with asymptotic phase 0, that is, we find the solutions $x(t)$ with $x(0)$ near $u(0,\mu)$ such that $|x(t) - u(t,\mu)| \to 0$ as $t \to -\infty$. We state the proposition without proof as it is similar to that of Proposition 3.5.

PROPOSITION 3.9. *As in Proposition* 3.5, *let the function* $f(x,\mu)$ *be* C^{r+2} $(r \geqslant 0)$ *and suppose*

$$\dot{x} = f(x,0)$$

has a hyperbolic T*-periodic solution* $u_0(t)$. *Then* (i) *and* (ii) *hold as in Proposition* 3.5. *Next suppose* $0 < \gamma < \alpha$. *Then for* $\Delta > 0$ *sufficiently small, there are positive numbers* μ_0, η_0 *depending on* Δ *such that if* $|\mu| < \mu_0$ *and* $\eta \in \mathcal{R}(P_-)$ *satisfies* $|\eta| < \eta_0$, *there exists a unique solution* $x(t) = x^-(t,\eta,\mu)$ *of* (20) *such that*

$$\big|x(t) - u(t,\mu)\big|e^{-\gamma t} \leqslant \Delta \quad (t \leqslant 0), \qquad P_-\big[x(0) - u(0,\mu)\big] = \eta.$$

Moreover,

$$x^-(t,0,\mu) = u(t,\mu),$$

$x^-(t,\xi,\mu)$ *is* C^r *in its arguments with* $x^-_\eta(t,\eta,\mu)$ *as the unique solution restricted to* $\mathcal{R}(P_-)$ *of*

$$\dot{Z} = f_x\big(x^+(t,\eta,\mu),\mu\big)Z$$

such that $\sup_{t\leqslant 0} e^{-\gamma t}|Z(t)| < \infty$ *and* $P_-Z(0) = P_-$ *with, in particular,*

$$x^-_\eta(t,0,0) = X(t)P_-$$

and there is a constant N such that when $r \geqslant 1$,

$$\left|x_\eta^-(t,\eta,\mu)\right| \leqslant N\mathrm{e}^{\gamma t}, \quad \left|x_\mu^-(t,\xi,\mu) - u_\mu(t,\mu)\right| \leqslant N\mathrm{e}^{\gamma t} \quad (t \leqslant 0)$$

and when $r \geqslant 2$,

$$\left|x_{\eta\eta}^-(t,\eta,\mu)\right| \leqslant N\mathrm{e}^{\gamma t} \quad (t \leqslant 0).$$

REMARK 3.10. It is clear from the proof that when $f(x,\mu)$ does not depend on μ the relevant conclusions of Propositions 3.5 and 3.9 hold under the weaker assumption that $f(x)$ is C^r $(r \geqslant 1)$.

3.3. *Global stable and unstable manifolds and asymptotic phase*

Consider now the system

$$\dot{x} = f(x), \tag{30}$$

with flow ϕ^t, where f is C^r $(r \geqslant 1)$ with $|f(x)| \leqslant M_0$ and $|f'(x)| \leqslant M_1$ for all $x \in \mathbf{R}^n$. We suppose the system has a hyperbolic periodic orbit $u_0(t)$ with minimal period T.

From Proposition 3.5(i), we know that

$$\dot{x} = f'\big(u_0(t)\big)x$$

has a trichotomy with projections P_0, P_+, P_- and constants K, α, where α can be any positive number such that $|\lambda| > \mathrm{e}^{\alpha T}$ for all Floquet multipliers λ with $|\lambda| > 1$ and $|\lambda| < \mathrm{e}^{-\alpha T}$ for those with $|\lambda| < 1$.

Also, given γ such that $0 < \gamma < \alpha$, there exist $\Delta > 0$ and ξ_0 such that if $\xi \in \mathcal{R}(P_+)$ with $|\xi| < \xi_0$ then (30) has a unique solution $x(t) = x^+(t,\xi)$ such that

$$\left|x(t) - u_0(t)\right| \leqslant \Delta\mathrm{e}^{-\gamma t} \quad \text{for } t \geqslant 0, \qquad P_+\big[x(0) - u_0(0)\big] = \xi.$$

Moreover, in view of Remark 3.10, $x_\xi^+(t,\xi)$ exists, is continuous and is the unique solution restricted to $\mathcal{R}(P_+)$ of

$$\dot{Z} = f'\big(x^+(t,\xi)\big)Z \tag{31}$$

such that $\|Z\|_\gamma < \infty$ and $P_+Z(0) = P_+$.

Similarly, if $\eta \in \mathcal{R}(P_-)$ with $|\eta| < \eta_0$ then (30) has a unique solution $x(t) = x^-(t,\eta)$ such that

$$\left|x(t) - u_0(t)\right| \leqslant \Delta\mathrm{e}^{\gamma t} \quad \text{for } t \leqslant 0, \qquad P_-\big[x(0) - u_0(0)\big] = \eta.$$

Moreover $x_\eta^-(t,\eta)$ exists, is continuous and is the unique solution restricted to $\mathcal{R}(P_-)$ of

$$\dot{Z} = f'\bigl(x^-(t,\eta)\bigr)Z$$

such that $\sup_{t\leqslant 0}|Z(t)|e^{-\gamma t} < \infty$ and $P_-Z(0) = P_-$.

Note that $x^+(t,\xi)$ (resp. $x_-(t,\eta)$) can be continued backwards (resp. forwards) so as to be defined for all t.

Now we define the global stable and unstable manifolds.

DEFINITION 3.11. The *global stable manifold* is defined as

$$W^s(u_0) = \bigl\{x \in \mathbf{R}^n\colon\ \operatorname{dist}\bigl(\phi^t(x), u_0(\cdot)\bigr) \to 0 \text{ as } t \to \infty\bigr\}$$

and the *global unstable manifold* as

$$W^u(u_0) = \bigl\{x \in \mathbf{R}^n\colon\ \operatorname{dist}\bigl(\phi^t(x), u_0(\cdot)\bigr) \to 0 \text{ as } t \to -\infty\bigr\},$$

where

$$\operatorname{dist}\bigl(y, u_0(\cdot)\bigr) = \min_{0\leqslant t\leqslant T}\bigl|y - u_0(t)\bigr|.$$

In the next proposition, we show that the global stable and unstable manifolds can be defined in terms of the functions $x^+(t,\xi)$ and $x^-(t,\eta)$.

PROPOSITION 3.12. $x \in W^s(u_0)$ *if and only if there exists* γ^+ *and* ξ *such that*

$$\phi^t(x) = x^+(t+\gamma^+,\xi)$$

for all t. *This* γ^+ *has the property that* $|\phi^t(x) - u_0(t+\gamma^+)| \to 0$ *as* $t \to \infty$ *and is called the* (*forward*) *asymptotic phase of* x. *It is unique up to a multiple of* T. *It also follows that*

$$W^s(u_0) = \bigl\{x^+(t,\xi)\colon\ -\infty < t < \infty, |\xi| < \xi_0\bigr\}$$

and is an immersed submanifold of $\mathbf{R}^n$. *Moreover, if* $x \in W^s(u_0)$, *then*

$$T_xW^s(u_0) = \Bigl\{\xi \in \mathbf{R}^n\colon\ \sup_{t\geqslant 0}\bigl|D\phi^t(x)\xi\bigr| < \infty\Bigr\}.$$

PROOF. Denote by C the hyperplane passing through $u_0(0)$ and orthogonal to $f(u_0(0))$. We choose an open neighbourhood N of $u_0(0)$ in C with the following properties:

(a) $|f(x) - f(u_0(0))| < |f(u_0(0))|$ if $x \in N$;

(b) the Poincaré map $P: N \to P(N) \subset C$ associated with the cross-section C as in Proposition 3.4 is a well-defined C^1 diffeomorphism such that

$$P(x) = \phi^{\tau(x)}(x) \quad \text{with } \bigl|\tau(x) - T\bigr| \leqslant T/2$$

and

$$\phi^t(x) \notin N \quad \text{if } 0 < t < \tau(x);$$

(c) there exists a constant L such that if $P^k(x) \in N$ for $k \geqslant 0$, then $|P^k(x) - u_0(0)| \leqslant Le^{-k\gamma T}|x - u_0(0)|$ for $k \geqslant 0$.

(c) can be arranged by Proposition 1.8 in [19, p. 5], since the eigenvalues λ of $DP(0)$ satisfy $|\lambda| > e^{\alpha T} > e^{\gamma T}$ if $|\lambda| > 1$ and $|\lambda| < e^{-\alpha T} < e^{-\gamma T}$ if $|\lambda| < 1$. Later another condition will be added to N.

Now consider the intersections of the orbit $x(t) = \phi^t(x)$ with N. By (a) the tangent vector $\dot{x}(t) = f(x(t))$ is transverse to C when $x(t) \in N$ and so the set $\{t\colon x(t) \in N\}$ is isolated. It follows from (b) and the fact that

$$\operatorname{dist}\bigl(x(t), u_0(\cdot)\bigr) \to 0$$

as $t \to \infty$ that the set $\{t \geqslant 0\colon x(t) \in N\}$ is an increasing sequence $\{t_k\}_{k=0}^{\infty}$, where $t_k \to \infty$ as $k \to \infty$, and there is a positive integer k_0 such that

$$x(t_{k+1}) = P\bigl(x(t_k)\bigr)$$

for $k \geqslant k_0$. Then from (c) it follows that

$$\bigl|x(t_k) - u_0(0)\bigr| \leqslant Le^{-\gamma(k-k_0)T}\bigl|x(t_{k_0}) - u_0(0)\bigr|$$

for $k \geqslant k_0$.

Now we define

$$s_k = t_k - kT$$

so that for $k \geqslant k_0$

$$\begin{aligned} |s_{k+1} - s_k| &= |t_{k+1} - t_k - T| \\ &= \bigl|\tau\bigl(x(t_k)\bigr) - \tau\bigl(u_0(0)\bigr)\bigr| \\ &\leqslant L_1 L\bigl|x(t_{k_0}) - u_0(0)\bigr|e^{-\gamma(k-k_0)T}, \end{aligned}$$

where L_1 is a Lipschitz constant for $\tau(x)$ in N. It follows that s_k is a Cauchy sequence. Let the limit be $-\gamma_1$. Then if $k \geqslant k_0$

$$|s_k + \gamma_1| = \left|\sum_{\ell=k+1}^{\infty}(s_\ell - s_{\ell-1})\right| \leqslant L_2\bigl|x(t_{k_0}) - u_0(0)\bigr|e^{-\gamma(k-k_0)T},$$

where

$$L_2 = L_1 L(1 - e^{-\gamma T})^{-1}.$$

Now it follows from Gronwall's lemma that

$$\left|x(t) - u_0(t+\beta)\right| \leqslant e^{M_1|t-s|}\left|x(s) - u_0(s+\beta)\right|$$

for all t, s and β. Then if $k \geqslant k_0$

$$\begin{aligned}\left|x(kT-\gamma_1) - u_0(0)\right| &\leqslant e^{M_1|s_k+\gamma_1|}\left|x(t_k) - u_0(s_k+\gamma_1)\right| \\ &\leqslant e^{M_1|s_k+\gamma_1|}\left[\left|x(t_k) - u_0(0)\right| + \left|u_0(0) - u_0(s_k+\gamma_1)\right|\right] \\ &\leqslant e^{M_1|s_k+\gamma_1|}\left[\left|x(t_k) - u_0(0)\right| + M_0|s_k+\gamma_1|\right] \\ &\leqslant e^{M_1|s_k+\gamma_1|}[L + M_0L_2]\left|x(t_{k_0}) - u_0(0)\right|e^{-\gamma(k-k_0)T} \\ &\leqslant L_3\left|x(t_{k_0}) - u_0(0)\right|e^{-\gamma kT},\end{aligned}$$

where

$$L_3 = e^{M_1L_2|x(t_{k_0})-u_0(0)|}[L + M_0L_2]e^{\gamma k_0T}.$$

Suppose now $t \geqslant k_0T - \gamma_1$. Then there exists $k \geqslant k_0$ such that

$$kT - \gamma_1 \leqslant t < (k+1)T - \gamma_1.$$

Then

$$\begin{aligned}\left|x(t) - u_0(t+\gamma_1)\right| &\leqslant e^{M_1(t-kT+\gamma_1)}\left|x(kT-\gamma_1) - u_0(kT)\right| \\ &= e^{M_1(t-kT+\gamma_1)}\left|x(kT-\gamma_1) - u_0(0)\right| \\ &\leqslant e^{M_1T}L_3\left|x(t_{k_0}) - u_0(0)\right|e^{-\gamma kT} \\ &\leqslant e^{M_1T}L_3\left|x(t_{k_0}) - u_0(0)\right|e^{-\gamma(t-T+\gamma_1)}\end{aligned}$$

so that for $t \geqslant 0$

$$\begin{aligned}\left|x(t+k_0T-\gamma_1) - u_0(t)\right| &= \left|x(t+k_0T-\gamma_1) - u_0(t+k_0T)\right| \\ &\leqslant e^{M_1T}L_3\left|x(t_{k_0}) - u_0(0)\right|e^{-\gamma(t+(k_0-1)T)} \\ &= e^{(M_1-\gamma(k_0-1))T}L_3\left|x(t_{k_0}) - u_0(0)\right|e^{-\gamma t},\end{aligned}$$

where

$$e^{(M_1-\gamma(k_0-1))T}L_3 = e^{M_1L_2|x(t_{k_0})-u_0(0)|}[L + M_0L_2]e^{(M_1+\gamma)T}.$$

Then if N is also chosen so that for all $x \in N$

$$e^{M_1L_2|x-u_0(0)|}[L + M_0L_2]e^{(M_1+\gamma)T}\left|x - u_0(0)\right| < \min\{\Delta, K^{-1}\xi_0\},$$

it follows by uniqueness in Proposition 3.5 (see the remarks at the beginning of this subsection) that

$$x(t + k_0T - \gamma_1) = x^+(t, \xi),$$

with

$$\xi = P_+\bigl[x(k_0T - \gamma_1) - u_0(0)\bigr].$$

It follows that

$$x(t) = x^+(t + \gamma_1 - k_0T, \xi)$$

for $t \geqslant k_0T - \gamma_1$ and hence for all t. Taking $\gamma^+ = k_0T - \gamma_1$, this proves the first part of the proposition.

Now we examine the extent to which the asymptotic phase is unique. Suppose x has asymptotic phases γ_1 and γ_2. Then

$$\bigl|u(t + \gamma_1) - u(t + \gamma_2)\bigr| \to 0 \quad \text{as } t \to \infty. \tag{32}$$

Let

$$\gamma_1 = \bar{\gamma}_1 + k_1T, \qquad \gamma_2 = \bar{\gamma}_2 + k_2T$$

where $0 \leqslant \bar{\gamma}_1, \bar{\gamma}_2 < T$ and k_1, k_2 are integers. Then it follows from (32) that

$$\bigl|u(\bar{\gamma}_1) - u(\bar{\gamma}_2)\bigr| = \bigl|u\bigl((k - k_1)T + \gamma_1\bigr) - u\bigl((k - k_1)T + \gamma_2\bigr)\bigr| \to 0 \quad \text{as } k \to \infty.$$

That is, $u(\bar{\gamma}_1) = u(\bar{\gamma}_2)$ and so $\bar{\gamma}_1 = \bar{\gamma}_2$. Hence if γ_1, γ_2 are two asymptotic phases, $\gamma_1 - \gamma_2$ is a multiple of T. Conversely, it is obvious that if γ^+ is an asymptotic phase, then so also is $\gamma^+ + kT$ for any integer k.

It is clear from the first part of the proposition that

$$W^s(u_0) = \bigl\{x^+(t, \xi)\colon -\infty < t < \infty, |\xi| < \xi_0\bigr\}.$$

Now we know that $x_t^+(t, \xi) = f(x(t, \xi))$ is a solution of (31) bounded on $\mathcal{R}_+$ whose norm is bounded below by a positive number as $t \to \infty$ and $x(t) = x_\xi^+(t, \xi)\xi_1$ for $\xi_1 \in \mathcal{R}(P_+)$ is a solution $\to 0$ as $t \to \infty$ such that $P_+x(0) = \xi_1$. It follows that the matrix

$$\bigl[\, x_t^+(t, \xi) \quad x_\xi^+(t, \xi) \,\bigr]$$

is of full rank so that $(t, \xi) \mapsto x^+(t, \xi)$ is an immersion.

If $x = x^+(t, \xi) \in W^s(u_0)$, then we define the tangent space

$$T_xW^s(u_0) = \text{the range of } \bigl[\, x_t^+(t, \xi) \quad x_\xi^+(t, \xi) \,\bigr].$$

We show this is independent of (t,ξ) such that $x = x^{+}(t,\xi)$. Suppose

$$x = x^{+}(t_1,\xi_1) = x^{+}(t_2,\xi_2)$$

with $t_1 \geqslant t_2$. Then, for all t,

$$\phi^{t}(x) = x^{+}(t+t_1,\xi_1) = x^{+}(t+t_2,\xi_2) \tag{33}$$

so that t_1 and t_2 are both asymptotic phases for x. It follows that $t_1 = t_2 + kT$ for some nonnegative integer $k \geqslant 0$ and taking $t = -t_2$ in (33) we find that

$$x^{+}(0,\xi_2) = x^{+}(kT,\xi_1).$$

It follows that

$$\bigl|P_{+}\bigl[x^{+}(kT,\xi_1) - u_0(0)\bigr]\bigr| = \bigl|P_{+}\bigl[x^{+}(0,\xi_2) - u_0(0)\bigr]\bigr| = |\xi_2| < \xi_0$$

so that

$$\bigl|P_{+}\bigl[x^{+}(kT,\xi) - u_0(0)\bigr]\bigr| < \xi_0$$

if ξ is sufficiently close to ξ_1. Thus, by uniqueness in Proposition 3.5

$$x^{+}(t+kT,\xi) = x^{+}\bigl(t,\psi(\xi)\bigr)$$

for ξ close to ξ_1, where

$$\psi(\xi) = P_{+}\bigl[x^{+}(kT,\xi) - u_0(0)\bigr].$$

It follows that

$$x_t^{+}(t_1,\xi_1) = x_t^{+}(t_2+kT,\xi_1) = x_t^{+}\bigl(t_2,\psi(\xi_1)\bigr) = x_t^{+}(t_2,\xi_2)$$

and

$$\begin{aligned}
x_\xi^{+}(t_1,\xi_1) &= x_\xi^{+}(t_2+kT,\xi_1)\\
&= \frac{\partial}{\partial\xi} x^{+}\bigl(t_2,\psi(\xi)\bigr)\Big|_{\xi=\xi_1}\\
&= x_\xi^{+}\bigl(t_2,\psi(\xi_1)\bigr)\psi'(\xi_1)\\
&= x_\xi^{+}(t_2,\xi_2)\psi'(\xi_1),
\end{aligned}$$

which show that the range of the matrix $[\,x_t^{+}(t_1,\xi_1) \quad x_\xi^{+}(t_1,\xi_1)\,]$ is contained in and hence coincides with the range of $[\,x_t^{+}(t_2,\xi_2) \quad x_\xi^{+}(t_2,\xi_2)\,]$. So $T_x W^{s}(u_0)$ is well-defined.

Now, as we mentioned earlier, $Z(t) = x_\xi^+(t, \xi_1)$ is the unique solution restricted to $\mathcal{R}(P_+)$ of (31) with $\xi = \xi_1$ such that $\|Z\|_\gamma < \infty$ and $P_+Z(0) = P_+$. However, since $|f'(x^+(t, \xi_1)) - f'(u_0(t))| \to 0$ as $t \to \infty$, it follows from Remarks 2.7 and 2.3 that

$$\dot{y} = f'\bigl(x^+(t, \xi_1)\bigr)y \tag{34}$$

has a trichotomy on $[t_1, \infty)$ with projections having the same rank as those for $\dot{x} = f'(u_0(t))x$. So this equation has a subspace W_t of solutions bounded on $[t_1, \infty)$ with dimension equal to $\operatorname{rank}(P_0 + P_+)$. It follows that

$$\operatorname{span}\bigl\{f(x^+(t, \xi_1)\bigr\} \oplus \mathcal{R}\bigl(x_\xi^+(t, \xi_1)\bigr),$$

which has dimension equal to $\operatorname{rank}(P_0 + P_+)$, coincides with W_t. Next note that (34) can be written as $\dot{y} = f'(\phi^{t-t_1}(x))y$. So, by a time shift, it follows that $\operatorname{span}\{f(x^+(t + t_1, \xi_1)\} \oplus \mathcal{R}(x_\xi^+(t + t_1, \xi_1))$ is the subspace of solutions of

$$\dot{y} = f'\bigl(\phi^t(x)\bigr)y \tag{35}$$

which are bounded on $[0, \infty)$. Hence

$$T_x W^s(u_0) = \operatorname{span}\bigl\{f\bigl(x^+(t_1, \xi_1)\bigr)\bigr\} \oplus \mathcal{R}\bigl(x_\xi^+(t_1, \xi_1)\bigr)$$

consists of the values at $t = 0$ of the subspace of solutions of (35) which are bounded on $[0, \infty)$, that is,

$$\Bigl\{\xi \in \mathbf{R}^n\colon \sup_{t \geqslant 0}\bigl|D\phi^t(x)\xi\bigr| < \infty\Bigr\}.$$

This completes the proof of the proposition. □

Similarly, we can prove the following proposition.

PROPOSITION 3.13. *$x \in W^u(u_0)$ if and only if there exists γ^- and η such that*

$$\phi^t(x) = x^-(t + \gamma^-, \eta)$$

for all t. This γ^- has the property that $|\phi^t(x) - u_0(t + \gamma^-)| \to 0$ as $t \to -\infty$ and is called the (*backward*) *asymptotic phase of x. It is unique up to a multiple of T. It also follows that*

$$W^u(u_0) = \bigl\{x^-(t, \eta)\colon -\infty < t < \infty, |\eta| < \eta_0\bigr\}$$

and is an immersed submanifold of $\mathbf{R}^n$. Moreover, if $x \in W^u(u_0)$, then

$$T_x W^u(u_0) = \Bigl\{\eta \in \mathbf{R}^n\colon \sup_{t \leqslant 0}\bigl|D\phi^t(x)\eta\bigr| < \infty\Bigr\}.$$

Finally we define the local stable and unstable fibres.

DEFINITION 3.14. Corresponding to γ^+ in $0 \leqslant \gamma^+ < T$, we define the *local stable fibre*:

$$W_{\mathrm{loc}}^{s,\gamma^+}(u_0) = \bigl\{x^+(\gamma^+, \xi)\colon\ \xi \in \mathcal{R}(P_+),\ |\xi| < \xi_0\bigr\}$$

and we define the *global stable fibre* $W^{s,\gamma^+}(u_0)$ as

$$\bigcup_{m\in Z} \phi^{mT}\bigl(W_{\mathrm{loc}}^{s,\gamma^+}(u_0)\bigr) = \bigl\{x^+(\gamma^+ + mT, \xi)\colon\ \xi \in \mathcal{R}(P_+), |\xi| < \xi_0, m \in Z\bigr\}.$$

It is easy to see that $x \in W^{s,\gamma^+}(u_0)$ if and only if $|\phi^t(x) - u_0(t + \gamma^+)| \to 0$ as $t \to \infty$ and that

$$W^s(u_0) = \bigcup_{0\leqslant\gamma^+<T} W^{s,\gamma^+}(u_0).$$

DEFINITION 3.15. Corresponding to γ^- in $0 \leqslant \gamma^- < T$, we define the *local unstable fibre*:

$$W_{\mathrm{loc}}^{u,\gamma^-}(u_0) = \bigl\{x^-(\gamma^-, \eta)\colon\ \eta \in \mathcal{R}(P_-), |\eta| < \eta_0\bigr\}$$

and we define the *global unstable fibre* $W^{u,\gamma^-}(u_0)$ as

$$\bigcup_{m\in Z} \phi^{mT}\bigl(W_{\mathrm{loc}}^{u,\gamma^-}(u_0)\bigr) = \bigl\{x^-(\gamma^- + mT, \eta)\colon\ \eta \in \mathcal{R}(P_-), |\eta| < \eta_0, m \in Z\bigr\}.$$

It is easy to see that $x \in W^{u,\gamma^-}(u_0)$ if and only if $|\phi^t(x) - u_0(t + \gamma^-)| \to 0$ as $t \to -\infty$ and that

$$W^u(u_0) = \bigcup_{0\leqslant\gamma^-<T} W^{u,\gamma^-}(u_0).$$

4. Homoclinic orbits

In this section we study properties of periodic-to-periodic homoclinic orbits. First we define a Poincaré map associated with such an orbit. Then we give equivalent conditions for transversality of the periodic-to-periodic homoclinic orbits including its equivalence with the transversality of the corresponding homoclinic orbit of the Poincaré map. Then Smale's theorem enables us to deduce the existence of chaotic behaviour in the neighbourhood of a transversal periodic-to-periodic homoclinic orbit.

So let $u_0(t)$ be a hyperbolic T-periodic orbit of the C^1 system

$$\dot{x} = f(x), \tag{36}$$

where, as usual, we assume that $|f(x)| \leqslant M_0$ and $|f'(x)| \leqslant M_1$. We make the following definition.

DEFINITION 4.1. If $u_0(t)$ is a hyperbolic T-periodic orbit of (36), then a solution $p(t)$ is said to be a *homoclinic orbit* if dist $(p(t), u_0(\cdot)) \to 0$ as $|t| \to \infty$ or, equivalently, $p(t) \in W^s(u_0) \cap W^u(u_0)$.

4.1. *Poincaré map associated with a homoclinic orbit*

Consider system (36) with hyperbolic T-periodic orbit $u_0(t)$ and associated homoclinic orbit $p(t)$. Denote by C the hyperplane passing through $u_0(0)$ and orthogonal to $f(u_0(0))$. In a neighbourhood of $u_0(0)$ in C, we will define a map F which has $u_0(0)$ as a hyperbolic fixed point and an associated homoclinic orbit which consists of points on the solution $p(t)$ which lie on C.

First we choose an open neighbourhood N of $u_0(0)$ in C with the following properties:

(a) $|f(y) - f(u_0(0))| \leqslant \frac{1}{2}|f(u_0(0))|$ if $y \in N$;

(b) the Poincaré map $P : N \mapsto C$ (see Proposition 3.4) has the properties that for $x \in N$

$$\phi^t(x) \notin N \quad \text{if } 0 < t < \tau_0(x)$$

and

$$\left|\tau_0(x) - T\right| \leqslant T/2,$$

where $P(x) = \phi^{\tau_0(x)}(x)$;

(c) $P^k(x) \in N$ for all integers k if and only if $x = u_0(0)$.

That N can be chosen so that (c) is satisfied follows from the fact that $u_0(0)$ is an isolated invariant set for P (see [19, p. 44]).

Consider the intersections of the homoclinic orbit $p(t)$ with N. Similarly to the beginning of the proof of Proposition 3.12, it follows that the set $\{t \in \mathbf{R}\colon\ p(t) \in N\}$ is an increasing sequence $\{t_k\}_{k=-\infty}^{\infty}$, where $t_k \to \pm\infty$ as $k \to \pm\infty$ and

$$p(t_{k+1}) = P\big(p(t_k)\big)$$

if $|k|$ is sufficiently large. It is also clear that $p(t_k) \to u_0(0)$ as $k \to \pm\infty$. For short, we write

$$p_k = p(t_k).$$

Because of property (c) of P, there exists an integer K^- such that

$$p_{k+1} = P(p_k) \quad \text{if } k < K^-$$

but

$$P(p_{K^-}) \neq p_{K^-+1}.$$

We set

$$v_1 = P(p_{K^-}).$$

It follows from (b) that $\phi^t(p_{K^-}) \notin N$ if $0 < t < \tau_0(p_{K^-})$ and so if v_1 were in N, we would have $v_1 = p_{K^-+1}$. So $v_1 \notin N$.

Next we take K^+ to be an integer greater than K^- such that $p_{k+1} = P(p_k)$ if $k \geqslant K^+$. Then we set

$$v_2 = p_{K^+} = \phi^{T_1}(v_1).$$

Now consider the subset

$$I_0 = \{\dots, p_{K^--1}, p_{K^-}\} \cup \bigl\{u(0)\bigr\} \cup \{p_{K^++1}, p_{K^++2}, \dots\}$$

of N and choose an open neighbourhood $V_2 \subset N$ of $v_2 = p_{K^+}$ such that the closure of V_2 does not intersect $I_0 \cup \{v_1\}$. Then we can choose an open neighbourhood $V_0 \subset N$ of I_0 such that v_1 is not in the closure of V_0 and $P(V_0)$ does not intersect V_2.

Next we let $V_1 \subset C$ be an open ball with centre v_1 such that $V_0 \cap V_1$ is empty and the Poincaré map $Q = P(v_1, 0, T_1, v, v)$ with $v = u_0'(0)/|u_0'(0)|$ is defined on V_1, has range in V_2 and satisfies $|\tau_1(x) - T_1| \leqslant T_1/2$, where $Q(x) = \phi^{\tau_1(x)}(x)$. Note that $Q(v_1) = \phi^{T_1}(v_1) = v_2 = p_{K^+}$.

Now we define a diffeomorphism $F : V_0 \cup V_1 \mapsto C$ by

$$F(x) = \begin{cases} P(x) & \text{if } x \in V_0, \\ Q(x) & \text{if } x \in V_1. \end{cases}$$

F is well-defined since $V_0 \cap V_1$ is empty and is one to one since

$$P(V_0) \cap Q(V_1) \subset P(V_0) \cap V_2,$$

which is empty, and it is a diffeomorphism on to an open set since P and Q are diffeomorphisms. Since $u_0(0)$ is a hyperbolic fixed point for P, it is also a hyperbolic fixed point for F. Next we see that

$$\bigl\{\dots, p_{K^--1}, p_{K^-}, v_1 = P(p_{K^-}), v_2 = Q(v_1) = p_{K^+}, p_{K^++1}, p_{K^++2}, \dots\bigr\}$$

is a homoclinic orbit to $u_0(0)$.

4.2. *Transversal periodic-to-periodic homoclinic orbits and hyperbolicity*

We now define the notion of transversality for a periodic-to-periodic homoclinic orbit and show it is equivalent to hyperbolicity and also to transversality of the corresponding homoclinic orbit of the Poincaré map defined in the previous subsection.

DEFINITION 4.2. A homoclinic orbit $p(t)$ of (36) is said to be *transversal* if the tangent spaces $T_{p(t)}W^s(u_0)$ and $T_{p(t)}W^u(u_0)$ intersect in the subspace spanned by $p'(t) = f(p(t))$.

DEFINITION 4.3. A bounded solution $p(t)$ of (36) is said to be *hyperbolic* if $|f(p(t))|$ is bounded below by a positive number and there is a splitting

$$\mathbf{R}^n = \operatorname{span}\{f(p(t))\} \oplus E^s_{p(t)} \oplus E^u_{p(t)} \quad \text{for } t \in \mathbf{R}$$

with the invariance property

$$D\phi^\tau(p(t))(E^s_{p(t)}) = E^s_{p(t+\tau)}, \quad D\phi^\tau(p(t))(E^u_{p(t)}) = E^u_{p(t+\tau)}, \quad \text{for } \tau, t \in \mathbf{R},$$

and there are positive constants K and α such that for all t and $\tau \geqslant 0$

$$|D\phi^\tau(p(t))\xi| \leqslant K\mathrm{e}^{-\alpha\tau}|\xi|, \quad \xi \in E^s_{p(t)},$$

and

$$|D\phi^{-\tau}(p(t))\xi| \leqslant K\mathrm{e}^{-\alpha\tau}|\xi|, \quad \xi \in E^u_{p(t)}.$$

Note that a periodic solution is hyperbolic in the sense of Definition 3.3 if and only if it is hyperbolic in the sense of Definition 4.3 (see, for example, [19]). Now we give equivalent conditions for transversality of a periodic-to-periodic homoclinic orbit.

PROPOSITION 4.4. *Suppose $u_0(t)$ is a hyperbolic T-periodic orbit of* (36) *and $p(t)$ is an associated homoclinic orbit. Then the following conditions are equivalent*:

(i) *$p(t)$ is transversal*;
(ii) *$p(t)$ is hyperbolic*;
(iii) *$p'(t)$ is, up to a scalar multiple, the unique bounded solution of*

$$\dot{x} = f'(p(t))x; \tag{37}$$

(iv) *the variational system* (37) *has a trichotomy on* $\mathbf{R}$;
(v) *the corresponding homoclinic orbit of the associated Poincaré map F is transversal.*

PROOF. Note that by Proposition 3.12, $T_{p(t)}W^s(u_0)$ is the subspace of solutions of (37) bounded on $[0, \infty)$ and by Proposition 3.13, $T_{p(t)}W^u(u_0)$ is the subspace of solutions of

(37) bounded on $(-\infty, 0]$. Hence their intersection is the subspace of solutions of (37) bounded on $(-\infty, \infty)$. It follows at once that (i) and (iii) are equivalent.

By Remark 2.4, (iv) implies (iii). To prove that (iii) implies (iv), let γ^+ and γ^- be the forward and backward asymptotic phases of $p(0)$ so that $|p(t) - u_0(t + \gamma^+)| \to 0$ as $t \to \infty$ and $|p(t) - u_0(t + \gamma^-)| \to 0$ as $t \to -\infty$. Then since $\dot{x} = f'(u_0(t))x$ has a trichotomy on $\mathbf{R}$, it follows from Remarks 2.7 and 2.3 that (37) has trichotomies on both half-lines with projections of the same rank. By Remark 2.4 this together with (iii) implies (iv). So (iii) and (iv) are equivalent.

To show that (iv) and (v) are equivalent, we apply Proposition 2.9 to (37). Note that $|f'(p(t))| \leqslant M_1$ and $p'(t) = f(p(t))$ is a bounded solution of (37) such that $\inf|p'(t)| > 0$ since $\operatorname{dist}(p(t), u_0(\cdot)) \to 0$ as $|t| \to \infty$. Now, as in 4.1, the corresponding homoclinic orbit of the Poincaré map F can be written as a sequence $p(s_k)$ where $p(s_{-1}) = p_{K^-}$, $p(s_0) = v_1 = P(p_{K^-})$, $p(s_1) = v_2 = p_{K^+}$, etc., where we see from the properties of $\tau_0(x)$ and $\tau_1(x)$ that $T/2 \leqslant s_{k+1} - s_k \leqslant 3T/2$ if $k \neq 0$ and $s_1 - s_0 > 0$. So

$$0 < \inf(s_{k+1} - s_k) \leqslant \sup(s_{k+1} - s_k) < \infty.$$

We define the projection Q_k to be the projection with range W, the orthogonal complement of $\dot{u}_0(0) = f(u_0(0))$, and nullspace spanned by $f(p(s_k))$. Since $f(p(s_k)) \to f(u_0(0))$ as $|k| \to \infty$, Q_k converges as $|k| \to \infty$ to an orthogonal projection and so $|Q_k|$ is bounded. Next we see from the definition of F and (19) that

$$DF\big(p(s_k)\big) = Q_k D\phi^{s_{k+1}-s_k}\big(p(s_k)\big) = Q_k X(s_{k+1}) X^{-1}(s_k),$$

where $X(t) = D\phi^t(p(0))$ is the fundamental matrix of (37) with $X(0) = I$.

Then it follows from Proposition 2.9 that (37) has a trichotomy on $\mathbf{R}$ if and only if the difference equation

$$u_{k+1} = DF\big(p(s_k)\big)u_k, \quad u_k \in W,$$

has an exponential dichotomy on Z. However, this last statement is equivalent to the statement that $\{p(s_k)\}_{k=-\infty}^{\infty}$ is a transversal homoclinic orbit for F (see [19]). So (iv) and (v) are equivalent.

If (iv) holds, then there exist supplementary projections P_0, P_+, P_- with rank $P_0 = 1$ and positive constants K, α such that

$$\begin{aligned} &\big|X(t)P_0X^{-1}(s)\big| \leqslant K \quad \text{for all } t, s,\\ &\big|X(t)P_+X^{-1}(s)\big| \leqslant K\mathrm{e}^{-\alpha(t-s)} \quad \text{for } t \geqslant s,\\ &\big|X(t)P_-X^{-1}(s)\big| \leqslant K\mathrm{e}^{-\alpha(s-t)} \quad \text{for } t \leqslant s, \end{aligned}$$

where $X(t) = D\phi^t(p(0))$. The last two inequalities can be rewritten as

$$\begin{aligned} &\big|D\phi^\tau\big(p(s)\big)P_+(s)\big| \leqslant K\mathrm{e}^{-\alpha\tau} \quad \text{for } \tau \geqslant 0,\\ &\big|D\phi^{-\tau}\big(p(s)\big)P_-(s)\big| \leqslant K\mathrm{e}^{-\alpha\tau} \quad \text{for } \tau \geqslant 0, \end{aligned}$$

where $P_\sigma(t) = X(t)P_\sigma X^{-1}(t)$. Note also since $p'(t)$ is a bounded solution of (37), it follows that the range of $P_0(t)$ is spanned by $p'(t) = f(p(t))$. Note also, as observed above, that $\inf|p'(t)| > 0$. Hence if we define $E^s_{p(t)} = \mathcal{R}(P_+(t))$ and $E^u_{p(t)} = \mathcal{R}(P_-(t))$, we see that $p(t)$ is hyperbolic. Thus (iv) implies (ii).

If (ii) holds, then any solution $x(t)$ of (37) can be written as $x(t) = \lambda p'(t) + x_+(t) + x_-(t)$, where $x_\sigma(t) \in E^\sigma_{p(t)}$. If $x_+(t) \neq 0$, then $|x_+(t)| \to \infty$ as $t \to -\infty$ whereas $p'(t) + x_-(t)$ is bounded so that $|x(t)| \to \infty$. Hence, if $x(t)$ is bounded, we must have $x_+(t) = 0$. Similarly, $x_-(t) = 0$. So $x(t)$ is a multiple of $p'(t)$. Thus (ii) implies (iii) and the proof is completed. □

4.3. *Chaos near a transversal periodic-to-periodic homoclinic orbit*

In this subsection we examine the chaotic behaviour in the neighbourhood of a transversal periodic-to-periodic homoclinic orbit.

Consider system (36) with hyperbolic T-periodic orbit $u_0(t)$ and associated transversal homoclinic orbit $p(t)$. In the previous two subsections, we defined an associated Poincaré map F and showed that the sequence of points

$$\left\{\ldots, p_{K^- -1}, p_{K^-}, v_1 = P(p_{K^-}), v_2 = Q(v_1) = p_{K^+}, p_{K^+ +1}, p_{K^+ +2}, \ldots\right\}$$

is a transversal homoclinic orbit bi-asymptotic to the hyperbolic fixed point $u_0(0)$.

Now we prove the following

PROPOSITION 4.5. *Suppose system* (36) *has a hyperbolic T-periodic orbit $u_0(t)$ with associated transversal homoclinic orbit $p(t)$. Then there exists a compact invariant set $\mathcal{S}$ such that*

(a) *the periodic points are dense*;
(b) *there is a dense orbit*;
(c) *there exists $\delta > 0$ such that if x is in $\mathcal{S}$ there exists y in $\mathcal{S}$ arbitrarily close to x such that for any continuous nonnegative function $\alpha(t)$ with $\alpha(0) = 0$ there exists $t \geqslant 0$ with $|\phi^{\alpha(t)}(y) - \phi^t(x)| > \delta$.*

REMARK 4.6. (c) is sensitive dependence on initial conditions in the flow context. To the author's knowledge, there is no definition of sensitive dependence directly for flows in the literature since usually chaos for flows is described in terms of a return map.

PROOF OF PROPOSITION 4.5. As shown, for example, in [19], there is a compact invariant subset $S \subset V_0 \cup V_1$ on which the Poincaré map F is chaotic in the sense of Devaney (see [8]), that is

(i) the periodic points are dense;
(ii) there is a dense orbit;
(iii) there exists $\delta > 0$ such that if $x \in S$ there is $y \in S$ arbitrarily close to x such that $|F^k(x) - F^k(y)| > \delta$ for some $k \geqslant 0$.

Then we define

$$\mathcal{S} = \{\phi^t(x) \colon x \in S, -\infty < t < \infty\}.$$

Clearly this set is invariant for system (36). To prove the compactness, first note that if $x \in S$, then there exist times $t_k(x)$ such that $\phi^{t_k}(x) = F^k(x)$ for all integers k. Now, using the notation of Section 4.1, $t_{k+1}(x) - t_k(x) = \tau_0(F^k(x))$ if $F^k(x) \in V_0$ or $\tau_1(F^k(x))$ if $F^k(x) \in V_1$ and so

$$c_1 \leqslant t_{k+1}(x) - t_k(x) \leqslant c_2,$$

where by properties of τ_0 and τ_1

$$c_1 = \frac{1}{2}\min\{T, T_1\}, \qquad c_2 = \frac{3}{2}\max\{T, T_1\}.$$

Now let $\phi^{s_m}(x_m)$ be a sequence in $\mathcal{S}$. Then there exists k_m such that $t_{k_m}(x_m) \leqslant s_m < t_{k_m+1}(x_m)$ and so it follows that $\phi^{s_m}(x_m) = \phi^{\sigma_m}(z_m)$, where $0 \leqslant \sigma_m = s_m - t_{k_m}(x_m) \leqslant c_2$ and $z_m = F^{k_m}(x_m) \in S$. Then there are subsequences $\sigma_{j_m} \to \sigma$ and $z_{j_m} \to z \in S$. By continuity of $\phi^t(x)$, it follows that $\phi^{s_{j_m}}(x_{j_m}) \to \phi^\sigma(z) \in \mathcal{S}$. Thus $\mathcal{S}$ is compact.

To prove (a) and (b), let $x = \phi^t(y)$ be in $\mathcal{S}$, where $y \in S$. Then there exists $t_k = t_k(y)$ such that $\phi^{t_k}(y) = F^k(y)$ and $t_k \leqslant t < t_{k+1}$. So $x = \phi^{t-t_k}(F^k(y))$. Given $\varepsilon > 0$, choose $\delta > 0$ so that $|z - F^k(y)| < \delta$ implies $|\phi^t(z) - \phi^t(F^k(y))| < \varepsilon$ for $0 \leqslant t \leqslant c_2$. Then we see that $|x - \phi^{t-t_k}(z)| < \varepsilon$. Since we can choose $z \in S$ to be periodic for F so that $\phi^{t-t_k}(z)$ is periodic for the flow, it follows that the periodic points are dense in $\mathcal{S}$. Since we can choose $z \in S$ to be on the orbit of F which is dense in S, it follows that the orbit of z in the flow is dense in $\mathcal{S}$.

To prove (c), suppose $x \in S$. Then we know there exists $y \in S$ arbitrarily close to x such that $|F^k(y) - F^k(x)| > \delta$ for some $k \geqslant 0$. In the following δ_1 will be a positive number which will be determined as we proceed. Also we write

$$\Delta = \inf_{x \in \mathcal{S}} |f(x)| > 0.$$

Let $\alpha(t)$ be a continuous nonnegative function with $\alpha(0) = 0$. Suppose

$$|\phi^{\alpha(t)}(y) - \phi^t(x)| \leqslant \delta_1$$

for all $t \geqslant 0$. Now we apply Lemma 3.1 with $\phi^{\alpha(t)}(y)$ as z, $\phi^t(x)$ as x, $v = f(\phi^t(x))/|f(\phi^t(x))|$ and $T = 0$ to deduce that if

$$\delta_1 < \min\left\{\frac{\Delta}{4M_1}, \frac{\Delta^2}{8M_0M_1}\right\},$$

there exists $\tau = \tau(t)$ with $|\tau| \leqslant 2\delta_1/\Delta$ such that

$$\langle \phi^{\alpha(t)+\tau}(y) - \phi^t(x), f(\phi^t(x)) \rangle = 0.$$

Then if we define

$$\beta(t) = \alpha(t) + \tau(t),$$

we see that

$$\big\langle \phi^{\beta(t)}(y) - \phi^t(x), f\big(\phi^t(x)\big)\big\rangle = 0$$

and

$$\begin{aligned} \big|\phi^{\beta(t)}(y) - \phi^t(x)\big| &\leqslant \big|\phi^{\beta(t)}(y) - \phi^{\alpha(t)}(y)\big| + \big|\phi^{\alpha(t)}(y) - \phi^t(x)\big| \\ &\leqslant M_0\big|\tau(t)\big| + \delta_1 \leqslant \delta_2, \end{aligned} \tag{38}$$

where

$$\delta_2 = (2M_0/\Delta + 1)\delta_1.$$

Note it also follows from Remark 3.2 that $\tau(t)$ and hence $\beta(t)$ is continuous. Now we show that $\beta(t)$ is, in fact, differentiable provided that

$$M_1\delta_2 < \Delta.$$

To this end, consider the continuously differentiable real function

$$g(\beta, t) = \big\langle \phi^{\beta}(y) - \phi^t(x), f\big(\phi^t(x)\big)\big\rangle. \tag{39}$$

We estimate

$$\begin{aligned} g_\beta\big(\beta(t), t\big) &= \big\langle f\big(\phi^{\beta(t)}(y)\big), f\big(\phi^t(x)\big)\big\rangle \\ &\geqslant \big|f\big(\phi^t(x)\big)\big|^2 - \big|f\big(\phi^{\beta(t)}(y)\big) - f\big(\phi^t(x)\big)\big|\big|f\big(\phi^t(x)\big)\big| \\ &\geqslant \big|f\big(\phi^t(x)\big)\big|[\Delta - M_1\delta_2] \\ &> 0. \end{aligned}$$

Then it follows from the continuity of $\beta(t)$ and the implicit function theorem that $\beta(t)$ is continuously differentiable. Moreover, if we differentiate the equation $g(\beta(t), t) = 0$ with respect to t, we obtain the equation

$$\begin{aligned} &\big\langle f\big(\phi^{\beta(t)}(y)\big), f\big(\phi^t(x)\big)\big\rangle\beta'(t) \\ &\quad - \big|f\big(\phi^t(x)\big)\big|^2 + \big\langle \phi^{\beta(t)}(y) - \phi^t(x), f'\big(\phi^t(x)\big) f\big(\phi^t(x)\big)\big\rangle = 0 \end{aligned}$$

so that $\beta'(t)$ is given by

$$\begin{aligned} &1 - \big[\big\langle \phi^{\beta(t)}(y) - \phi^t(x), f'\big(\phi^t(x)\big) f\big(\phi^t(x)\big)\big\rangle \\ &\quad + \big\langle f\big(\phi^{\beta(t)}(y)\big) - f\big(\phi^t(x)\big), f\big(\phi^t(x)\big)\big\rangle\big] \big/ \big\langle f\big(\phi^{\beta(t)}(y)\big), f\big(\phi^t(x)\big)\big\rangle \end{aligned}$$

from which we see, using the estimate after (39), that

$$\left|\beta'(t) - 1\right| \leqslant 2(\Delta - M_1\delta_2)^{-1}M_1\delta_2. \tag{40}$$

Now there is a sequence $t_k = t_k(x)$ such that $\phi^{t_k}(x) = F^k(x)$ for $k \geqslant 0$. We apply Lemma 3.1 with $\phi^{\beta(t_k)}(y)$ as z, $\phi^{t_k}(x)$ as x, $T = 0$ and $v = f(u_0(0))/|f(u_0(0))|$. Note that by (38), $|\phi^{\beta(t_k)}(y) - \phi^{t_k}(x)| \leqslant \delta_2$ and that since $\phi^{t_k}(x) \in N$,

$$\left|\left\langle f\left(\phi^{t_k}(x)\right), f\left(u_0(0)\right)\right\rangle\right| \geqslant \frac{1}{2}\left|f\left(u_0(0)\right)\right|^2$$

so that

$$\left|\left\langle f\left(\phi^{t_k}(x)\right), v\right\rangle\right| \geqslant \frac{1}{2}\Delta.$$

Then if

$$8M_1\delta_2 < \Delta, \qquad 32M_0M_1\delta_2 < \Delta^2,$$

it follows from Lemma 3.1 that there exists a h_k satisfying $|h_k| \leqslant 4\Delta^{-1}\delta_2$ which is unique in $|h_k| \leqslant \Delta/(8M_0M_1)$ such that

$$\left\langle \phi^{\beta(t_k)+h_k}(y) - \phi^{t_k}(x), f\left(u_0(0)\right)\right\rangle = 0.$$

Now we show that for $k \geqslant 0$, $\phi^{\beta(t_k)+h_k}(y) \in S$ and

$$\phi^{\beta(t_k)+h_k}(y) = F^k(y). \tag{41}$$

Note when $k = 0$ that the equation

$$\left\langle \phi^{\beta(t_0)+h}(y) - \phi^{t_0}(x), f\left(u_0(0)\right)\right\rangle = 0$$

has the solution $h = -\beta(0)$ since $t_0 = 0$ and $x, y \in S$ which is in the hyperplane orthogonal to $f(u_0(0)$. Since

$$\left|\beta(0)\right| = \left|\tau(0)\right| \leqslant 2\Delta^{-1}\delta_1 \leqslant 2\Delta^{-1}\delta_2 < \Delta/(8M_0M_1),$$

it follows from uniqueness that $h_0 = -\beta(t_0)$ so that $\beta(t_0) + h_0 = 0$. Hence (41) holds for $k = 0$ and, of course, $y \in S$.

Now we use induction to complete the proof. We suppose for some $k \geqslant 0$ that $\phi^{\beta(t_k)+h_k}(y) \in S$ and (41) holds. Note, using (40), that

$$\begin{aligned}&\left|\beta(t_{k+1}) + h_{k+1} - \left(\beta(t_k) + h_k\right) - (t_{k+1} - t_k)\right| \\ &\quad \leqslant \left|\beta(t_{k+1}) - \beta(t_k) - (t_{k+1} - t_k)\right| + |h_{k+1}| + |h_k|\end{aligned}$$

$$
\begin{aligned}
&\leqslant 2(\Delta - M_1\delta_2)^{-1}M_1\delta_2(t_{k+1}-t_k) + 8\Delta^{-1}\delta_2\\
&\leqslant 2(\Delta - M_1\delta_2)^{-1}M_1c_2\delta_2 + 8\Delta^{-1}\delta_2\\
&\leqslant \Delta/(16M_0M_1) + \Delta/(16M_0M_1)\\
&= \Delta/(8M_0M_1),
\end{aligned}
$$

provided that

$$64M_0M_1^2c_2\delta_2 < \Delta^2, \qquad 128M_0M_1\delta_2 < \Delta^2.$$

Note also that

$$\left|\phi^{\beta(t_k)+h_k}(y) - \phi^{t_k}(x)\right| \leqslant M_0|h_k| + \left|\phi^{\beta(t_k)}(y) - \phi^{t_k}(x)\right| \leqslant (4M_0\Delta^{-1}+1)\delta_2.$$

Next we assume δ_1 is so small that

$$(4M_0\Delta^{-1}+1)\delta_2 \leqslant \frac{\mathrm{e}^{-M_1c_2}|f(u_0(0))|}{8M_1}\min\left\{1, \frac{|f(u_0(0))|}{4M_0}\right\}$$

and such that

$$\left|\tau(x) - \tau(y)\right| \leqslant \Delta/(8M_0M_1)$$

if x, y are in S and $|x-y| \leqslant (4M_0\Delta^{-1}+1)\delta_2$, where $\tau(x)$ is the function such that $F(x) = \phi^{\tau(x)}(x)$ for $x \in V_0 \cup V_1$. From the latter we conclude that

$$\left|\tau\left(\phi^{\beta(t_k)+h_k}(y)\right) - (t_{k+1}-t_k)\right| = \left|\tau\left(\phi^{\beta(t_k)+h_k}(y)\right) - \tau\left(\phi^{t_k}(x)\right)\right| \leqslant \Delta/(8M_0M_1).$$

Then it follows by the uniqueness in Lemma 3.1 with $\phi^{t_k}(x)$ as x, $z = \phi^{\beta(t_k)+h_k}(y)$, $T = t_{k+1} - t_k = \tau(\phi^{t_k}(x))$, $v = f(u_0(0))/|f(u_0(0))|$ and

$$\alpha = \frac{\langle f(\phi^{t_{k+1}}(x)), f(u_0(0))\rangle}{4M_0M_1|f(u_0(0))|} \quad \left(\geqslant \Delta/(8M_0M_1)\right)$$

that

$$\beta(t_{k+1}) + h_{k+1} - \left(\beta(t_k) + h_k\right) = \tau\left(\phi^{\beta(t_k)+h_k}(y)\right).$$

Hence

$$\phi^{\beta(t_{k+1})+h_{k+1}}(y) = F\left(F^k(y)\right) = F^{k+1}(y) \in S.$$

This finishes the induction proof and so we know that $\phi^{\beta(t_k)+h_k}(y) \in S$ and (41) holds for all $k \geqslant 0$.

Hence we conclude that for $k \geqslant 0$

$$\left|F^k(y) - F^k(x)\right| = \left|\phi^{\beta(t_k)+h_k}(y) - \phi^{t_k}(x)\right| \leqslant (4M_0/\Delta + 1)\delta_2.$$

This is a contradiction if

$$(4M_0/\Delta + 1)\delta_2 \leqslant \delta.$$

So we conclude that if $x \in S$ and for some $y \in S$ there exists $k \geqslant 0$ such that $|F^k(y) - F^k(x)| > \delta$, then if $\alpha(t)$ is a continuous nonnegative function with $\alpha(0) = 0$ we must have

$$\left|\phi^{\alpha(t)}(y) - \phi^t(x)\right| > \delta_1$$

for some $t \geqslant 0$, where δ_1 is a positive number satisfying all the conditions given above.

Now suppose $x \in \mathcal{S}$. Then, as at the beginning of this proof, $x = \phi^s(z)$ where $z \in S$ and $0 \leqslant s \leqslant c_2$. Let the nonnegative continuous function $\alpha(t)$ with $\alpha(0) = 0$ be given. We define the continuous function

$$\tilde{\alpha}(t) = \begin{cases} t & \text{if } 0 \leqslant t \leqslant s, \\ \alpha(t-s) + s, & \text{if } t \geqslant s. \end{cases}$$

Now, applying the first part of this proof to $\tilde{\alpha}(t)$ instead of $\alpha(t)$, we deduce that there exists $w \in S$ arbitrarily close to z such that

$$\left|\phi^{\tilde{\alpha}(t)}(w) - \phi^t(z)\right| > \delta_1$$

for some $t \geqslant 0$. Note that if w is sufficiently close to z then we can assume $t \geqslant s$. So we have

$$\left|\phi^{\alpha(t-s)+s}(w) - \phi^{t-s}(x)\right| > \delta_1$$

for some $t \geqslant s$. Hence

$$\left|\phi^{\alpha(t)}\big(\phi^s(w)\big) - \phi^t(x)\right| > \delta_1$$

for some $t \geqslant 0$, where $\phi^s(w)$ is in $\mathcal{S}$ and can be taken arbitrarily close to x. This completes the proof of (c) and hence of the proposition. ☐

5. Robustness of transversal periodic-to-periodic homoclinic orbits

In this section we consider a system depending on a parameter

$$\dot{x} = f(x, \mu) \tag{42}$$

where we assume the unperturbed system

$$\dot{x} = f(x, 0) \tag{43}$$

has a hyperbolic T-periodic orbit $u_0(t)$ with associated transversal homoclinic orbit $p_0(t)$. Our aim is to show that the perturbed system has a similar structure.

THEOREM 5.1. *Let* $f(x, \mu)$ *be a* C^3 *function. Suppose system* (43) *has a hyperbolic* T*-periodic orbit* $u_0(t)$ *and a transversal homoclinic orbit* $p_0(t)$ *to* $u_0(t)$. *Then*

(i) *for* μ *sufficiently small, the perturbed system* (42) *has a hyperbolic periodic orbit* $u(t, \mu)$ *with* $u(t, 0) = u_0(t)$;

(ii) *for* μ *sufficiently small,* (42) *has a transversal homoclinic orbit* $x(t, \mu)$ *associated with the hyperbolic periodic orbit* $u(t, \mu)$ *with* $x(t, 0) = p_0(t)$.

PROOF. (i) follows from standard theorems [10].

As in Proposition 3.5, the variational system

$$\dot{x} = f_x\bigl(u_0(t), 0\bigr)x$$

has a trichotomy on $(-\infty, \infty)$ with projections P_0, P_+, P_- and constants $K \geqslant 1$ and $\alpha > 0$.

Denote by γ^+ $(0 \leqslant \gamma^+ < T)$ the forward asymptotic phase for $p_0(t)$ and by γ^- $(0 \leqslant \gamma^- < T)$ the backward asymptotic phase. Note there exists a positive integer m such that

$$\bigl|p_0(t) - u_0(t + \gamma^+)\bigr| < \min\{\Delta, \xi_0/K\}$$

if $t \geqslant mT - \gamma^+$, where Δ and ξ_0 are as in Proposition 3.5. It follows by uniqueness that

$$p_0(t + mT - \gamma^+) = x^+(t, \xi_0, 0),$$

where $\xi_0 = P_+[p_0(mT - \gamma^+) - u_0(0)]$. Then

$$p_0(0) = x^+(-mT + \gamma^+, \xi_0, 0).$$

Similarly,

$$p_0(0) = x^-(mT + \gamma^-, \eta_0, 0)$$

where $\eta_0 = P_-[p_0(-mT - \gamma^-) - u_0(0)]$, provided that m also satisfies

$$\bigl|p_0(t) - u_0(t + \gamma^-)\bigr| < \min\{\Delta, \eta_0/K\}$$

if $t \leqslant -mT - \gamma^-$, where η_0 is as in Proposition 3.9.

Now it follows from Proposition 4.4 that the variational system

$$\dot{x} = f_x\bigl(p_0(t), 0\bigr)x \tag{44}$$

has a trichotomy on $\mathbf{R}$ with (say) projections Q_0, Q_+, Q_- where the range of Q_0 is spanned by $f(p_0(0), 0)$.

To find homoclinic orbits of (42) near $p_0(t)$, we solve the equation

$$F(\gamma, \xi, \eta, \mu) = x^+(-mT + \gamma, \xi, \mu) - x^-(mT + \gamma^-, \eta, \mu) = 0 \tag{45}$$

for $(\gamma, \xi, \eta) \in \mathbf{R} \times \mathcal{R}(P_+) \times \mathcal{R}(P_-)$ near $(\gamma^+, \xi_0, \eta_0)$, where we note that as a consequence of Proposition 3.5, F is a C^1 function. Note that

$$F(\gamma^+, \xi_0, \eta_0, 0) = p_0(0) - p_0(0) = 0.$$

We use the implicit function theorem. Write $z_0 = (\gamma^+, \xi_0, \eta_0)$. Then

$$F(z_0, 0) = 0.$$

Now we show $F_z(z_0, 0)$ is invertible. Note that if $z = (\gamma, \xi, \eta)$

$$\begin{aligned} F_z(z_0, 0)z &= \gamma x_t^+(-mT + \gamma^+, \xi_0, 0) + x_\xi^+(-mT + \gamma^+, \xi_0, 0)\xi \\ &\quad - x_\eta^-(mT + \gamma^-, \eta_0, 0)\eta \\ &= \gamma f\big(p_0(0), 0\big) + x_\xi^+(-mT + \gamma^+, \xi_0, 0)\xi - x_\eta^-(mT + \gamma^-, \eta_0, 0)\eta. \end{aligned}$$

Now it follows from Proposition 3.5 that $Z(t) = x_\xi^+(t - mT + \gamma^+, \xi_0, 0)$ is a solution of (44) restricted to $\mathcal{R}(P_+)$ such that $\|Z\|_\gamma < \infty$ and $P_+Z(mT - \gamma^+) = P_+$. It follows that the range of $x_\xi^+(-mT + \gamma^+, \xi_0, 0)$ is contained in the range of Q_+. However the condition $P_+X(mT - \gamma^+) = P_+$ implies that rank $x_\xi^+(-mT + \gamma^+, \xi_0, 0) =$ rank $P_+ =$ rank Q_+ so that the range of $x_\xi^+(-mT + \gamma^+, \xi_0, 0)$ coincides with the range of Q_+. Similarly, the range of $x_\eta^-(mT + \gamma^-, \eta_0, 0)$ coincides with the range of Q_-. Hence

$$\mathcal{R}\big(F_z(z_0, 0)\big) = \operatorname{span}\big\{f\big(p_0(0), 0\big)\big\} \oplus \mathcal{R}(Q_+) \oplus \mathcal{R}(Q_-) = \mathbf{R}^n$$

and so $F_z(z_0, 0)$ is invertible.

Then we deduce from the implicit function theorem that Eq. (45) has a C^1 solution $(\gamma(\mu), \xi(\mu), \eta(\mu))$ such that $(\gamma(0), \xi(0), \eta(0)) = (\gamma^+, \xi_0, \eta_0)$. Then

$$x(t, \mu) = \begin{cases} x^+(t - mT + \gamma(\mu), \xi(\mu), \mu) & \text{for } t \geqslant 0, \\ x^-(t + mT + \gamma^-, \eta(\mu), \mu) & \text{for } t \leqslant 0 \end{cases}$$

is a solution of (42) such that

$$\big|x(t, \mu) - u\big(t + \gamma(\mu), \mu\big)\big| \leqslant \Delta \mathrm{e}^{-\gamma(t - mT + \gamma(\mu))} \quad \text{for } t \geqslant mT - \gamma(\mu)$$

and

$$\big|x(t, \mu) - u(t + \gamma^-, \mu)\big| \leqslant \Delta \mathrm{e}^{\gamma(t + mT + \gamma^-)} \quad \text{for } t \leqslant -mT - \gamma^-.$$

So $x(t, \mu)$ is a homoclinic orbit associated with the hyperbolic periodic orbit $u_\mu(t) = u(t, \mu)$. Note also that

$$x(t, 0) = p_0(t).$$

To verify that the homoclinic orbit is transversal, note it follows from the proofs of Propositions 3.12 and 3.13 that for

$$x(0, \mu) = x^+\big(-mT + \gamma(\mu), \xi(\mu), \mu\big) = x^-\big(mT + \gamma^-, \eta(\mu), \mu\big),$$

the tangent space $T_{x(0,\mu)}W^s(u_\mu)$ is spanned by $f(x(0, \mu), \mu)$ and the range of $x^+_\xi(-mT + \gamma(\mu), \xi(\mu), \mu)$, and the tangent space $T_{x(0,\mu)}W^u(u_\mu)$ is spanned by $f(x(0, \mu), \mu)$ and the range of $x^-_\eta(mT + \gamma^-, \eta(\mu), \mu)$. Then

$$w \in T_{x(0,\mu)}W^s(u_\mu) \cap T_{x(0,\mu)}W^u(u_\mu)$$

if and only if there exist $\gamma_1 \in \mathbf{R}$, $\xi \in \mathcal{R}(P_+)$, $\gamma_2 \in \mathbf{R}$ and $\eta \in \mathcal{R}(P_-)$ such that

$$\begin{aligned} w &= \gamma_1 f\big(x(0, \mu), \mu\big) + x^+_\xi\big(-mT + \gamma(\mu), \xi(\mu), \mu\big)\xi \\ &= \gamma_2 f\big(x(0, \mu), \mu\big) + x^-_\eta\big(mT + \gamma^-, \eta(\mu), \mu\big)\eta. \end{aligned}$$

Now note that

$$\begin{aligned} &F_z\big(\gamma(\mu), \xi(\mu), \eta(\mu), \mu\big)(\gamma_1 - \gamma_2, \xi, \eta) \\ &\quad = (\gamma_1 - \gamma_2) f\big(x(0, \mu), \mu\big) + x^+_\xi\big(-mT + \gamma(\mu), \xi(\mu), \mu\big)\xi \\ &\qquad - x^-_\eta\big(mT + \gamma^-, \eta(\mu), \mu\big)\eta \\ &\quad = w - w \\ &\quad = 0. \end{aligned}$$

However, when $\mu = 0$, this derivative is $F_z(z_0, 0)$, which is invertible, and hence still invertible for μ sufficiently small. So $(\gamma_1 - \gamma_2, \xi, \eta) = 0$ and $w = \gamma_1 f(x(0, \mu), \mu)$. Thus, for μ sufficiently small,

$$T_{x(0,\mu)}W^s(u_\mu) \cap T_{x(0,\mu)}W^u(u_\mu) = \operatorname{span}\big\{f\big(x(0, \mu), \mu\big)\big\}$$

and so $x(t, \mu)$ is a transversal homoclinic orbit. Thus we have proved the theorem. □

6. Finding transversal periodic-to-periodic homoclinic orbits through regular perturbation

Even though, as we found in the previous section, transversal periodic-to-periodic homoclinic orbits persist under perturbation, it does not seem possible to write down an explicit system with an explicit transversal periodic-to-periodic homoclinic orbit. However, it turns out to be quite easy to write down an explicit system with an explicit nontransversal homoclinic orbit and obtain a transversal orbit by perturbation. To this end, as in the previous section, we consider a system depending on a parameter

$$\dot{x} = f(x, \mu). \tag{46}$$

We assume the unperturbed system

$$\dot{x} = f(x, 0) \tag{47}$$

has a hyperbolic T-periodic orbit $u_0(t)$ with associated homoclinic orbit $p_0(t)$. As in Proposition 3.5, this means that the variational system

$$\dot{x} = f_x\big(u_0(t), 0\big)x$$

has a trichotomy with projections P_0, P_+, P_- and constants K, α. Note, again as in Proposition 3.5, it follows by standard theorems that the perturbed system (46) has a periodic orbit $u(t, \mu)$ with $u(t, 0) = u_0(t)$.

Denote by γ^+ $(0 \leqslant \gamma^+ < T)$ the forward asymptotic phase for $p_0(t)$ and by γ^- $(0 \leqslant \gamma^- < T)$ the backward asymptotic phase. Then

$$p_0(0) \in W^{s,\gamma^+}(u_0) \cap W^{u,\gamma^-}(u_0).$$

What we suppose here is that $W^{s,\gamma^+}(u_0)$ and $W^{u,\gamma^-}(u_0)$ intersect in a curve C containing $p_0(0)$ which certainly implies that $p_0(t)$ cannot be transversal. So we are supposing there is a function $p(t, \beta)$ defined for all real t and β near 0 with $p(t, 0) = p_0(t)$ such that for each β, $p(t, \beta)$ is homoclinic to $u_0(t)$ with asymptotic phases γ^+ and γ^-.

THEOREM 6.1. *Let $f(x, \mu)$ be a C^4 function. Suppose system* (47) *has a hyperbolic T-periodic orbit $u_0(t)$ with associated homoclinic orbit $p_0(t)$ with asymptotic phases γ^+ and γ^- such that the variational system*

$$\dot{x} = f_x\big(p_0(t), 0\big)x \tag{48}$$

has a two-dimensional subspace of bounded solutions. Suppose $W^{s,\gamma^+}(u_0)$ and $W^{u,\gamma^-}(u_0)$ intersect in a C^1 curve $p(\beta)$ with $p(0) = p_0(0)$ and denote by $p(t, \beta)$ the solution of (47) *with $p(0, \beta) = p(\beta)$. Then*

(i) *for μ sufficiently small, the perturbed system* (46) *has a hyperbolic periodic orbit $u_\mu(t) = u(t, \mu)$ with $u(t, 0) = u_0(t)$;*

(ii) *the adjoint system*

$$\dot{x} = -f_x\big(p(t, \beta), 0\big)^* x \tag{49}$$

has, up to a scalar multiple, a unique nontrivial solution $\psi(t,\beta)$ *with* $\psi(t,\beta)\to 0$ *as* $|t|\to\infty$ *and if the Melnikov function*

$$\int_{-\infty}^{\infty}\psi^*(t,\beta)f_\mu\bigl(p(t,\beta),0\bigr)\,\mathrm{d}t$$

has a simple zero at $\beta=0$, *then for* $\mu\neq 0$ *sufficiently small, the perturbed system* (46) *has a transversal homoclinic orbit* $y(t,\mu)$ *associated with the hyperbolic periodic orbit* $u(t,\mu)$ *with* $y(t,0)=p_0(t)$.

PROOF. (i) follows from standard theorems as in Proposition 3.5.

Now, as in Proposition 3.5, the variational system

$$\dot{x}=f_x\bigl(u_0(t),0\bigr)x$$

has a trichotomy on **R** with projections P_0, P_+, P_- and constants $K\geqslant 1$ and $\alpha>0$.

As in the proof of Theorem 5.1, there exists a positive integer m such that

$$p_0(0)=x^+(-mT+\gamma^+,\xi_0,0),\quad \text{where } \xi_0=P_+\bigl[p_0(mT-\gamma^+)-u_0(0)\bigr].$$

So, in a neighbourhood of $p_0(0)$, $W^{s,\gamma^+}(u_\mu)$ can be parametrized as $x^+(-mT+\gamma^+,\xi,\mu)$, where $\xi\in\mathcal{R}(P_+)$ is near ξ_0. Similarly, in the neighbourhood of $p_0(0)$, $W^{u,\gamma^-}(u_\mu)$ can be parametrized as $x^-(mT+\gamma^-,\eta,\mu)$, where η is near

$$\eta_0=P_-\bigl[p_0(-mT-\gamma^-)-u_0(0)\bigr].$$

So we can write

$$p(\beta)=x^+\bigl(-mT+\gamma^+,\xi(\beta),0\bigr)=x^-\bigl(mT+\gamma^-,\eta(\beta),0\bigr),$$

where $\xi(\beta)$ and $\eta(\beta)$ are C^1 functions with $\xi(0)=\xi_0$ and $\eta(0)=\eta_0$.

Note that since

$$p_\beta(t,\beta)=x_\xi^+\bigl(t-mT+\gamma^+,\xi(\beta),0\bigr)\xi'(\beta)=x_\eta^-\bigl(t+mT+\gamma^-,\eta(\beta),0\bigr)\eta'(\beta),$$

it follows from Propositions 3.5 and 3.9 that $p_\beta(t,\beta)\to 0$ as $|t|\to\infty$. Next, by Proposition 3.5 again,

$$\begin{aligned}\bigl|p(t,\beta)-u_0(t+\gamma^+)\bigr|&=\bigl|x^+\bigl(t-mT+\gamma^+,\xi(\beta),0\bigr)-u_0(t-mT+\gamma^+)\bigr|\\&\leqslant\Delta\mathrm{e}^{-\gamma(t-mT+\gamma^+)}\end{aligned}\tag{50}$$

if $t\geqslant mT-\gamma^+$ and using Proposition 3.9 we get a similar inequality for $|p(t,\beta)-u_0(t+\gamma^-)|$ in $t\leqslant -mT-\gamma^-$. This implies that the norm of $p_t(t,\beta)=f(p(t,\beta))$ has a uniform upper bound and positive lower bound independent of β. In particular, this means that $p_t(t,\beta)$ and $p_\beta(t,\beta)$ are independent bounded solutions of the variational system

$$\dot{x} = f_x\bigl(p(t,\beta),0\bigr)x. \tag{51}$$

Next the last inequality above implies that for $t \geqslant t_1$,

$$\bigl|f_x\bigl(p(t,\beta),0\bigr) - f_x\bigl(u_0(t+\gamma^+),0\bigr)\bigr| \leqslant \delta(t_1),$$

where $\delta(t_1) \to 0$ as $t_1 \to \infty$ uniformly with respect to β. It follows from the roughness result in Remark 2.7 that if t_1 is large enough, not depending on β, system (51) has a trichotomy on $[t_1,\infty)$ with projections $P_0(\beta)$, $P_+(\beta)$, $P_-(\beta)$ such that $\mathcal{R}(P_0(\beta)) = \operatorname{span}\{p_t(0,\beta)\}$ and with exponent and constant independent of β. Next by Remark 2.3, Eq. (51) has a trichotomy on $[0,\infty)$ with the same projections and exponent but a different constant perhaps.

Similarly, (51) has a trichotomy on $\mathbf{R}_-$ with projections $Q_0(\beta)$, $Q_+(\beta)$, $Q_-(\beta)$. Note that the range of $P_0(\beta)$ and $Q_0(\beta)$ coincide as the span of $p_t(0,\beta)$. Also from Remark 2.8, we know that the projections $P_\sigma(\beta)$ and $Q_\sigma(\beta)$ are C^1 in β. Now the subspace of initial values of bounded solutions of (51) is

$$\mathcal{R}\bigl(P_0(\beta)+P_+(\beta)\bigr) \cap \mathcal{R}\bigl(Q_0(\beta)+Q_-(\beta)\bigr).$$

By hypothesis, when $\beta = 0$ this subspace has dimension 2. Since the projections are continuous, for small β the dimension is at most 2. However, $p_t(0,\beta)$ and $p_\beta(t,\beta)$ are linearly independent bounded solutions. So the dimension is exactly 2.

Next the adjoint equation (49) has a trichotomy on $\mathbf{R}_+$ with projections $P_0^*(\beta)$, $P_-^*(\beta)$, $P_+^*(\beta)$ and on $\mathbf{R}_-$ with projections $Q_0^*(\beta)$, $Q_-^*(\beta)$, $Q_+^*(\beta)$. The subspace of initial values at $t = 0$ of solutions $x(t)$ with $|x(t)| \to 0$ as $|t| \to \infty$ is

$$\mathcal{R}\bigl(P_-^*(\beta)\bigr) \cap \mathcal{R}\bigl(Q_+^*(\beta)\bigr) = V(\beta)^\perp,$$

where

$$V(\beta) = \mathcal{R}\bigl(P_0(\beta)+P_+(\beta)\bigr) + \mathcal{R}\bigl(Q_0(\beta)+Q_-(\beta)\bigr).$$

$V(\beta)^\perp$ is one-dimensional since the subspaces $\mathcal{R}(P_0(\beta)+P_+(\beta))$ and $\mathcal{R}(Q_0(\beta)+Q_-(\beta))$ have a two-dimensional intersection and their dimensions sum to $n+1$.

To get a C^1 basis for $V(\beta)$, we begin with bases for $\mathcal{R}(P_0(0)+P_+(0))$ and $\mathcal{R}(Q_0(0)+Q_-(0))$ containing $p_t(0,0)$ and $p_\beta(t,0)$. Then we obtain a C^1 basis for $V(\beta)$ by taking $p_t(0,\beta)$ and $p_\beta(t,\beta)$ and by applying the projections $P_0(\beta)+P_+(\beta)$ and $Q_0(\beta)+Q_+(\beta)$ to the other vectors. We assemble these vectors into an $n \times (n-1)$ matrix. Then the vector whose ith component is $(-i)^i$ times the determinant of the matrix obtained by deleting the ith row is nonzero, depends C^1 on β and is orthogonal to all the vectors in the basis. Dividing by its norm, we obtain a C^1 unit vector $\psi(\beta)$ which spans $V(\beta)^\perp$. Denote by

$\psi(t,\beta)$ the solution of the adjoint equation (49) with $\psi(0,\beta)=\psi(\beta)$. Up to a scalar multiple, it is the unique solution of (49) which tends to zero as $|t|\to\infty$.

To find homoclinic orbits near $p_0(t)$, we solve the equation

$$F(\gamma,\xi,\eta,\mu)=x^+(-mT+\gamma,\xi,\mu)-x^-(mT+\gamma^-,\eta,\mu)=0$$

for (γ,ξ,η) near (γ^+,ξ_0,η_0). Note that by Propositions 3.5 and 3.9, F is a C^2 function. Also we have

$$F\big(\gamma^+,\xi(\beta),\eta(\beta),0\big)=p(0,\beta)-p(0,\beta)=0.$$

We use the Crandall–Rabinowitz theorem [6] as given in Theorem 4.1 in [17], a suitable version of which can be stated as follows:

Let $F:E\times\mathbf{R}\mapsto G$, $(z,\mu)\mapsto F(z,\mu)$, *be a* C^2 *mapping, where* E *and* G *are Banach spaces. Suppose that there exists a* C^2 *function* $\phi(\beta)$ *defined on an interval* I *such that* $\phi'(\beta)\neq 0$ *and such that*

$$F\big(\phi(\beta),0\big)=0$$

and $L(\beta)=F_z(\phi(\beta),0)$ *is Fredholm of index zero with nullspace spanned by* $\phi'(\beta)$. *Define*

$$d(\beta)=\psi(\beta)\big(F_\mu\big(\phi(\beta),0\big)\big),$$

where $\psi(\beta)\in G^*$ *is a* C^1 *function such that* $\mathcal{N}(\psi(\beta))=\mathcal{R}(L(\beta))$. *Then if* $d(\beta)$ *has a simple zero at* β_0, *the equation*

$$F(z,\mu)=0$$

has a solution $z(\mu)$ *for* μ *sufficiently small. Moreover,* $z(0)=\phi(\beta_0)$, $z(\mu)$ *is a* C^1 *function and* $F_z(z(\mu),\mu)$ *is invertible if* $\mu\neq 0$.

We apply this theorem with $E=\mathbf{R}\times\mathcal{R}(P_+)\times\mathcal{R}(P_-)$ and $G=\mathbf{R}^n$. Write $z=(\gamma,\xi,\eta)$ and $\phi(\beta)=(\gamma^+,\xi(\beta),\eta(\beta))$. Then

$$F\big(\phi(\beta),0\big)=0.$$

Now we examine $F_z(\phi(\beta),0)$. Note if $z=(\gamma,\xi,\eta)$ that

$$\begin{aligned}F_z\big(\phi(\beta),0\big)z&=\gamma x_t^+\big(-mT+\gamma^+,\xi(\beta),0\big)+x_\xi^+\big(-mT+\gamma^+,\xi(\beta),0\big)\xi\\&\quad-x_\eta^-\big(mT+\gamma^-,\eta(\beta),0\big)\eta\\&=\gamma f\big(p(0,\beta),0\big)+x_\xi^+\big(-mT+\gamma^+,\xi(\beta),0\big)\xi\\&\quad-x_\eta^-\big(mT+\gamma^-,\eta(\beta),0\big)\eta.\end{aligned}$$

Reasoning as in the proof of Theorem 5.1, we find that

$$\mathcal{R}\big(F_z\big(\phi(\beta),0\big)\big)=\operatorname{span}\big\{f\big(p(0,\beta),0\big)\big\}\oplus\big[\mathcal{R}\big(P_+(\beta)\big)+\mathcal{R}\big(Q_-(\beta)\big)\big]$$

so that

$$\mathcal{R}\big(F_z\big(\phi(\beta),0\big)\big)=\operatorname{span}\big\{\psi(\beta)\big\}^{\perp}.$$

(Note that $\psi(\beta)$ can be regarded as the linear functional $x\mapsto\psi(\beta)^*x$ and the last subspace is its nullspace.) From this it follows that $\dim\mathcal{N}(F_z(\phi(\beta),0))=1$ and so

$$\mathcal{N}\big(F_z\big(\phi(\beta),0\big)\big)=\operatorname{span}\big\{\phi'(\beta)\big\}.$$

Next we calculate the Melnikov function

$$d(\beta)=\psi(\beta)^*F_\mu\big(\phi(\beta),0\big).$$

Now

$$F_\mu\big(\phi(\beta),0\big)=z_\mu^+\big(-mT+\gamma^+,\xi(\beta),0\big)-z_\mu^-\big(mT+\gamma^-,\eta(\beta),0\big),$$

where $z^+(t,\xi,\mu)=x^+(t,\xi,\mu)-u(t,\mu)$ and $z^-(t,\eta,\mu)=x^-(t,\eta,\mu)-u(t,\mu)$. By differentiation with respect to μ at $\mu=0$ and using the fact that $x^+(t,\xi(\beta),0)=p(t+mT-\gamma^+,\beta)$, we find that $w(t)=z_\mu^+(t-mT+\gamma^+,\xi(\beta),0)$ is a solution of

$$\dot w=f_x\big(p(t,\beta),0\big)w+h(t),$$

where

$$\begin{aligned}h(t)&=f_x\big(p(t,\beta),0\big)u_\mu(t+\gamma^+,0)-u_{\mu t}(t+\gamma^+,0)+f_\mu\big(p(t,\beta),0\big)\\&=\big[f_x\big(p(t,\beta),0\big)-f_x\big(u_0(t+\gamma^+),0\big)\big]u_\mu(t+\gamma^+,0)\\&\quad+f_\mu\big(p(t,\beta),0\big)-f_\mu\big(u_0(t+\gamma^+),0\big)\end{aligned}$$

so that, by (50), $\|h\|_\gamma<\infty$. We also know from Proposition 3.5 that $\|w\|_\gamma<\infty$. So, by Lemma 3.6,

$$\begin{aligned}w(t)=Y(t,\beta)P_+(\beta)w(0)&+\int_0^t Y(t,\beta)P_+(\beta)Y^{-1}(s,\beta)h(s)\,\mathrm{d}s\\&-\int_t^\infty Y(t,\beta)\big(I-P_+(\beta)\big)Y^{-1}(s,\beta)h(s)\,\mathrm{d}s,\end{aligned}$$

where $Y(t,\beta)$ is the fundamental matrix of (51) with $Y(0,\beta)=I$. Thus

$$z_\mu^+\big(-mT+\gamma^+,\xi(\beta),0\big)=P_+(\beta)w(0)-\int_0^\infty\big(I-P_+(\beta)\big)Y^{-1}(t,\beta)h(t)\,\mathrm{d}t$$

and so since $\psi(\beta)^*P_+(\beta)=0$

$$\psi(\beta)^*z_\mu^+\big(-mT+\gamma^+,\xi(\beta),0\big)=-\int_0^\infty\psi(t,\beta)^*h(t)\,\mathrm{d}t.$$

Similarly,

$$\psi(\beta)^* z_\mu^-\big(mT+\gamma^+,\eta(\beta),0\big)=\int_{-\infty}^{0}\psi(t,\beta)^* h(t)\,\mathrm{d}t$$

and hence

$$d(\beta)=-\int_{-\infty}^{\infty}\psi(t,\beta)^* h(t)\,\mathrm{d}t.$$

However, integrating by parts,

$$\begin{aligned}\int_{-\infty}^{\infty}\psi(t,\beta)^* u_{\mu t}(t+\gamma^+,0)\,\mathrm{d}t &= -\int_{-\infty}^{\infty}\psi_t(t,\beta)^* u_\mu(t+\gamma^+,0)\,\mathrm{d}t\\ &= \int_{-\infty}^{\infty}\psi(t,\beta)^* f_x\big(p(t,\beta),0\big)u_\mu(t+\gamma^+,0)\,\mathrm{d}t\end{aligned}$$

and so

$$d(\beta)=-\int_{-\infty}^{\infty}\psi(t,\beta)^* f_\mu\big(p(t,\beta),0\big)\,\mathrm{d}t.$$

Since, by hypothesis, $d(\beta)$ has a simple zero at $\beta=0$ it follows from the Crandall–Rabinowitz theorem as stated above that the equation

$$F(\gamma,\xi,\eta,\mu)=0$$

has a solution $(\gamma(\mu),\xi(\mu),\eta(\mu))$ such that $(\gamma(0),\xi(0),\eta(0))=(\gamma^+,\xi_0,\eta_0)$. Then

$$x(t,\mu)=\begin{cases}x^+(t-mT+\gamma(\mu),\xi(\mu),\mu) & \text{for } t\geqslant 0,\\ x^-(t+mT+\gamma^-,\eta(\mu),\mu) & \text{for } t\leqslant 0\end{cases}$$

is a solution of (46) such that

$$\big|x(t,\mu)-u\big(t+\gamma(\mu),\mu\big)\big|\leqslant \Delta\mathrm{e}^{-\beta(t-mT+\gamma(\mu))}\quad \text{for } t\geqslant mT-\gamma(\mu)$$

and

$$\big|x(t,\mu)-u(t+\gamma^-,\mu)\big|\leqslant \Delta\mathrm{e}^{\beta(t+mT+\gamma^-)}\quad \text{for } t\leqslant -mT-\gamma^-.$$

So $x(t,\mu)$ is a homoclinic orbit associated with the hyperbolic periodic orbit $u(t,\mu)$.

To verify that the homoclinic orbit is transversal when $\mu\neq 0$, we use the fact that the Crandall–Rabinowitz theorem also tells us that the operator $F_z(\gamma(\mu),\xi(\mu),\eta(\mu),\mu)$ is invertible when $\mu\neq 0$.

As in the proof of Theorem 5.1, we find that

$$w \in T_{x(0,\mu)}W^s(u_\mu) \cap T_{x(0,\mu)}W^u(u_\mu)$$

if and only if there exist $(\gamma_1, \xi) \in \mathbf{R} \times \mathcal{R}(P_+)$ and $(\gamma_2, \eta) \in \mathbf{R} \times \mathcal{R}(Q_-)$ such that

$$\begin{aligned} w &= \gamma_1 f\bigl(x(0,\mu),\mu\bigr) + x_\xi^+\bigl(-mT + \gamma(\mu), \xi(\mu), \mu\bigr)\xi \\ &= \gamma_2 f\bigl(x(0,\mu),\mu\bigr) + x_\eta^-\bigl(mT + \gamma^-, \eta(\mu), \mu\bigr)\eta \end{aligned}$$

and that then

$$F_z\bigl(\gamma(\mu), \xi(\mu), \eta(\mu), \mu\bigr)(\gamma_1 - \gamma_2, \xi, \eta) = 0.$$

However, when $\mu \neq 0$, this derivative is invertible and hence one to one. So $(\gamma_1 - \gamma_2, \xi, \eta) = 0$ and $w = \gamma_1 f(x(0,\mu),\mu)$. Thus, when $\mu \neq 0$,

$$T_{x(0,\mu)}W^s(u_\mu) \cap T_{x(0,\mu)}W^u(u_\mu) = \operatorname{span}\bigl\{f\bigl(x(0,\mu),\mu\bigr)\bigr\}$$

and so $x(t,\mu)$ is a transversal homoclinic orbit. Thus we have proved the theorem. □

EXAMPLE 6.2. Suppose $F : U \to \mathbf{R}^2$ is a C^4 vector field such that

$$\dot{x} = F(x)$$

has a hyperbolic periodic solution $r(t)$ with minimal period T and $g : \mathbf{R}^2 \to \mathbf{R}^2$ is a C^4 vector field such that 0 is a saddle point of

$$\dot{y} = g(y)$$

and $\zeta(t)$ is an associated homoclinic orbit. Then $(r(t), 0)$ is a hyperbolic periodic orbit of the system

$$\dot{x} = F(x), \qquad \dot{y} = g(y)$$

and $p(t,\beta) = (r(t), \zeta(t+\beta))$ is a nontransversal homoclinic orbit with asymptotic phases $\gamma^+ = \gamma^- = 0$ such that the subspace of bounded solutions of the variational system

$$\dot{x} = F'\bigl(r(t)\bigr)x, \qquad \dot{y} = g'\bigl(\zeta(t)\bigr)y$$

is spanned by $(r'(t), 0)$ and $(0, \zeta'(t))$.

Let $\psi(t)$ be, up to a scalar multiple, the unique bounded solution of

$$\dot{y} = -g'\bigl(\zeta(t)\bigr)^* y.$$

Then $(0, \psi(t))$ is, up to a scalar multiple, the unique solution of

$$\dot{x} = -F'\big(r(t)\big)^* x, \qquad \dot{y} = -g'\big(\zeta(t)\big)^* y$$

which tends to 0 as $|t| \to \infty$.

Then it follows from Theorem 6.1 that if the Melnikov function

$$d(\beta) = \int_{-\infty}^{\infty} \psi^*(t+\beta) g_1\big(r(t), \zeta(t+\beta), 0\big)\, \mathrm{d}t$$

has a simple zero at $\beta = 0$, then for $\mu \neq 0$ sufficiently small, the perturbed C^4 system

$$\dot{x} = F(x) + \mu F_1(x, y, \mu), \qquad \dot{y} = g(y) + \mu g_1(x, y, \mu)$$

has a transversal homoclinic orbit associated with a hyperbolic periodic orbit.

Now we give an explicit example of this situation.

EXAMPLE 6.3. The planar autonomous system $\dot{x} = F(x)$ given by

$$\dot{x}_1 = -x_1 + x_2 + \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \qquad \dot{x}_2 = -x_1 - x_2 + \frac{x_2}{\sqrt{x_1^2 + x_2^2}}$$

has the periodic solution $r(t) = (\cos t, -\sin t)$. One calculates the trace $\operatorname{Tr} DF(r(t)) = -1$. Then, by Jacobi's formula, the product of the Floquet multipliers is

$$\mathrm{e}^{\int_0^{2\pi} \operatorname{Tr} DF(r(t))\, \mathrm{d}t} = \mathrm{e}^{-2\pi}.$$

Since one Floquet multiplier is 1, the other must be $\mathrm{e}^{-2\pi}$. So $r(t)$ is hyperbolic.

Next $(0, 0)$ is a saddle for the system

$$\dot{y}_1 = y_2, \qquad \dot{y}_2 = y_1 - 2y_1^3$$

with associated saddle connexion $\zeta(t) = (\xi(t), \dot{\xi}(t))$ where $\xi(t) = \operatorname{sech} t$. Note that $\psi(t) = (\ddot{\xi}(t), -\dot{\xi}(t))$ is a bounded solution of the system adjoint to the variational equation along $\zeta(t)$.

Following Example 6.2, we apply Theorem 6.1 to the system

$$\begin{aligned}
\dot{x}_1 &= -x_1 + x_2 + \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \\
\dot{x}_2 &= -x_1 - x_2 + \frac{x_2}{\sqrt{x_1^2 + x_2^2}}, \\
\dot{y}_1 &= y_2 + \mu x_1, \\
\dot{y}_2 &= y_1 - 2y_1^3 + \mu x_2.
\end{aligned} \tag{52}$$

The Melnikov function is

$$\begin{aligned} d(\beta) &= \int_{-\infty}^{\infty} \dot{\xi}(t+\beta)\sin t + \ddot{\xi}(t+\beta)\cos t \, \mathrm{d}t \\ &= \int_{-\infty}^{\infty} \dot{\xi}(t)\sin(t-\beta) + \ddot{\xi}(t)\cos(t-\beta)\, \mathrm{d}t \\ &= \int_{-\infty}^{\infty} \dot{\xi}(t)(\sin t\cos\beta - \cos t\sin\beta) + \ddot{\xi}(t)(\cos t\cos\beta + \sin t\sin\beta)\, \mathrm{d}t \\ &= I_1\cos\beta + I_2\sin\beta, \end{aligned}$$

where

$$I_1 = \int_{-\infty}^{\infty} \dot{\xi}(t)\sin t + \ddot{\xi}(t)\cos t\, \mathrm{d}t \quad \text{and} \quad I_2 = \int_{-\infty}^{\infty} -\dot{\xi}(t)\cos t + \ddot{\xi}(t)\sin t\, \mathrm{d}t.$$

Note that $I_2 = 0$ since its integrand is odd. We calculate

$$\begin{aligned} I_1 &= \int_{-\infty}^{\infty} -\frac{\mathrm{d}}{\mathrm{d}t}\big(\dot{\xi}(t)\cos t\big) + 2\ddot{\xi}(t)\cos t\, \mathrm{d}t \\ &= 2\int_{-\infty}^{\infty} \ddot{\xi}(t)\cos t\, \mathrm{d}t \\ &= 2\bigg(\dot{\xi}(t)\cos t|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \dot{\xi}(t)(-\sin t)\, \mathrm{d}t\bigg) \\ &= 2\int_{-\infty}^{\infty} \dot{\xi}(t)\sin t\, \mathrm{d}t \\ &= 2\bigg(\xi(t)\sin t|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \xi(t)\cos t\, \mathrm{d}t\bigg) \\ &= -2\int_{-\infty}^{\infty} \xi(t)\cos t\, \mathrm{d}t \end{aligned}$$

and so

$$I_1 = -2\int_{-\infty}^{\infty} \operatorname{sech} t\cos t\, \mathrm{d}t \neq 0.$$

Hence we have shown that

$$d(\beta) = I_1\cos\beta,$$

where $I_1 \neq 0$. It is clear that this function has simple zeros and hence we may conclude that, for μ nonzero and sufficiently small, system (52) has a hyperbolic periodic orbit with associated transversal homoclinic orbit.

7. Finding transversal periodic-to-periodic homoclinic orbits through numerical shadowing

In this last section, which is based on [4], we show how transversal periodic-to-periodic homoclinic orbits can be constructed via numerical shadowing. As usual, we consider the differential equation

$$\dot{x} = f(x) \tag{53}$$

where now f is a C^2 vector field in $\mathbf{R}^n$ with

$$\left|f(x)\right| \leqslant M_0, \quad \left|f'(x)\right| \leqslant M_1, \quad \left|f''(x)\right| \leqslant M_2.$$

Denote by $\phi^t(x)$ the corresponding flow.

First we define pseudo (or approximate) orbits, which are obtained when an orbit is computed numerically.

DEFINITION 7.1. A sequence $\{w_k\}_{k=-\infty}^{+\infty}$ with associated times $\{h_k\}_{k=-\infty}^{+\infty}$ is said to be a δ *pseudo orbit* of (53) if

$$\sup_{k\in Z} h_k < \infty, \qquad \inf_{k\in Z} h_k > 0$$

and for all k

$$f(w_k) \neq 0, \qquad \left|w_{k+1} - \phi^{h_k}(w_k)\right| \leqslant \delta.$$

We look for a true orbit near a pseudo orbit. Such a true orbit is said to shadow the pseudo orbit.

DEFINITION 7.2. A sequence $\{x_k\}_{k=-\infty}^{+\infty}$ with associated times $\{t_k\}_{k=-\infty}^{+\infty}$ is said to ε *shadow* a δ pseudo orbit $\{w_k\}_{k=-\infty}^{+\infty}$ with associated times $\{h_k\}_{k=-\infty}^{+\infty}$ if for all k, $x_{k+1} = \phi^{t_k}(x_k)$ and

$$|x_k - w_k| \leqslant \varepsilon, \qquad |t_k - h_k| \leqslant \varepsilon.$$

We are interested in finding homoclinic orbits associated to periodic orbits. There are particular kinds of pseudo orbits associated with these. First we define pseudo periodic orbits.

DEFINITION 7.3. A δ pseudo orbit $\{y_k\}_{k=-\infty}^{+\infty}$ with associated times $\{\ell_k\}_{k=-\infty}^{+\infty}$ is said to be a δ *pseudo periodic orbit* with period N of (53) if for all k

$$y_{k+N} = y_k, \qquad \ell_{k+N} = \ell_k.$$

Normally we expect pseudo periodic orbits to be shadowed by discrete true periodic orbits, that is, we expect x_k and t_k to satisfy

$$x_{k+N} = x_k, \qquad t_{k+N} = t_k$$

for all k.

Next we define pseudo homoclinic orbits.

DEFINITION 7.4. Let $\{y_k\}_{k=-\infty}^{+\infty}$ be a δ pseudo periodic orbit with associated times $\{\ell_k\}_{k=-\infty}^{+\infty}$ and period N. A δ pseudo orbit $\{w_k\}_{k=-\infty}^{+\infty}$ with associated times $\{h_k\}_{k=-\infty}^{+\infty}$ is said to be a δ *pseudo homoclinic orbit* if for some τ, $0 \leqslant \tau < N$,

$$w_k = y_{k+\tau}, \quad h_k = \ell_{k+\tau} \quad \text{for } k \leqslant p, \qquad w_k = y_k, \quad h_k = \ell_k \quad \text{for } k \geqslant q$$

for some integers $p < q$.

Note that τ allows for the possibility of homoclinic orbits with different forward and backward asymptotic phases.

In preparation for the statement of the main theorem, we define a linear operator associated with a pseudo orbit. Consider a fixed pseudo orbit $\{w_k\}_{k=-\infty}^{+\infty}$ with associated times $\{h_k\}_{k=-\infty}^{+\infty}$. For each k, let Y_{w_k} be the subspace of $\mathbf{R}^n$ consisting of the vectors orthogonal to $f(w_k)$. Let Y_w be the Banach space of bounded sequences $\mathbf{v} = \{v_k\}_{k\in Z}$ with $v_k \in Y_{w_k}$, and equip Y_w with the norm

$$\|\mathbf{v}\| = \sup_{k\in Z} |v_k|.$$

Also, let $\tilde{Y}_w$ be a similar Banach space except that $v_k \in Y_{w_{k+1}}$.

Then let $L_w : Y_w \to \tilde{Y}_w$ be the linear operator defined by

$$(L_w \mathbf{v})_k = v_{k+1} - P_{w_{k+1}} D\phi^{h_k}(w_k) v_k,$$

where if $f(x) \neq 0$, the operator $P_x : \mathbf{R}^n \to \mathbf{R}^n$ is the orthogonal projection defined by

$$P_x v = v - \frac{f(x)^* v}{|f(x)|^2} f(x).$$

L_w represents the action of the derivative of the flow along the pseudo orbit projected on to the subspace orthogonal to the vector field.

Now we are ready to state the main theorem in this section, which gives conditions under which a pseudo homoclinic orbit is shadowed by a true homoclinic orbit.

THEOREM 7.5. *Suppose* $f : \mathbf{R}^n \to \mathbf{R}^n$ *is a* C^2 *vectorfield. Let* $\{y_k\}_{k=-\infty}^{+\infty}$ *be a* δ *pseudo periodic orbit with period* N *of* (53) *and let* $\{w_k\}_{k=-\infty}^{+\infty}$ *be a* δ *pseudo homoclinic* *orbit with associated times* $\{h_k\}_{k=-\infty}^{+\infty}$ *connecting* $\{y_{k+\tau}\}_{k=-\infty}^{\infty}$ *to* $\{y_k\}_{k=-\infty}^{+\infty}$ *such that* $\Delta = \inf_{k\in Z} |f(w_k)| > 0$ *and the operator* L_w *is invertible.*

Then there exist constants C *and* δ_0 *depending only on* f, $\|L_w^{-1}\|$, Δ, $h_{\max}$ *and* $h_{\min}$, *where*

$$h_{\max} = \sup_{k\in Z} h_k, \qquad h_{\min} = \inf_{k\in Z} h_k,$$

such that if $\delta < \delta_0$,

(i) *the pseudo periodic orbit* $\{y_k\}_{k=-\infty}^{+\infty}$ *is shadowed by a true discrete periodic orbit* $\{x_k\}_{k=-\infty}^{\infty}$ *such that for all* k

$$f(y_k)^*(x_k - y_k) = 0, \quad |x_k - y_k| \leqslant C\delta.$$

Moreover, the periodic orbit $\phi^t(x_0)$ *is hyperbolic*;

(ii) *The pseudo connecting orbit* $\{w_k\}_{k=-\infty}^{+\infty}$ *is shadowed by a true orbit* $\{z_k\}_{k=-\infty}^{\infty}$ *such that for all* k

$$f(w_k)^*(z_k - w_k) = 0, \quad |z_k - w_k| \leqslant C\delta.$$

Moreover, $\phi^t(z_0)$ *is hyperbolic as in Definition* 4.3 *and* $\operatorname{dist}(\phi^t(z_0), \mathrm{O}(x_0)) \to 0$ *as* $|t| \to \infty$, *where* $\mathrm{O}(x_0)$ *is the orbit of* x_0.

Furthermore, suppose there is r *such that for* $k = 0, \dots, N-1$

$$|w_r - y_k| > \bigl(\bigl|f(y_k)\bigr| + M_1 C\delta\bigr)\frac{\mathrm{e}^{M_1(h_{\max}+C\delta)} - 1}{M_1} + 2C\delta, \tag{54}$$

where $M_1 = \sup_{x\in\mathbf{R}^n} |f'(x)|$. *Then* z_r *and* x_r *lie on distinct orbits and so we may conclude that* $\phi^t(z_0)$ *is a transversal homoclinic orbit associated with the periodic orbit* $\phi^t(x_0)$.

To prove the theorem, we need the following lemmas. The first is a Newton–Kantorovich type of result.

LEMMA 7.6. *Let* X, Y *be Banach spaces and* $\mathcal{G}: X \to Y$ *a* C^2 *function. Suppose* $\mathbf{y}$ *is an element of* X *for which*

$$\bigl\|\mathcal{G}(\mathbf{y})\bigr\| \leqslant \delta$$

and $D\mathcal{G}(\mathbf{y}) = L$ *is invertible with* $\|L^{-1}\| \leqslant K$. *Set*

$$M = \sup\bigl\{\bigl\|D^2\mathcal{G}(\mathbf{x})\bigr\|: \mathbf{x} \in X,\ \|\mathbf{x} - \mathbf{y}\| \leqslant 2K\delta\bigr\}.$$

Then if

$$2MK^2\delta < 1,$$

there is a unique solution $\mathbf{x}$ *of the equation*

$$\mathcal{G}(\mathbf{x}) = \mathbf{0}$$

satisfying $\|\mathbf{x} - \mathbf{y}\| \leqslant 2K\delta$.

PROOF. Define the operator $F : X \to X$ by

$$F(\mathbf{x}) = \mathbf{y} - L^{-1}\big[\mathcal{G}(\mathbf{x}) - D\mathcal{G}(\mathbf{y})(\mathbf{x} - \mathbf{y})\big].$$

Clearly $\mathcal{G}(\mathbf{x}) = \mathbf{0}$ if and only if $F(\mathbf{x}) = \mathbf{x}$. Moreover, if $\|\mathbf{x} - \mathbf{y}\| \leqslant \varepsilon = 2K\delta$,

$$\begin{aligned} \big\|F(\mathbf{x}) - \mathbf{y}\big\| &\leqslant \|L^{-1}\| \big\|\mathcal{G}(\mathbf{x}) - \mathcal{G}(\mathbf{y}) - D\mathcal{G}(\mathbf{y})(\mathbf{x} - \mathbf{y}) + \mathcal{G}(\mathbf{y})\big\| \\ &\leqslant K\left[\frac{1}{2}M\|\mathbf{x} - \mathbf{y}\|^2 + \delta\right] \\ &\leqslant MK\varepsilon \cdot \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon \\ &< \varepsilon. \end{aligned}$$

Furthermore, if $\|\mathbf{x} - \mathbf{y}\| \leqslant \varepsilon$ and $\|\mathbf{z} - \mathbf{y}\| \leqslant \varepsilon$, then

$$\big\|F(\mathbf{x}) - F(\mathbf{z})\big\| \leqslant \|L^{-1}\| \big\|\mathcal{G}(\mathbf{x}) - \mathcal{G}(\mathbf{z}) - D\mathcal{G}(\mathbf{y})(\mathbf{x} - \mathbf{z})\big\| \leqslant KM\varepsilon\|\mathbf{x} - \mathbf{z}\|.$$

Since $KM\varepsilon < 1$, F is a contraction on the closed ball of radius ε, centre $\mathbf{y}$, and thus the lemma follows from the contraction mapping principle. ☐

We use this lemma to prove a second lemma, which gives a condition under which a pseudo orbit is shadowed by a true orbit.

LEMMA 7.7. *Let* $\{w_k\}_{k=-\infty}^{+\infty}$ *be a bounded* δ *pseudo orbit of* (53) *with associated times* $\{h_k\}_{k=-\infty}^{+\infty}$ *such that* $\Delta = \inf_{k\in Z} |f(w_k)| > 0$ *and* L_w *is invertible with* $\|L_w^{-1}\| \leqslant K$. *Then there exist constants* C, δ_0 *depending only on* f, K, Δ, $h_{\max}$ *and* $h_{\min}$, *where*

$$h_{\max} = \sup_{k\in Z} h_k, \qquad h_{\min} = \inf_{k\in Z} h_k,$$

such that if $\delta < \delta_0$ *there is a unique true orbit* $\{x_k\}_{k=-\infty}^{+\infty}$ *of* (53) *with associated times* $\{t_k\}_{k=-\infty}^{+\infty}$ *such that for* $k \in Z$

$$f(w_k)^*(x_k - w_k) = 0, \quad |x_k - w_k| \leqslant C\delta, \quad |t_k - h_k| \leqslant C\delta.$$

Moreover, the orbit $\phi^t(x_0)$ *is not an equilibrium and is hyperbolic in the sense of Definition* 4.3.

PROOF. In order to prove the existence and uniqueness of the orbit $\{x_k\}_{k=-\infty}^{+\infty}$, let X be the Banach space $Y_w \times \ell^\infty(Z, \mathbf{R})$ with the maximum norm (note: we use the Euclidean norm in $\mathbf{R}^n$). So we define

$$\big\|(\mathbf{z}, \mathbf{t})\big\| = \max\big\{\|\mathbf{z}\|, \|\mathbf{t}\|\big\},$$

for $(\mathbf{z}, \mathbf{t}) = (\{z_k\}_{k=-\infty}^{+\infty}, \{t_k\}_{k=-\infty}^{+\infty}) \in X$, where

$$\|\mathbf{z}\| = \sup_{k\in Z} |z_k|, \qquad \|\mathbf{t}\| = \sup_{k\in Z} |t_k|.$$

Then we define $\mathcal{G} : X \to \ell^\infty(Z, \mathbf{R}^n)$ by

$$\big[\mathcal{G}(\mathbf{z}, \mathbf{t})\big]_k = w_{k+1} + z_{k+1} - \phi^{t_k}(w_k + z_k).$$

We seek a solution $(\mathbf{z}, \mathbf{t})$ of the equation

$$\mathcal{G}(\mathbf{z}, \mathbf{t}) = \mathbf{0}$$

such that

$$\|\mathbf{z}\| \leqslant C\delta, \qquad \|\mathbf{t} - \mathbf{h}\| \leqslant C\delta,$$

where $\mathbf{h} = \{h_k\}_{k=-\infty}^{\infty}$ and C is to be determined. We apply Lemma 7.6. To this end, note that

$$\big\|\mathcal{G}(\mathbf{0}, \mathbf{h})\big\| \leqslant \delta.$$

Next note that the derivative of $\mathcal{G}$ at $(\mathbf{0}, \mathbf{h})$ is given for $(\mathbf{u}, \mathbf{s}) \in X$ by

$$\big[D\mathbf{G}(\mathbf{0}, \mathbf{h})(\mathbf{u}, \mathbf{s})\big]_k = u_{k+1} - f\big(\phi^{h_k}(w_k)\big)s_k - D\phi^{h_k}(w_k)u_k.$$

We approximate this operator by the operator T given by

$$\big[T(\mathbf{u}, \mathbf{s})\big]_k = u_{k+1} - f(w_{k+1})s_k - D\phi^{h_k}(w_k)u_k.$$

First we show that T is invertible. To do this, we must show that for all $\mathbf{g} = \{g_k\}_{k=-\infty}^{+\infty}$ in $\ell^\infty(Z, \mathbf{R}^n)$, there is a unique solution of the equation

$$T(\mathbf{u}, \mathbf{s}) = \mathbf{g},$$

that is,

$$u_{k+1} - f(w_{k+1})s_k - D\phi^{h_k}(w_k)u_k = g_k, \quad k \in Z. \tag{55}$$

We can solve (55) for s_k by multiplying both sides by $f(w_{k+1})^*$ to get

$$s_k = -\frac{f(w_{k+1})^*}{|f(w_{k+1})|^2}\bigl\{D\phi^{h_k}(w_k)u_k + g_k\bigr\}, \quad k \in Z \tag{56}$$

and we obtain the estimate

$$\|\mathbf{s}\| \leqslant \Delta^{-1}\bigl[\mathrm{e}^{M_1 h_{\max}}\|\mathbf{u}\| + \|\mathbf{g}\|\bigr], \tag{57}$$

where

$$M_1 = \sup_{x\in\mathbf{R}^n}\bigl|f'(x)\bigr|.$$

Substituting (56) in (55) and rearranging, we obtain the equation

$$(L_w\mathbf{u})_k = \bar{g}_k, \tag{58}$$

where $\mathbf{u} = \{u_k\}_{k=-\infty}^{\infty} \in Y_w$ and the sequence $\bar{\mathbf{g}} = \{\bar{g}_k\}_{k=-\infty}^{\infty}$ is given by

$$\bar{g}_k = g_k - \frac{f(w_{k+1})^* g_k}{|f(w_{k+1})|^2} f(w_{k+1}) = P_{w_{k+1}} g_k, \quad k \in Z,$$

and so is in $\tilde{Y}_w$. Since L_w is invertible, we may solve (58) to get

$$\mathbf{u} = L_w^{-1}\bar{\mathbf{g}}. \tag{59}$$

So (55) has a unique solution $(\mathbf{u}, \mathbf{s})$ given by (59) and (56) with

$$\|\mathbf{u}\| \leqslant \|L_w^{-1}\|\|\bar{\mathbf{g}}\| \leqslant K\|\mathbf{g}\|$$

and $\|\mathbf{s}\|$ satisfying the inequality in (57). It follows that T is invertible with

$$\|T^{-1}\| \leqslant C_1 = \max\bigl\{\Delta^{-1}(\mathrm{e}^{M_1 h_{\max}}K + 1), K\bigr\}.$$

Now observe that

$$\bigl\|D\mathcal{G}(\mathbf{0}, \mathbf{h}) - T\bigr\| \leqslant M_1\delta.$$

Then if

$$M_1 C_1 \delta \leqslant 1/2,$$

it follows that $D\mathcal{G}(\mathbf{0}, \mathbf{h})$ is also invertible and

$$\bigl\|D\mathcal{G}(\mathbf{0}, \mathbf{h})^{-1}\bigr\| \leqslant C/2,$$

where

$$C = 4C_1.$$

Next if $\mathbf{v} = (\{z_k\}_{k=-\infty}^{+\infty}, \{t_k\}_{k=-\infty}^{+\infty})$, $\mathbf{u} = (\{\tau_k\}_{k=-\infty}^{\infty}, \{\xi_k\}_{k=-\infty}^{\infty})$, and $\bar{\mathbf{u}} = (\{\sigma_k\}_{k=-\infty}^{\infty}, \{\eta_k\}_{k=-\infty}^{\infty})$, one calculates, for $k = 0, \ldots, N$, that

$$\begin{aligned}\left[D^2\mathcal{G}(\mathbf{v})\mathbf{u}\bar{\mathbf{u}}\right]_k = {} & -\tau_k\sigma_k f'\left(\varphi^{t_k}(w_k + z_k)\right) f\left(\varphi^{t_k}(w_k + z_k)\right) \\ & - \tau_k f'\left(\varphi^{t_k}(w_k + z_k)\right) D\varphi^{t_k}(w_k + z_k)\eta_k \\ & - \sigma_k f'\left(\varphi^{t_k}(w_k + z_k)\right) D\varphi^{t_k}(w_k + z_k)\xi_k \\ & - D^2\varphi^{t_k}(w_k + z_k)\xi_k\eta_k.\end{aligned}$$

We use the following estimates, which follow easily from Gronwall's lemma and the variation of constants formula

$$\left\|D\varphi^t(\mathbf{x})\right\| \leqslant \mathrm{e}^{M_1 t}, \qquad \left\|D^2\varphi^t(\mathbf{x})\right\| \leqslant M_2 t \mathrm{e}^{2M_1 t}$$

where

$$M_2 = \sup_{x \in \mathbf{R}^n} \left|f''(x)\right|.$$

It follows that

$$\left\|D^2\mathcal{G}(\mathbf{z}, \mathbf{t})\right\| \leqslant M_3, \tag{60}$$

where

$$M_3 = M_0 M_1 + 2M_1 \mathrm{e}^{M_1(h_{\max} + C\delta)} + M_2(h_{\max} + C\delta)\mathrm{e}^{2M_1(h_{\max} + C\delta)}.$$

Then provided

$$M_3 C^2 \delta < 2,$$

Lemma 7.6 can be applied and we get a unique solution $(\mathbf{z}, \mathbf{t})$ of $\mathcal{G}(\mathbf{z}, \mathbf{t}) = 0$ with $\|\mathbf{z}\| \leqslant C\delta$, $\|\mathbf{t}\| \leqslant C\delta$. Then with $x_k = w_k + z_k$, the first part of the lemma follows.

In the rest of the proof we also assume that

$$M_1 C\delta < \Delta.$$

Then we observe that

$$\left|f(x_k)\right| \geqslant \left|f(w_k)\right| - \left|f(w_k) - f(x_k)\right| \geqslant \Delta - M_1 C\delta > 0$$

so that the orbit $\phi^t(x_0)$ is not an equilibrium.

Now we show that $\phi^t(x_0)$ is hyperbolic. First note that the derivative of $\mathcal{G}$ at $(\mathbf{z},\mathbf{t})$, where $\mathbf{z}=\{x_k-w_k\}_{k=-\infty}^{\infty}$, is given for $(\mathbf{u},\mathbf{s})\in X$ by

$$\big[D\mathbf{G}(\mathbf{z},\mathbf{t})(\mathbf{u},\mathbf{s})\big]_k=u_{k+1}-f(x_{k+1})s_k-D\phi^{t_k}(x_k)u_k.$$

In view of (60),

$$\big\|D\mathbf{G}(\mathbf{z},\mathbf{t})-D\mathbf{G}(\mathbf{0},\mathbf{h})\big\|\leqslant M_3\big\|(\mathbf{z},\mathbf{t})-(\mathbf{0},\mathbf{h})\big\|\leqslant M_3C\delta.$$

Since $\|D\mathbf{G}(\mathbf{0},\mathbf{h})^{-1}\|\leqslant C/2$ and $M_3C\delta(C/2)<1$, it follows that $D\mathbf{G}(\mathbf{z},\mathbf{t})$ is invertible. In particular, this implies that for all bounded sequences $g_k\in Y_{w_{k+1}}$, the equation

$$v_{k+1}-f(x_{k+1})s_k-D\phi^{t_k}(x_k)v_k=g_k,\quad k\in Z \tag{61}$$

has a unique bounded solution $v_k\in Y_{w_k}$ and $s_k\in\mathbf{R}$. Note that

$$f(w_{k+1})^*f(x_{k+1})\geqslant\big|f(w_{k+1})\big|\big[\big|f(w_{k+1})\big|-M_1C\delta\big]>0.$$

Multiplying (61) by $f(w_{k+1})^*$, we obtain

$$s_k=-\frac{f(w_{k+1})^*D\phi^{t_k}(x_k)v_k}{f(w_{k+1})^*f(x_{k+1})},\quad k\in Z. \tag{62}$$

Then substituting this back in (61), we obtain

$$v_{k+1}=P_{x_{k+1}}D\phi^{t_k}(x_k)v_k+g_k,\quad k\in Z.$$

The upshot is that for all bounded sequences $g_k\in Y_{w_{k+1}}$, the equation

$$v_{k+1}-P_{x_{k+1}}D\phi^{t_k}(x_k)v_k=g_k \tag{63}$$

has a bounded solution $v_k\in Y_{w_k}$. Moreover, it must be unique for if v_k is a bounded solution of (63) and we define s_k by (62), then we see that s_k and v_k form a bounded solution of (61). So it follows that the difference equation

$$v_{k+1}=P_{x_{k+1}}D\phi^{t_k}(x_k)v_k \tag{64}$$

has an exponential dichotomy on Z (see [12,24]).

Now we want to apply Proposition 2.9 with $A(t)=f'(\phi^t(x_0))$ so that $X(t)=D\phi^t(x_0)$, $x_0(t)=f(\phi^t(x_0))$ and with $Q_k=P_{x_k}$ and s_k defined recursively by $s_0=0$ and $s_{k+1}=s_k+t_k$. Now, provided that

$$C\delta<h_{\min},$$

we have

$$0 < h_{\min} - C\delta \leqslant t_k \leqslant h_{\max} + C\delta < \infty.$$

Also note that $|f(\phi^t(x_0))| \leqslant M_0$ for all t. Next given any real t there exists k such that $s_k \leqslant t < s_{k+1}$. Then

$$\begin{aligned}\left|f(x_k)\right| = \left|f\left(\phi^{s_k}(x_0)\right)\right| &= \left|D\phi^{s_k-t}\left(\phi^t(x_0)\right)f\left(\phi^t(x_0)\right)\right| \\ &\leqslant \mathrm{e}^{M_1(h_{\max}+C\delta)}\left|f\left(\phi^t(x_0)\right)\right|\end{aligned}$$

so that since $0 < \Delta - M_1C\delta \leqslant |f(x_k)|$ it follows that for all t

$$\left|f\left(\phi^t(x_0)\right)\right| \geqslant \mathrm{e}^{-M_1(h_{\max}+C\delta)}(\Delta - M_1C\delta) > 0.$$

Then, since (64) can be written as $v_{k+1} = Q_{k+1}X(s_{k+1})X^{-1}(s_k)v_k$ with $v_k \in \mathcal{R}(Q_k)$, it follows from Proposition 2.9 that

$$\dot{x} = f'\left(\phi^t(x_0)\right)x$$

has a trichotomy and so by Proposition 4.4 the solution $\phi^t(x_0)$ is hyperbolic.

This completes the proof of the lemma with δ_0 as the largest $\delta > 0$ such that

$$M_3C^2\delta \leqslant 2, \quad M_1C_1\delta \leqslant 1/2, \quad M_1C\delta \leqslant \Delta, \quad C\delta \leqslant h_{\min}. \tag{65}$$

□

In the next lemma we show under a uniqueness condition that an orbit which shadows a pseudo homoclinic orbit is asymptotic to it.

LEMMA 7.8. *Let* $\{w_k\}_{k=-\infty}^{+\infty}$ *be a pseudo orbit of* (53) *with associated times* $\{h_k\}_{k=-\infty}^{+\infty}$ *and suppose it is* ε *shadowed by a true orbit* $\{z_k\}_{k=-\infty}^{+\infty}$ *with associated times* $\{s_k\}_{k=-\infty}^{+\infty}$ *such that for all* k

$$f(w_k)^*(z_k - w_k) = 0.$$

Next let $\{y_k\}_{k=-\infty}^{+\infty}$ *be a pseudo periodic orbit with associated times* $\{\ell_k\}_{k=-\infty}^{+\infty}$ *and suppose it is* ε *shadowed by a unique true orbit* $\{x_k\}_{k=-\infty}^{+\infty}$ *with associated times* $\{t_k\}_{k=-\infty}^{+\infty}$ *such that for all* k

$$f(y_k)^*(x_k - y_k) = 0.$$

Suppose also that

$$w_k = y_k, \quad h_k = \ell_k \quad \textit{for } k \geqslant q.$$

Then

$$|z_k - x_k| + |s_k - t_k| \to 0 \quad \text{as } k \to \infty.$$

PROOF. Let $\{y_k\}_{k=-\infty}^{+\infty}$ have period N. Since for all k

$$f(y_k)^*(x_{k+N} - y_k) = f(y_{k+N})^*(x_{k+N} - y_{k+N}) = 0,$$

and

$$|x_{k+N} - y_k| = |x_{k+N} - y_{k+N}| \leqslant \varepsilon, \qquad |t_{k+N} - \ell_k| = |t_{k+N} - \ell_{k+N}| \leqslant \varepsilon$$

also, it follows by uniqueness that $x_{k+N} = x_k$ and $t_{k+N} = t_k$ for all k.

Define for $j \geqslant 0$ and all k

$$z_k^{(j)} = z_{k+jN}, \qquad s_k^{(j)} = s_{k+jN}.$$

We claim that for any fixed k

$$z_k^{(j)} \to x_k, \qquad s_k^{(j)} \to t_k$$

as $j \to \infty$. Suppose not. Then there exists $i, \eta > 0$ and a sequence $j_r \to \infty$ as $r \to \infty$ such that

$$|z_i^{(j_r)} - x_i| + |s_i^{(j_r)} - t_i| \geqslant \eta \tag{66}$$

for all r. By taking a further subsequence if necessary, we can assume that for each k

$$z_k^{(j_r)} \to \tilde{x}_k, \qquad s_k^{(j_r)} \to \tilde{t}_k$$

as $r \to \infty$. Now, for all k and r, $\phi^{s_k^{(j_r)}}(z_k^{(j_r)}) = z_{k+1}^{(j_r)}$ and for fixed k, if r is large enough,

$$|z_k^{(j_r)} - y_k| = |z_k^{(j_r)} - y_{k+j_rN}| = |z_{k+j_rN} - w_{k+j_rN}| \leqslant \varepsilon$$

and, similarly,

$$|s_k^{(j_r)} - \ell_k| \leqslant \varepsilon.$$

Also for fixed k, if r is large enough,

$$f(y_k)^*(z_k^{(j_r)} - y_k) = 0.$$

So, letting $r \to \infty$, we obtain that for all k, $\phi^{\tilde{t}_k}(\tilde{x}_k) = \tilde{x}_{k+1}$ and

$$|\tilde{x}_k - y_k| \leqslant \varepsilon, \quad |\tilde{t}_k - \ell_k| \leqslant \varepsilon, \quad f(y_k)^*(\tilde{x}_k - y_k) = 0.$$

By uniqueness, it follows that for all k, $\tilde{x}_k = x_k$ and $\tilde{t}_k = t_k$. Hence for each k

$$z_k^{(j_r)} \to x_k, \qquad s_k^{(j_r)} \to t_k$$

as $r \to \infty$. This contradicts (66). So we deduce that for each k

$$z_{k+jN} = z_k^{(j)} \to x_k \quad \text{and} \quad s_{k+jN} = s_k^{(j)} \to t_k \quad \text{as } j \to \infty.$$

It follows that

$$z_k - x_k \to 0 \quad \text{and} \quad s_k - t_k \to 0 \quad \text{as } k \to \infty. \qquad \square$$

PROOF OF THEOREM 7.5. We take δ_0 to be the largest δ satisfying the inequalities in (65) with $K = \|L_w^{-1}\|$.

Now let Y_y be the Banach space of bounded sequences $\mathbf{v} = \{v_k\}_{k\in Z}$ with $v_k \in \mathbf{R}^n$ and orthogonal to $f(y_k)$, Y_y being equipped with the norm

$$\|\mathbf{v}\| = \sup_{k\in Z} |v_k|$$

and let $\tilde{Y}_y$ be a similar Banach space but with v_k orthogonal to $f(y_{k+1})$. Then let $L_y : Y_y \to \tilde{Y}_y$ be the linear operator defined by

$$(L_y\mathbf{v})_k = v_{k+1} - P_{y_{k+1}} D\phi^{\ell_k}(y_k)v_k,$$

where $\{\ell_k\}_{k=-\infty}^{+\infty}$ are the times associated with $\{y_k\}_{k=-\infty}^{+\infty}$. By hypothesis, L_w is invertible. We now show that L_y is invertible and that

$$\|L_y^{-1}\| \leqslant \|L_w^{-1}\|.$$

Let $\mathbf{g} = \{g_k\}_{-\infty}^{\infty} \in \tilde{Y}_y$ and r a natural number. Then, since L_w is invertible, the difference equation

$$v_{k+1} = P_{w_{k+1}} D\phi^{h_k}(w_k)v_k + P_{w_{k+1}} g_{k-rN}, \quad k \in Z$$

has a solution $v_k^{(r)} \in Y_{w_k}$ such that for all k

$$|v_k^{(r)}| \leqslant \|L_w^{-1}\| \|\mathbf{g}\|.$$

Take

$$u_k^{(r)} = v_{k+rN}^{(r)}.$$

Then $u_k^{(r)} \in Y_{w_{k+rN}}$ is a solution of

$$u_{k+1}^{(r)} = P_{w_{k+rN+1}} D\phi^{h_{k+rN}}(w_{k+rN}) u_k^{(r)} + P_{w_{k+rN+1}} g_k, \quad k \in Z \tag{67}$$

such that for all k

$$|u_k^{(r)}| \leqslant \|L_w^{-1}\| \|\mathbf{g}\|. \tag{68}$$

Next, by Cantor's diagonalization procedure, we can find a subsequence $u_k^{(j_r)} \to \bar{u}_k$, where $\bar{u}_k$ is orthogonal to $f(y_k)$, as $r \to \infty$ for each k. Then letting $r \to \infty$ in (67) and (68) (with r replaced by j_r), we obtain

$$\bar{u}_{k+1} = P_{y_{k+1}} D\phi^{\ell_k}(y_k)\bar{u}_k + g_k, \quad k \in Z$$

and

$$|\bar{u}_k| \leqslant \|L_w^{-1}\| \|\mathbf{g}\| \tag{69}$$

for all k. In particular, since $\mathbf{g}$ is arbitrary, this means that the difference equation

$$u_{k+1} = P_{y_{k+1}} D\phi^{\ell_k}(y_k) u_k + g_k$$

has a solution u_k bounded on $[0, \infty)$, where u_k is orthogonal to $f(y_k)$, whenever g_k is a sequence bounded on $[0, \infty)$ with g_k orthogonal to $f(y_{k+1})$. It follows from the discrete analogue of Proposition 3 of Lecture 3 in [5] that the difference equation

$$u_{k+1} = P_{y_{k+1}} D\phi^{\ell_k}(y_k) u_k,$$

where u_k is orthogonal to $f(y_k)$, has an exponential dichotomy on $[0, \infty)$. However, then it follows from the discrete analogue of the periodic case of Proposition 3 in Lecture 8 of [5] that the equation has an exponential dichotomy on Z. This, in turn, means that L_y is invertible so that the equation

$$u_{k+1} = P_{y_{k+1}} D\phi^{\ell_k}(y_k) u_k + g_k, \quad k \in Z$$

has the unique bounded solution $(L_y^{-1}\mathbf{g})_k$ orthogonal to $f(y_k)$. Thus $\bar{u}_k = (L_y^{-1}\mathbf{g})_k$ and it follows from (69) that

$$\|L_y^{-1}\| \leqslant \|L_w^{-1}\|.$$

Then it follows from Lemma 7.7 applied to $\{y_k\}_{k=-\infty}^{+\infty}$ and $\{\ell_k\}_{k=-\infty}^{+\infty}$ (note that $h_{\min} \leqslant \ell_k \leqslant h_{\max}$ and that $|f(y_k)| \geqslant \Delta$) that provided $\delta < \delta_0$ there is a unique orbit $\{x_k\}_{k=-\infty}^{+\infty}$ of (53) with associated times $\{t_k\}_{k=-\infty}^{+\infty}$ such that for all k

$$f(y_k)^*(x_k - y_k) = 0, \quad |x_k - y_k| \leqslant C\delta, \quad |t_k - \ell_k| \leqslant C\delta.$$

However, since for all k

$$f(y_k)^*(x_{k+N} - y_k) = f(y_{k+N})^*(x_{k+N} - y_{k+N}) = 0,$$

and

$$|x_{k+N} - y_k| = |x_{k+N} - y_{k+N}| \leqslant C\delta, \qquad |t_{k+N} - \ell_k| = |t_{k+N} - \ell_{k+N}| \leqslant C\delta$$

also, it follows by uniqueness that $x_{k+N} = x_k$ and $t_{k+N} = t_k$ for all k. Hence $\phi^t(x_0)$ has period $t_0 + t_1 + \cdots + t_{N-1}$. Also note it follows from Lemma 7.7 that x_0 is not an equilibrium and that $\phi^t(x_0)$ is hyperbolic.

Next, by Lemma 7.7 applied to $\{w_k\}_{k=-\infty}^{+\infty}$ and $\{h_k\}_{k=-\infty}^{+\infty}$, provided $\delta < \delta_0$, there is a unique orbit $\{z_k\}_{k=-\infty}^{\infty}$ of (53) with associated times $\{s_k\}_{k=-\infty}^{+\infty}$ such that for all k

$$f(w_k)^*(z_k - w_k) = 0, \quad |z_k - w_k| \leqslant C\delta, \quad |s_k - h_k| \leqslant C\delta.$$

Again $\phi^t(z_0)$ is hyperbolic.

Next it follows from Lemma 7.8 that

$$z_k - x_k \to 0 \quad \text{and} \quad s_k - t_k \to 0 \quad \text{as } k \to \infty.$$

Now for all k

$$s_k = h_k + s_k - h_k \geqslant h_{\min} - C\delta > 0.$$

We denote by $O(x_0)$ the orbit of x_0. Then, if $0 \leqslant t \leqslant s_k$,

$$\begin{aligned} \operatorname{dist}\bigl(\phi^t(z_k), O(x_0)\bigr) &\leqslant \bigl|\phi^t(z_k) - \phi^t(x_k)\bigr| \\ &\leqslant \mathrm{e}^{M_1 t}|z_k - x_k| \\ &\leqslant \mathrm{e}^{M_1(h_{\max}+C\delta)}|z_k - x_k|. \end{aligned}$$

It follows that $\operatorname{dist}(\phi^t(z_0), O(x_0)) \to 0$ as $t \to \infty$.

Similarly, we prove that $\operatorname{dist}(\phi^t(z_0), O(x_0)) \to 0$ as $t \to -\infty$.

Finally to show that z_0 does not lie on the periodic orbit, we first note that for $0 \leqslant t \leqslant t_k$

$$\phi^t(x_k) - x_k = \int_0^t f\bigl(\phi^s(x_k)\bigr)\,\mathrm{d}s$$

so that

$$\bigl|\phi^t(x_k) - x_k\bigr| \leqslant \bigl|f(x_k)\bigr|t + \int_0^t \bigl|f\bigl(\phi^s(x_k)\bigr) - f(x_k)\bigr|\,\mathrm{d}s$$

and hence for $0 \leqslant t \leqslant t_k$

$$\left|\phi^t(x_k) - x_k\right| \leqslant \left|f(x_k)\right|t + M_1 \int_0^t \left|\phi^s(x_k) - x_k\right| \mathrm{d}s.$$

Then it follows from Gronwall's lemma that for $0 \leqslant t \leqslant t_k$

$$\left|\phi^t(x_k) - x_k\right| \leqslant \left|f(x_k)\right| \frac{\mathrm{e}^{M_1 t} - 1}{M_1}.$$

Then for $0 \leqslant t \leqslant t_k$, $k = 0, \dots, N-1$,

$$\begin{aligned}\left|z_r - \phi^t(x_k)\right| &\geqslant |w_r - y_k| - \left|\phi^t(x_k) - x_k\right| - |x_k - y_k| - |w_r - z_r| \\ &> \left(\left|f(y_k)\right| + M_1 C\delta\right) \frac{\mathrm{e}^{M_1(h_{\max} + C\delta)} - 1}{M_1} + 2C\delta \\ &\quad - \left|f(x_k)\right| \frac{\mathrm{e}^{M_1 t} - 1}{M_1} - 2C\delta \\ &\geqslant 0,\end{aligned}$$

where we have used (54). Hence z_r does not lie on the periodic orbit. So the point z_0 is indeed a transversal homoclinic point to the periodic orbit $\phi^t(x_0)$, where we note that the transversality follows from the hyperbolicity.

Thus the proof of the theorem is completed. □

EXAMPLE 7.9. In [4], Theorem 7.5 (with some refinement of the estimates) is used to show that the *Lorenz Equations*

$$\begin{aligned}\dot{x}_1 &= \sigma(x_2 - x_1), \\ \dot{x}_2 &= \rho x_1 - x_2 - x_1 x_3, \\ \dot{x}_3 &= x_1 x_2 - \alpha x_3\end{aligned}$$

with the classic parameter values

$$\sigma = 10.0, \quad \rho = 28.0, \quad \beta = 8/3$$

possess a hyperbolic periodic orbit with an associated transversal homoclinic orbit.

References

[1] F. Battelli and K.J. Palmer, *Transversal periodic-to-periodic homoclinic orbits in singularly perturbed systems*, in preparation.

[2] W.-J. Beyn, *On well-posed problems for connecting orbits in dynamical systems*, Chaotic Dynamics, P. Kloeden and K. Palmer, eds., Contemporary Mathematics, Vol. 172, Amer. Math. Soc., Providence, RI (1994), 131–168.

[3] G. Birkhoff, *Nouvelles recherches sur les systèmes dynamiques*, Mem. Pont. Acad. Sci. Novi. Lyncaei **1** (1935), 85–216.

[4] B.A. Coomes, H. Koçak and K.J. Palmer, *Transversal connecting orbits from shadowing*, Numer. Math. **106** (2007), 427–469.

[5] W.A. Coppel, *Dichotomies in Stability Theory*, Lecture Notes in Math., Vol. 629, Springer-Verlag, Berlin (1978).

[6] M.G. Crandall and P.H. Rabinowitz, *Bifurcation from simple eigenvalues*, J. Funct. Anal. **8** (1971), 321–340.

[7] B. Deng and K. Sakamoto, *Sil'nikov–Hopf bifurcation*, J. Differential Equations **119** (1995), 1–23.

[8] R.L. Devaney, *An Introduction to Chaotic Dynamical Systems*, second ed., Addison–Wesley, Redwood City (1989).

[9] L. Dieci and J. Rebaza, *Point-to-periodic and periodic-to-periodic connections*, Bit Numer. Math. **44** (2004), 41–62. Erratum **45**, 617–618.

[10] J.K. Hale, *Ordinary Differential Equations*, Wiley-Interscience (1969).

[11] J.K. Hale and X.B. Lin, *Symbolic dynamics and nonlinear semiflows*, Ann. Mat. Pura Appl. **144** (1986), 229–259.

[12] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Math., Vol. 840, Springer-Verlag, New York (1981).

[13] R.A. Johnson and G.R. Sell, *Smoothness of spectral subbundles and reducibility of quasiperiodic linear differential systems*, J. Differential Equations **41** (1981), 262–288.

[14] U. Kirchgraber and D. Stoffer, *Possible chaotic motion of comets in the Sun–Jupiter system – a computer-assisted approach based on shadowing*, Nonlinearity **17** (2004), 281–300.

[15] V.K. Melnikov, *On the stability of the center for time periodic solutions*, Trans. Moscow Math. Soc. **12** (1963), 3–52.

[16] K.J. Palmer, *Exponential separation, exponential dichotomy and spectral theory for linear systems of ordinary differential equations*, J. Differential Equations **46** (1982), 324–345.

[17] K.J. Palmer, *Exponential dichotomies and transversal homoclinic points*, J. Differential Equations **55** (1984), 225–256.

[18] K.J. Palmer, *Transverse heteroclinic orbits and Cherry's example of a nonintegrable Hamiltonian system*, J. Differential Equations **65** (1986), 321–360.

[19] K.J. Palmer, *Shadowing in Dynamical Systems*, Kluwer Academic Publishers, Dordrecht (2000).

[20] T. Pampel, *Numerical approximation of connecting orbits with asymptotic rate*, Numer. Math. **90** (2001), 309–348.

[21] H. Poincaré, *Sur le problème des trois corps et les équations de la dynamique*, Acta Math. **13** (1890), 1–270.

[22] L.P. Sil'nikov, *On a Poincaré–Birkhoff problem*, Math. USSR-Sb. **3** (1967), 353–371.

[23] L.P. Sil'nikov, *Homoclinic orbits: since Poincaré till today*, Preprint, ISSN 0946-8633, Weierstrass-Institut für Angewandte Analysis und Stochastik (2000).

[24] V.E. Slyusarchuk, *Exponential dichotomy for solutions of discrete systems*, Ukrain. Math. J. **35** (1983), 98–103.

[25] S. Smale, *Diffeomorphisms with many periodic points*, Differential and Combinatorial Topology, Princeton University Press (1965).

[26] P. Szmolyan, *Transversal heteroclinic and homoclinic orbits in singular perturbation problems*, J. Differential Equations **92** (1991), 252–281.

This page intentionally left blank

CHAPTER 5

Successive Approximation Techniques in Non-Linear Boundary Value Problems for Ordinary Differential Equations

A. Rontó*

Institute of Mathematics, Academy of Sciences of Czech Republic, Žižkova 22, CZ-61662 Brno, Czech Republic

M. Rontó†

Department of Analysis, University of Miskolc, 3515 Miskolc-Egyetemváros, Hungary
E-mail: *matronto@gold.uni-miskolc.hu*

Contents

1. Notation 443
2. Introduction 443
3. Periodic successive approximations for non-autonomous systems 446
    3.1. Properties of the limit function 461
    3.2. Existence theorems for periodic solutions 470
4. Successive approximations for autonomous systems 487
    4.1. Non-linear substitution of variables 490
    4.2. Linear substitution of variables 501
5. Periodic solutions of differential systems with symmetries 509
    5.1. Symmetry properties of periodic functions 509
    5.2. Assumptions 517
    5.3. Auxiliary statements. The operator H_σ 519
    5.4. Successive approximations and their convergence 523
    5.5. Case of "large" Lipschitz constants 529
    5.6. Symmetries (τ, E) of solutions of the integral equation *(5.58)* 532
    5.7. Periodic solutions with property (τ, E) 535
6. Two-point boundary value problem with non-linear non-separated conditions 539

---

[*] The research of the first author was supported in part by AS CR, Institutional Research Plan No. AV0Z10190503, and by GA CR, Grant No. 201/06/0254.

[†] The research of the second author was partially supported by the Hungarian Scientific Research Fund OTKA through Grant No. K68311.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 4
Edited by F. Battelli and M. Fečkan
© 2008 Elsevier B.V. All rights reserved

6.1. Reduction to a problem with linear conditions 539
6.2. A separated non-linear boundary condition 547
7. A non-linear problem with another type of separated boundary conditions 551
7.1. Subsidiary lemmata 551
7.2. Successive approximations for a problem with separated boundary conditions 556
8. Parametrisation method for three-point boundary value problems 567
8.1. Problem setting 568
8.2. Transformation to a two-point problem with a parameter in boundary conditions 569
8.3. Convergence of successive approximations 570
8.4. Practical realisation 576
8.5. Example of a two-dimensional three-point problem 576
9. Historical remarks 584
10. Exercises 587
References 588

Abstract

In this work, we investigate the solvability and the approximate construction of solutions of certain types of regular non-linear boundary value problems for systems of ordinary differential equations on a compact interval. For this purpose, we construct analytically a uniformly convergent parametrised sequence of functions depending on the properties of the concrete boundary conditions and non-linearities in the given systems. The value of the parameter introduced artificially into the scheme is to be determined by solving a certain system of algebraic or transcendental equations.

The text is divided into 10 sections. Sections 1 and 2 contain the notation used in what follows and provide a short introduction. In Section 3, the successive approximation techniques are treated for the investigation of periodic solutions of non-autonomous periodic systems. In Section 4, we apply the method for the study the periodic solutions of autonomous systems by using the appropriate reduction to a non-autonomous system. In Section 5, we establish conditions under which a system of non-linear non-autonomous ordinary differential equations has a family of solutions that are periodic with a common period and possess a certain symmetry property. Sections 6 and 7 deal with the investigation of non-linear two-point boundary value conditions by using a parametrisation that leads one to a family of problem with linear two-point conditions considered together with certain additional algebraic or transcendental equations with respect to certain parameters. In Section 8, we use the parametrisation approach to study some three-point non-linear boundary value problem which, as a result, can be investigated through auxiliary two-point problems. Most of theoretical results are illustrated by examples. Section 9 contains some historical remarks concerning the development and application of the method. Finally, in Section 10, we give several exercises concerning the successive approximation technique under consideration.

1. Notation

(1) $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R}_+ = [0, \infty)$, $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$, $\mathbb{N} = \{1, 2, \ldots\}$.

(2) C_T^n is the Banach space of continuous vector-functions $x = (x_k)_{k=1}^n : (-\infty, \infty) \to \mathbb{R}^n$ periodic with period T (i.e., $x(t) = x(t+T)$ for any $t \in (-\infty, \infty)$), with the usual norm

$$x \longmapsto \max_{k=1,2,\ldots,n} \max_{t \in [0,T]} \left| x_k(t) \right|. \tag{1.1}$$

(3) $C([0, T] \times D, \mathbb{R}^n)$, where $D \subset \mathbb{R}^n$ is a compact set, stands for the Banach space of continuous vector-functions $x = (x_k)_{k=1}^n : [0, T] \to \mathbb{R}^n$ with norm (1.1).

(4) $L_1([0, T], \mathbb{R}^n)$ is the Banach space of the Lebesgue integrable functions $x = (x_k)_{k=1}^n : (-\infty, \infty) \to \mathbb{R}^n$ with the norm

$$x \longmapsto \max_{k=1,2,\ldots,n} \int_0^T \left| x_k(t) \right| \mathrm{d}t.$$

(5) $\mathrm{GL}_n(\mathbb{R})$ is the algebra of square real matrices of dimension n.

(6) $\mathbb{1}_n$ is the n-dimensional unit matrix.

(7) The symbol $r(A)$ stands for the spectral radius of a square matrix A.

(8) I is the identity operator in various spaces.

(9) $\partial\Omega$ is the boundary of a set $\Omega \subset \mathbb{R}^n$.

(10) If $\Omega_1 \subset \mathbb{R}^n$, $\Omega_2 \subset \mathbb{R}^n$, and $\{\alpha_1, \alpha_2\} \subset \mathbb{R}$, then $\alpha_1\Omega_1 + \alpha_2\Omega_2 := \{\alpha_1 x_1 + \alpha_2 x_2 \mid x_1 \in \Omega_1,\ x_2 \in \Omega_2\}$.

(11) For a point $x = \mathrm{col}(x_{1,} x_2, \ldots, x_n) \in \mathbb{R}^n$, we denote by $|x|$ the vector $|x| = \mathrm{col}(|x_1|, |x_2|, \ldots, |x_n|)$. Similarly, the signs $|\cdot|$, max, min, and $\leqslant$ for vectors are understood componentwise.

(12) Let $\beta = \mathrm{col}(\beta_1, \ldots, \beta_n)$ be a vector from $\mathbb{R}_+^n$. By the β-neighbourhood of a point $y \in \mathbb{R}^n$ we understand the set $B(y, \beta)$ of points defined as follows:

$$B(y, \beta) := \left\{ x \in \mathbb{R}^n \mid |x - y| \leqslant \beta \right\}. \tag{1.2}$$

(13) For any vector $\beta \in \mathbb{R}_+^n$, by D_β, we mean the set

$$D_\beta := \left\{ x \in D \mid B(x, \beta) \subset D \right\} \tag{1.3}$$

that consists of points $x \in \mathbb{R}^n$ lying in D together with their β-neighbourhoods.

The symbol $\square$ denotes the end of a proof.

2. Introduction

In this work, we investigate several types of regular boundary value problems for the system of ordinary differential equations

$$x'(t) = f\big(t, x(t)\big), \quad t \in [0, T], \tag{2.1}$$

where $T \in (0, +\infty)$, $f \in C([0, T] \times D, \mathbb{R}^n)$. Here, D is the closure of a connected and bounded domain in $\mathbb{R}^n$. We are interested in continuously differentiable solutions of (2.1) which satisfy either the periodic conditions

$$x(0) = x(T) \tag{2.2}$$

or certain other, more complicated kinds of two or three-point boundary conditions, both linear and non-linear ones.

The singular problems are not dealt with in this text. We refer, in particular, to the work of Agarwal and O'Regan [2] for a survey of recent results on problems for Eqs. (2.1) having singularities with respect to the independent variable.

Boundary value problems for higher order scalar equations with both time and space singularities are discussed in the work of Rachůnková, Staněk, and Tvrdý [59]. The techniques used there are based on construction of certain sequences of regular problems that are assumed to be solvable and whose solutions can be regarded as approximations of the solution in question. Certain singular two-point problems with separated boundary conditions are studied in [59] by using the method of upper and lower functions.

Singular boundary value problems for systems (2.1) in various complicated cases are studied by Kiguradze [36] by using techniques based upon properties of differential inequalities.

The solvability of various boundary value problems by using fixed point theorems is studied, in particular, by Ntouyas [52]. The works cited also contain an extensive bibliography on the subject.

In this text, we investigate systems (2.1) subjected to the boundary conditions

$$g\big(x(0), x(T)\big) = 0, \tag{2.3}$$

which are, generally speaking, non-separated and non-linear. Conditions (2.3), obviously, cover the periodic conditions (2.2), the separated non-linear conditions

$$x(0) = b\big(x(T)\big), \tag{2.4}$$

and so on. We also study the three-point problems with the conditions

$$Ax(0) + Bx(\xi) + Cx(T) = d. \tag{2.5}$$

In order to show the existence of a solution and provide an efficient algorithm for its approximate finding, we construct a certain uniformly convergent sequence of functions that depends essentially on the character of concrete non-linearities in system (2.1) and on the form of the given boundary conditions. This approach may be regarded as a generalisation of the Picard successive approximation scheme.

In the original works of Samoilenko [82,83] and Samoilenko and Ronto [88–90], the approach used and described here had been referred to as the numerical-analytic method based upon successive approximations. The idea of the method, originally aimed at the investigation of periodic solutions only, had been later applied in studies of numerous types of boundary value problems. The monograph of Farkas [17], in particular, refers to it as to a relatively new, powerful, and popular means for the investigation of periodic solutions of non-linear periodic systems of ordinary differential equations. The merit of this method is that it gives an opportunity to solve two main questions of the theory of periodic motions, namely, to establish existence theorems and to construct approximate solutions.

It should be noted that appropriate versions of the method considered can be applied in many situations for handling periodic or two-point non-linear boundary value problems in the case of systems of first or second order ordinary differential equations, integro-differential equations, equations with retarded argument and equations with more general argument deviations, systems containing unknown parameters, and countable systems of differential equations. A survey of the investigations on the subject can be found in the series of papers by Ronto, Samoilenko, and Trofimchuk [74,76,77,75,78–80] where the achievements and trends in the development of the method for the period of about 40 years are described. This text is devoted to applications of the method to the existence analysis and approximate finding of solutions of periodic and some other boundary value problems for ordinary differential equations.

In what follows, it is convenient to use the term “periodic boundary value problem” rather than “T-periodic solution of a T-periodic system of differential equations” and, instead of $\mathbb{R}$, consider the segment $[0, T]$ as the domain of definition of the equation.

The scheme of the method under consideration is rather simple. According to the basic idea of the method, the given periodic boundary value problem is replaced by the Cauchy problem for a “perturbed” differential equation containing some artificially introduced parameter, whose value is to be determined later. The solution of this family of systems of differential equations parametrised by a single vector parameter is sought for in the analytic form by the method of successive approximations. Thus, all the iterations depend upon the parameter mentioned.

As regards the way how the auxiliary problem is constructed, it is essential that the form of the “perturbation term” generates a certain system of (algebraic or transcendental) “determining equations,” which give the numerical values of the parameter corresponding to the solutions sought for. By studying these determining equations, we can establish existence results for the original problem.

Thus, the method consists of two main parts, namely, the “analytic” part, i.e., solving the auxiliary problem for the “perturbed” equation by the method of successive approximations, and then the “numerical ”part, when the determining equations are solved.

The text is organised as follows. Sections 1 and 2 contain the notation used in what follows and provide a short introduction.

In Section 3, the successive approximation techniques are treated for the investigation of periodic solutions of non-autonomous periodic systems.

In Section 4, we apply the method for the study the periodic solutions of autonomous systems by using the appropriate reduction to a non-autonomous system.

In Section 5, we establish conditions under which a system of non-linear non-autonomous ordinary differential equations has a family of solutions that are periodic with a common period and possess a certain generalised symmetry property.

Sections 6 and 7 deal with the investigation of non-linear boundary problems by using a parametrisation that leads one to a family of problems with linear two-point conditions considered together with certain additional algebraic or transcendent equations.

In Section 8, we use the parametrisation approach to study some three-point boundary value problems, which, as a result, can be investigated through auxiliary two-point problems.

Section 9 contains some historical remarks concerning the development and application of the method.

In Section 10, we pose several problems concerning the method of successive approximations.

3. Periodic successive approximations for non-autonomous systems

Let us consider the system of ordinary differential equations

$$x'(t) = f\big(t, x(t)\big), \quad t \in (-\infty, \infty), \tag{3.1}$$

where the right-hand side function f is such that $f \in C([0,T] \times D, \mathbb{R}^n)$ and

$$f(t,x) = f(t+T, x) \tag{3.2}$$

for all $x \in D$ and all real t. Here, T is a positive number and D is the closure of a bounded and connected domain in $\mathbb{R}^n$.

We are interested in continuously differentiable solutions of system (3.1) which are periodic with period T (i.e., belonging to C_T^n). It follows from Lemma 3.1 below that this problem is equivalent to finding a solution of the periodic boundary value problem (2.1), (2.2) on the bounded interval $[0,T]$:

$$x'(t) = f\big(t, x(t)\big), \quad t \in [0,T], \tag{3.3}$$

$$x(0) = x(T). \tag{3.4}$$

LEMMA 3.1. (See, e.g., [17, Lemma 2.2.1].) *If a function* $x : (-\infty, \infty) \to D$ *is a* T*-periodic solution of system* (3.1)*, then its restriction* $x|_{[0,T]}$ *to* $[0,T]$ *satisfies system* (3.3), (3.4). *Conversely, if* x: $[0,T] \to D$ *is a solution of the boundary value problem* (3.3), (3.4)*, then its continuous* T*-periodic extension to* $(-\infty, \infty)$ *is a* T*-periodic solution of system* (3.1).

The solution of the periodic boundary value problem (3.3), (3.4) can be considered as a restriction of the T-periodic solution of the T-periodic system (3.1) onto the segment $[0,T]$. Thus, below we consider the segment $[0,T]$ instead of $\mathbb{R}$ in the domain of definition (3.2): $f \in C([0,T] \times D, \mathbb{R}^n)$.

It is assumed that the following conditions hold for the periodic boundary value problem (3.3), (3.4):

(1) For any $t \in [0, T]$ fixed, the function $f(t, \cdot)$ for all $\{x_1, x_2\} \subset D$ satisfies the Lipschitz condition

$$\bigl|f(t, x_1) - f(t, x_2)\bigr| \leqslant K|x_1 - x_2| \tag{3.5}$$

with some non-negative matrix $K = \{K_{ij}\}_{i,j=1}^{n} \in \mathrm{GL}_n(\mathbb{R})$ and, furthermore, the estimate

$$\frac{1}{2}\Bigl[\max_{(t,x)\in[0,T]\times D} f(t, x) - \min_{(t,x)\in[0,T]\times D} f(t, x)\Bigr] \leqslant M \tag{3.6}$$

is true with some $M \in \mathbb{R}_{+}^{n}$.

(2) The maximal eigenvalue $r(K)$ of the matrix K in (3.5) is such that

$$r(K) < \frac{10}{3T}. \tag{3.7}$$

(3) The set $D_{\frac{T}{2}M}$ is non-empty:

$$D_{\frac{T}{2}M} \neq \varnothing \tag{3.8}$$

Recall that the set $D_{\frac{T}{2}M}$ is defined according to notation (13), Section 1. It is clear that the form of this set essentially depends on the properties of the function f and the structure of the domain D.

REMARK 3.2. Inequality (3.7) is equivalent to the condition that the greatest eigenvalue of the matrix

$$Q = \frac{T}{q}K, \tag{3.9}$$

where

$$q = \frac{10}{3}, \tag{3.10}$$

is less than 1.

REMARK 3.3. By virtue of the non-negativity of K, it follows from the Perron theorem (see, e.g., Krein and Rutman [38]) that the largest eigenvalue of this matrix is real, non-negative, and is estimated from above by the number

$$\min\Biggl\{\max_{i=1,\dots,n}\sum_{j=1}^{n} K_{ij}, \max_{j=1,\dots,n}\sum_{i=1}^{n} K_{ij}\Biggr\}.$$

In (1.2), as well as in (3.6), (3.5) and similar relations below, the signs $|\cdot|$, max, min, and $\leqslant$ are understood componentwise. Note that the componentwise estimate of form (3.5), (3.6), etc., in general, are more accurate than the corresponding estimates in terms of the norm because they use the algebraic structure of the space $\mathbb{R}^n$ in a more subtle way.

REMARK 3.4. One can verify (see Remark 2 in Rontó and Mészáros [69]) that the inequality

$$\frac{1}{2}\Big[\max_{(t,x)\in[0,T]\times D} f(t,x) - \min_{(t,x)\in[0,T]\times D} f(t,x)\Big] \leqslant \max_{(t,x)\in[0,T]\times D}\big|f(t,x)\big|, \tag{3.11}$$

is true. Moreover, the equality

$$\frac{1}{2}\Big[\max_{(t,x)\in[0,T]\times D} f(t,x) - \min_{(t,x)\in[0,T]\times D} f(t,x)\Big] = \max_{(t,x)\in[0,T]\times D}\big|f(t,x)\big|$$

holds if, and only if

$$\max_{(t,x)\in[0,T]\times D} f(t,x) = -\min_{(t,x)\in[0,T]\times D} f(t,x) = \max_{(t,x)\in[0,T]\times D}\big|f(t,x)\big|.$$

EXAMPLE 3.5. For the function

$$f(t,x) = \frac{1}{2}\ln x \sin t + 12$$

defined for (t,x) from the domain $[0,T]\times D = [0,2\pi]\times[\mathrm{e},\mathrm{e}^2] =: G$, estimate (3.6) is sharper than the inequality $|f(t,x)| \leqslant M$. Indeed, it easy to see that

$$\max_{(t,x)\in[0,T]\times D}\big|f(t,x)\big| = 13$$

and

$$\frac{1}{2}\Big[\max_{(t,x)\in[0,T]\times D} f(t,x) - \min_{(t,x)\in[0,T]\times D} f(t,x)\Big] = \frac{1}{2}[13-11] = 1.$$

Taking into account that the vector M and the matrix K in estimates (3.6), (3.5) depend on the domain D, it may happen that conditions (3.6)–(3.8) for system (3.1) are satisfied in some domain D (or some subdomains of D) and does not hold in another one. Let us consider some examples to illustrate this fact.

EXAMPLE 3.6. Assume that the function f in (3.1) satisfies conditions (3.6), (3.5) for $(t,x)\in\mathbb{R}\times\mathbb{R}^n$.

If inequality (3.7) is satisfied with some Lipschitz matrix K, then condition (3.8) also holds in any closed bounded domain D containing an n-dimensional rectangle with the side length componentwise greater than $\frac{T}{2}M$.

EXAMPLE 3.7. Consider the case where (3.1) is a linear system,

$$x'(t) = P(t)x + q(t), \tag{3.12}$$

where $q \in C_T^n$ and $P = \{P_{ij}\}_{i,j=1}^n$, P_{ij}, $i, j = 1, 2, \dots, n$, are continuous T-periodic functions.

Let us put $\hat{P} := \{\max_{t\in[0,T]} |P_{ij}(t)|\}_{i,j=1}^n$,

$$\hat{q} := \operatorname{col}\Big(\max_{t\in[0,T]}\big|q_1(t)\big|, \dots, \max_{t\in[0,T]}\big|q_n(t)\big|\Big)$$

and consider system (3.12) in the domain

$$D := \big\{x \mid |x| \leqslant b\big\}, \tag{3.13}$$

where $b = \operatorname{col}(b_1, \dots, b_n)$ is such that conditions (3.6)–(3.8) hold for system (3.12).

We can put $K = \hat{P}$ in (3.5). According to (3.11), condition (3.6) takes the form

$$\big|f(t, x)\big| \leqslant \big|P(t)x + q(t)\big| \leqslant \hat{P}b + \hat{q} =: M.$$

Condition (3.8) holds in domain (3.13) if the inequality

$$\hat{q} \leqslant \frac{2}{T}\left(\mathbb{1}_n - \frac{T}{2}\hat{P}\right)b.$$

is true. In particular, when $r(\hat{P}) < 2T^{-1}$, we can put

$$b = \frac{T}{2}\left(\mathbb{1}_n - \frac{T}{2}\hat{P}\right)\hat{q}.$$

Condition (3.7) holds whenever the greatest eigenvalue of the matrix $\hat{P}$ does not exceed $\frac{10}{3T}$.

EXAMPLE 3.8. Let us consider the system of equations

$$\begin{aligned} x_1'(t) &= x_2(t), \\ x_2'(t) &= -x_1(t) - \sin(20t)\big(x_2(t)\big)^s + \cos(20t) \end{aligned} \tag{3.14}$$

for $(t, x) \in [0, \frac{\pi}{10}] \times D$, where

$$D := \big\{(x_1, x_2) \mid |x_1| \leqslant a, |x_2| \leqslant 1\big\}. \tag{3.15}$$

Then, for $s \geqslant 0$, the function

$$f(t, x_1, x_2) := \begin{pmatrix} x_2 \\ -x_1 - (\sin 20t)x_2^s + \cos(20t) \end{pmatrix} \tag{3.16}$$

defined on the set $[0, T] \times D$ with $T = \frac{\pi}{10}$ satisfies the relation

$$\bigl|f(t,x)\bigr| \leqslant \operatorname{col}(1, a+2) =: M$$

for all $(t, x) \in [0, \frac{\pi}{10}] \times D$. Furthermore, the Lipschitz condition (3.5) holds with $K = \left[\begin{smallmatrix} 0 & 1 \\ 1 & s \end{smallmatrix}\right]$. The set (3.8) corresponding to system (3.14) is defined as follows:

$$D_{\frac{\pi}{20}M} = \left\{ (x_1, x_2) \;\middle|\; |x_1| \leqslant a - \frac{\pi}{20},\ |x_2| \leqslant 1 - \frac{\pi}{20}(a+2) \right\}.$$

This subdomain is non-empty if

$$a - \frac{\pi}{20} \geqslant 0,$$
$$1 - \frac{\pi}{20}(a+2) \geqslant 0,$$

whence we see that $D_{\frac{\pi}{20}M} \neq \varnothing$ whenever a satisfies the relations

$$\frac{\pi}{20} \leqslant a \leqslant \frac{20}{\pi} - 2. \tag{3.17}$$

It is easy to see that λ is an eigenvalue of the matrix K if, and only if $\lambda^2 - \lambda s - 1 = 0$, whence

$$r(K) = \frac{s + \sqrt{s^2+4}}{2}.$$

Therefore, condition (3.7) holds whenever $\frac{3\pi}{100}\frac{s+\sqrt{s^2+4}}{2} < 1$, $\sqrt{s^2+4} < \frac{200}{3\pi} - s$, or, finally, when

$$s < \frac{100}{3\pi} - \frac{3\pi}{100}. \tag{3.18}$$

Thus, the conditions (3.6)–(3.8) for the system (3.14) are fulfilled in the domain (3.15) if the values a and s satisfy the inequalities (3.17), (3.18).

EXAMPLE 3.9. Let us consider the damped Duffing equation with periodic forcing term (see (3.3.8) and (4.3.15) in [17])

$$x''(t) + dx' + ax(t) + b\bigl(x(t)\bigr)^3 = H \sin\left(\frac{2\pi}{T}t\right), \tag{3.19}$$

where $a > 0$, $d > 0$, $b \in \mathbb{R}$ (“soft” or “hard spring”), $H \in \mathbb{R}$. Equation (3.19) is equivalent to the Cauchy normal form system

$$\begin{aligned} x_1'(t) &= x_2(t), \\ x_2'(t) &= -dx_2(t) - ax_1(t) - bx_1^3(t) + H\sin\left(\frac{2\pi}{T}t\right). \end{aligned} \tag{3.20}$$

For the domain of definition of this system, we take

$$(t, x) \in [0, T] \times D, \quad D := \bigl\{(x_1, x_2)\colon |x_1| \leqslant m, |x_2| \leqslant 1\bigr\}. \tag{3.21}$$

In this case, the right-hand side function, f, in (3.20) satisfies the relation

$$\bigl|f(t, x)\bigr| \leqslant \operatorname{col}\bigl(1, d + am + |b|m^3 + |H|\bigr) =: M$$

and the Lipschitz condition (3.5) with

$$K = \begin{bmatrix} 0 & 1 \\ a + 3|b|m & d \end{bmatrix}.$$

The set (3.8) for system (3.20) is determined by the formula

$$D_{\frac{T}{2}M} = \left\{(x_1, x_2) \mid |x_1| \leqslant m - \frac{T}{2}, |x_2| \leqslant 1 - \frac{T}{2}\bigl(d + am + |b|m^3 + |H|\bigr)\right\}.$$

This subdomain is non-empty if

$$\begin{aligned} &m - \frac{T}{2} \geqslant 0, \\ &1 - \frac{T}{2}\bigl(d + am + b|m^3 + |H|\bigr) \geqslant 0. \end{aligned}$$

Thus, condition (3.8) holds whenever

$$m \geqslant \frac{T}{2}, \qquad am + |b|m^3 \leqslant \frac{2}{T} - d. \tag{3.22}$$

Since the eigenvalues λ of the matrix K are determined by the equation

$$\lambda^2 - d\lambda - \bigl(a + 3|b|m^2\bigr) = 0,$$

it is obvious that

$$r(K) = \frac{d + \sqrt{d^2 + 4(a + 3|b|m^2)}}{2}$$

and, therefore, (3.21) holds if

$$d + \sqrt{d^2 + 4\big(a + 3|b|m^2\big)} < \frac{20}{3T}. \tag{3.23}$$

Conditions (3.6)–(3.8) for system (3.20) are satisfied in domain (3.21) when the values of the parameters a, b, d, H, and m satisfy inequalities (3.22), (3.23).

We can now return to the investigation of T-periodic solution of system (3.1), which, as we have mentioned above, is equivalent to the periodic boundary value problem (3.3), (3.4).

Let us put

$$(Lx)(t) := \int_0^t x(s)\,\mathrm{d}s - \frac{t}{T}\int_0^T x(s)\,\mathrm{d}s, \quad t \in [0, T]. \tag{3.24}$$

Obviously, formula (3.24) determines a continuous linear operator in the space $C([0, T], \mathbb{R}^n)$. Moreover, one can easily verify that the following lemma holds.

LEMMA 3.10. *For an arbitrary x from $C([0, T], \mathbb{R}^n)$,*

$$(Lx)(0) = (Lx)(T).$$

Consider the parametrised integral equation

$$x(t) = z + \int_0^t \left(f\big(s, x(s)\big) - \frac{1}{T}\int_0^T f\big(\tau, x(\tau)\big)\,\mathrm{d}\tau \right) \mathrm{d}s$$

or, which is the same,

$$x(\cdot) = z + Lf\big(\cdot, x(\cdot)\big) \tag{3.25}$$

where $z \in \mathbb{R}^n$ is a vector parameter.

REMARK 3.11. It is clear from Lemma 3.10 that every solution x of Eq. (3.25) satisfies the T-periodic boundary condition (3.4); see also the discussion of Eq. (4.1.19) in [17, Section 4.1].

We have the following simple

PROPOSITION 3.12. *If $x(\cdot, z)$ is a solution of the integro-differential equation*

$$x'(t) = f\big(t, x(t)\big) - \frac{1}{T}\int_0^T f\big(s, x(s)\big)\,\mathrm{d}s, \quad t \in [0, T], \tag{3.26}$$

with the initial value condition

$$x(0) = z \tag{3.27}$$

and, moreover,

$$\int_0^T f\bigl(s, x(s,z)\bigr)\,\mathrm{d}s = 0, \tag{3.28}$$

then $x(\cdot, z)$ is also a solution of the T-periodic boundary value problem (3.3), (3.4).

Conversely, if a function x satisfies (3.3), (3.4), *then it is also a solution of the initial value problem* (3.26), (3.27) *with $z = x(0)$.*

PROOF. It is obvious that a continuously differentiable function x satisfying the initial value problem (3.26), (3.27) is a continuous solution of Eq. (3.25) and *vice versa*. Therefore, by virtue of Remark 3.11, a solution x of (3.26), (3.27) satisfies the T-periodic boundary condition (3.4) for an arbitrary z. Thus, it follows that (3.26), (3.27), and (3.28) imply (3.3), (3.4).

Conversely, if x is a solution of the T-periodic problem (3.3), (3.4), then

$$\int_0^T f\bigl(s, x(s)\bigr)\,\mathrm{d}s = 0$$

and, hence, (3.26), (3.27) hold with $z = x(0)$. □

Equation (3.26) can be considered as a perturbed version of Eq. (3.1) with the perturbation term determined by the expression in the left-hand side of (3.28).

We shall show that, under conditions (3.6)–(3.8), the integral equation (3.25) can be solved by the method of successive approximations. This is the "analytic" part of our method.

With the integral equation (3.25), we associate the following sequence of T-periodic functions $\{x_m(\cdot, z)\}_{m=0}^{+\infty}$ depending on the parameter $z \in \mathbb{R}^n$:

$$x_m(t,z) = z + \int_0^t \left[f\bigl(s, x_{m-1}(s,z)\bigr) - \frac{1}{T}\int_0^T f\bigl(\tau, x_{m-1}(\tau,z)\bigr)\,\mathrm{d}\tau \right] \mathrm{d}s, \tag{3.29}$$

where $m = 1, 2, \ldots$ and $x_0(t,z) = z$.

It is obvious from Lemma 3.10 that

$$x_m(0,z) = x_m(T,z), \qquad x_m(0,z) = z$$

for every natural m.

Let us establish two subsidiary statements needed below for the convergence analysis of sequence (3.29).

LEMMA 3.13. *Let* $f:[0,T]\to\mathbb{R}^n$ *be a continuous function. Then, for an arbitrary* $t\in[0,T]$, *the inequality*

$$\left|(Lf)(t)\right| \leqslant \frac{1}{2}\alpha_{1,T}(t)\Big[\max_{t\in[0,T]} f(t) - \min_{t\in[0,T]} f(t)\Big], \tag{3.30}$$

holds, where the operator L *is given by formula* (3.24) *and*

$$\alpha_{1,T}(t) := 2t\left(1-\frac{t}{T}\right), \quad t\in[0,T]. \tag{3.31}$$

REMARK 3.14. It is easy to see that the estimate

$$0 \leqslant \alpha_{1,T}(t) \leqslant \frac{T}{2}, \quad t\in[0,T] \tag{3.32}$$

holds for function (3.31).

PROOF OF LEMMA 3.13. It is obvious that

$$\begin{aligned}
&\int_0^t \left[f(\tau) - \frac{1}{T}\int_0^T f(s)\,\mathrm{d}s\right]\mathrm{d}\tau \\
&\quad = \frac{1}{T}\int_0^t\int_0^T \big[f(\tau)-f(s)\big]\,\mathrm{d}s\,\mathrm{d}\tau \\
&\quad = \frac{1}{T}\int_0^t\int_0^t \big[f(\tau)-f(s)\big]\,\mathrm{d}s\,\mathrm{d}\tau + \frac{1}{T}\int_0^t\int_t^T \big[f(\tau)-f(s)\big]\,\mathrm{d}s\,\mathrm{d}\tau.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\int_0^t\int_0^t \big[f(\tau)-f(s)\big]\,\mathrm{d}s\,\mathrm{d}\tau &= \int_0^t \left[t f(\tau) - \int_0^t f(s)\,\mathrm{d}s\right]\mathrm{d}\tau \\
&= t\int_0^t f(\tau)\,\mathrm{d}\tau - t\int_0^t f(s)\,\mathrm{d}s = 0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left|\int_0^t \left[f(\tau) - \frac{1}{T}\int_0^T f(s)\,\mathrm{d}s\right]\mathrm{d}\tau\right| &\leqslant \frac{1}{T}\int_0^t\int_t^T \big|f(\tau)-f(s)\big|\,\mathrm{d}s\,\mathrm{d}\tau \\
&\leqslant \frac{1}{T}\max_{\substack{\tau\in[0,t]\\ s\in[t,T]}} \big|f(\tau)-f(s)\big| t(T-t)
\end{aligned}$$

$$= \frac{1}{2}\alpha_{1,T}(t) \max_{\substack{\tau\in[0,t]\\ s\in[t,T]}} \left|f(\tau) - f(s)\right|$$

$$\leqslant \frac{1}{2}\alpha_{1,T}(t)\Big[\max_{t\in[0,T]} f(t) - \min_{t\in[0,T]} f(t)\Big].$$

The last estimate leads us to the required inequality (3.30). □

REMARK 3.15. Estimate (3.30) from Lemma 3.13 improves the corresponding estimate

$$\left|\int_0^t \left[f(\tau) - \frac{1}{T}\int_0^T f(s)\,\mathrm{d}s\right]\mathrm{d}\tau\right| \leqslant \alpha_{1,T}(t) \max_{t\in[0,T]}\left|f(t)\right| \tag{3.33}$$

of [70, Lemma 2.1].

Indeed, let us consider the function f from Example 3.5. Taking into account the fact that

$$\max_{(t,x)\in[0,T]\times D}\left|f(t,x)\right| = \max_{(t,x)\in[0,2\pi]\times[e,e^2]}\left|\frac{1}{2}\ln x\sin t + 12\right| = 13,$$

according to (3.33) and estimate (3.32) from Remark 3.14, we have

$$\left|\int_0^t \left[f(\tau,x) - \frac{1}{T}\int_0^T f(s,x)\,\mathrm{d}s\right]\mathrm{d}\tau\right| \leqslant \alpha_{1,T}(t)\cdot 13 \leqslant 13\pi \approx 40.82.$$

In view of estimate (3.30),

$$\left|\int_0^t \left[f(\tau,x) - \frac{1}{T}\int_0^T f(s,x)\,\mathrm{d}s\right]\mathrm{d}\tau\right| \leqslant \frac{\pi}{2}\cdot 1 \approx 1.57.$$

On the other hand, by direct computation, we obtain that

$$\begin{aligned}
&\int_0^t \left[\left(\frac{1}{2}\ln x\sin\tau + 12\right) - \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2}\ln x\sin s + 12\right)\mathrm{d}s\right]\mathrm{d}\tau\\
&\quad = \int_0^t \left[\left(\frac{1}{2}\ln x\sin\tau + 12\right) - \frac{1}{2\pi}24\pi\right]\mathrm{d}\tau\\
&\quad = \int_0^t \frac{1}{2}\ln x\sin\tau\,\mathrm{d}\tau = \frac{1}{2}\ln x(1-\cos t)
\end{aligned}$$

and

$$\max_{(t,x)\in[0,2\pi]\times[e,e^2]}\left|\frac{1}{2}\ln x(1-\cos t)\right| = 1.$$

LEMMA 3.16. *Let the sequence of continuous functions*

$$\alpha_{0,T}(t) = 1, \quad \alpha_{1,T}(t) = 2t\left(1 - \frac{t}{T}\right), \quad \ldots, \quad \alpha_{m,T}(t), \ldots \tag{3.34}$$

be defined by the recurrence relation

$$\begin{aligned}\alpha_{m+1,T}(t) = {} & \left(1 - \frac{t}{T}\right)\int_0^t \alpha_{m,T}(s)\,\mathrm{d}s \\ & + \frac{t}{T}\int_t^T \alpha_{m,T}(s)\,\mathrm{d}s, \quad m = 0, 1, 2, \ldots.\end{aligned} \tag{3.35}$$

Then, for all $m = 0, 1, 2, \ldots$, *the estimate*

$$\alpha_{m+1,T}(t) \leqslant \left(\frac{3T}{10}\right)^m \bar{\alpha}_{1,T}(t), \quad t \in [0, T] \tag{3.36}$$

holds, where

$$\bar{\alpha}_{1,T}(t) := \frac{10}{9}\alpha_{1,T}(t). \tag{3.37}$$

PROOF. By direct computation, from the iterative formula (3.35) we obtain

$$\alpha_{2,T}(t) = \alpha_{1,T}(t)\left[\frac{T}{6} + \frac{1}{3}\alpha_{1,T}(t)\right]$$

and

$$\alpha_{3,T}(t) = \alpha_{1,T}(t)\left[\frac{T^2}{20} + \frac{T}{15}\alpha_{1,T}(t) + \frac{1}{15}\alpha_{1,T}^2(t)\right]. \tag{3.38}$$

Taking estimate (3.32) into account, we get

$$\alpha_{2,T}(t) \leqslant \frac{T}{3}\alpha_{1,T}(t) = \frac{3}{10}T\left(\frac{10}{9}\alpha_{1,T}(t)\right) = \left(\frac{3}{10}T\right)\bar{\alpha}_{1,T}(t).$$

It follows from (3.38) that

$$\begin{aligned}\alpha_{3,T}(t) &= \frac{3}{10}T\alpha_{1,T}(t)\left[\frac{T}{6} + \frac{1}{3}\alpha_{1,T}(t)\right] - \frac{1}{30}T\alpha_{1,T}^2(t) + \frac{1}{15}\alpha_{1,T}^3(t) \\ &= \left(\frac{3}{10}T\right)\alpha_{2,T}(t) - \frac{1}{30}\alpha_{1,T}^2(t)\left[T - 2\alpha_{1,T}(t)\right] \\ &\leqslant \left(\frac{3}{10}T\right)\alpha_{2,T}(t) \leqslant \left(\frac{3}{10}T\right)^2\bar{\alpha}_{1,T}(t).\end{aligned}$$

By induction, the last inequality yields

$$\begin{aligned}\alpha_{m+1,T}(t) &\leqslant \left(1-\frac{t}{T}\right)\int_0^t \left(\frac{3}{10}T\right)\alpha_{m-1,T}(s)\,\mathrm{d}s \\ &\quad + \frac{t}{T}\int_t^T \left(\frac{3}{10}T\right)\alpha_{m-1,T}(s)\,\mathrm{d}s \\ &\leqslant \left(\frac{3}{10}T\right)\alpha_{m,T}(t) \leqslant \left(\frac{3}{10}T\right)^m \bar{\alpha}_{1,T}(t).\end{aligned}$$

Thus, we have obtained the required estimate (3.36). □

The last lemma allows us to establish a statement on the convergence of sequence (3.29).

THEOREM 3.17. *Let the T-periodic boundary-value problem* (3.3), (3.4) *be such that conditions* (3.5), (3.6), *and* (3.8) *are satisfied. Then, for an arbitrary value of the parameter $z \in D_\beta$, the following assertions are true*:

(i) *The sequence $\{x_m(\cdot, z)\}_{m\geqslant 0}$ of the T-periodic functions* (3.29) *converges uniformly in (t, z) from*

$$[0, T] \times D_{\frac{T}{2}M}; \tag{3.39}$$

(ii) *The limit function*

$$x^*(t, z) := \lim_{m\to\infty} x_m(t, z) \tag{3.40}$$

of sequence (3.29) *is the unique T-periodic solution of the integral equation* (3.25)

$$x^*(t, z) = z + \int_0^t f\bigl(s, x^*(s, z)\bigr)\,\mathrm{d}s - \frac{t}{T}\int_0^T f\bigl(s, x^*(s, z)\bigr)\,\mathrm{d}s \tag{3.41}$$

in the domain D or, which is the same, it is the unique T-periodic solution of the integro-differential equation (3.26)

$$\frac{\partial x(t, z)}{\partial t} = f\bigl(t, x(t, z)\bigr) - \Delta(z), \quad t \in [0, T], \tag{3.42}$$

where the "perturbation term" has the form

$$\Delta(z) := \frac{1}{T}\int_0^T f\bigl(t, x^*(t, z)\bigr)\,\mathrm{d}t; \tag{3.43}$$

(iii) *The estimate*

$$\bigl|x^*(t, z) - x_m(t, z)\bigr| \leqslant \bar{\alpha}_{1,T}(t) Q^m (I - Q)^{-1} M, \tag{3.44}$$

holds for all $m = 1, 2, \dots$ *and* $t \in [0, T]$, *where the function* $\bar{\alpha}_{1,T}$ *is given by* (3.37),

$$Q = \frac{3T}{10} K, \tag{3.45}$$

and K *is the matrix from the Lipschitz condition* (3.5).

REMARK 3.18. It is obvious from (3.37) and Remark 3.14 that

$$\bar{\alpha}_{1,T}(t) \leqslant \frac{5}{9} T, \quad t \in [0, T]. \tag{3.46}$$

PROOF OF THEOREM 3.17. We shall show that $\{x_m(\cdot, z)\}_{m \geqslant 0}$ given by (3.29) is a Cauchy sequence in the space $C([0, T], \mathbb{R}^n)$ and, therefore, this sequence is uniformly convergent in domain (3.39).

First we prove that if z belongs to the set $D_{\frac{T}{2}M} \subset D$, then $x_m(\cdot, z) \in D$ for all t and m. Indeed, by virtue of estimate (3.30) of Lemma 3.13 and relation (3.6), equality (3.29) implies

$$\begin{aligned}
\left|x_1(t, z) - z\right| &= \left|x_1(t, z) - x_0(t, z)\right| \\
&\leqslant \left| \int_0^t \left[f(s, z)\,\mathrm{d}s - \frac{1}{T} \int_0^T f(\tau, z)\,\mathrm{d}\tau \right] \mathrm{d}s \right| \mathrm{d}s \\
&\leqslant \frac{1}{2} \alpha_{1,T}(t) \Big[\max_{(t,z) \in [0,T] \times D_{\frac{T}{2}M}} f(t, z) - \min_{(t,x) \in [0,T] \times D_{\frac{T}{2}M}} f(t, z) \Big] \\
&\leqslant \alpha_{1,T}(t) M \leqslant \frac{T}{2} M,
\end{aligned} \tag{3.47}$$

where $\alpha_{1,T}$ is given by (3.31). Therefore, $x_1(t, z) \in D$ for $t \in [0, T]$ and $z \in D_{\frac{T}{2}M}$. Arguing by induction, one can show that

$$\left|x_m(t, z) - z\right| \leqslant \frac{T}{2} M, \tag{3.48}$$

i.e., the values of all the members $x_m(\cdot, z)$ of sequence (3.29) are contained in the domain D for all $t \in [0, T]$ and arbitrary $z \in D_{\frac{T}{2}M}$.

Let us estimate the differences between neighbouring members of sequence (3.29). We have

$$\begin{aligned}
x_{m+1}(t, z) - x_m(t, z) = \int_0^t \Big[&\big(f\big(s, x_m(s, z)\big) - f\big(s, x_{m-1}(s, z)\big)\big) \\
&- \frac{1}{T} \int_0^T \big(f\big(\tau, x_m(\tau, z)\big) - f\big(s, x_{m-1}(\tau, z)\big)\big)\,\mathrm{d}\tau \Big]\,\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^t \big[f\big(s, x_m(s,z)\big) - f\big(s, x_{m-1}(s,z)\big)\big]\,\mathrm{d}s \\
&\quad - \frac{t}{T}\int_0^T \big[f\big(s, x_m(s,z)\big) - f\big(s, x_{m-1}(s,z)\big)\big]\,\mathrm{d}s \\
&= \left(1-\frac{t}{T}\right)\int_0^t \big[f\big(s, x_m(s,z)\big) - f\big(s, x_{m-1}(s,z)\big)\big]\,\mathrm{d}s \\
&\quad - \frac{t}{T}\int_t^T \big[f\big(s, x_m(s,z)\big) - f\big(s, x_{m-1}(s,z)\big)\big]\,\mathrm{d}s. \qquad (3.49)
\end{aligned}
$$

Putting

$$r_m(t,z) := \big|x_{m+1}(t,z) - x_m(t,z)\big|, \quad t \in [0,T],$$

and taking the Lipschitz condition (3.5) into account, we see that (3.49) yields

$$r_{m+1}(t,z) \leqslant K\left[\left(1-\frac{t}{T}\right)\int_0^t r_m(s,z)\,\mathrm{d}s + \frac{t}{T}\int_t^T r_m(s,z)\,\mathrm{d}s\right]. \tag{3.50}$$

According to (3.47), we have

$$r_1(t,z) \leqslant \alpha_{1,T}(t)M \leqslant \frac{T}{2}M.$$

Therefore, for $m=1$, on base of the recurrence formula (3.35), relation (3.50) implies

$$
\begin{aligned}
r_2(t,z) &\leqslant K\left[\left(1-\frac{t}{T}\right)\int_0^t r_1(s,z)\,\mathrm{d}s + \frac{t}{T}\int_t^T r_1(s,z)\,\mathrm{d}s\right] \\
&\leqslant KM\left[\left(1-\frac{t}{T}\right)\int_0^t \alpha_{1,T}(s)\,\mathrm{d}s + \frac{t}{T}\int_t^T \alpha_{1,T}(s)\,\mathrm{d}s\right] \\
&\leqslant KM\alpha_{2,T}(t).
\end{aligned}
$$

Arguing by induction, we easily obtain

$$r_{m+1}(t,z) \leqslant K^m M\alpha_{m+1,T}(t), \quad m = 0, 1\ldots,$$

where $\alpha_{m+1,T}(t)$ is given by formula (3.33). By virtue of estimate (3.36) from Lemma 3.16, we have

$$r_{m+1}(t,z) \leqslant \bar{\alpha}_{1,T}(t)K^m\left(\frac{3T}{10}\right)^m M = \bar{\alpha}_{1,T}(t)Q^m M \tag{3.51}$$

for all $m = 1, 2\ldots$, where Q is the matrix given by (3.45).

Furthermore, in view of (3.51), the relation

$$\begin{aligned} x_{m+j}(t,z) - x_m(t,z) = {} & \big(x_{m+j}(t,z) - x_{m+j-1}(t,z)\big) \\ & + \big(x_{m+j}(t,z) - x_{m+j-2}(t,z)\big) \\ & + \cdots + \big(x_{m+1}(t,z) - x_m(t,z)\big) \end{aligned}$$

yields

$$\begin{aligned} \big|x_{m+j}(t,z) - x_m(t,z)\big| & \leqslant \sum_{i=1}^{j} r_{m+i} \leqslant \bar{\alpha}_{1,T}(t) \sum_{i=1}^{j} Q^{m+i-1} M \\ & \leqslant \bar{\alpha}_{1,T}(t) \sum_{i=0}^{j-1} Q^{m+i-1} M = \bar{\alpha}_{1,T}(t) Q^m \sum_{i=0}^{j-1} Q^i M. \end{aligned} \tag{3.52}$$

Taking into account the fact that, according to assumption (3.7), the maximal eigenvalue of matrix (3.45) is less than 1, we see that

$$\sum_{i=0}^{j-1} Q^{m+i} \leqslant \sum_{i=0}^{\infty} Q^{m+i} = Q^m (I - Q)^{-1} \tag{3.53}$$

and

$$\lim_{m \to \infty} Q^m = 0. \tag{3.54}$$

Consequently, it follows from (3.52), (3.53), and (3.54) that, for an arbitrary $t \in [0, T]$,

$$\big|x_{m+j}(t,z) - x_m(t,z)\big| \leqslant \bar{\alpha}_{1,T}(t) Q^m (I - Q)^{-1} M. \tag{3.55}$$

Inequality (3.29) implies immediately that (3.29) is a Cauchy sequence in the space $C([0,T], \mathbb{R}^n)$ and, therefore, it uniformly converges to a continuous limit function $x^*(\cdot, z)$, i.e., (3.40) holds.

Since all the functions $x_m(\cdot, z)$ of the sequence (3.29) are T-periodic, it follows its uniform limit $x^*(\cdot, z)$ is also a T-periodic function. Passing to the limit as $m \to +\infty$ in equality (3.29), we conclude that the limit function $x^*(\cdot, z)$ is indeed a unique T-periodic solution of the integral equation (3.25) with the initial value

$$x^*(0, z) = z. \tag{3.56}$$

Thus, equalities (3.41), (3.42) hold. When $j \to +\infty$ and $m = 1, 2, \ldots$, relation (3.55) yields the required estimate (3.43) for the difference between the exact solution $x^*(\cdot, z)$ of Eq. (3.41) and its mth approximation $x_m(\cdot, z)$. □

Equality (3.56) shows that, in fact, every point z from the subset $D_{\frac{T}{2}M}$ is the initial value of the unique T-periodic solution of Eqs. (3.25) and (3.26).

3.1. *Properties of the limit function*

Along with system (3.1), we shall also consider the system of equations

$$x'(t) = f\bigl(t, x(t)\bigr) - \mu, \quad t \in [0, T], \tag{3.57}$$

where $\mu = \operatorname{col}(\mu_1, \ldots, \mu_n)$ is a control parameter. Let us show that the parameter μ can always be chosen so that the solution of Eq. (3.57) having the initial value (3.27) with

$$z \in D_{\frac{T}{2}M} \tag{3.58}$$

also satisfies the T-periodic boundary condition (3.4).

In other words, this means that we can always introduce an additive perturbation of the right-hand side of the T-periodic system (3.1) such that, for an arbitrary $z \in D_{\frac{T}{2}M}$, the solution $x = x(t, 0, z, \mu)$ of the resulting initial value problem (3.57), (3.27) is, at the same time, a solution of the T-periodic problem (3.57), (3.4).

THEOREM 3.19. *Assume that the T-periodic system of differential Eqs.* (3.57) *satisfies the conditions of Theorem* 3.17.

Then, for an arbitrary $z \in D_\beta$, there exists a unique value of the control parameter μ,

$$\mu = \frac{1}{T}\int_0^T f\bigl(t, x^*(t, z)\bigr)\,\mathrm{d}t, \tag{3.59}$$

such that the solution $x = x(t, 0, z, \mu)$ of the initial value problem (3.57), (3.58), *with μ given by* (3.59), *is a T-periodic function and*

$$x(t, 0, z, \mu) = x^*(t, z) = \lim_{m\to\infty} x_m(t, z), \tag{3.60}$$

where $\{x_m(\cdot, z)\}_{m\geqslant 0}$ is the sequence of functions defined by formula (3.29).

PROOF. Since (3.1) satisfies conditions (3.5), (3.6), (3.61), and (3.62), the sequence of T-periodic functions given by (3.29) converges to a T-periodic function $x^*(\cdot, z)$ uniformly in domain (3.39). It follows from Theorem 3.17 that, according (3.42), (3.43), if (3.59) holds, then the T-periodic solution $x = x^*(t, z)$ of system (3.57) also satisfies the initial value condition (3.56). Thus, we have found the value of the parameter μ of the form (3.59) for which (3.60) holds.

It can be shown that this parameter value is unique because, for any other value $\mu = \nu$, the solution $x = x(t, 0, z, \nu)$ of the initial value problem (3.27) for the differential system

$$x'(t) = f\bigl(t, x(t)\bigr) - \nu, \quad t \in [0, T], \tag{3.61}$$

where $\nu \neq \Delta(z)$ and z satisfies (3.58), does not satisfy the T-periodic boundary condition (3.4). Indeed, assume the contrary. Then there exist at least two values μ, ν $(\mu \neq \nu)$ such that the solutions $x = x(t, 0, z, \mu)$ and $x = x(t, 0, z, \nu)$ of the Cauchy problems (3.57), (3.58) and (3.61) respectively also satisfy the T-periodic boundary condition (3.4). Then, by using (3.25), we obtain the following identity for the difference of these functions:

$$\begin{aligned}
&\bigl|x(t,0,z,\mu) - x(t,0,z,\nu)\bigr| \\
&= \biggl|\int_0^t \bigl[f\bigl(s, x(s,0,z,\mu)\bigr) - f\bigl(s, x(s,0,z,\nu)\bigr)\bigr]\,\mathrm{d}s \\
&\quad - \frac{t}{T}\int_0^T \bigl[f\bigl(s, x(s,0,z,\mu)\bigr) - f\bigl(s, x(s,0,z,\nu)\bigr)\bigr]\,\mathrm{d}s\biggr| \\
&= \biggl|\left(1 - \frac{t}{T}\right)\int_0^t \bigl[f\bigl(s, x(s,0,z,\mu)\bigr) - f\bigl(s, x(s,0,z,\nu)\bigr)\bigr]\,\mathrm{d}s \\
&\quad - \frac{t}{T}\int_t^T \bigl[f\bigl(s, x(s,0,z,\mu)\bigr) - f\bigl(s, x(s,0,z,\nu)\bigr)\bigr]\,\mathrm{d}s\biggr| \\
&\leqslant \left(1 - \frac{t}{T}\right)\int_0^t \bigl|f\bigl(s, x(s,0,z,\mu)\bigr) - f\bigl(s, x(s,0,z,\nu)\bigr)\bigr|\,\mathrm{d}s \\
&\quad + \frac{t}{T}\int_t^T \bigl|f\bigl(s, x(s,0,z,\mu)\bigr) - f\bigl(s, x(s,0,z,\nu)\bigr)\bigr|\,\mathrm{d}s. \qquad (3.62)
\end{aligned}$$

Putting

$$r(t) := \bigl|x(t,0,z,\mu) - x(t,0,z,\nu)\bigr|, \quad t \in [0, T], \tag{3.63}$$

and $\hat{r} := \operatorname{col}(\max_{t\in[0,T]} r_1(t), \dots, \max_{t\in[0,T]} r_n(t))$, and using the Lipschitz condition (3.5), from (3.62) we find

$$r(t) \leqslant K\left[\left(1 - \frac{t}{T}\right)\int_0^t r(s)\,\mathrm{d}s + \frac{t}{T}\int_t^T r(s)\,\mathrm{d}s\right]. \tag{3.64}$$

In view of (3.64), application of Lemma 3.16 yields

$$r(t) \leqslant K\left[\left(1 - \frac{t}{T}\right)\int_0^t \hat{r}\,\mathrm{d}s + \frac{t}{T}\int_t^T \hat{r}\,\mathrm{d}s\right] \equiv \alpha_{1,T}(t)K\hat{r}. \tag{3.65}$$

Taking into account inequality (3.65) and Lemma 3.16 and arguing by induction, from (3.64) we obtain the estimate

$$r(t) \leqslant \alpha_{m+1,T}(t)K^{m+1}\hat{r}, \quad m = 0, 1, 2, \dots, \tag{3.66}$$

where $\alpha_{m,T}(\cdot)$ are positive functions defined by formula (3.35). By virtue of inequality (3.36), relation (3.66) implies

$$r(t) \leqslant \left(\frac{3T}{10}K\right)^m K\bar{\alpha}_{1,T}(t)r_0 = Q^m K\bar{\alpha}_{1,T}(t)\hat{r}, \quad t \in [0, T],$$

where $\bar{\alpha}_{1,T}$ is given by (3.37). Hence, by virtue of estimate (3.46) from Remark 3.18, we obtain

$$\hat{r} \leqslant \frac{5T}{9} Q^m K \hat{r}, \quad m = 1, 2, \ldots. \tag{3.67}$$

On the assumption that the eigenvalues of matrix (3.45) are less than 1, inequality (3.67) is possible only when $\hat{r} = 0$ and, hence, we have $\mu = \nu$ in (3.63). This contradiction proves that (3.59) is the only suitable value of the parameter μ. □

EXAMPLE 3.20. Let us consider the differential system (3.14) from Example 3.8 in domain (3.15) and try to determine approximately the control parameter $\mu = \mu(z)$, for which the solution of the system

$$\begin{aligned} x_1'(t) &= x_2(t) - \mu_1, \\ x_2'(t) &= -x_1(t) - \big(x_2(t)\big)^s \sin 20t + \cos 20t - \mu_2, \quad t \in [0, T], \end{aligned} \tag{3.68}$$

such that

$$x_1(0) = z_1, \qquad x_2(0) = z_2, \tag{3.69}$$

where $(z_1, z_2) \in D_{\frac{T}{2}M}$, is periodic with period $T = \frac{\pi}{10}$.

Assume that s in (3.14) and a in (3.15) satisfy inequalities (3.17) and (3.18). Then we can seek for the control

$$\mu(z) = \operatorname{col}\big(\mu_1(z), \mu_2(z)\big)$$

as the limit

$$\mu(z) = \lim_{m\to\infty} \mu_m(z) \tag{3.70}$$

of the successive approximations

$$\mu_m(z) = \operatorname{col}\big(\mu_{1m}(z), \mu_{2m}(z)\big), \quad m = 0, 1, 2, \ldots,$$

defined by the equality

$$\mu_m(z) = \frac{1}{T} \int_0^T f\big(t, x_m(t, z)\big)\, \mathrm{d}t$$

with f given by formula (3.16). Thus, we have

$$\begin{aligned} \mu_1(z) &= \lim_{m\to\infty} \frac{10}{\pi} \int_0^{\frac{\pi}{10}} x_{2m}(t, z)\, \mathrm{d}t, \\ \mu_2(z) &= \lim_{m\to\infty} \frac{10}{\pi} \int_0^{\frac{\pi}{10}} \big[-x_{1m}(t, z) - \sin 20t \big(x_{2m}(t, z)\big)^s + \cos 20t\big]\, \mathrm{d}t, \end{aligned}$$

where $x_m(t,z) = \operatorname{col}(x_{1m}(t,z), x_{2m}(t,z))$,

$$x_{10}(t,z) = z_1, \qquad x_{20}(t,z) = z_2,$$

and, for all $m = 1, 2, \dots$ and $t \in [0, T]$,

$$x_{1m}(t,z) = z_1 + \int_0^t x_{2,m-1}(t,z)\,\mathrm{d}t - \frac{10t}{\pi}\int_0^{\frac{\pi}{10}} x_{2,m-1}(t,z)\,\mathrm{d}t, \tag{3.71}$$

$$\begin{aligned} x_{2m}(t,z) &= z_2 + \int_0^t \bigl[-x_{1,m-1}(t,z) - \sin 20t\bigl(x_{2,m-1}(t,z)\bigr)^s + \cos 20t\bigr]\,\mathrm{d}t, \\ &\quad - \frac{10t}{\pi}\int_0^{\frac{\pi}{10}} \bigl[-x_{1,m-1}(t,z) - \sin 20t\bigl(x_{2,m-1}(t,z)\bigr)^s + \cos 20t\bigr]\,\mathrm{d}t. \end{aligned} \tag{3.72}$$

Let us compute the value of the control parameter for system (3.14) with $s = 1$ (i.e., when (3.14) is a linear system with periodic coefficients) and $a = 2$. In this case (see Example 3.8), we have

$$\begin{aligned} D &= \bigl\{(x_1, x_2) \mid |x_1| \leqslant 2, |x_2| \leqslant 1\bigr\}, \\ D_{\frac{\pi}{20}M} &= \left\{(x_1, x_2) \;\middle|\; |x_1| \leqslant 2 - \frac{\pi}{20}, |x_2| \leqslant 1 - \frac{\pi}{5}\right\}. \end{aligned} \tag{3.73}$$

Let us put, e.g.,

$$z_1 = \frac{3}{2}, \qquad z_2 = \frac{1}{4} \tag{3.74}$$

in the initial value condition (3.69). It is easy to see that the vector (z_1, z_2) defined by (3.74) belongs to set (3.73).

According to (3.71) with $m = 1$, the first coordinate of the first approximation $x_1(\cdot, z)$ has the form

$$x_{11}(t,z) = \frac{3}{2} + \int_0^t \frac{1}{4}\,\mathrm{d}t - \frac{10t}{\pi}\int_0^{\frac{\pi}{10}} \frac{1}{4}\,\mathrm{d}t = \frac{3}{2}.$$

Quite similarly, (3.72) yields

$$\begin{aligned} x_{21}(t,z) &= \frac{1}{4} + \int_0^t \left[-\frac{3}{2} - \frac{1}{4}\sin 20t + \cos 20t\right]\mathrm{d}t \\ &\qquad - \frac{10t}{\pi}\int_0^{\frac{\pi}{10}} \left[-\frac{3}{2} - \frac{1}{4}\sin 20t + \cos 20t\right]\mathrm{d}t \\ &= \frac{19}{80} + \frac{1}{80}\cos 20t + \frac{1}{20}\sin 20t. \end{aligned} \tag{3.75}$$

Therefore, in view of (3.70), the control parameters μ_1, μ_2, in the first approximation, have the following values:

$$\mu_{11} = \frac{10}{\pi}\int_0^{\frac{\pi}{10}}\left[\frac{19}{80} + \frac{1}{80}\cos 20t + \frac{1}{20}\sin 20t\right]\mathrm{d}t = \frac{19}{80}$$
$$\approx 0.2125,$$
$$\mu_{21} = \frac{10}{\pi}\int_0^{\frac{\pi}{10}}\left[-\frac{3}{2} - \sin 20t\left(\frac{19}{80} + \frac{1}{80}\cos 20t + \frac{1}{20}\sin 20t\right)\right]\mathrm{d}t$$
$$= -\frac{3}{2} - \frac{\pi}{40}.$$

Analogously, by using (3.71), we can easily compute the second approximations to the limit function ($m = 2$):

$$x_{12}(t,z) = \frac{3}{2} + \int_0^t\left[\frac{19}{80} + \frac{1}{80}\cos 20t + \frac{1}{20}\sin 20t\right]\mathrm{d}t$$
$$- \frac{10t}{\pi}\int_0^{\frac{\pi}{10}}\left[\frac{19}{80} + \frac{1}{80}\cos 20t + \frac{1}{20}\sin 20t\right]\mathrm{d}t$$
$$= \frac{601}{400} + \frac{1}{1600}\sin 20t - \frac{1}{400}\cos 20t \tag{3.76}$$

and

$$x_{22}(t,z) = \frac{1}{4} + \int_0^t\left[-\frac{3}{2} - \sin 20t\left(\frac{19}{80} + \frac{1}{80}\cos 20t\right.\right.$$
$$\left.\left. + \frac{1}{20}\sin 20t + \cos 20t\right) + \cos 20t\right]\mathrm{d}t$$
$$- \frac{10t}{\pi}\int_0^{\frac{\pi}{10}}\left[-\frac{3}{2} - \sin 20t\left(\frac{19}{80} + \frac{1}{80}\cos 20t\right.\right.$$
$$\left.\left. + \frac{1}{20}\sin 20t + \cos 20t\right) + \cos 20t\right]\mathrm{d}t$$
$$= \frac{1523}{6400} + \frac{19}{1600}\cos 20t + \frac{1}{20}\sin 20t + \frac{1}{6400}\cos 40t + \frac{1}{1600}\sin 40t.$$

The control parameters μ_1, μ_2, in the second approximation, are given by the formulae

$$\mu_{12} = \frac{10}{\pi}\int_0^{\frac{\pi}{10}}\left[\frac{1523}{6400} + \frac{19}{1600}\cos 20t + \frac{1}{20}\sin 20t\right.$$
$$\left. + \frac{1}{6400}\cos 40t + \frac{1}{1600}\sin 40t\right] = \frac{1523}{6400} \approx 0.2379,$$

and

$$\mu_{22} = \frac{10}{\pi}\int_0^{\frac{\pi}{10}}\Bigg[-\frac{3}{2} - \frac{1}{1600}\sin 20t + \frac{1}{400}\cos 20t$$
$$- \sin 20t\left(\frac{1523}{6400} + \frac{19}{1600}\cos 20t + \frac{1}{20}\sin 20t\right.$$
$$\left.+ \frac{1}{6400}\cos 40t + \frac{1}{1600}\sin 40t\right)\Bigg]\,\mathrm{d}t = -\frac{3}{2} - \frac{\pi}{400}.$$

Note that (3.75) and (3.76) are the first and second approximations to the exact solution of the initial value problem (3.68), (3.74) with z_1, z_2 given by (3.74). One can verify that an exact solution of this problem exists for control parameter

$$\mu = \mathrm{col}(\mu_1, \mu_2) = \mathrm{col}\left(\frac{1}{4}, -\frac{3}{2}\right).$$

This solution, at the same time, is a $\frac{\pi}{10}$-periodic function as well as $x^*(t, z)$ and satisfies the initial value conditions (3.69) with z_1, z_2 given by (3.74). Higher order approximations can be found in a similar way by using formulae (3.70), (3.71).

Let us find necessary and sufficient conditions for the limit function $x^*(\cdot, z)$ of sequence (3.29) to be a solution of the original T-periodic boundary value problem n (3.3), (3.4).

THEOREM 3.21. *For the T-periodic boundary value problem* (3.3), (3.4), *let conditions* (3.5)–(3.7), *and* (3.8) *hold. Then the solution $x = x(\cdot, 0, z)$ of the Cauchy problem* (3.3), (3.27) *with $z \in D_{\frac{T}{2}M}$ is a solution of the T-periodic problem* (3.3), (3.4) *if, and only if the equality*

$$\int_0^T f\big(t, x^*(t, z)\big)\,\mathrm{d}t = 0, \tag{3.77}$$

holds, where $x^*(\cdot, z)$ is the limit function of sequence (3.29). *Moreover, in this case*

$$x(t, 0, z) \equiv x^*(t, z) \tag{3.78}$$

and for all $m = 1, 2, \ldots$ inequality (3.44) *holds for the deviation of the exact solution* (3.78) *of the T-periodic boundary value problem* (3.3), (3.4) *from its approximate T-periodic solution $x = x_m(\cdot, z)$ given by* (3.29) .

PROOF. The sufficiency of condition (3.77) follows immediately from the fact that the limit function $x^*(\cdot, z)$ satisfying the periodic boundary condition (3.4). and the initial value condition (3.27), is a solution of the integral equation (3.41) or, which is the same, of the integro-differential equations (3.42), (3.43). Obviously, if (3.43) is true, this condition is sufficient for $x^*(\cdot, z)$ to be a solution both of the periodic boundary value problem (3.3), (3.4) and the initial value problem (3.3), (3.27). In this case, (3.78) holds.

The necessity of condition (3.77) follows from the fact that if the solution $x = x(\cdot, 0, x_0)$ of (3.3), (3.27) is a solution of the T-periodic boundary value problem (3.3), (3.4), then the solution $x = x(\cdot, 0, z, \Delta)$ of system (3.42) with the same initial value $x(0, 0, z, \Delta) = z$ is T-periodic provided that $\Delta(z) = 0$ because, in this case,

$$x(t, 0, z, 0) \equiv x(t, 0, z) \equiv x^*(t, z) \tag{3.79}$$

and estimate (3.44) holds.

According to Theorem 3.19, a solution $x = x(t, 0, z, \mu)$ of the Cauchy problem (3.57), (3.58) is, at the same time, a T-periodic function only for the parameter value μ of the form (3.59). However, we already have this value of the parameter when$\mu = \Delta(z) = 0$. The theorem is proved. □

Let us show that the function $x^*(\cdot, z)$ is Lipschitzian with respect to z.

THEOREM 3.22. *Under the assumptions of Theorem* 3.17, *the function* $x^*(\cdot, z)$ *satisfies the following Lipschitz condition for all* $\{y, z\} \subset D_{\frac{T}{2}M}$ *and* $t \in [0, T]$:

$$\left|x^*(t, z) - x^*(t, y)\right| \leqslant \left[I + \bar{\alpha}_{1,T}(t) K (I - Q)^{-1}\right] |z - y|, \tag{3.80}$$

where $\bar{\alpha}_{1,T}$ *is the function* (3.37) *and the matrix* Q *is given by* (3.45).

PROOF. It follows immediately from (3.29) that

$$\begin{aligned} x_1(t, z) - x_1(t, y) &= (z - y) + \int_0^t \left[f(s, z) - f(s, y)\right] \mathrm{d}s \\ &\quad - \frac{t}{T} \int_0^T \left[f(s, z) - f(s, y)\right] \mathrm{d}s \\ &= (z - y) + \left(1 - \frac{t}{T}\right) \int_0^t \left[f(s, z) - f(s, y)\right] \mathrm{d}s \\ &\quad - \frac{t}{T} \int_t^T \left[f(s, z) - f(s, y)\right] \mathrm{d}s. \end{aligned} \tag{3.81}$$

Taking into account identity (3.81), the Lipschitz condition (3.5), and formula (3.35) of Lemma 3.16, we obtain

$$\begin{aligned} \left|x_1(t, z) - x_1(t, y)\right| &\leqslant |z - y| + \left[\left(1 - \frac{t}{T}\right) \int_0^t \mathrm{d}s + \frac{t}{T} \int_t^T \mathrm{d}s\right] K |z - y| \\ &= \left[I + \alpha_{1,T}(t) K\right] |z - y|. \end{aligned} \tag{3.82}$$

Similarly,

$$\begin{aligned}\left|x_2(t,z)-x_2(t,y)\right| \leqslant |z-y| + K\Bigg[\left(1-\frac{t}{T}\right)\int_0^t \left|x_1(s,z)-x_1(s,y)\right| \mathrm{d}s \\ + \frac{t}{T}\int_t^T \left|x_1(s,z)-x_1(s,y)\right| \mathrm{d}s\Bigg].\end{aligned}$$

Taking into account (3.82) and the recurrence relation (3.35), we get

$$\begin{aligned}\left|x_2(t,z)-x_2(t,y)\right| &\leqslant |z-y| + K\Bigg[\left(1-\frac{t}{T}\right)\int_0^t \left|\left(I+\alpha_{1,T}(s)K\right)\mathrm{d}s\right| \mathrm{d}s \\ &\quad + \frac{t}{T}\int_t^T \left(I+\alpha_{1,T}(s)K\right)\mathrm{d}s\Bigg] \\ &\leqslant \left[I+\alpha_{1,T}(t)K+\alpha_{2,T}(t)K^2\right]|z-y|.\end{aligned}$$

From these relations we conclude by induction that

$$\begin{aligned}\left|x_m(t,z)-x_m(t,y)\right| \leqslant \big(I+\alpha_{1,T}(t)K+\alpha_{2,T}(t)K^2 \\ + \cdots + \alpha_{m,T}(t)K^m\big)|z-y|. \qquad (3.83)\end{aligned}$$

Condition (3.7) means that the spectral radius $r(Q)$ of the matrix Q defined by (3.45) is less than 1. Passing to the limit as $m \to +\infty$ in (3.83) and using condition (3.7) together with estimate (3.36) of Lemma 3.16, we obtain

$$\begin{aligned}\left|x^*(t,z)-x^*(t,y)\right| &\leqslant \left(I+\bar{\alpha}_{1,T}(t)\sum_{i=0}^{\infty} K\left(\frac{3T}{10}K\right)^i\right)|z-y| \\ &= \left[I+\bar{\alpha}_{1,T}(t)K(I-Q)^{-1}\right]|z-y|,\end{aligned}$$

i.e., inequality (3.80) holds. □

When the set $D_{\frac{T}{2}M}$ given by (3.8) is not a singleton, to our T-periodic problem (3.3), (3.4) one can associate the “determining function”

$$\Delta : D_{\frac{T}{2}M} \to \mathbb{R}^n \qquad (3.84)$$

defined by formula (3.43).

Assuming that $D_{\frac{T}{2}M}$ contains more than one point, we now establish some properties of the determining function (3.84).

THEOREM 3.23. *Under the conditions of Theorem* 3.17, *the function*

$$D_{\frac{T}{2}M} \ni z \longmapsto \Delta(z) := \frac{1}{T}\int_0^T f\left(t, x^*(t,z)\right)\mathrm{d}t, \qquad (3.85)$$

where $x^*(\cdot, z)$ *is the limit function of sequence* (3.29), *is well-defined, continuous, and bounded inside the set* $D_{\frac{T}{2}M}$. *Furthermore, for all* $\{z, y\} \subset D_{\frac{T}{2}M}$, *it satisfies the Lipschitz condition*

$$\left|\Delta(z) - \Delta(y)\right| \leqslant K\left[I + \frac{10}{27}TK(I - Q)^{-1}\right]|z - y|. \tag{3.86}$$

PROOF. It follows from Theorem 3.17 that, for every $z \in D_{\frac{T}{2}M}$, the limit function $x^*(\cdot, z)$ of the uniformly convergent sequence (3.29) exists and is continuous. Therefore, when z varies over $D_{\frac{T}{2}M}$, the function Δ is also continuous and bounded,

$$\left|\Delta(z)\right| = \frac{1}{T}\left|\int_0^T f\big(t, x^*(t, z)\big)\,\mathrm{d}t\right| \leqslant \max_{(t,x)\in[0,T]\times D}\left|f(t, x)\right|.$$

Equality (3.85) implies

$$\begin{aligned}\left|\Delta(z) - \Delta(y)\right| &= \frac{1}{T}\left|\int_0^T \big[f\big(t, x^*(t, z)\big) - f\big(t, x^*(t, y)\big)\big]\,\mathrm{d}t\right| \\ &\leqslant \frac{1}{T}K\int_0^T \left|x^*(t, z) - x^*(t, y)\right|.\end{aligned} \tag{3.87}$$

Substituting (3.80) into (3.87), we get

$$\left|\Delta(z) - \Delta(y)\right| \leqslant \frac{1}{T}K\int_0^T \big[I + \bar{\alpha}_{1,T}(t)K(I - Q)^{-1}\big]\,\mathrm{d}t|z - y|. \tag{3.88}$$

Since

$$\int_0^T \bar{\alpha}_{1,T}(t)\,\mathrm{d}t = \frac{10}{9}\int_0^T 2t\left(1 - \frac{t}{T}\right)\mathrm{d}t = \frac{10}{27}T^2, \tag{3.89}$$

inequality (3.88) yields

$$\begin{aligned}\left|\Delta(z) - \Delta(y)\right| &\leqslant K|z - y| + \frac{10T}{27}K^2(I - Q)^{-1}|z - y| \\ &\leqslant K\left[I + \frac{10}{27}TK(I - Q)^{-1}\right]|z - y|,\end{aligned}$$

which completes the proof of our theorem. □

REMARK 3.24. Theorem 3.21 gives the following numerical-analytic algorithm for the construction of the T-periodic solutions of the non-autonomous T-periodic system (3.1).

(1) For a $z \in D_{\frac{T}{2}M}$, according to (3.29), we analytically construct the sequence of T-periodic functions $x_m(\cdot, z)$ depending on the parameter z.

(2) We find the limit $x^*(\cdot, z)$ of the sequence $x_m(\cdot, z)$.
(3) We construct the determining function Δ of the form (3.43).
(4) Using a suitable method for the numerical solution of non-linear algebraic or transcendental equations, we (approximately) find the solution(s) $z = z^*$ of the determining equation

$$\Delta(z) = 0 \tag{3.90}$$

lying in the set $D_{\frac{T}{2}M}$.
(5) After the substitution of the obtained value(s) $z = z^*$ into $x^*(\cdot, z)$, we get the (approximate) T-periodic solution(s) in the form $x^*(\cdot, z^*)$. This solution can also be obtained by solving the Cauchy problem $x(0) = z^*$ for the given differential system (3.3).

REMARK 3.25. Under our assumptions, the method under consideration allows one to find all the T-periodic solutions of (3.1) with values in the domain D because, by virtue of Theorem 3.21, the roots of the determining equation (3.90) “detect” the initial values of all the T-periodic solutions.

The principal difficulty in the realisation of the method is related with the analytic construction of the limit function $x^*(\cdot, z)$. However, in a number of cases, this problem can be avoided because, as can be shown, it is possible to prove the existence of a T-periodic solution on base of the properties of a certain approximation $x_m(\cdot, z)$ known in the analytic form.

3.2. *Existence theorems for periodic solutions*

In this section, we establish certain sufficient conditions for the solvability of periodic boundary value problem (3.3), (3.4) using the properties of successive approximations $\{x_m(\cdot, z)\}_{m\geqslant 0}$ of the form (3.29). Then, we indicate some necessary conditions for the existence of T-periodic solutions, i.e., conditions necessary for a certain subset of the set $D_{\frac{T}{2}M}$ to contain a point z^* determining the initial value $x^*(0) = z^*$ of the solution $x^*(\cdot)$ of the T-periodic problem (3.3), (3.4). We describe also the numerical algorithm for selecting these points.

To investigate the solvability of the periodic boundary value problem (3.3), (3.4), in addition to the *determining function* (3.85), we introduce the mth *approximate determining function*

$$D_{\frac{T}{2}M} \ni z \longmapsto \Delta_m(z) := \frac{1}{T} \int_0^T f\big(t, x_m(t, z)\big)\, \mathrm{d}t, \tag{3.91}$$

replacing $x^*(\cdot, z)$ in (3.85) by the function $x_m(\cdot, z)$ given by (3.29).

It is natural to expect that, under suitable conditions, the functions Δ and Δ_m are “close enough” to one another for m sufficiently large.

LEMMA 3.26. *Let us suppose that the function f satisfies assumptions* (3.5)–(3.7), *and* (3.8). *Then, for arbitrary $z \in D_{\frac{T}{2}M}$ and $m \geqslant 0$,*

$$\left|\Delta(z) - \Delta_m(z)\right| \leqslant \frac{10}{27} T K Q^m (I - Q)^{-1} M, \tag{3.92}$$

where the matrix Q and vector M are given by (3.45) *and* (3.6), *respectively.*

PROOF. By virtue of estimate (3.44) and relations (3.85), (3.91), we have

$$\begin{aligned}\left|\Delta(z) - \Delta_m(z)\right| &\leqslant \left|\frac{1}{T}\int_0^T \left[f\big(t, x^*(t,z)\big) - \int_0^T f\big(t, x_m(t,z)\big)\right] \mathrm{d}t\right| \\ &\leqslant \frac{1}{T} K \int_0^T \left|x^*(t,z) - x_m(t,z)\right| \mathrm{d}t \\ &\leqslant \frac{1}{T} K Q^m (I - Q)^{-1} M \int_0^T \bar{\alpha}_{1,T}(t)\, \mathrm{d}t.\end{aligned}$$

Taking (3.89) into account, we arrive at (3.92). □

REMARK 3.27. It is easy to see from Theorem 3.23, Lemma 3.26, and inequality (3.44) that the mappings (3.85), (3.91) are completely continuous.

Let us formulate a statement that establishes sufficient conditions for the solvability of the of the periodic boundary value problem (3.3), (3.4).

THEOREM 3.28. *Let us suppose that, in addition to assumptions* (3.5)–(3.7), *and* (3.8), *there exist a bounded closed subdomain $\Omega \subset D_{\frac{T}{2}M}$ and an number $m \in \mathbb{N}$ such that the approximate determining function* (3.91) *satisfies the following conditions*:

(a) *The relation*

$$\inf_{z \in \partial\Omega} \left|\Delta_m(z)\right| > \frac{10T}{27} K Q^m (I - Q)^{-1} M, \tag{3.93}$$

holds, where Q is the matrix given by (3.45);

(b) *The Brouwer degree*[1] *of Δ_m over Ω with respect to* 0 *satisfies the inequality*

$$\deg(\Delta_m, \Omega, 0) \neq 0. \tag{3.94}$$

Then the periodic boundary value problem (3.3), (3.4) *has a T-periodic solution $x = x^*(\cdot)$ with the initial value $x^*(0)$ belonging to Ω.*

---

[1] See, e.g., [17, Definition A2.5, Definition A2.9] and [81, Section 5].

PROOF. Recall that [17, Theorem A2.5] establishes the invariance of the topological degree by homotopy. Thus, if we show that the vector fields (3.85) and (3.91) are homotopic, then, in view of assumption (3.94),

$$\deg(\Delta, \Omega, 0) = \deg(\Delta_m, \Omega, 0), \tag{3.95}$$

whence we see that

$$\deg(\Delta, \Omega, 0) \neq 0. \tag{3.96}$$

Let us prove that the fields Δ and Δ_m are homotopic. For this purpose, we consider the "linear deformation" determined by the family of mappings

$$P(\theta, z) := \Delta_m(z) + \theta\bigl[\Delta(z) - \Delta_m(z)\bigr], \quad z \in \partial\Omega, \tag{3.97}$$

where $\theta \in [0, 1]$. Obviously, $P(\theta, \cdot)$ is a continuous mapping on $\partial\Omega$ for every θ from $[0, 1]$ and, furthermore,

$$P(0, z) = \Delta_m(z), \qquad P(1, z) = \Delta(z) \tag{3.98}$$

for all $z \in \partial\Omega$.

For an arbitrary z from $\partial\Omega$, in view of (3.98), (3.92), and (3.93), we have

$$\begin{aligned} |P(\theta, z)| &= \bigl|\Delta_m(z) + \theta\bigl[\Delta(z) - \Delta_m(z)\bigr]\bigr| \\ &\geqslant \bigl|\Delta_m(z)\bigr| - \bigl|\Delta(z) - \Delta_m(z)\bigr| > 0, \end{aligned} \tag{3.99}$$

i.e., $P(\theta, \cdot)$ does not vanish on $\partial\Omega$. We see that, being a non-degenerate deformation, (3.97) guarantees that (3.85) and (3.91) are homotopic. Therefore, (3.96) is true.

By [17, Theorem A2.4], inequality (3.96) guarantees the existence of a point $z^* \in \Omega$ such that $\Delta(z^*) = 0$. Consequently, in view of Theorem 3.21, the solution $x = x^*(\cdot)$ of the Cauchy problem

$$x(0) = z^*$$

for Eq. (3.3) is also a solution of the T-periodic boundary value problem (3.3), (3.4). Moreover,

$$x^*(t) = x^*\bigl(t, z^*\bigr), \quad t \in [0, T],$$

where $x^*(\cdot, z^*)$ is the limit function of the T-periodic sequence of functions defined by formula (3.29). □

REMARK 3.29. To verify condition (a) of Theorem 3.28 in concrete cases, one has to use the recurrence formula (3.29) to compute the function $x_m(\cdot, z)$ depending on $z \in \Omega$ as a parameter, construct the corresponding mapping (3.91), and verify whether its topological degree is non-zero.

This is a rather difficult problem in general. However, in a number of important cases, there are relatively simple criteria to study this question. In particular, when Δ_m is an odd mapping, i.e.,

$$\Delta_m(-z) = -\Delta_m(z),$$

for any z, then, according to the Borsuk theorem (see, e.g., [81, p. 174] or [17, Theorem A2.12]), the Brouwer degree $\deg(\Delta_m, \Omega, 0)$ is an odd number and, therefore, is not equal to zero.

Alternatively, it follows directly from [17, Definition A2.1] of the topological degree that if the Jacobian matrix of the function (3.91) is non-degenerate at its isolated zero $z_{m0} \in \Omega$, i.e., if

$$\det \frac{\partial \Delta_m}{\partial z}(z_{m0}) \neq 0,$$

then inequality (3.96) holds.

REMARK 3.30. The assertion of Theorem 3.28 is valid, in particular, for $m = 0$. In this case, condition (3.93) is replaced by the relation

$$\inf_{z \in \partial\Omega} \left| \int_0^T f(t, z)\, \mathrm{d}t \right| > \frac{10T^2}{27} K(I - Q)^{-1} M. \tag{3.100}$$

Here, we do not have to compute any approximations at all. Due to its simplicity, inequality (3.100) may prove convenient for the investigation of periodic solutions with "small" periods.

To seek for points z^* through which the T-periodic solutions of (3.3) pass at $t = 0$, i.e., to search for zeroes of the determining function (3.85), is by no means an easy problem, because the analytic form of this mapping is, in general, unknown. For this reason, one has to use approximate determining functions and apply statements similar to Theorem 3.28. However, there are particular cases where the existence of zeroes of function (3.85) is clear from some other considerations.

Let us consider one of such cases, where the right-hand side function of Eq. (3.1) has a certain symmetric property.

We recall the classical Arzelà–Ascoli theorem (see, e.g., Arzelà [5], Petrovski [58]).

THEOREM 3.31. (See [5].) *Let* $\{x_m(\cdot)\}_{m=1}^{\infty}$ *be a uniformly bounded and equicontinuous sequence of functions on the interval* $[\alpha, \beta]$. *Then it contains a subsequence* $\{x_{m_k}(\cdot)\}_{k=1}^{\infty}$ *uniformly convergent on* $[\alpha, \beta]$.

The uniform boundedness of the sequence means that there exists a positive constant K for which $\max_{t \in [\alpha,\beta]} |x_m(t)| \leqslant K$ independently of $m = 1, 2, \ldots$, whereas its equicontinuity means that, for an arbitrary $\varepsilon > 0$, there exists a $\delta > 0$ such that if $\{t_1, t_2\} \subset [\alpha, \beta]$ satisfy the inequality $|t_1 - t_2| < \delta$, then

$$\sup_{m \geqslant 1} \left| x_m(t_1) - x_m(t_2) \right| < \varepsilon. \quad m = 1, 2, \ldots.$$

The following theorem is a modified version of [89, Theorem 8.1].

THEOREM 3.32. *Let the right-hand side function* $f : \mathbb{R} \times D \to \mathbb{R}^n$ *of the differential system* (3.1) *be defined on* $\mathbb{R} \times D$, *where* D *is the closure of a bounded domain in* $\mathbb{R}^n$, *and let* f *satisfy the following conditions*:

(1) *The function* f *is continuous and periodic of period* T *in the first argument and, with a certain vector* $M \in \mathbb{R}^n_+$, *satisfies condition* (3.6);

(2) *The equality*

$$f(-t, x) = -f(t, x) \tag{3.101}$$

holds for all $t \in (-\infty, \infty)$ *and* $x \in D$;

(3) *Condition* (3.8) *is satisfied*: $D_{\frac{T}{2}M} \neq \varnothing$.

Then, for every $z \in D_{\frac{T}{2}M}$, *system* (3.1) *possesses an even* T*-periodic solution satisfying the initial condition*

$$x(0) = z. \tag{3.102}$$

If, moreover, there exists a non-negative matrix $K \in \mathrm{GL}_n(\mathbb{R})$ *such that* f *satisfies the Lipschitz* (3.5) *for all* $t \in \mathbb{R}$ *and* $\{x_1, x_2\}$ *from* D, *then, for any* $z \in D_{\frac{T}{2}M}$, *system* (3.1) *has no other solutions with property* (3.102) *and graph lying in* D.

In other words, under conditions of Theorem 3.32, the set $D_{\frac{T}{2}M}$ consists of the initial values of T-periodic solutions of system (3.1).

PROOF OF THEOREM 3.32. Let us fix a vector $z \in D_{\frac{T}{2}M}$ and consider the Picard successive approximations

$$x_{m+1}(t) = z + \int_0^t f\bigl(s, x_m(s)\bigr)\, \mathrm{d}s, \quad m = 0, 1, 2, \ldots, \tag{3.103}$$

where $x_0(t) = z$ for all t from $(-\infty, \infty)$.

By assumption, $t \mapsto f(t, z)$ is an odd function for any fixed z and, hence,

$$\int_0^T f(t, z)\, \mathrm{d}t = 0.$$

Therefore, formula (3.103) and the T-periodicity of sequence (3.29) yield

$$x_1(t) = z + \int_0^t \left(f(s, z) - \frac{1}{T} \int_0^T f(\tau, z)\, \mathrm{d}\tau \right) \mathrm{d}s = x_1(t + T) \tag{3.104}$$

for any $t \in (-\infty, \infty)$, i.e., the function x_1 is T-periodic. Let us argue by analogy with the proof of Theorem 3.17. Taking assumption (3.6) into account, we arrive at estimate (3.47), whence, in view of (3.104), it follows that, for all t,

$$\left|x_1(t) - z\right| \leqslant \frac{T}{2} M,$$

i.e., the function x_1 takes values in D. Furthermore, being the integral of an odd function, x_1 is even, i.e., $x_1(t) = x_1(-t)$ for any t from $(-\infty, \infty)$.

Taking assumption (3.6) into account and arguing by induction, one can show that, for an arbitrary $m \geqslant 1$, the function $x_m(\cdot)$ given by formula (3.103) is well-defined on $\mathbb{R}$, periodic with period T, and satisfies the relations

$$\int_0^T f\big(t, x_m(t)\big)\,\mathrm{d}t = 0, \tag{3.105}$$

$$x_m(t) = x_m(-t), \tag{3.106}$$

and (3.48) for all t from $(-\infty, \infty)$.

Inequality (3.48) implies the uniform boundedness of sequence (3.103). The form of sequence (3.103) and the continuity of f on the bounded set $[0, T] \times D$ guarantees the equicontinuity of sequence (3.103). Thus, in view of the Arzelà–Ascoli theorem (Theorem 3.31), we can select a uniformly convergent subsequence $\{x_{m_k} \mid k \geqslant 0\}$ of the sequence of T-periodic functions $\{x_m \mid m \geqslant 0\}$. Hence, the function

$$\lim_{k\to\infty} x_{m_k} =: x \tag{3.107}$$

is well-defined on the entire $\mathbb{R}$. Setting $m = m_k$, $k = 1, 2, \ldots,$ in (3.103), (3.106) and passing to the limit as $k \to +\infty$, we obtain that x is an even function satisfying the differential system (3.1). Moreover, x is periodic with period T because each of the functions x_{m_k}, $k = 1, 2, \ldots,$ has this property and the limit in (3.107) is uniform. Thus, x is an even T-periodic solution of (3.1) satisfying condition (3.102).

Let now f satisfy the Lipschitz condition (3.5) for all $t \in \mathbb{R}$ and $\{x_1, x_2\}$ from D. According to assumption (3.5), the conditions of the Picard–Lindelöf existence and uniqueness theorem (see, for instance [17, Theorem 1.1.1]) hold, whence the uniqueness of the solution in question follows. □

EXAMPLE 3.33. Let us apply the results above to study the 1-periodic solutions of the system

$$\begin{aligned} x_1'(t) &= a x_2(t) \sin 4\pi t, \\ x_2'(t) &= b\big(x_1(t)\big)^2 \sin 2\pi t + \sin 4\pi t, \qquad t \in [0, T], \end{aligned} \tag{3.108}$$

where $a \neq 0$, $b \in \mathbb{R}$, $T = 1$.

It is obvious that property (3.101) holds for the right-hand side of (3.108). Let the system be defined in the domain

$$(t,x) \in [0,T] \times D, \quad D\colon |x_1| \leqslant A, |x_2| \leqslant B. \tag{3.109}$$

Then

$$\bigl|f(t,x)\bigr| \leqslant \operatorname{col}\bigl(|a|B, 1+|b|A^2\bigr). \tag{3.110}$$

The set (3.8) for system (3.108) is defined as follows:

$$D_{\frac{T}{2}M} = \left\{(x_1,x_2) \,\middle|\, |x_1| \leqslant A - \frac{1}{2}|a|B, |x_2| \leqslant B - \frac{1}{2}\bigl[1+|b|A^2\bigr]\right\}. \tag{3.111}$$

This set is non-empty if

$$\begin{aligned} &A - \frac{1}{2}|a|B \geqslant 0, \\ &B - \frac{1}{2}\bigl[1+|b|A^2\bigr] \geqslant 0, \end{aligned}$$

i.e., when

$$\frac{1}{2} + \frac{1}{2}|b|A^2 \leqslant B \leqslant \frac{2A}{|a|}. \tag{3.112}$$

The Lipschitz condition (3.5) holds with the matrix

$$K = \begin{bmatrix} 0 & |a| \\ 2b|A| & 0 \end{bmatrix}.$$

Consequently, if our domain satisfies condition (3.112), by virtue of Theorem 3.32, every solution $x(t)$ of system (3.108) with the initial value $x(0)$ lying in domain (3.111) is periodic with period 1. For instance, if

$$a = \frac{1}{2}, \quad b = -1, \quad A = 1,$$

and, according to (3.112),

$$1 \leqslant B \leqslant 4,$$

then the set

$$D_{\frac{T}{2}M} = \left\{(x_1,x_2) \,\middle|\, |x_1| \leqslant 1 - \frac{B}{4}, |x_2| \leqslant B - 1\right\}$$

consists of the initial values of 1-periodic solutions of system (3.108). In particular, for $A = 1$ and $B = 3$, all points from the set

$$D_{\frac{T}{2}M} = \left\{ (x_1, x_2) \;\middle|\; |x_1| \leqslant \frac{1}{4}, |x_2| \leqslant 2 \right\}$$

can serve as the initial values of 1-periodic solutions of system (3.108).

Let us find the first approximation of the 1-periodic solutions of the given system (3.108), supposing that the parameters a, b and the domain (3.109) such that the inequalities (3.112) are true and the vector of initial values $z = \operatorname{col}(z_1, z_2)$ belongs to subset $D_{\frac{T}{2}M}$ of the form (3.111). According to formula (3.103) for $m = 0$ we get:

$$\begin{aligned} x_{11}(t) &= z_1 + \int_0^t a z_2 \sin 4\pi s \,\mathrm{d}s = \left(z_1 + \frac{a z_2}{4\pi} \right) - \frac{a z_2}{4\pi} \cos 4\pi t, \\ x_{21}(t) &= z_2 + \int_0^t \left[b z_1^2 \sin 2\pi s + \sin 4\pi s \right] \mathrm{d}s \\ &= \left(z_2 + \frac{b z_1^2}{2\pi} + \frac{1}{4\pi} \right) - \frac{b z_1^2}{2\pi} \cos 2\pi t - \frac{\cos 4\pi t}{4\pi}. \end{aligned} \tag{3.113}$$

Similarly to (3.113) one can easily get the second approximation.

EXAMPLE 3.34. Consider the damped Duffing equation with periodic forcing term (see (4.3.15) in [17])

$$x''(t) + d x'(t) + a x + b\big(x(t)\big)^3 = H \sin \frac{2\pi t}{T}, \quad t \in [0, T], \tag{3.114}$$

for some $T > 0$, $d \in \mathbb{R}$, $a \in \mathbb{R}$, $H \in \mathbb{R}$, $b \in \mathbb{R}$ (if $b > 0$ this is the "hard spring case," and if $b < 0$ this is the "soft spring case"). Rewriting the differential equation (3.114) in the Cauchy normal form, we get

$$\begin{aligned} x_1'(t) &= x_2(t), \\ x_2'(t) &= H \sin \frac{2\pi t}{T} - d x_2(t) - a x_1(t) - b\big(x_1(t)\big)^3. \end{aligned} \tag{3.115}$$

Let the system (3.115) be defined for (t, x) from $[0, T] \times D$, where

$$D = \big\{ (x_1, x_2) \mid |x_1| \leqslant A, |x_2| \leqslant B \big\}. \tag{3.116}$$

Then

$$\big| f(t, x) \big| \leqslant \operatorname{col}\big(B, |H| + |d| B + |a| A + |b| A^3 \big).$$

The subset (3.8) for system (3.115) is given by the formula

$$
\begin{aligned}
D_{\frac{T}{2}M} = \Big\{ (x_1, x_2) \,\Big|\, |x_1| \leqslant A - \frac{T}{2} B, \\
|x_2| \leqslant B - \frac{T}{2}\big[|H| + |d|B + |a|A + |b|A^3\big] \Big\},
\end{aligned}
\tag{3.117}
$$

which is non-empty if

$$
\begin{aligned}
&A - \frac{T}{2} B \geqslant 0, \\
&B - \frac{T}{2}\big[|H| + |d|B + |a|A + |b|A^3\big] \geqslant 0,
\end{aligned}
$$

i.e., when

$$
\begin{aligned}
&\frac{T}{2 - T|d|}\big[|H| + |a|A + |b|A^3\big] \leqslant B \leqslant \frac{2}{T} A, \\
&T|d| < 2.
\end{aligned}
\tag{3.118}
$$

The Lipschitz condition (3.5) is fulfilled for system (3.115) with the matrix

$$
K = \begin{bmatrix} 0 & 1 \\ |a| + 3|b|A^2 & |d| \end{bmatrix},
\tag{3.119}
$$

whose greatest eigenvalue is equal to

$$
r(K) = \frac{|d|}{2} + \sqrt{\frac{d^2}{4} + |a| + 3|b|A^2}.
$$

Therefore, the maximal eigen-value of the matrix Q of the form (3.45) is less that 1 if

$$
\frac{|d|}{2} + \sqrt{\frac{d^2}{4} + |a| + 3|b|A^2} < \frac{10}{3T}.
\tag{3.120}
$$

Thus, for system (3.115), conditions (3.5)–(3.7), and (3.8) hold in domain (3.116) if the parameters in Eq. (3.114) satisfy inequalities (3.117)–(3.120). If, for instance,

$$
T = \frac{1}{2}, \quad b = 2, \quad d = 1, \quad a = 1, \quad H = 1, \quad A = 1
\tag{3.121}
$$

(the “hard spring case”), the domain D is

$$
D = \big\{ (x_1, x_2) \mid |x_1| \leqslant 1, |x_2| \leqslant B \big\}
\tag{3.122}
$$

and, according to (3.118),

$$\frac{4}{3} \leqslant B \leqslant 4.$$

In particular, for $B = 2$ the subset $D_{\frac{T}{2}M}$ has the form

$$D_{\frac{T}{2}M} = \left\{ (x_1, x_2) \;\middle|\; |x_1| \leqslant \frac{1}{2}, |x_2| \leqslant \frac{1}{2} \right\} \tag{3.123}$$

and, thus, the initial value of the $\frac{1}{2}$-periodic solution must belong to the set (3.123). The successive approximations (3.29) are well-defined. So, the first approximation has the form:

$$\begin{aligned} x_{11}(t, z) &= z_1 + \int_0^t \left[z_2 - \frac{1}{T} \int_0^T z_2 \, \mathrm{d}\tau \right] \mathrm{d}s = z_1, \\ x_{21}(t, z) &= z_2 + \int_0^t \left[H \sin \frac{2\pi}{T} s - d z_2 - a z_1 - b z_1^3 \right. \\ &\qquad \left. - \frac{1}{T} \int_0^T \left(H \sin \frac{2\pi}{T} \tau - d z_2 - a z_1 - b z_1^3 \right) \mathrm{d}\tau \right] \mathrm{d}s \\ &= z_2 + \frac{HT}{2\pi} - \frac{HT}{2\pi} \cos \frac{2\pi}{T} t = z_2 + \frac{1}{4\pi} - \frac{1}{4\pi} \cos 4\pi t. \end{aligned} \tag{3.124}$$

The first approximation to the determining equation of the form (3.91) is

$$\begin{aligned} \Delta_{11}(z) &= \frac{1}{T} \int_0^T \left[z_2 + \frac{1}{4\pi} - \frac{1}{4\pi} \cos 4\pi t \right] \mathrm{d}t \\ &= z_2 + \frac{1}{4\pi} = 0, \\ \Delta_{21}(z) &= \frac{1}{T} \int_0^T \left[\sin 4\pi t - \left(-\frac{1}{4\pi} + \frac{1}{4\pi} - \frac{1}{4\pi} \cos 4\pi t \right) - z_1 - 2 z_1^3 \right] \mathrm{d}t \\ &= -z_1 - 2 z_1^3 = 0, \end{aligned} \tag{3.125}$$

whence

$$z_2 = -\frac{1}{4\pi}, \qquad z_1 = 0.$$

Thus, the initial value of the $\frac{1}{2}$-periodic solution of system (3.115), in the first approximation, is

$$\operatorname{col}(z_1, z_2) = \operatorname{col}\left(0, -\frac{1}{4\pi}\right).$$

Let us take, instead of (3.121), the following values:

$$T = \frac{1}{2}, \quad b = -1, \quad d = 1, \quad a = \frac{1}{9}, \quad H = 1, \quad A = 1 \tag{3.126}$$

(the "soft spring case"). Then, in view of (3.118), the domain D is given by (3.122), where

$$\frac{19}{27} < B \leqslant 4. \tag{3.127}$$

In particular, for $A = 1$ and $B = 2$ the subset $D_{\frac{T}{2}M}$ is the following:

$$D_{\frac{T}{2}M} = \left\{ (x_1, x_2) \,\middle|\, |x_1| \leqslant \frac{1}{2}, |x_2| \leqslant \frac{19}{28} \right\}. \tag{3.128}$$

Taking into account (3.29) and performing the computation, we obtain that the first approximation is also given by formula (3.124). However, instead of (3.125), we have

$$\begin{aligned} \Delta_{21}(z) &= \frac{1}{T} \int_0^T \left[\sin 4\pi t + \frac{1}{4\pi} \cos 4\pi t - \frac{1}{9} z_1 + z_1^3 \right] \mathrm{d}t \\ &= -\frac{1}{9} z_1 + z_1^3 = 0, \end{aligned}$$

whence

$$z_1 \left(-\frac{1}{9} + z_1^2 \right) = 0, \quad z_1^{(1)} = 0, \quad z_1^{(2)} = \frac{1}{3}, \quad z_1^{(3)} = -\frac{1}{3}.$$

Thus, for the Duffing equation (3.114), (3.126) (the "soft spring case"), in the domain (3.128), on base of the first approximation, we find three initial values for the $\frac{1}{2}$-periodic solutions:

$$\begin{aligned} &\operatorname{col}(z_1, z_2) = \operatorname{col}\left(0, -\frac{1}{4\pi} \right), \\ &\operatorname{col}(z_1, z_2) = \operatorname{col}\left(\frac{1}{3}, -\frac{1}{4\pi} \right), \qquad \operatorname{col}(z_1, z_2) = \operatorname{col}\left(-\frac{1}{3}, -\frac{1}{4\pi} \right). \end{aligned} \tag{3.129}$$

Note that it is impossible to apply the results of Rouche and Mawhin [81] (see also Theorem 4.3.4 in Farkas [17]) to study the Duffing differential equation (3.114) with parameters (3.121), (3.126) in the domain D: $|x_1| \leqslant 1$, $|x_2| \leqslant 2$, because we assume that the "hard spring case" ($b > 0$) may take place and, moreover, the inequalities

$$aA + bA^3 + H \sin \omega t < 0, \quad -aA - bA^3 + H \sin \omega t > 0$$

of the mentioned [17, Theorem 4.3.4] (see also Corollary 5.6 from [81, p. 195]) do not hold in the domain D under consideration. These inequalities may hold in a "smaller" domain. So, this approach adds to the possibilities of investigation of the Duffing type equation (3.114).

EXAMPLE 3.35. Let us investigate the T-periodic solution of the non-autonomous Van der Pol equation with a periodic forcing term of the form

$$x''(t) + a\big(\big(x(t)\big)^2 - 1\big)x'(t) + x(t) = H \sin \omega t, \quad t \in [0, T], \tag{3.130}$$

where $T = 2\pi\omega^{-1}$, $a \in \mathbb{R}$, $H \in \mathbb{R}$. Equation (3.130) can be rewritten as the system

$$\begin{aligned} x_1'(t) &= x_2(t), \\ x_2'(t) &= H \sin \omega t - a\big(\big(x_1(t)\big)^2 - 1\big)x_2(t) - x_1(t), \quad t \in [0, T], \end{aligned} \tag{3.131}$$

which is considered in $[0, T] \times D$ with

$$D = \big\{(x_1, x_2) \mid |x_1| \leqslant A, |x_2| \leqslant B\big\}. \tag{3.132}$$

It is easy to see that f, the right-hand side function in (3.131), satisfies the estimate

$$\big|f(t, x)\big| \leqslant \operatorname{col}\big(B, |H| + |a|\big(A^2 + 1\big)B + A\big) \tag{3.133}$$

for any $(t, x) \in [0, T] \times D$. For M, we can thus take the term from the right-hand side of (3.133). Then

$$\begin{aligned} D_{\frac{T}{2}M} = \bigg\{(x_1, x_2) \,\bigg|\, & |x_1| \leqslant A - \frac{T}{2}B, \\ & |x_2| \leqslant B - \frac{T}{2}\big[|H| + |a|(A^2 + 1)B + A\big]\bigg\}. \end{aligned} \tag{3.134}$$

The set $D_{\frac{T}{2}M}$ is non-empty if

$$A - \frac{T}{2}B \geqslant 0, \tag{3.135}$$

$$B - \frac{T}{2}\big[|H| + |a|(A^2 + 1)B + A\big] \geqslant 0. \tag{3.136}$$

It is easy to see that inequalities (3.135) and (3.136) hold whenever

$$\frac{T[|H| + A]}{2 - T|a|(A^2 + 1)} \leqslant B \leqslant \frac{2}{T}A, \tag{3.137}$$

for a and A such that $T|a|(A^2 + 1) < 2$.

For the system (3.131) the Lipschitz condition (3.5) holds with the matrix

$$K = \begin{bmatrix} 0 & 1 \\ 1+2|a|AB & |a|\bigl(A^2+1\bigr) \end{bmatrix},$$

whose eigenvalues λ are given by the equalities

$$\lambda = \frac{|a|(A^2+1)}{2} \pm \sqrt{\frac{a^2(A^2+1)^2}{4} + 1 + 2|a|AB}.$$

The greatest eigenvalue of the matrix Q of the form (3.45) is, therefore, less then 1 whenever

$$|a|\bigl(A^2+1\bigr) + \sqrt{a^2\bigl(A^2+1\bigr)^2 + 4 + 8|a|AB} < \frac{20}{3T}. \tag{3.138}$$

Thus, conditions (3.5), (3.6), (3.7), and (3.8) are fulfilled for system (3.131) in domain (3.132) if (3.134), (3.137), and (3.138) are true. For instance, if

$$T = 1, \quad \omega = 2\pi, \quad a = -\frac{1}{2}, \quad A = 1, \quad H = \frac{1}{2}, \tag{3.139}$$

then (3.137) means that $\frac{3}{2} \leqslant B \leqslant 2$. Specifically, for the domain

$$D = \left\{ (x_1, x_2) \;\middle|\; |x_1| \leqslant 1 = A, |x_2| \leqslant \frac{7}{4} \right\} \tag{3.140}$$

corresponding to $B = \frac{7}{4}$, set (3.134) is given by the formula

$$D_{\frac{T}{2}M} = \left\{ (x_1, x_2) \;\middle|\; |x_1| \leqslant \frac{1}{8}, |x_2| \leqslant \frac{1}{8} \right\}. \tag{3.141}$$

Thus, for the Van der Pol system (3.131) conditions (3.5)–(3.140) with parameters (3.139), (3.141) and, hence, the successive approximations (3.29) are well-defined.

According to (3.29), the first approximation is

$$\begin{aligned} x_{11}(t,z) &= z_1 + \int_0^t \left[z_2 - \int_0^1 z_2 \,\mathrm{d}\tau \right] \mathrm{d}s = z_1, \\ x_{21}(t,z) &= z_2 + \int_0^t \left[\frac{1}{2}\sin 2\pi t + \frac{1}{2}\bigl(z_1^2 - 1\bigr) z_2 - z_1 \right. \\ &\qquad \left. - \int_0^1 \left(\frac{1}{2}\sin 2\pi\tau + \frac{1}{2}\bigl(z_1^2-1\bigr)z_2 - z_1 \right) \mathrm{d}\tau \right] \mathrm{d}s \\ &= z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi t. \end{aligned}$$

The first approximation to determining equation (3.91) has the form

$$\Delta_{11}(z) = \int_0^1 \left[z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 4\pi t \right] \mathrm{d}t = z_2 + \frac{1}{4\pi} = 0,$$

$$\Delta_{21}(z) = \int_0^1 \left[\frac{1}{2}\sin 2\pi t + \frac{1}{2}\big(z_1^2 - 1\big)\left(z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi t \right) - z_1 \right] \mathrm{d}t$$
$$= \int_0^1 \left[-\frac{1}{8\pi} z_1^2 + \frac{1}{8\pi} + \frac{1}{8\pi} z_1^2 - \frac{1}{8\pi} - z_1 \right] \mathrm{d}t = 0,$$

whence we find that

$$z_2 = -\frac{1}{4\pi}, \qquad z_1 = 0.$$

So, the approximate initial value of the 1-periodic solution of the Van der Pol system (3.131), (3.139) is given by the formula

$$\operatorname{col}(z_1, z_2) = \operatorname{col}\left(0, -\frac{1}{4\pi} \right).$$

The components of the second approximation are:

$$x_{12}(t, z) = z_1 + \int_0^t \left[z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi t - \int_0^1 \left(z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi s \right) \right] \mathrm{d}t = z_1 - \frac{1}{8\pi^2}\sin 2\pi t,$$

$$x_{22}(t, z) = z_2 + \int_0^t \left[\frac{1}{2}\sin 2\pi t + \frac{1}{2}\big(z_1^2 - 1\big)\left(z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi t \right) - z_1 - \int_0^1 \left[\frac{1}{2}\sin 2\pi s + \frac{1}{2}\big(z_1^2 - 1\big)\left(z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi s \right) - z_1 \right] \mathrm{d}s \right] \mathrm{d}t$$
$$= z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi t - \frac{1}{16\pi^2}\big(z_1^2 - 1\big)\sin 2\pi t,$$

whereas the second approximate determining system

$$\Delta_{12}(z) = \int_0^1 \left[z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi t - \frac{1}{16\pi^2}(z_1^2 - 1)\sin 2\pi t \right] \mathrm{d}t$$
$$= z_2 + \frac{1}{4\pi} = 0,$$

$$\begin{aligned}\Delta_{22}(z) &= \int_0^1 \biggl[\frac{1}{2}\sin 2\pi t + \frac{1}{2}\biggl(z_1^2 - \frac{1}{4\pi^2}\sin 2\pi t + \frac{1}{64\pi^4}\sin^2 2\pi t - 1\biggr) \\ &\quad \times \biggl(z_2 + \frac{1}{4\pi} - \frac{1}{4\pi}\cos 2\pi t - \frac{1}{16\pi^2}\bigl(z_1^2 - 1\bigr)\sin 2\pi t\biggr) \\ &\quad - z_1 + \frac{1}{8\pi^2}\sin 2\pi t\biggr]\,\mathrm{d}t \\ &= \frac{1}{2}\int_0^1 \biggl[\frac{1}{64\pi^4}\bigl(z_1^2 - 1\bigr)\sin^2 2\pi t - z_1\biggr]\,\mathrm{d}t \\ &= \frac{1}{128\pi^4}\bigl(z_1^2 - 1\bigr) - 2z_1 = 0,\end{aligned}$$

has the solution

$$z_2 = -\frac{1}{4\pi}, \qquad z_1 = 128\pi^4 \pm \sqrt{\bigl(128\pi^4\bigr)^2 + 1},$$

because z_1 satisfies the equality $z_1^2 - 256\pi^4 z_1 - 1 = 0$.

Taking into account the fact that $\operatorname{col}(z_1, z_2)$ must belong to domain (3.141), we obtain that, in the second approximation, the initial value of the 1-periodic solution of (3.131), (3.139) has the form

$$\operatorname{col}(z_1, z_2) = \operatorname{col}\biggl(128\pi^4 - \sqrt{\bigl(128\pi^4\bigr)^2 + 1}, -\frac{1}{4\pi}\biggr).$$

Again we note that Corollary 5.6 from [81, p. 195] cannot be applied to Van der Pol's system (3.131), (3.139) in domain (3.140), because the inequalities

$$-1 + \frac{1}{2}\sin\omega t > 0, \qquad 1 + \frac{1}{2}\sin\omega t < 0$$

from the assumptions of the corollary mentioned do not hold for all $t \in [0, T]$.

The following theorem gives a condition necessary for $G \subset D_{\frac{T}{2}M}$ to contain the initial value z^* of a T-periodic solution $x^*(t) = x^*(t, 0, z^*)$ of system (3.3).

THEOREM 3.36. *Suppose that the T-periodic boundary value problem* (3.3), (3.4) *satisfies conditions* (3.5), (3.6), (3.61), *and* (3.62). *Then in order that a subset $G \subset D_{\frac{T}{2}M}$ contain the initial value $z^* = x^*(0)$ of the T-periodic solution $x = x^*(t)$ of* (3.1) *at $t = 0$, it is necessary that, for arbitrary $w \in G$ and integer m, the following inequality be satisfied*:

$$\begin{aligned}\Delta_m(w) &\leqslant \sup_{z\in G} K\biggl[I + \frac{10}{27}TK(I-Q)^{-1}\biggr]|z - w| \\ &\quad + \frac{10}{27}TKQ^m(I-Q)^{-1}M, \end{aligned}\tag{3.142}$$

where Δ_m *is the* m*-th approximation* (3.91) *of the determining function* (3.85) *and* Q *is the matrix given by* (3.45).

PROOF. Assume that $z = z^*$ is a solution of the determining equation (3.85), i.e., $\Delta(z^*) = 0$. Consequently, according to Theorem 3.21, $z^* = x^*(0)$ is the initial value at $t = 0$ of the periodic solution $x^*(t) = x^*(t, z^*)$, $t \in [0, T]$, where $x^*(\cdot, z)$ is the limit function of the sequence (3.29).

Let apply estimate (3.86) from Theorem 3.23 with $z = z^*$, $y = w$. Then it follows from (3.86) that

$$\left|\Delta(w)\right| = K\left[I + \frac{10}{27}TK(I - Q)^{-1}\right]|z^* - w|. \tag{3.143}$$

In view of inequality (3.92) of Lemma 3.26, we obtain

$$\begin{aligned}
&\left|\Delta(w) - \Delta_m(w)\right| \leqslant \frac{10}{27}TKQ^m(I - Q)^{-1}M,\\
&\left|\Delta_m(w)\right| \leqslant \left|\Delta(w)\right| + \frac{10}{27}TKQ^m(I - Q)^{-1}M.
\end{aligned} \tag{3.144}$$

Combining (3.144) with (3.143), we prove the validity of inequality (3.142):

$$\begin{aligned}
\left|\Delta_m(w)\right| &\leqslant K\left[I + \frac{10}{27}TK(I - Q)^{-1}\right]|z^* - w| + \frac{10}{27}TKQ^m(I - Q)^{-1}M\\
&\leqslant \sup_{z \in G} K\left[I + \frac{10}{27}TK(I - Q)^{-1}\right]|z - w| + \frac{10}{27}TKQ^m(I - Q)^{-1}M.
\end{aligned}$$

The theorem is proved. □

REMARK 3.37. According to Theorem 3.36, we can indicate an algorithm for approximate determination of the initial value z^* of a solution of the T-periodic boundary value problem (3.3), (3.4). For this purpose, we represent the subset $D_{\frac{T}{2}M}$ defined in (3.8) as the union of finitely many subsets Ω_i, i.e.,

$$D_{\frac{T}{2}M} = \bigcup_{i=1}^{N} \Omega_i.$$

In every Ω_i, we choose an arbitrary point $w^i \in \Omega_i$ and, for a certain m, we compute the value of the mth successive approximation $x_m(t, w^i)$ according to formula (3.29) and the value of the corresponding approximate determining function, $\Delta_m(w^i)$ according to (3.91). Then, by using (3.142), we exclude the subsets Ω_i for an arbitrary point w^i of which the inequality inverse to (3.142), i.e.,

$$\left|\Delta_m(w^i)\right| > K\left[I + \frac{10}{27}TK(I - Q)^{-1}\right]\sup_{z \in G}|z - w^i| + \frac{10}{27}TKQ^m(I - Q)^{-1}M$$

holds. The reason for this is provided by Theorem 3.36, which guarantees that such subsets cannot contain initial values of solution of the periodic boundary value problem (3.3), (3.4).

The remaining subsets $\Omega_{i_1}, \Omega_{i_2}, \dots, \Omega_{i_s}$ form a certain set $\Omega_{m,N}$,

$$\Omega_{m,N} := \bigcup_{\nu=1}^{s} \Omega_{i_\nu},$$

such that only through its points the T-periodic solutions of system (3.1) can pass. When N and m tend to ∞, the set $\Omega_{m,N}$ "tends" (see Problem 10.1) to the set Ω^* of *all* the initial values of the T-periodic solutions with values in $D_{\frac{T}{2}M}$. Every point u from $\Omega_{m,N}$ can be regarded as an approximate value of an exact initial value $z^* \in D^*$. In this case, obviously,

$$\left|u - z^*\right| \leqslant \sup_{u \in \Omega_{m,N}} |u - z|, \tag{3.145}$$

and the function $x_m(\cdot, u)$ given by the recurrence formula (3.29) can naturally be taken as an approximate T-periodic solution of the T-periodic boundary value problem (3.3), (3.4).

Let us estimate the absolute error of the approximate T-periodic solution, i.e., the deviation of the limit function $x^*(\cdot, z^*)$ from the approximate T-periodic solution $x_m(\cdot, u)$.

THEOREM 3.38. *Let conditions* (3.5), (3.6), (3.61), *and* (3.62) *be satisfied for the periodic boundary value problem* (3.3), (3.4); z^* *be a solution of the exact determining equation* (3.85), *and* u *an arbitrary point from the subset* $\Omega_{m,N}$ *satisfying condition* (3.145). *Then for the deviation of the exact* T*-periodic solution* $x^*(\cdot, z^*)$ *from its approximation* $x_m(\cdot, u)$ *given by* (3.29), *the following inequality is true*:

$$\begin{aligned}\left|x^*(t, z^*) - x_m(t, u)\right| &\leqslant \bar{\alpha}_{1,T}(t) Q^m (I - Q)^{-1} M \\ &\quad + \left(I + \bar{\alpha}_{1,T}(t) K (I - Q)^{-1}\right) \sup_{z \in \Omega_{m,N}} |z - u|,\end{aligned} \tag{3.146}$$

where Q *is given by* (3.45) *and* $\bar{\alpha}_{1,T}$ *is the function* (3.37).

PROOF. Let us use the obvious inequality

$$\begin{aligned}\left|x^*(t, z^*) - x_m(t, u)\right| &\leqslant \left|x^*(t, z^*) - x_m(t, z^*)\right| \\ &\quad + \left|x_m(t, z^*) - x_m(t, u)\right|.\end{aligned} \tag{3.147}$$

We estimate the first term on the right-hand side of (3.147) by using (3.44):

$$\left|x^*(t, z^*) - x_m(t, z^*)\right| \leqslant \bar{\alpha}_{1,T}(t) Q^m (I - Q)^{-1} M, \tag{3.148}$$

whereas the second term is estimated according to (3.83):

$$\begin{aligned}
&\bigl|x_m(t,z^*)-x_m(t,u)\bigr|\\
&\quad\leqslant\bigl[I+\alpha_{1,T}(t)K+\alpha_{2,T}(t)K^2+\cdots+\alpha_{m,T}(t)K^m\bigr]|z^*-u|\\
&\quad\leqslant\left[I+\bar\alpha_{1,T}(t)\sum_{i=0}^{m-1}K\left(\frac{3T}{10}K\right)^i\right]|z^*-u|\\
&\quad\leqslant\bigl[I+\bar\alpha_{1,T}(t)K(I-Q)^{-1}\bigr]|z^*-u|.
\end{aligned}\tag{3.149}$$

Consequently, in view of (3.145), (3.146), (3.149) inequality (3.147) yields

$$\begin{aligned}
\bigl|x^*(t,z^*)-x_m(t,u)\bigr|\leqslant{}&\bar\alpha_{1,T}(t)Q^m(I-Q)^{-1}M\\
&+\bigl[I+\bar\alpha_{1,T}(t)K(I-Q)^{-1}\bigr]|z^*-u|,
\end{aligned}$$

or

$$\begin{aligned}
\bigl|x^*(t,z^*)-x_m(t,u)\bigr|\leqslant{}&\bar\alpha_{1,T}(t)Q^m(I-Q)^{-1}M\\
&+\bigl[I+\bar\alpha_{1,T}(t)K(I-Q)^{-1}\bigr]\sup_{z\in\Omega_{m,N}}|z-u|,
\end{aligned}\tag{3.150}$$

i.e., estimate (3.146) is true. □

4. Successive approximations for autonomous systems

In this section, we present some results concerning the application of the successive approximation techniques to autonomous systems of differential equations. Consider the case where the right-hand side of system (3.1) is independent of t:

$$x'(t)=f\bigl(x(t)\bigr),\quad -\infty<t<\infty.\tag{4.1}$$

Here, $f:D\to\mathbb{R}^n$ is a continuous function and $D\subset\mathbb{R}^n$ is a closed bounded set. We are interested in T-periodic solutions of system (4.1).

Unlike the non-autonomous systems, where the period T of the solution is known, in the autonomous case, besides the initial values of the periodic solution of (4.1), one should also find its period. Moreover, the T-periodic solution $x=x^*(t)$ of autonomous system (4.1) is not isolated in the extended phase space, which means that every member of the one-parameter family of functions

$$[0,T]\ni t\longmapsto x^*(t+\varphi),\quad \varphi\in[0,T],\tag{4.2}$$

is also a T-periodic solution. The latter circumstance, in particular, excludes every possibility of studying this problem via topological degree methods.

In the autonomous case (4.1), the direct application of the successive approximation techniques with the main equations (3.29), (3.41), (3.85), described above for the non-autonomous case, yields

$$\begin{aligned} x_m(t,z) &= z + \int_0^t f\big(x_{m-1}(s,z)\big)\,\mathrm{d}s - \frac{t}{T}\int_0^T f\big(x_{m-1}(s,z)\big)\,\mathrm{d}s, \\ \Delta(z) &= \frac{1}{T}\int_0^T f\big(x_{m-1}(t,z)\big)\,\mathrm{d}t, \end{aligned} \tag{4.3}$$

which implies that

$$x_m(t,z) = x_m^*(t,z) = z, \qquad \Delta(z) \equiv f(z) = 0.$$

Therefore, the successive approximation scheme determined by (4.3) “detects” only the stationary (i.e., constant) periodic solutions of the autonomous system (4.1). In order to find the non-constant T-periodic solutions of (4.1), the following modification can be used [90,91].

The change of variables

$$t = \frac{T}{\tilde{T}}\theta \tag{4.4}$$

allows one to study the periodic solution of (4.1) with a certain fixed period $\tilde{T} \in (0, +\infty)$, i.e., to investigate the following parametrised $\tilde{T}$-periodic boundary value problem

$$x'(\theta) = \mu f\big(x(\theta)\big), \tag{4.5}$$

$$x(0) = x(\tilde{T}), \tag{4.6}$$

where $x : [0, \tilde{T}] \to D$ and $\mu = T\tilde{T}^{-1}$.

Let us suppose that, for some fixed

$$\mu = \mu_0 > 0, \tag{4.7}$$

the $\tilde{T}$-periodic boundary value problem (4.5), (4.6) has a non-constant solution $x = \bar{x}(\theta)$. As in (4.2), the members of the one-parameter family

$$[0, T] \ni \theta \longmapsto \bar{x}(\theta + \varphi), \quad \varphi \in [0, \tilde{T}] \tag{4.8}$$

are also $\tilde{T}$-periodic solutions. Due to the periodicity and continuity of functions (4.8), the components of the solution

$$\bar{x}(\theta) = \mathrm{col}\big(\bar{x}_1(\theta), \bar{x}_2(\theta), \ldots, \bar{x}_n(\theta)\big) \tag{4.9}$$

assume their extremal values over the interval $[0, \tilde{T}]$ at some points θ_{i1}, θ_{i2}:

$$\max_{\theta\in[0,\tilde{T}]} \bar{x}_i(\theta) = \bar{x}_i(\theta_{i1}), \qquad \min_{\theta\in[0,\tilde{T}]} \bar{x}_i(\theta) = \bar{x}_i(\theta_{i2})$$

for $i = 1, 2, \dots, n$.

Since all the periodic solutions (4.8) represent one and the same trajectory, it follows that for all $i = 1, 2, \dots, n$

$$\max_{\theta\in[0,\tilde{T}]} \bar{x}_i(\theta+\varphi) = \bar{x}_i(\theta_{i1}), \qquad \min_{\theta\in[0,\tilde{T}]} \bar{x}_i(\theta+\varphi) = \bar{x}_i(\theta_{i2}).$$

Therefore, without loss of generality, having replaced x^0 by $x^0(\cdot+\varphi)$ with a suitable φ, we can assume in the subsequent consideration that a certain fixed, say jth, component of function (4.8) takes its extremal value over $[0, \tilde{T}]$ at the point $\theta = 0$:

$$\max_{\theta\in[0,\tilde{T}]} \bar{x}_j(\theta) = \bar{x}_j(0) \quad \text{or} \quad \min_{\theta\in[0,\tilde{T}]} \bar{x}_j(\theta) = \bar{x}_j(0).$$

Thus, for the boundary value problem (4.5), (4.6) to be solvable it is necessary that

$$\frac{\mathrm{d}\bar{x}_j}{\mathrm{d}\theta}(0) = 0,$$

i.e., the initial value $x_0 = \bar{x}(0)$ of the solution should lie on the hypersurface

$$\Gamma_j = \left\{x \in \mathbb{R}^n \mid f_j(x) = 0\right\},$$

where $f = \mathrm{col}(f_1, f_2, \dots, f_n)$.

Let us suppose that, in D, the equation

$$f_j(x) = 0 \tag{4.10}$$

has an $(n-1)$-parametric family of solutions

$$x = x_0(c), \tag{4.11}$$

where $c = \mathrm{col}(c_1, c_2, \dots, c_{n-1}) \in \mathbb{R}^{n-1}$ is a vector parameter. Therefore, according to our approach, the $\tilde{T}$-periodic solution of the boundary value problem (4.5), (4.6) will be determined on the set of solutions of the Cauchy problem of (4.5), (4.12):

$$x(0) = x_0(c). \tag{4.12}$$

4.1. *Non-linear substitution of variables*

The next step is to change the variable x according to the formula

$$x(\theta) = \Phi\big(\theta, y(\theta)\big), \quad \theta = \frac{\tilde{T}}{T}t,\ t \in [0, T], \tag{4.13}$$

which allows us to transform the autonomous system (4.5) into a non-autonomous one and then study problem (4.5), (4.6) near the hypothetical solution. The function

$$\Phi : \mathbb{R} \times \mathcal{B} \to D, \tag{4.14}$$

where $\mathcal{B} \subset \mathbb{R}^n$ is a bounded set, is $\tilde{T}$-periodic in the first variable and satisfies all conditions which are usually imposed on the change of variables, namely, $\Phi(\theta, y)$ is continuously differentiable with respect to the variables θ and y and satisfies the condition

$$\det\left[\frac{\partial\Phi(\theta, y)}{\partial y}\right] \neq 0, \quad (\theta, y) \in [0, \tilde{T}] \times \mathcal{B}. \tag{4.15}$$

As a result of the transformation of variables (4.13), Eq. (4.5) takes the form

$$y'(\theta) = F\big(\theta, y(\theta), \mu\big), \tag{4.16}$$

where

$$F(\theta, y, \mu) = \left[\frac{\partial\Phi(\theta, y)}{\partial y}\right]^{-1}\left[\mu f\big(\Phi(\theta, y)\big) - \frac{\partial\Phi(\theta, y)}{\partial\theta}\right]. \tag{4.17}$$

After the coordinate transformation (4.13), at $\theta = 0$, the hypersurface (4.9) turns into the hypersurface

$$\bar{\Gamma}_j = \big\{y_0(c) \mid c \in \mathbb{R}^n\big\},$$

where $y = y_0(c)$ is the solution of the equation

$$x_0(c) = \Phi(0, y). \tag{4.18}$$

This means that the initial value $x = x_0(c)$ that may determine a $\tilde{T}$-periodic solution of (4.12), turns into the initial value

$$y = y_0(c), \tag{4.19}$$

of the $\tilde{T}$-periodic solution of Eq. (4.16).

In order to use the successive approximation techniques with the main equations (3.29), (3.41), (3.85) for system (4.16), it will be assumed that for the values of parameter

$$\mu = \frac{T}{\tilde{T}} \tag{4.20}$$

from some interval

$$J = [\mu_1, \mu_2], \tag{4.21}$$

and for $(\theta, y) \in [0, \tilde{T}] \times \mathcal{B}$, conditions analogous to (3.5)–(3.7), and (3.8) are satisfied for system (4.16). More precisely, we suppose that:

(1) In the domain

$$G = [0, \tilde{T}] \times \mathcal{B} \times J, \tag{4.22}$$

the function F is bounded by the vector $M \in \mathbb{R}^n_+$ in the sense that

$$\frac{1}{2}\left[\max_{(\theta,y,\mu)\in G} F(\theta, y, \mu) - \min_{(\theta,y,\mu)\in G} F(\theta, y, \mu)\right] \leqslant M, \tag{4.23}$$

and, furthermore, satisfies the Lipschitz condition in the second argument with some matrix $K = \{K_{ij}\}_{i,j=1}^n \in \mathrm{GL}_n(\mathbb{R})$:

$$\bigl|F(\theta, u, \mu) - F(\theta, v, \mu)\bigr| \leqslant K|u - v| \tag{4.24}$$

for all $\theta \in [0, \tilde{T}]$, $u, v \in \mathcal{B}$, $\mu \in J$.

(2) The maximal eigenvalue $r(K)$ of the matrix K is such that

$$r(K) < \frac{10}{3\tilde{T}}. \tag{4.25}$$

(3) The set

$$\mathcal{B}_{\frac{\tilde{T}}{2}M} = \left\{ y \in \mathbb{R}^n \;\middle|\; B\left(y, \frac{\tilde{T}}{2}M\right) \subset \mathcal{B} \right\},$$

determined by the non-negative vector M in (4.23), is non-empty:

$$\mathcal{B}_{\frac{\tilde{T}}{2}M} \neq \varnothing. \tag{4.26}$$

REMARK 4.1. Similarly to (3.11), the following estimate holds:

$$\max_{(\theta,y,\mu)\in G} F(\theta, y, \mu) - \min_{(\theta,y,\mu)\in G} F(\theta, y, \mu) \leqslant 2 \max (\theta, y, \mu) \in G\bigl|F(\theta, y, \mu)\bigr|.$$

We note also that if the relations (4.25) and (4.26) are true for some $\mu = \mu_0$, then they also hold for all $\mu > \mu_0$.

In connection with Eq. (4.16), we introduce the sequence of $\tilde{T}$-periodic functions $\{y_m(\theta, c, \mu)\}$ defined by the recurrence relation

$$y_m(\theta, c, \mu) = y_0(c) + \int_0^\theta \Big[F\big(s, y_{m-1}(s, c, \mu), \mu\big) - \frac{1}{\tilde{T}} \int_0^{\tilde{T}} F\big(\tau, y_{m-1}(\tau, c, \mu), \mu\big)\, \mathrm{d}\tau \Big] \mathrm{d}s, \quad m = 1, 2, \ldots. \tag{4.27}$$

These functions depend on the n real parameters

$$c = \operatorname{col}(c_1, c_2, \ldots, c_{n-1}) \in \mathbb{R}^{n-1}, \quad \mu \in J.$$

We define the starting member of sequence (4.27) as

$$y_0(\theta, c, \mu) = y_0(c)$$

with c such that $y_0(c) \in \mathcal{B}_{\frac{\tilde{T}}{2}M}$. The set of all such c from $\mathbb{R}^{n-1}$ will be denoted by H_f. Recall that $y_0(c)$ is given by formulae (4.18) and (4.19)

The following statement establishes the convergence of the sequence (4.27).

THEOREM 4.2. *Let the function f in the autonomous system* (4.5) *and the coordinate transformation* (4.13) *be such that for the non-autonomous $\tilde{T}$-periodic system* (4.16) *with F given by* (4.17)*, conditions* (4.23)–(4.26) *hold. Then, for arbitrary values of the parameters*

$$c \in H_f, \qquad \mu \in J,$$

the following assertions are true:

(i) *The sequence $\{y_m(\theta, c, \mu)\}$ of $\tilde{T}$-periodic functions* (4.27) *is uniformly convergent,*

$$\lim_{m \to \infty} y_m(\theta, c, \mu) = y^*(\theta, c, \mu) \tag{4.28}$$

with respect to

$$(\theta, c, \mu) \in (-\infty, \infty) \times H_f \times J;$$

(ii) *The limit function $y^*(\theta, c, \mu)$ is the unique $\tilde{T}$-periodic solution of the integral equation*

$$y(\theta) = y_0(c) + \int_0^\theta F\big(s, y(s), \mu\big)\, \mathrm{d}s - \frac{\theta}{\tilde{T}} \int_0^{\tilde{T}} F\big(s, y(s), \mu\big)\, \mathrm{d}s, \tag{4.29}$$

with values in the domain $\mathcal{B}$; *alternatively, it is the unique* $\tilde{T}$*-periodic solution of the Cauchy problem for the integro-differential equation*

$$\frac{\partial y(\theta, c, \mu)}{\partial \theta} = F\bigl(\theta, y(s, c, \mu), \mu\bigr) - \Delta(c, \mu), \tag{4.30}$$

$$y(0, c, \mu) = y_0(c), \tag{4.31}$$

where

$$\Delta(c, \mu) = \frac{1}{\tilde{T}} \int_0^{\tilde{T}} F\bigl(\theta, y^*(\theta, c, \mu), \mu\bigr)\, \mathrm{d}\theta; \tag{4.32}$$

(iii) *The estimate*

$$\bigl|y^*(\theta, c, \mu) - y_m(\theta, c, \mu)\bigr| \leqslant \bar{\alpha}_{1,\tilde{T}}(\theta)\left(\frac{3\tilde{T}}{10}K\right)^m \left(I - \frac{3\tilde{T}}{10}K\right)^{-1} M \tag{4.33}$$

holds for an arbitrary $\theta \in (-\infty, \infty)$, *where the function* $\bar{\alpha}_{1,\tilde{T}}$ *is given by formulae* (3.37), (3.31).

PROOF. It is easy to see that Theorem 4.2 is rather similar to Theorem 3.17 and, hence, its proof is omitted. □

Let us establish a general condition necessary and sufficient for the existence of a T-periodic solution of the autonomous system (4.1).

THEOREM 4.3. *Let us assume that, for the autonomous system* (4.5), *for the coordinate transformation* (4.13), *and for the non-autonomous* $\tilde{T}$*-periodic system* (4.16), (4.17), *conditions* (4.23)–(4.26) *hold. Then the solution* $y = y(\theta, 0, y_0(c^*), \mu^*)$ *of the Cauchy problem*

$$y'(\theta) = F\bigl(\theta, y(\theta), \mu^*\bigr), \tag{4.34}$$

$$y(0) = y_0(c^*), \tag{4.35}$$

where $c \in H_f$ *and, hence* $y_0(c^*) \in \mathcal{B}_{\frac{T}{2}M}$, *is a* $\tilde{T}$*-periodic solution of* (4.34) *if and only if* $(c^*, \mu^*) \in H_f \times J$ *is a zero of the determining function* (4.32), *i.e.*

$$\Delta(c^*, \mu^*) := \frac{1}{\tilde{T}} \int_0^{\tilde{T}} F\bigl(\theta, y^*(\theta, c^*, \mu^*), \mu^*\bigr)\, \mathrm{d}\theta = 0, \tag{4.36}$$

and in this case $y = y(\theta, 0, y_0(c^*), \mu^*) = y^*(\theta, c^*, \mu^*)$, *where*

$$y^*(\theta, c^*, \mu^*) = \lim_{m \to \infty} y_m(\theta, c^*, \mu^*). \tag{4.37}$$

According to the change of variables (4.13), *the values* (c^*, μ^*) *and the* $\tilde{T}$*-periodic solution* $y^*(\theta, c^*, \mu^*)$ *of the non-autonomous system* (4.34) *determine the* T*-periodic solution of the autonomous system* (4.1)

$$x = x^*(t) = \Phi\left(\frac{t}{\mu^*}, y^*\left(\frac{t}{\mu^*}, c^*, \mu^*\right)\right) \tag{4.38}$$

and its period

$$T^* = \mu^* \tilde{T}. \tag{4.39}$$

PROOF. The sufficiency of condition (4.36) follows immediately from the fact that the function $y^*(\theta, c, \mu)$ is a $\tilde{T}$-periodic solution of the initial value problem (4.30), (4.31). It is obvious that if (4.36) holds, then this condition is sufficient for (4.37) to be a $\tilde{T}$-periodic solution of (4.16) and of the initial value problem (4.34), (4.35). Our change of variables,

$$\frac{t}{\mu} = \theta, \quad \tilde{T}\mu = T, \quad x = \Phi(\theta, y),$$

yields (4.38), (4.39).

The necessity of condition (4.36) can be established by analogy with the proof of Theorem 3.21. □

Now we establish some properties of the limit function (4.28) and the determining function (4.32).

THEOREM 4.4. *Let us suppose that, in domain* (4.22), *conditions* (4.23), (4.25), (4.26) *hold and, moreover, there exist some non-negative matrix* $K = \{K_{ij}\}_{i,j=1}^n \in \mathrm{GL}_n(\mathbb{R})$ *and vector* $\tilde{M} = \{\tilde{M}_i\}_{i=1}^n \in \mathbb{R}_+^n$ *such that the function* $(\theta, y, \mu) \mapsto F(\theta, y, \mu)$ *satisfies the following Lipschitz condition in* y *and* μ:

$$\left|F(\theta, u, \mu) - F(\theta, v, \xi)\right| \leqslant K|u - v| + |\mu - \xi|\tilde{M}. \tag{4.40}$$

Then the function $y^*(\theta, c, \mu)$ *given by* (4.28), (4.27) *and the determining function* $\Delta(c, \mu)$ *of the form* (4.32) *are Lipschitzian in the sense that, for all*

$$\{c, a\} \subset H_f, \quad \{\mu, \xi\} \subset J, \quad \theta \in [0, \tilde{T}], \tag{4.41}$$

we have

$$\begin{aligned} &\left|y^*(\theta, c, \mu) - y^*(\theta, a, \xi)\right| \\ &\quad \leqslant \left[I + \bar{\alpha}_{1,\tilde{T}}(\theta) K \left(I - \frac{3\tilde{T}}{10} K\right)^{-1}\right] \left|y_0(c) - y_0(a)\right| \\ &\qquad + \left[I + \left(I - \frac{3\tilde{T}}{10} K\right)^{-1}\right] \bar{\alpha}_{1,\tilde{T}}(\theta) |\mu - \xi| \tilde{M}, \end{aligned} \tag{4.42}$$

and the relations

$$\left|\Delta(c,\mu)-\Delta(a,\xi)\right| \leqslant K\left[I+\frac{10\tilde{T}}{27}K\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]\left|y_0(c)-y_0(a)\right| \\ + K\left[I+\frac{10\tilde{T}}{27}K\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]|\mu-\xi|\tilde{M}, \tag{4.43}$$

$$\left|\Delta(c,\mu)-\Delta_m(c,\mu)\right| \leqslant \frac{10\tilde{T}}{27}K\left(\frac{3\tilde{T}}{10}K\right)^m\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}M, \tag{4.44}$$

hold true, where

$$\Delta_m(c,\mu) := \frac{1}{\tilde{T}}\int_0^{\tilde{T}} F\bigl(\theta, y_m(\theta,c,\mu),\mu\bigr)\,\mathrm{d}\theta \tag{4.45}$$

is the m*th approximation to function* (4.32) *and* $y_0(c)$, *and* $y_0(a)$ *are determined by* (4.11), (4.18), *and* (4.19).

PROOF. The uniform convergence of sequence (4.27) in domain (4.41) guarantees the continuity of the function $y^*(\theta,c,\mu)$ with respect to all the variables. Equality (4.32) ensures the continuity of the function $\Delta(c,\mu)$ in domain (4.41) and the fulfilment of the inequality

$$\left|\Delta(c,\mu)\right| \leqslant \max_{(\theta,y,\mu)\in[0,\tilde{T}]\times\mathcal{B}\times J}\left|F(\theta,y,\mu)\right|.$$

It follows directly from (4.27) that

$$\begin{aligned} & y_1(\theta,c,\mu)-y_1(\theta,a,\xi) \\ &\quad = y_0(c)-y_0(a)+\int_0^{\theta}\bigl[F\bigl(s,y_0(c),\mu\bigr)-F\bigl(s,y_0(a),\xi\bigr)\bigr]\,\mathrm{d}s \\ &\qquad - \frac{\theta}{\tilde{T}}\int_0^{\tilde{T}}\bigl[F\bigl(s,y_0(c),\mu\bigr)-F\bigl(s,y_0(a),\xi\bigr)\bigr]\,\mathrm{d}s \\ &\quad = y_0(c)-y_0(a)+\left(1-\frac{\theta}{\tilde{T}}\right)\int_0^{\theta}\bigl[F\bigl(s,y_0(c),\mu\bigr)-F\bigl(s,y_0(a),\xi\bigr)\bigr]\,\mathrm{d}s \\ &\qquad - \frac{\theta}{\tilde{T}}\int_{\theta}^{\tilde{T}}\bigl[F\bigl(s,y_0(c),\mu\bigr)-F\bigl(s,y_0(a),\xi\bigr)\bigr]\,\mathrm{d}s. \end{aligned} \tag{4.46}$$

Due to the Lipschitz condition (4.40) and equality (3.33) of Lemma 3.16, from (4.46) we get

$$\begin{aligned} & \left|y_1(\theta,c,\mu)-y_1(\theta,a,\xi)\right| \\ &\quad \leqslant \left|y_0(c)-y_0(a)\right|+\left(1-\frac{\theta}{\tilde{T}}\right)\int_0^{\theta}\bigl[K\left|y_0(c)-y_0(a)\right|+|\mu-\xi|\tilde{M}\bigr]\,\mathrm{d}s \end{aligned}$$

$$
\begin{aligned}
&+\frac{\theta}{\tilde{T}}\int_{\theta}^{\tilde{T}}\left[K\left|y_0(c)-y_0(a)\right|+|\mu-\xi|\tilde{M}\right]\mathrm{d}s\\
&=\left[I+\alpha_{1,\tilde{T}}(\theta)K\right]\left|y_0(c)-y_0(a)\right|+\alpha_{1,\tilde{T}}(\theta)|\mu-\xi|\tilde{M}.
\end{aligned}
\tag{4.47}
$$

Similarly, taking into account (4.47), we obtain

$$
\begin{aligned}
&\left|y_2(\theta,c,\mu)-y_2(\theta,a,\xi)\right|\\
&\quad\leqslant\left|y_0(c)-y_0(a)\right|\\
&\qquad+\left(1-\frac{\theta}{\tilde{T}}\right)\int_0^{\theta}\left[K\left|y_1(s,c,\mu)-y_1(s,a,\xi)\right|+|\mu-\xi|\tilde{M}\right]\mathrm{d}s\\
&\qquad+\frac{\theta}{\tilde{T}}\int_{\theta}^{\tilde{T}}\left[K\left|y_1(s,c,\mu)-y_1(s,a,\xi)\right|+|\mu-\xi|\tilde{M}\right]\mathrm{d}s\\
&\quad=\left[I+\alpha_{1,\tilde{T}}(\theta)K+\alpha_{2,\tilde{T}}(\theta)K^2\right]\left|y_0(c)-y_0(a)\right|\\
&\qquad+\left[\alpha_{1,\tilde{T}}(\theta)K+\alpha_{2,\tilde{T}}(\theta)K^2\right]|\mu-\xi|\tilde{M}.
\end{aligned}
$$

Proceeding analogously and arguing by induction, we finally conclude that

$$
\begin{aligned}
&\left|y_m(\theta,c,\mu)-y_m(\theta,a,\xi)\right|\\
&\quad\leqslant\left[I+\alpha_{1,\tilde{T}}(\theta)K+\alpha_{2,\tilde{T}}(\theta)K^2+\cdots+\alpha_{m,\tilde{T}}(\theta)K^m\right]\left|y_0(c)-y_0(a)\right|\\
&\qquad+\left[\alpha_{1,\tilde{T}}(\theta)K+\alpha_{2,\tilde{T}}(\theta)K^2+\cdots+\alpha_{m,\tilde{T}}(\theta)K^{m-1}\right]|\mu-\xi|\tilde{M},
\end{aligned}
\tag{4.48}
$$

where $\alpha_{m,\tilde{T}}$ is the function given by (3.33), i.e.,

$$
\alpha_{m+1,\tilde{T}}(\theta)=\left(1-\frac{\theta}{\tilde{T}}\right)\int_0^{\theta}\alpha_{m,\tilde{T}}(s)\,\mathrm{d}s+\frac{\theta}{\tilde{T}}\int_{\theta}^{\tilde{T}}\alpha_{m,\tilde{T}}(s)\,\mathrm{d}s.
\tag{4.49}
$$

In view of inequality (4.25) the maximal eigenvalue of the matrix $\frac{3\tilde{T}}{10}K$ is less than one and, therefore, by letting $m\to\infty$ in (4.48) and using estimates (3.36) of Lemma 3.16 and the inequality

$$
\alpha_{1,\tilde{T}}(\theta)\leqslant\frac{10}{9}\alpha_{1,\tilde{T}}(\theta)=\bar{\alpha}_{1,\tilde{T}}(t),
$$

we get

$$
\begin{aligned}
\left|y^*(\theta,c,\mu)-y^*(\theta,a,\xi)\right|\leqslant&\left[I+\bar{\alpha}_{1,\tilde{T}}(\theta)K\sum_{i=0}^{\infty}\left(\frac{3\tilde{T}}{10}K\right)^i\right]\left|y_0(c)-y_0(a)\right|\\
&+\left[I+\sum_{i=0}^{\infty}\left(\frac{3\tilde{T}}{10}K\right)^i\right]\bar{\alpha}_{1,\tilde{T}}(\theta)|\mu-\xi|\tilde{M}.
\end{aligned}
$$

Thus, we have

$$\begin{aligned}\big|y^*(\theta,c,\mu)-y^*(\theta,a,\xi)\big| \leqslant{}& \left[I+\bar{\alpha}_{1,\tilde{T}}(\theta)K\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]\big|y_0(c)-y_0(a)\big|\\ &+\left[I+\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]\bar{\alpha}_{1,\tilde{T}}(\theta)|\mu-\xi|\tilde{M}.\end{aligned} \tag{4.50}$$

Equality (4.32), together with the Lipschitz condition (4.24), yields

$$\begin{aligned}\big|\Delta(c,\mu)-\Delta(a,\mu)\big| &= \frac{1}{\tilde{T}}\left|\int_0^{\tilde{T}}\big[F\big(\theta,y^*(\theta,c,\mu),\mu\big)-F\big(\theta,y^*(\theta,a,\xi),\xi\big)\big]\,\mathrm{d}\theta\right|\\ &\leqslant \frac{1}{\tilde{T}}K\int_0^{\tilde{T}}\big|y^*(\theta,c,\mu)-y^*((\theta,a,\xi),\xi)\big|\,\mathrm{d}\theta.\end{aligned} \tag{4.51}$$

Substituting (4.42) into (4.51) and taking into account the equality

$$\int_0^{\tilde{T}}\bar{\alpha}_{1,T}(\theta)\,\mathrm{d}\theta=\frac{10}{9}\int_0^{\tilde{T}}2\theta\left(1-\frac{\theta}{\tilde{T}}\right)\mathrm{d}\theta=\frac{10}{27}\tilde{T}^2, \tag{4.52}$$

we get

$$\begin{aligned}&\big|\Delta(c,\mu)-\Delta(a,\mu)\big|\\ &\quad\leqslant\frac{1}{\tilde{T}}K\int_0^{\tilde{T}}\left[I+\bar{\alpha}_{1,\tilde{T}}(\theta)K\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]\mathrm{d}\theta\big|y_0(c)-y_0(a)\big|\\ &\qquad+\frac{1}{\tilde{T}}K\int_0^{\tilde{T}}\left[I+\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]\bar{\alpha}_{1,\tilde{T}}(\theta)|\mu-\xi|\tilde{M}\\ &\quad\leqslant K\left[I+\frac{10\tilde{T}}{27}K\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]\big|y_0(c)-y_0(a)\big|\\ &\qquad+K\left[I+\frac{10\tilde{T}}{27}K\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]|\mu-\xi|\tilde{M}.\end{aligned}$$

In view of estimate (4.33), it follows directly from (4.32), (4.45), and (4.52) that

$$\begin{aligned}\big|\Delta(c,\mu)-\Delta_m(c,\mu)\big| = \bigg|\frac{1}{\tilde{T}}\int_0^{\tilde{T}}&\big[F\big(\theta,y^*(\theta,c,\mu),\mu\big)\\ &-F\big(\theta,y_m(\theta,c,\mu),\mu\big)\big]\,\mathrm{d}\theta\bigg|\end{aligned}$$

$$\leqslant \frac{1}{\tilde{T}} K \int_0^{\tilde{T}} \left| y^*(\theta, c, \mu) - y_m(\theta, c, \mu) \right| \mathrm{d}\theta$$
$$\leqslant \frac{1}{\tilde{T}} K \left(\frac{3\tilde{T}}{10} K \right)^m \left(I - \frac{3\tilde{T}}{10} K \right)^{-1} M \int_0^{\tilde{T}} \overline{\alpha}_{1,\tilde{T}}(\theta) \, \mathrm{d}\theta$$
$$= \frac{10\tilde{T}}{27} K \left(\frac{3\tilde{T}}{10} K \right)^m \left(I - \frac{3\tilde{T}}{10} K \right)^{-1} M.$$

The theorem is proved. □

Theorem 4.3 provides the following numerical-analytic algorithm for finding a T-periodic solution of the autonomous system (4.1) approximately and for the determination of its unknown period T.

(1) We choose a $\tilde{T}$-periodic coordinate transformation (4.13) with properties (4.14) and (4.15) such that the transformed equation satisfies conditions (4.23)–(4.26) in domain (4.22);

(2) We find a $\tilde{T}$-periodic limit function $y^*(\theta, c, \mu)$ of sequence (4.27) starting from $y_0(\theta, c, \mu) = y_0(c) \in \mathcal{B}_{\frac{\tilde{T}}{2} M}$ given by formulae (4.18) and (4.19);

(3) Given the function $y^*(\theta, c, \mu)$, we construct the determining equation

$$\Delta(c, \mu) \equiv \frac{1}{\tilde{T}} \int_0^{\tilde{T}} F\big(\theta, y^*(\theta, c, \mu), \mu\big) \, \mathrm{d}\theta = 0; \tag{4.53}$$

(4) By using certain numerical methods for solving non-linear algebraic or transcendental equations, we find approximately the solution(s) $c = c^*$ of the determining equation (4.53) belonging to the set $D_{\frac{T}{2} M}$ and the value $\mu = \mu^*$ from J;

(5) According to the coordinate transformations (4.4) and (4.13), the function

$$x^*(t) = \Phi\left(\frac{t}{\mu^*}, y^*\left(\frac{t}{\mu^*}, c^*, \mu^* \right) \right) \tag{4.54}$$

gives the T-periodic solution of the autonomous system (4.1) with the period

$$T^* = \mu^* \tilde{T}. \tag{4.55}$$

This solution, by virtue of (4.12), can also be obtained by solving the Cauchy problem with the values of parameters given by (4.53).

The functions

$$x_m(t) = \Phi\left(\frac{t}{\mu^*}, y_m\left(\frac{t}{\mu^*}, c^*, \mu^* \right) \right), \tag{4.56}$$

$$\tilde{x}_m(t) = \Phi\left(\frac{t}{\mu_m}, y_m\left(\frac{t}{\mu_m}, c_m, \mu_m \right) \right) \tag{4.57}$$

are natural to be taken for the mth approximation of the exact T-periodic solution (4.54), where the pair (c_m, μ_m) satisfies the mth approximate determining equation

$$\Delta_m(c,\mu) \equiv \frac{1}{\tilde{T}} \int_0^{\tilde{T}} F\big(\theta, y_m(\theta, c, \mu), \mu\big)\, \mathrm{d}\theta = 0. \tag{4.58}$$

Let us estimate the differences $|x^*(t) - x_m(t)|$ and $|x^*(t) - \tilde{x}_m(t)|$. Assume that, for all $\{\theta_1, \theta_2\} \subset \mathbb{R}$ and $\{y, z\} \subset \mathcal{B}$, function (4.13) satisfies the Lipschitz condition

$$\big|\Phi(\theta_1, y) - \Phi(\theta_2, z)\big| \leqslant |\theta_1 - \theta_2| M_\Phi + K_\Phi |y - z| \tag{4.59}$$

with some non-negative vector $M_\Phi \in \mathbb{R}_+^n$ and matrix $K_\Phi \in \mathrm{GL}_n(\mathbb{R})$. Then, taking into account (4.54), (4.56), and (4.33), we obtain the estimate

$$\begin{aligned} \big|x^*(t) - x_m(t)\big| &= \left|\Phi\left(\frac{t}{\mu^*}, y^*\left(\frac{t}{\mu^*}, c^*, \mu^*\right)\right) - \Phi\left(\frac{t}{\mu^*}, y_m\left(\frac{t}{\mu^*}, c^ast, \mu^*\right)\right)\right| \\ &\leqslant K_\Phi \left| y^*\left(\frac{t}{\mu^*}, c^*, \mu^*\right) - y_m\left(\frac{t}{\mu^*}, c^*, \mu^*\right)\right| \\ &\leqslant \bar{\alpha}_{1,\tilde{T}}\left(\frac{t}{\mu^*}\right) K_\Phi \left(\frac{3\tilde{T}}{10} K\right)^m \left(I - \frac{3\tilde{T}}{10} K\right)^{-1} M. \end{aligned} \tag{4.60}$$

Furthermore, due to (4.54), (4.57), and (4.59),

$$\begin{aligned} \big|x^*(t) - \widetilde{x}_m(t)\big| &= \left|\Phi\left(\frac{t}{\mu^*}, y^*(\frac{t}{\mu^*}, c^*, \mu^*)\right) - \Phi\left(\frac{t}{\mu_m}, y_m\left(\frac{t}{\mu_m}, c_m, \mu_m\right)\right)\right| \\ &\leqslant \left|\frac{1}{\mu^*} - \frac{1}{\mu_m}\right| t M_\Phi + K_\Phi \left| y^*\left(\frac{t}{\mu^*}, c^*, \mu^*\right)\right. \\ &\qquad \left. - y_m\left(\frac{t}{\mu_m}, c_m, \mu_m\right)\right|. \end{aligned} \tag{4.61}$$

Obviously,

$$\begin{aligned} &\left| y^*\left(\frac{t}{\mu^*}, c^*, \mu^*\right) - y_m\left(\frac{t}{\mu_m}, c_m, \mu_m\right)\right| \\ &\quad \leqslant \left| y^*\left(\frac{t}{\mu^*}, c^*, \mu^*\right) - y_m\left(\frac{t}{\mu^*}, c_m, \mu_m\right)\right| \\ &\qquad + \left| y_m\left(\frac{t}{\mu^*}, c_m, \mu_m\right) - y_m\left(\frac{t}{\mu_m}, c_m, \mu_m\right)\right|. \end{aligned} \tag{4.62}$$

The first term in the right-hand side of (4.62) can be estimated by using (4.42):

$$\begin{aligned}&\left|y^*\left(\frac{t}{\mu^*}, c^*, \mu^*\right) - y_m\left(\frac{t}{\mu^*}, c_m, \mu_m\right)\right|\\&\quad\leqslant \left[I + \bar{\alpha}_{1,\tilde{T}}(\theta)K\left(I - \frac{3\tilde{T}}{10}K\right)^{-1}\right]\left|y_0(c^*) - y_0(c_m)\right|\\&\qquad + \left[I + \left(I - \frac{3\tilde{T}}{10}K\right)^{-1}\right]\bar{\alpha}_{1,\tilde{T}}(\theta)|\mu^* - \mu_m|\tilde{M}. \qquad (4.63)\end{aligned}$$

Let us put

$$\frac{1}{\mu^*} = \frac{1}{\mu_m} + \frac{1}{\gamma_m}. \qquad (4.64)$$

By virtue of Lemma 3.13, it then follows from (4.27) that

$$\begin{aligned}&\left|y_m\left(\frac{t}{\mu^*}, c_m, \mu_m\right) - y_m\left(\frac{t}{\mu_m}, c_m, \mu_m\right)\right|\\&\quad= \Bigg|y_0(c_m) + \int_0^{\frac{t}{\mu^*}}\Bigg[F\big(s, y_{m-1}(s, c_m, \mu_m), \mu_m\big)\\&\qquad - \frac{1}{\tilde{T}}\int_0^{\tilde{T}} F\big(\tau, y_{m-1}(\tau, c_m, \mu_m), \mu_m\big)\,\mathrm{d}\tau\Bigg]\mathrm{d}s\\&\qquad - y_0(c_m) - \int_0^{\frac{t}{\mu_m}}\Bigg[F\big(s, y_{m-1}(s, c_m, \mu_m), \mu_m\big)\\&\qquad - \frac{1}{\tilde{T}}\int_0^{\tilde{T}} F\big(\tau, y_{m-1}(\tau, c_m, \mu_m), \mu_m\big)\,\mathrm{d}\tau\Bigg]\mathrm{d}s\Bigg|\\&\quad= \Bigg|\int_{\frac{t}{\mu_m}}^{\frac{t}{\gamma_m}}\Bigg[F\big(s, y_{m-1}(s, c_m, \mu_m), \mu_m\big)\\&\qquad - \frac{1}{\tilde{T}}\int_0^{\tilde{T}} F\big(\tau, y_{m-1}(\tau, c_m, \mu_m), \mu_m\big)\,\mathrm{d}\tau\Bigg]\mathrm{d}s\Bigg|\\&\quad\leqslant \alpha_{1,\tilde{T}}(t)M, \quad t \in [0, \tilde{T}]. \qquad (4.65)\end{aligned}$$

Combining (4.63) and (4.65), we see that (4.62), (4.61) yield the following estimate:

$$\begin{aligned}\left|x^*(t) - \tilde{x}_m(t)\right| \leqslant{}& \left|\frac{1}{\mu^*} - \frac{1}{\mu_m}\right| t M_\Phi\\&+ K_\Phi\Bigg\{\left[I + \bar{\alpha}_{1,\tilde{T}}\left(\frac{t}{\mu_m}\right)K\left(I - \frac{3\tilde{T}}{10}K\right)^{-1}\right]\\&\times \left|y_0(c^*) - y_0(c_m)\right|\end{aligned}$$

$$+\left[I+\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}\right]\bar{\alpha}_{1,\tilde{T}}\left(\frac{t}{\mu_m}\right)|\mu^*-\mu_m|\tilde{M}\Bigg\}$$
$$+\alpha_{1,\tilde{T}}(t)M, \quad t\in[0,T]. \tag{4.66}$$

REMARK 4.5. By using inequality (4.44) of Theorem 4.4, one can obtain a condition sufficient for the absence of $\tilde{T}$-periodic solutions of system (4.16). Indeed, (4.44) implies

$$\left|\Delta(c,\mu)\right|\geqslant\frac{10\tilde{T}}{27}K\left(\frac{3\tilde{T}}{10}K\right)^{m}\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}M-\left|\Delta_m(c,\mu)\right|.$$

Therefore, the determining equation (4.53) has no solutions (c,μ) in the set $H_f\times J$ whenever

$$\sup_{(c,\mu)\in H_f\times J}\left|\frac{1}{\tilde{T}}\int_0^{\tilde{T}}F\big(\theta,y_m(\theta,c_m,\mu_m),\mu_m\big)\,\mathrm{d}\theta\right|$$
$$>\frac{10\tilde{T}}{27}K\left(\frac{3\tilde{T}}{10}K\right)^{m}\left(I-\frac{3\tilde{T}}{10}K\right)^{-1}M \tag{4.67}$$

for some $m\in\{0,1,2,\ldots\}$. Consequently, it follows from Theorem 4.3 that, under condition (4.67), the $\tilde{T}$-periodic system (4.16) has no $\tilde{T}$-periodic solutions and, hence, the autonomous system (4.1) has no T-periodic solutions.

4.2. *Linear substitution of variables*

It is clear that the simplest case of substitution (4.13) is the linear change of variables

$$x(\theta)=A(\theta)y(\theta)+b(\theta), \quad \theta\in\mathbb{R}, \tag{4.68}$$

where A and b are $\tilde{T}$-periodic matrix and vector-valued functions, respectively. We assume that these functions are continuously differentiable and such that

$$\det A(\theta)\neq 0, \quad \theta\in[0,\tilde{T}]. \tag{4.69}$$

The change of variable (4.68) transforms the autonomous system (4.1) into the non-autonomous system

$$\frac{\mathrm{d}y}{\mathrm{d}\theta}=F(\theta,y,\mu), \tag{4.70}$$

where μ is given by (4.20) and

$$F(\theta,y,\mu):=A(\theta)^{-1}\big[\mu f\big(A(\theta)y+b(\theta)\big)-A'(\theta)y-b'(\theta)\big].$$

The use of the linear substitution (4.68) instead of (4.13) simplifies the transformation of the autonomous system (4.1) to a non-autonomous one and the verification of conditions (4.23)–(4.26). In this case, the initial data (4.11) and (4.19), for which there may exist $\tilde{T}$-periodic solutions of (4.5) and (4.70), are related by the formula

$$x_0(c) = A(0)y_0(c) + b(0), \tag{4.71}$$

whence, in view of (4.69),

$$y_0(c) = A(0)^{-1}\big[x_0(c) - b(0)\big]. \tag{4.72}$$

Therefore, we can start our iteration (4.27) from $y_0(c)$ given by (4.72).

The T-periodic solution of the autonomous system (4.1) is given by the formula

$$x^*(t) = A\left(\frac{t}{\mu^*}\right)y^*\left(\frac{t}{\mu^*}, c^*, \mu^*\right) + b\left(\frac{t}{\mu^*}\right), \tag{4.73}$$

where the pair (c^*, μ^*) is a solution of the determining equation (4.36). Similarly, the approximate solutions (4.56), (4.57) take the form

$$x_m(t) = A\left(\frac{t}{\mu^*}\right)y_m\left(\frac{t}{\mu^*}, c^*, \mu^*\right) + b\left(\frac{t}{\mu^*}\right), \tag{4.74}$$

and

$$\tilde{x}_m(t) = A\left(\frac{t}{\mu_m}\right)y_m\left(\frac{t}{\mu_m}, c_m, \mu_m\right) + b\left(\frac{t}{\mu_m}\right), \tag{4.75}$$

where the pair (c_m, μ_m) is a solution of the approximate determining equation (4.58).

For the linear change of variable (4.68), the vector M_Φ and the matrix K_Φ in the Lipschitz condition (4.59) as well as in the error estimates (4.60), (4.66) are given by the formulae $K_\Phi = \max_{\theta\in[0,\tilde{T}]} |A(\theta)|$,

$$M_\Phi = \max_{(\theta,y)\in[0,\tilde{T}]\times\Omega} \big|A'(\theta)y + f'(\theta)\big|.$$

EXAMPLE 4.6. Let us suppose that the system (4.1) has the form

$$\begin{aligned} x_1'(t) &= f_1\big(x_1(t), x_2(t)\big),\\ x_2'(t) &= f_2\big(x_1(t), x_2(t)\big). \end{aligned} \tag{4.76}$$

Consider (4.76), (4.77) together with the T-periodic boundary conditions

$$x_1(0) = x_1(T), \qquad x_2(0) = x_2(T), \tag{4.77}$$

where $f_1 : \mathbb{R}^n \to \mathbb{R}^p$, $f_2 : \mathbb{R}^n \to \mathbb{R}^{n-p}$, p is even, and T is an positive constant.

Let us use the special case of the linear transformation (4.68) where

$$A(\theta) = \operatorname{diag}\big(\exp(A\theta), I_{n-p}\big), \qquad b(\theta) \equiv 0, \tag{4.78}$$

where A is a real skew-symmetric matrix with purely imaginary simple eigenvalues.

In this case, substitution (4.68) is rewritten as follows:

$$x_1(t) = \exp(A\theta)y_1(\theta), \qquad x_2(\theta) = y_2(\theta), \tag{4.79}$$

and the transformed system (4.16) corresponding to (4.76) has the form

$$\begin{aligned} y_1'(\theta) &= \mu \exp[-\theta A] f_1\big(\exp(\theta A)y_1(\theta), y_2(\theta)\big) - Ay_1(\theta),\\ y_2'(\theta) &= \mu f_2\big(\exp(\theta A)y_1(\theta), y_2(\theta)\big), \end{aligned}$$

where μ is given by (4.20).

EXAMPLE 4.7. Consider the autonomous system of the form

$$\begin{aligned} x_1'(t) &= Ax_1(t) + f_1\big(x_1(t), x_2(t)\big),\\ x_2'(t) &= f_2\big(x_1(t), x_2(t)\big), \end{aligned} \tag{4.80}$$

where $f_1 : \mathbb{R}^n \to \mathbb{R}^p$, $f_2 : \mathbb{R}^n \to \mathbb{R}^{n-p}$, p is even, $T > 0$, and A is a square matrix of dimension p having the same properties as in Example 4.6. Let us assume that the functions f_1 and f_2 are continuous in the domain

$$G = \big\{(x_1, x_2) \mid \|x_1\| \leqslant d, x_2 \in D_2\big\}, \tag{4.81}$$

where $\|x_1\|^2 \equiv \sum_{k=1}^{p} x_{1k}^2$ and D_2 is a closed bounded domain.

Suppose that the functions f_1 and f_2 satisfy the following conditions in domain (4.81): there exist some non-negative constants k_1, m_1, a non-negative vector $M_2 \in \mathbb{R}^{n-p}$, and a non-negative $(n-p) \times n$ matrix K_2 such that

$$\big\|f_1(x)\big\| \leqslant m_1, \qquad \big|f_2(x)\big| \leqslant M_2 \tag{4.82}$$

and

$$\big\|f_1(x_1, x_2) - f_1(z_1, z_2)\big\| \leqslant k_1\|x - z\|, \tag{4.83}$$

$$\big|f_2(x_1, x_2) - f_2(z_1, z_2)\big| \leqslant K_2|x - z| \tag{4.84}$$

for all (x_1, x_2) and (z_1, z_2) from domain (4.81).

The change of variable (4.68) with A given by (4.78) transforms system (4.80) to the form

$$\begin{aligned}\frac{dy_1}{d\theta} &= (\mu - 1)Ay_1 + \mu \exp[-\theta A] f_1\big(\exp(\theta A)y_1, y_2\big)\\ &=: \mu F_1(\theta, y_1, y_2), \\ \frac{dy_2}{d\theta} &= \mu f_2\big(\exp(\theta A)y_1, y_2\big) =: \mu F_2(\theta, y_1, y_2).\end{aligned} \tag{4.85}$$

Since A is skew-symmetric, it follows that the matrices $\exp(\pm\theta A)$ are orthogonal, i.e., $e^{\pm\theta A}(e^{\pm\theta A})^* = I_{p\times p}$ and, furthermore, x_1 and x_2 in (4.79) satisfy the relations

$$\|x_1\| = \|e^{\theta A} y_1\| = \|y_1\|, \tag{4.86}$$

$$\big\|e^{-\theta A} f_1(e^{\theta A} y_1, y_2)\big\| = \big\|f_1(e^{\theta A} y_1, y_2)\big\|. \tag{4.87}$$

It follows from (4.86), (4.87) that the right-hand side of system (4.85) is well defined and continuous in domain (4.81). Moreover, the functions $F_1(\theta, y_1, y_2)$ and $F_2(\theta, y_1, y_2)$ in (4.85) satisfy the inequalities

$$\begin{aligned}&\big|F_1(\theta, y_1, y_2)\big| \leqslant \bar{m}_1 e_p, \qquad \big|F_2(\theta, y_1, y_2)\big| \leqslant M_2, \\ &\big|F_1(\theta, y_1, y_2) - F_1(\theta, z_1, z_2)\big| \leqslant \bar{k}_1 J_{p\times n}|y - z|\end{aligned} \tag{4.88}$$

for all (y_1, y_2) and (z_1, z_2) from domain (4.81), where

$$\begin{aligned}\bar{m}_1 &= m_1 + (1 - \mu^{-1}) \max_{\|y_1\|\leqslant d} \|Ay_1\|, \\ \bar{k}_1 &= k_1 + (1 - \mu^{-1})\|A\|,\end{aligned} \tag{4.89}$$

and m_1 and k_1 are the same as in (4.82), (4.83). Here, e_p and $J_{p\times n}$ are the vector and matrix of dimension $p \times 1$ and $p \times n$, respectively, with all elements equal to 1.

It follows from inequalities (4.88), (4.89) that the functions in the right-hand side of system (4.85) satisfy conditions (4.23) and (4.24) with the vector

$$M = \mu \operatorname{col}(\bar{m}_1 e_p, M_2) \tag{4.90}$$

and matrix

$$K = \mu \begin{pmatrix} \bar{k}_1 J_{p\times n} \\ K_2 \end{pmatrix}, \tag{4.91}$$

where K_2 is the matrix from (4.84).

We thus see that, in some cases (in particular, for system (4.80)) one can verify conditions (4.23), (4.24) in terms of the constants m_1, M_2, k_1, and K_2 involved in conditions (4.82), (4.83), and (4.84) for the original system.

REMARK 4.8. Quite similarly, one can show that, for the autonomous system

$$x'(t) = Ax + f(x)$$

with A having the same properties as in Example 4.7, the linear change of variables

$$x(\theta) = \mathrm{e}^{\theta A} y(\theta) \tag{4.92}$$

brings the $T\mu^{-1}$-periodic problem (4.5), (4.6)

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \mu\bigl[Ax + f(x)\bigr], \tag{4.93}$$

$$x(0) = x(\tilde{T}) \tag{4.94}$$

to the non-autonomous $\tilde{T}$-periodic problem (4.94) for the equation

$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = (\mu - 1)Ay + \mu\mathrm{e}^{-\theta A} f(\mathrm{e}^{\theta A} y). \tag{4.95}$$

EXAMPLE 4.9. Let us consider the second-order autonomous differential equation

$$\ddot{x} + \dot{x}^2 + (x - \varepsilon)^2 + x = \varepsilon^2 + \varepsilon, \tag{4.96}$$

where ε is a real parameter, and try do determine the period, T, which a periodic solution of (4.96) may have.

To Eq. (4.96), we associate the first-order system

$$\begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}t} &= x_2, \\ \frac{\mathrm{d}x_2}{\mathrm{d}t} &= (2\varepsilon - 1)x_1 - x_1^2 - x_2^2 + \varepsilon \end{aligned} \tag{4.97}$$

and the following T-periodic boundary conditions:

$$x_1(0) = x_1(T), \qquad x_2(0) = x_2(T). \tag{4.98}$$

System (4.97) will be considered for $x_1^2 + x_2^2 \leqslant \varrho^2$, where $\varrho \geqslant 2|\varepsilon|$.

It is clear that (4.97), (4.98) is equivalent to the problem on T-periodic C^2 solutions of (4.96).

Applying the coordinate transformation (4.4) with $\tilde{T} = 2\pi$, we transform (4.97), (4.98) into the system

$$\begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}\theta} &= \mu x_2, \\ \frac{\mathrm{d}x_2}{\mathrm{d}\theta} &= \mu\bigl((2\varepsilon - 1)x_1 - x_1^2 - x_2^2 + \varepsilon\bigr) \end{aligned} \tag{4.99}$$

with the 2π-periodic conditions

$$x_1(0)=x_1(2\pi), \qquad x_2(0)=x_2(2\pi). \tag{4.100}$$

In (4.99), according to (4.20), $\mu = T(2\pi)^{-1}$.

Rewriting (4.99), (4.100) as (4.93) by putting

$$A=\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad f(x)=\begin{pmatrix} 0 \\ -x_1^2-x_2^2+2\varepsilon x_1+\varepsilon \end{pmatrix} \tag{4.101}$$

and using transformation (4.92), we arrive at system (4.95) of the form

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}\theta} &= \mu\Biggl[\left(1-\frac{1}{\mu}\right)\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} y \\ &\quad + \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} f\left(\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} y\right)\Biggr]. \end{aligned} \tag{4.102}$$

Here, we have used the equality

$$\exp(-A\theta)=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{4.103}$$

easily derived from the series expansion of $\exp(-A\theta)$ in a neighbourhood of 0.

It follows from (4.103) that, in our case, the variable transformation (4.92) is given by the equalities

$$x_1=\cos\theta y_1+\sin\theta y_2, \qquad x_2=-\sin\theta y_1+\cos\theta y_2 \tag{4.104}$$

and, hence, system (4.102) has the form

$$\begin{aligned} \frac{\mathrm{d}y_1}{\mathrm{d}\theta} &= (\mu-1)y_2+\mu\sin\theta\bigl[y_1^2+y_2^2-2\varepsilon(\cos\theta y_1+\sin\theta y_2)-\varepsilon\bigr], \\ \frac{\mathrm{d}y_2}{\mathrm{d}\theta} &= (1-\mu)y_1-\mu\cos\theta\bigl[y_1^2+y_2^2-2\varepsilon(\cos\theta y_1+\sin\theta y_2)-\varepsilon\bigr]. \end{aligned} \tag{4.105}$$

It is obvious from (4.104) that system (4.105) should be considered in the same ϱ-neighbourhood of zero as system (4.97).

We need to verify whether conditions (4.23)–(4.25) hold in our case.

As it has been already pointed out above, the vector M in (4.90) and matrix K in (4.91) are determined directly from the terms in the right-hand side of the original system (4.97), i.e., on base of the properties of function f given by (4.101).

It is easy to see that the function

$$W(x_1,x_2)=-x_1^2-x_2^2+2\varepsilon x_1+\varepsilon$$

takes its maximal value $W_{\max} = -|\varepsilon|(1 + |\varepsilon|)$ when $x_1 = \varepsilon$, $x_2 = 0$, whereas its minimal value $W_{\min} = -[\varrho^2 + 2\varepsilon^2 + (2\varrho + 1)|\varepsilon|]$ is achieved for $x_1 = \varrho \operatorname{sign} \varepsilon$, $x_2 = 0$. Therefore, according to (4.90), we may take the following vector M for system (4.105):

$$M = \mu_2\big[\varrho^2 + 2\varepsilon^2 + (2\varrho + 1)|\varepsilon| + |1 - \mu_2^{-1}|\varrho\big]e_2,$$

where $e_2 = \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$.

Furthermore, as can readily be verified, for the number k_1 appearing in (4.83) and (4.89), we may take

$$k_1 = \max_{(x_1,x_2)\in B(0,\varrho)} \sqrt{r\big(\Gamma(x_1, x_2)\big)}, \tag{4.106}$$

where

$$\Gamma(x_1, x_2) = 4\begin{bmatrix} 0 & 0 \\ 0 & (x_1 - \varepsilon)^2 + x_2^2 \end{bmatrix}. \tag{4.107}$$

Indeed, it is easy to see that

$$\Gamma(x) = \frac{\partial f(x)}{\partial x}\left[\frac{\partial f(x)}{\partial x}\right]^*.$$

The maximal eigenvalue of matrix (4.107) is, obviously, given by the formula

$$r\big(\Gamma(x_1, x_2)\big) = 4\big[(x_1 - \varepsilon)^2 + x_2^2\big],$$

whence it follows that, according to (4.106), we may put

$$k_1 = 2\big(\varrho + |\varepsilon|\big).$$

Hence, since $\|\left[\begin{smallmatrix}0 & 1\\-1 & 0\end{smallmatrix}\right]\| = 1$, we see that the constant $\bar{k}_1$ in (4.89) is given by the formula

$$\bar{k}_1 = 2\big(\varrho + |\varepsilon|\big) + \left|1 - \frac{1}{\mu_2}\right|.$$

Therefore, we have the following matrix K in (4.91):

$$K = \mu_2\left[2\big(\varrho + |\varepsilon|\big) + \left|1 - \frac{1}{\mu_2}\right|\right]J_{2\times2},$$

where $J_{2\times2} = \left[\begin{smallmatrix}1 & 1\\1 & 1\end{smallmatrix}\right]$.

Since $r(J_{2\times2}) = 2$, it is not difficult to verify that conditions (4.25) and (4.26) are satisfied in $B(0, \varrho)$ whenever

$$2(\varrho + |\varepsilon|) + \left|1 - \frac{1}{\mu_2}\right| < \frac{5}{6\pi\mu_2} \tag{4.108}$$

and

$$\varrho^2 + 2\varepsilon^2 + (2\varrho + 1)|\varepsilon| + \left|1 - \frac{1}{\mu_2}\right|\varrho < \frac{\varrho}{\pi\mu_2}. \tag{4.109}$$

Inequalities (4.108) and (4.109) hold, for example, when

$$\varepsilon = 0.02, \quad \varrho = 0.06, \quad 0.95 < \mu < 1.05. \tag{4.110}$$

Thus, the techniques described above are applicable to our system (4.97). For system (4.97), Eq. (4.10) with $j = 1$ has the form

$$x_2 = 0$$

and, hence, the one-parametric family of solutions

$$x_1 = c, \qquad x_2 = 0,$$

where $c \in \mathbb{R}$. According to the coordinate transformation (4.92), we have

$$x_1(0) = y_1(0), \qquad x_2(0) = y_2(0).$$

Therefore, we may start the successive approximations (4.27) from the constant function

$$y_0(\theta, c, \mu) = \begin{pmatrix} c \\ 0 \end{pmatrix}.$$

Consider the approximate determining equation (4.58) with $m = 0$ and $\tilde{T} = 2\pi$:

$$\mu \int_0^{2\pi} \sin\theta[c^2 - 2\varepsilon c\cos\theta - \varepsilon]\,\mathrm{d}\theta = 0,$$
$$\int_0^{2\pi} \big[(1-\mu)c - \mu\cos\theta[c^2 - 2\varepsilon c\cos\theta - \varepsilon]\big]\,\mathrm{d}\theta = 0.$$

The first of these two equations is, in fact, an identity for an arbitrary c, whereas the second one means that, for all real c,

$$2\pi(1-\mu)c + 2\varepsilon\mu c\int_0^{2\pi}\cos^2\theta\,\mathrm{d}\theta = 0$$

whence it follows that

$$\mu = \frac{1}{1+\varepsilon},$$

and therefore, in the zero approximation, the period, T, of a solution of Eq. (4.96) is given by the formula

$$T = \frac{2\pi}{1-\varepsilon}.$$

For instance, in the case where ε, ϱ, and μ are given by (4.110), we have $T = 2.04\pi$.

Note that the autonomous equation (4.96) has the periodic solution

$$x(t) = \varepsilon(1+\sin t), \tag{4.111}$$

whose period is equal to 2π. Thus, for ε sufficiently small, even the zero approximation provides a reasonable degree of accuracy for the unknown value of the period.

5. Periodic solutions of differential systems with symmetries

In this section, we study a class of systems of non-linear non-autonomous ordinary differential equations that have "sufficiently many" periodic solutions possessing certain symmetry properties. The symmetry properties considered here (see Definition 5.1) are close to those studied by Hale [28] but, generally speaking, do not coincide with them; see Remark 5.43 in this relation. The theorem obtained in Section 5.7 contains, in particular, Theorem 6 from [70, p. 326] (see also [89]). Similar topics related to symmetries in differential equations were addressed to in the works of Fečkan [18,21–23], Mawhin and Walter [49] Vanderbauwhede [105].

We consider the problem on the T-periodic solutions of the system of n ordinary differential equations

$$x'(t) = f\big(t, x(t)\big), \quad -\infty < t < \infty, \tag{5.1}$$

where $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is a certain function periodic in the first variable with the period T.

By a solution of the problem on the T-periodic solutions of system (5.1), we mean a continuously differentiable element of the Banach space C_T^n of continuous vector-functions $(-\infty,\infty) \to \mathbb{R}^n$ periodic with period T. The set of all continuous n-dimensional vector-functions $(-\infty,\infty)$ is denoted by the symbol C^n.

5.1. *Symmetry properties of periodic functions*

We are interested in those properties of the equations of the differential system (5.1) under which the T-periodic solutions of the system are "symmetric" in a certain sense. We restrict ourselves to the consideration of the following symmetry property.

DEFINITION 5.1. (See [63, Definition 3.1].) We say that a function $x \in C_T^n$ *possesses property* (τ, E) (or, alternatively, *is* (τ, E)*-proper*) for certain number τ and non-singular matrix $E \in \mathrm{GL}_n(\mathbb{R})$ if, for all real t, the equality

$$x(t) = Ex(-t-\tau) \tag{5.2}$$

is true.

The set of all functions from C_T^n possessing property (τ, E) will be denoted by $C_T^n(\tau, E)$:

$$C_T^n(\tau, E) := \left\{x \in C_T^n \mid \text{(5.2) holds for all } t \in (-\infty, \infty)\right\}. \tag{5.3}$$

The following lemma is obvious.

LEMMA 5.2. *For arbitrary point* $\tau \in (-\infty, \infty)$ *and non-singular* n*-dimensional matrix* E*, the set* $C_T^n(\tau, E)$ *forms a closed linear subspace in* C_T^n*.*

REMARK 5.3. It is clear from (5.2) that, for $x \in C_T^n(\tau, E)$, the vectors $x(kT)$, where $k = 0, \pm 1, \pm 2, \dots$, are fixed points of the linear operator determined by the matrix E.

The property dealt with in Definition 5.1 above can be regarded as a certain geometrical symmetry of the graph of the function under consideration. In particular, for $E \in \{-\mathbb{1}_n, \mathbb{1}_n\}$, where $\mathbb{1}_n$ stands for the unit matrix of dimension n, Definition 5.1 describes the very natural properties of τ-evenness and τ-oddness understood in the following sense.

DEFINITION 5.4. A function $x \in C_T^n$ will be called τ*-odd* (resp., τ*-even*) if, for an arbitrary t from $(-\infty, \infty)$, equality (5.2) holds with $E = -\mathbb{1}_n$ (resp., with $E = \mathbb{1}_n$).

Clearly, the usual notions of evenness and oddness of a vector-valued function of a scalar variable are obtained immediately from Definition 5.4 when $\tau = 0$.

REMARK 5.5. Properties similar to the τ-evenness and τ-oddness conditions arise in the theory of Fourier series [103]. Similar notions are used, e.g., in [62].

It is easy to construct examples of τ-even and τ-odd functions by using the following

PROPOSITION 5.6. *If a function* $x \in C_T^n$ *is even* (*resp., odd*), *then, for an arbitrary real* τ, *the corresponding function*

$$(-\infty, \infty) \ni t \longmapsto x\left(t + \frac{\tau}{2}\right) \tag{5.4}$$

is τ*-even* (*resp.,* τ*-odd*).

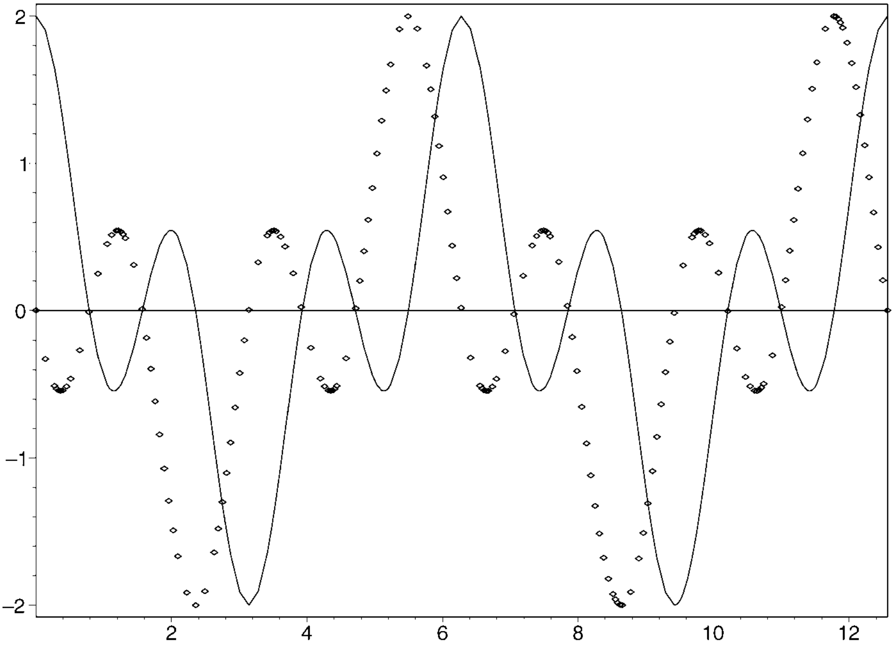


Fig. 1. Functions (5.5) and (5.6).

Proposition 5.6 is easily proved by using Corollary 5.11 of Proposition 5.10 below.

EXAMPLE 5.7. The function

$$u(t) := \cos\left(t + \frac{\pi}{4}\right) - \sin\left(3t + \frac{\pi}{4}\right), \quad t \in (-\infty, \infty), \tag{5.5}$$

is $\frac{\pi}{2}$-even.

Indeed, for the odd function x defined by the formula

$$x(t) := \cos t + \cos 3t, \quad t \in (-\infty, \infty), \tag{5.6}$$

the corresponding function (5.4) is given by equality (5.5). It remains to apply Proposition 5.6 for $T = 2\pi$ and $\tau = \frac{\pi}{2}$.

On Fig. 1, we draw with the “diamond” symbols the graph of the $\frac{\pi}{2}$-even function (5.5) obtained from the even function (5.6) through the shift of the argument by $\frac{\pi}{4}$.

EXAMPLE 5.8. The function

$$u(t) := \sin\left(t + \frac{\pi}{4}\right) - \cos\left(3t + \frac{\pi}{4}\right) - \sin\left(5t + \frac{\pi}{4}\right), \quad t \in (-\infty, \infty), \tag{5.7}$$

is $\frac{\pi}{2}$-odd.

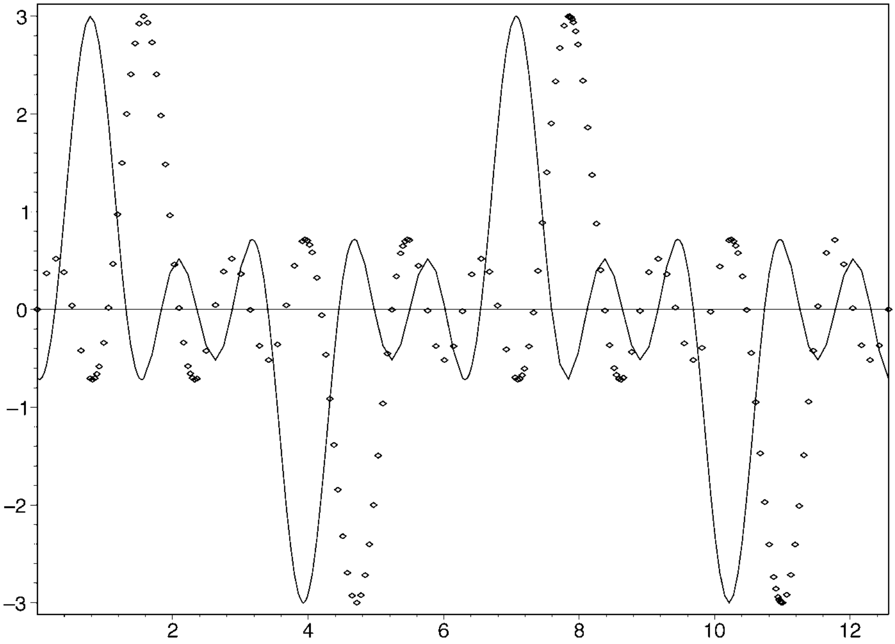


Fig. 2. Functions (5.7) and (5.8).

Indeed, for the even function x given by the formula

$$x(t) := \sin t - \sin 3t + \sin 5t, \quad t \in (-\infty, \infty), \tag{5.8}$$

the corresponding function (5.4) is defined by equality (5.7). Therefore, Proposition 5.6 with $T = 2\pi$ and $\tau = \frac{\pi}{2}$ implies the $\frac{\pi}{2}$-oddness of function (5.7).

On Fig. 2, the “diamond” symbols represent the graph of the $\frac{\pi}{2}$-odd function (5.7) obtained from the even function (5.8) through the shift of the argument by $\frac{\pi}{4}$.

REMARK 5.9. Proposition 5.6 is, in a sense, reversible (see Corollary 5.11 below).

It turns out that the functions with property (τ, E) can be relatively easily constructed from the functions possessing property $(0, E)$, and vice versa. This fact is a consequence of the following

PROPOSITION 5.10. *For arbitrary real σ and τ, the inclusion*

$$S_{-\frac{\tau}{2}} C^n_T(\sigma, E) \subset C^n_T(\sigma + \tau, E), \tag{5.9}$$

holds, where S_h stands for the operator of shift by $h \in [-T, T]$ acting in C^n_T, i.e.,

$$(S_h x)(t) := x(t - h), \quad t \in (-\infty, \infty). \tag{5.10}$$

PROOF. Assume that a certain function x belongs to $C_T^n(\sigma, E)$ and, hence, according to Definition 5.1, satisfies the condition

$$x(t) = Ex(-t-\sigma), \quad t \in (-\infty, \infty). \tag{5.11}$$

Let us set

$$x_\tau := S_{-\frac{\tau}{2}} x. \tag{5.12}$$

By virtue of the definition (5.10) of the operator $S_{-\frac{\tau}{2}}$, the function x_τ thus constructed satisfies the relations

$$\begin{aligned} x_\tau(-t-\sigma-\tau) &= x\left(-t-\sigma-\tau+\frac{\tau}{2}\right) = x\left(-t-\sigma-\frac{\tau}{2}\right) \\ &= x\left(-\left[t+\frac{\tau}{2}\right]-\sigma\right), \end{aligned}$$

whence, by (5.11), we have the relation

$$Ex_\tau(-t-\tau-\sigma) = x\left(t+\frac{\tau}{2}\right), \quad t \in (-\infty, \infty).$$

According to (5.12), the last equality can be rewritten as

$$Ex_\tau(-t-\tau-\sigma) = x_\tau(t), \quad t \in (-\infty, \infty),$$

whence, recalling Definition 5.4, we conclude that x_τ belongs to $C_T^n(\tau+\sigma, E)$. Since x is chosen arbitrarily from $C_T^n(\sigma, E)$, this proves the validity of the required inclusion (5.9). □

COROLLARY 5.11. *If a function* $x \in C_T^n$ *has property* $(0, E)$, *then, for an arbitrary real* τ, *the corresponding function*

$$(-\infty, \infty) \ni t \longmapsto x\left(t+\frac{\tau}{2}\right) \tag{5.13}$$

possesses property (τ, E). *Conversely, if a certain function* $x \in C_T^n$ *has property* (τ, E), *then the corresponding function*

$$(-\infty, \infty) \ni t \longmapsto x\left(t-\frac{\tau}{2}\right)$$

has property $(0, E)$.

PROOF. According to Proposition 5.10, inclusion (5.9) is true, which, for $\sigma = 0$, has the form

$$S_{-\frac{\tau}{2}} C_T^n(0, E) \subset C_T^n(\tau, E).$$

This means that, for an arbitrary function x from $C_T^n(0, E)$, the corresponding function (5.13) has property (τ, E), i.e., the first assertion is true.

On the other hand, for $\sigma = -\tau$, inclusion (5.9) takes the form

$$S_{-\frac{\tau}{2}} C_T^n(-\tau, E) \subset C_T^n(0, E),$$

whence it follows that, for $x \in C_T^n(\tau, E)$, function (5.13) belongs to $C_T^n(0, E)$, Replacing here τ by $-\tau$, we arrive at the second assertion of the corollary. □

The assertions of Corollary 5.11 are illustrated by the diagram

$$\begin{tikzcd} C_T^n(\tau, E) \arrow[r, "S_{\frac{\tau}{2}}"] & C_T^n(0, E) \arrow[r, "S_{-\frac{\tau}{2}}"] & C_T^n(\tau, E). \end{tikzcd}$$

A similar idea can sometimes be applied for the construction of (τ, E)-proper functions from even functions. Namely, the following statement holds.

PROPOSITION 5.12. *If $x \in C_T^n$ is an even function and E is a non-singular matrix of dimension n, then, for an arbitrary n-dimensional square matrix G whose columns belong to the kernel of the matrix $\mathbb{1}_n - E$, the function*

$$(-\infty, \infty) \ni t \longmapsto Gx\left(t + \frac{\tau}{2}\right), \tag{5.14}$$

has property (τ, E).

PROOF. For the function

$$u(t) := Gx\left(t + \frac{\tau}{2}\right)$$

given by formula (5.14), we have

$$Eu(-t - \tau) = EGx\left(-t + \frac{\tau}{2} - \tau\right) = EGx\left(-t - \frac{\tau}{2}\right). \tag{5.15}$$

Since, by assumption, the matrix G is such that

$$EG = G,$$

we see that, for the even functions x, relation (5.15) for an arbitrary $t \in (-\infty, \infty)$ implies the equality

$$Eu(-t - \tau) = u(t),$$

which means that the validity of the assertion required. □

REMARK 5.13. In contrast to Proposition 5.10, the recipe described by Proposition 5.12 is only efficient when $\det(\mathbb{1}_n - E) = 0$ because in the contrary case formula (5.14) defines the function equal identically to zero.

Some further information on the structure of the functions with property (τ, E) is contained in the following

PROPOSITION 5.14. *For an arbitrary function $x \in C_T^n$ having property (τ, E), its value $x(t)$ at every point $t \in (-\infty, \infty)$ belongs to the kernel of the matrix $\mathbb{1}_n - E^2$.*

PROOF. Indeed, let x belong to $C_T^n(\tau, E)$. According to Definition 5.1, this means that, for all real t, equality (5.2) is true, whence

$$x(-t-\tau) = Ex\bigl(-(-t-\tau)-\tau\bigr) = Ex(t), \quad t \in (-\infty, \infty). \tag{5.16}$$

Substituting (5.16) into (5.2), we get the equality

$$x(t) = E^2 x(t), \quad t \in (-\infty, \infty), \tag{5.17}$$

i.e., $x(t) \in \ker(\mathbb{1}_n - E^2)$ for all $t \in (-\infty, \infty)$, as required. □

REMARK 5.15. The peculiar property (5.17) of the functions from $C_T^n(\tau, E)$ that is dealt with in Proposition 5.14, is, obviously, absent when E satisfies the equality

$$E^2 = \mathbb{1}_n. \tag{5.18}$$

Proposition 5.14 yields the following characterisation of the property (τ, E) for the scalar functions.

COROLLARY 5.16. *For an arbitrary point $\tau \in (-\infty, \infty)$, the set $C_T^1(\tau, E)$ consists of the zero function when $E \neq 1$ and coincides with the set of all τ-even functions when $E = 1$.*

PROOF. It follows from Proposition 5.14 that, for an arbitrary (τ, E)-proper scalar function x, identity (5.17) is true. Since, in our case, $E \in (-\infty, \infty)$, it follows that either $E \neq 1$ and x is equal identically to 0, or $E = 1$ and x is τ-even in the sense of Definition 5.4. □

Corollary 5.16 implies that property (τ, E) with E different of $\mathbb{1}_n$ has essentially many-dimensional nature and may be of interest in the cases where $n \geqslant 2$ only.

EXAMPLE 5.17. Assume that τ is a fixed real number and $\theta : (-\infty, \infty) \to (-\infty, \infty)$ is an arbitrary function such that, for a certain $\sigma \in \{-1, 1\}$,

$$\theta(t) = \sigma\theta(-t-\tau), \quad t \in (-\infty, \infty), \tag{5.19}$$

i.e., θ is either τ-even ($\sigma = 1$) or τ-odd ($\sigma = -1$). Let us fix some real α, β, and γ such that $\alpha\beta \neq 0$, and construct the matrix

$$E = \begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix}. \tag{5.20}$$

This matrix is, obviously, non-singular.

The following assertions hold.

(i) For an arbitrary real μ, the function

$$u(t) = \mu\theta(t) \begin{pmatrix} \gamma(\sigma - \alpha)^{-1} \\ 1 \end{pmatrix}, \quad t \in (-\infty, \infty), \tag{5.21}$$

belongs to the class $C_T^2(\tau, E)$ with the matrix E given by

$$E = \begin{pmatrix} \alpha & \gamma \\ 0 & \sigma \end{pmatrix} \tag{5.22}$$

for $\alpha \neq \sigma$.

(ii) For all real λ and μ, the function

$$u(t) = \theta(t) \begin{pmatrix} \lambda \\ \mu \end{pmatrix}, \quad t \in (-\infty, \infty), \tag{5.23}$$

belongs to the class $C_T^2(\tau, E)$ if the matrix E is given by the formula

$$E = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}.$$

(iii) For arbitrary real λ, the function

$$u(t) = \theta(t) \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \quad t \in (-\infty, \infty),$$

belongs to the class $C_T^2(\tau, E)$ if matrix (5.20) has the form

$$E = \begin{pmatrix} \sigma & \gamma \\ 0 & \beta \end{pmatrix}$$

with $\beta \neq \sigma$.

It follows immediately from Proposition 5.14 that the functions u of the class $C_T^2(\tau, E)$ take the values in a subspace of a lower dimension. This dimension is either 0 or 1 in our case and, hence, every such function admits representation in form (5.23), where

$\theta:(-\infty,\infty)\to(-\infty,\infty)$ is continuous and T-periodic, and λ and μ are certain constants.

According to Definition 5.1, function (5.23) has property (τ, E) if, and only if, for all $t\in(-\infty,\infty)$,

$$\theta(t)\begin{pmatrix}\lambda\\ \mu\end{pmatrix}=\theta(-t-\tau)\begin{pmatrix}\alpha\lambda+\gamma\mu\\ \beta\mu\end{pmatrix}.$$

Therefore, analysing the system of linear algebraic equations

$$(\sigma-\alpha)\lambda=\gamma\mu,\qquad (\sigma-\beta)\mu=0$$

and taking into account our assumption (5.19) on the function θ, we arrive at the assertions desired.

EXAMPLE 5.18. The function

$$u(t)=\begin{bmatrix}-\cos(t+\frac{\pi}{4})+\sin(3t+\frac{\pi}{4})\\ \frac{1}{3}\cos(t+\frac{\pi}{4})-\frac{1}{3}\sin(3t+\frac{\pi}{4})\end{bmatrix},\quad t\in(-\infty,\infty),\tag{5.24}$$

has property $(\frac{\pi}{2}, E)$ for

$$E=\begin{bmatrix}2&3\\0&1\end{bmatrix}.\tag{5.25}$$

Indeed, it is easy to see that function (5.21) from Example 5.17 has form (5.24) if, for θ, we take the $\frac{\pi}{2}$-even function (5.5) from Example 5.7:

$$\theta(t)=\cos\left(t+\frac{\pi}{4}\right)-\sin\left(3t+\frac{\pi}{4}\right),\quad t\in(-\infty,\infty).\tag{5.26}$$

Matrix (5.25) coincides with (5.22) for $\sigma=1$, $\alpha=2$ and $\gamma=3$.

On Fig. 3, we present the graphs of the first (solid line) and second (“diamonds”) components of function (5.24). By the plus symbol, we draw the graph of the $\frac{\pi}{2}$-even function θ given by formula (5.26).

5.2. *Assumptions*

The implementation of the successive approximation method described below requires some technical assumptions on the function f in the right-hand side of Eq. (5.1). Namely, we assume the following conditions.

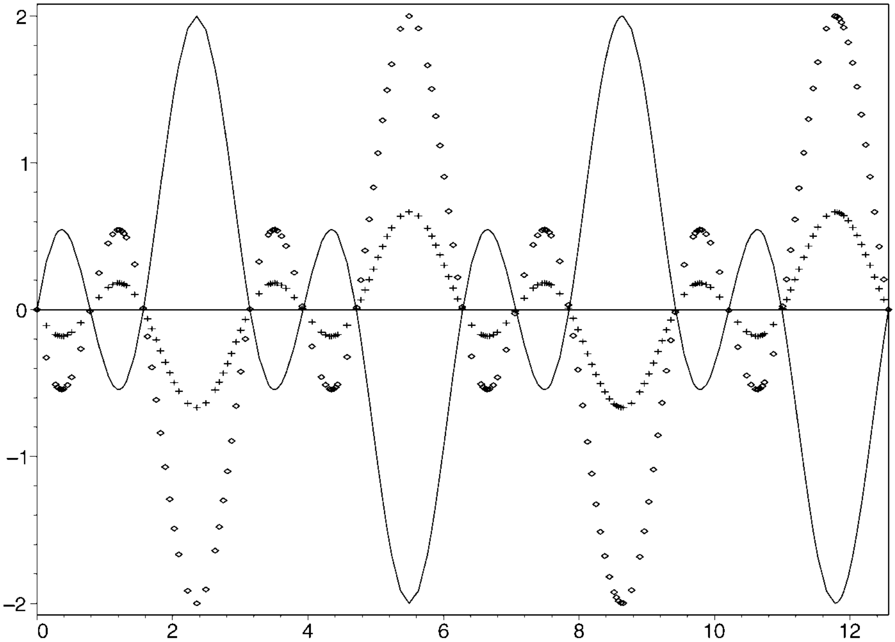


Fig. 3. The components of function (5.24), and function (5.26).

ASSUMPTION 5.19.

(1) The function $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is periodic with period T with respect to the first variable and is continuous on the set $\mathbb{R} \times D$, where D is the closure of a bounded domain in $\mathbb{R}^n$.

(2) There exists a matrix-valued function $K : \mathbb{R} \to \mathrm{GL}_n(\mathbb{R})$ with non-negative Riemann integrable elements such that

$$\left| f(t, x_1) - f(t, x_2) \right| \leqslant K(t)|x_1 - x_2| \tag{5.27}$$

for all $t \in (-\infty, \infty)$ and $\{x_1, x_2\} \subset D$.

(3) There exists an non-negative vector $M \in \mathbb{R}^n_+$ such that

$$\frac{1}{2}\left[\max_{(t,x)\in(-\infty,\infty)\times D} f(t,x) - \min_{(t,x)\in(-\infty,\infty)\times D} f(t,x)\right] \leqslant M.$$

(4) The set $D_{\frac{T}{2}M}$ that consists of all the vectors x from $\mathbb{R}^n$ lying in D together with their closed $\frac{T}{2}M$-neighbourhood is non-empty.

REMARK 5.20. Here and below, the inequality sign, the notion of a neighbourhood, and the symbols “max” and “min” are understood componentwise.

5.3. *Auxiliary statements. The operator H_σ*

Given some $\sigma \in (-\infty, \infty)$, for an arbitrary function u from C_T^n, we set

$$(H_\sigma u)(t) = \left(1 + k_\sigma(t) - \frac{t-\sigma}{T}\right) \int_\sigma^{t - Tk_\sigma(t)} u(s)\,\mathrm{d}s \\ + \left(\frac{t-\sigma}{T} - k_\sigma(t)\right) \int_{t - Tk_\sigma(t)}^{T+\sigma} u(s)\,\mathrm{d}s \tag{5.28}$$

at every point $t \in (-\infty, \infty)$. Here, by definition,

$$k_\sigma(t) := \left\lfloor (t-\sigma)T^{-1} \right\rfloor, \tag{5.29}$$

the integer part of the number $(t-\sigma)T^{-1}$.

LEMMA 5.21. *For any $u \in C_T^n$ and $t \in (-\infty, \infty)$, the equality $(H_\sigma u)(t+T) = (H_\sigma u)(t)$ holds.*

This statement is obtained by direct computation. In view of Lemma 5.21, the mapping H_σ given by (5.28) can be regarded as an operator from C_T^n to C_T^n.

Let us set

$$R_{K,T} := r(H_\sigma K), \tag{5.30}$$

where K in the right-hand side of the equality stands for the operator of multiplication by the continuous matrix-valued function $K : (-\infty, \infty) \to \mathrm{GL}_n(\mathbb{R})$ denoted by the same letter,

$$C_T^n \ni u(\cdot) \longmapsto K(\cdot)u(\cdot). \tag{5.31}$$

The symbol $r(A)$, as usual, denotes the spectral radius of a bounded linear operator A. It is clear from Lemma 5.21 that the composition mapping $H_\sigma K$ can be regarded as a linear operator transforming the space C_T^n into itself.

LEMMA 5.22. *For an arbitrary integrable matrix-valued function $K : \mathbb{R} \to \mathrm{GL}_n(\mathbb{R})$, the number $R_{K,T}$ is independent of the value of σ in* (5.30). *Moreover,*

$$R_{K,T} = 1 \Big/ \min\left\{ |\mu| \;\Big|\; \det\left(\frac{T}{\mu}\mathbb{1}_n - \Omega_{\mu K,0}(T) \int_0^T s\,\Omega_{\mu K,0}^{-1}(s)\,\mathrm{d}s \right) = 0 \right\}. \tag{5.32}$$

Here and in the proof below, for $P:[\sigma, T+\sigma]\to \mathrm{GL}_n(\mathbb{R})$, the symbol $\Omega_{P,\sigma}$ stands for the normalised, at the point σ, fundamental matrix of the homogeneous linear differential system

$$x'(t)=\left(1-2\frac{t-\sigma}{T}\right)P(t)x(t),\qquad t\in[\sigma, T+\sigma].$$

PROOF OF LEMMA 5.22. Obviously, it is sufficient to consider the case where not all the elements of the matrix-valued function K are equal identically to zero.

Being compact, the linear integral operator $H_\sigma K : C^n_T \to C^n_T$ has discrete spectrum only. Let, therefore, λ be a non-zero eigenvalue of $H_\sigma K$ and $u \in C^n_T$ be an eigenfunction corresponding to it.

According to (5.29), for $t\in[\sigma, T+\sigma]$, we have $k_\sigma(t)=0$ and, therefore, (5.28) yields

$$\left(1-\frac{t-\sigma}{T}\right)\int_\sigma^t K(s)u(s)\,\mathrm{d}s+\frac{t-\sigma}{T}\int_t^{T+\sigma} K(s)u(s)\,\mathrm{d}s=\lambda u(t). \tag{5.33}$$

If we now set

$$y(t)=\int_\sigma^t K(s)u(s)\,\mathrm{d}s,\qquad t\in[\sigma, T+\sigma], \tag{5.34}$$

then (5.33) implies that, for all t from $[\sigma, T+\sigma]$,

$$\left(1-\frac{t-\sigma}{T}\right)K(t)y(t)+\frac{t-\sigma}{T}K(t)\big[y(T+\sigma)-y(t)\big]=\lambda y'(t). \tag{5.35}$$

Thus, if u is an eigenfunction of the operator $H_\sigma K$ corresponding to the eigenvalue λ, then function (5.34) necessarily satisfies the initial-value problem

$$\begin{aligned}&y'(t)=\frac{1}{\lambda}\left(1-2\frac{t-\sigma}{T}\right)K(t)y(t)+\frac{t-\sigma}{\lambda T}K(t)y(T+\sigma),\\ &\quad t\in[\sigma, T+\sigma],\\ &y(\sigma)=0.\end{aligned} \tag{5.36}$$

The solution of this problem, as can be readily verified, satisfies the functional equation

$$\begin{aligned}&y(t)=\frac{1}{\lambda T}\Omega_{\frac{1}{\lambda}K,\sigma}(t)\int_\sigma^t \Omega^{-1}_{\frac{1}{\lambda}K,\sigma}(s)(s-\sigma)\,\mathrm{d}s\cdot y(T+\sigma),\\ &\quad t\in[\sigma, T+\sigma].\end{aligned} \tag{5.37}$$

Recall that the matrix-valued function $\Omega_{\frac{1}{\lambda}K,\sigma}:[\sigma, T+\sigma]\to \mathrm{GL}_n(\mathbb{R})$ have been defined immediately before the begin of present proof, whereas $\Omega^{-1}_{\frac{1}{\lambda}K,\sigma}$ denotes the matrix inverse for $\Omega_{\frac{1}{\lambda}K,\sigma}$.

Relation (5.37) implies, in particular, that the vector

$$c = y(T+\sigma) \tag{5.38}$$

is a solution of the system of homogeneous linear algebraic equations

$$\left[\lambda \mathbb{1}_n - \frac{1}{T}\Omega_{\frac{1}{\lambda}K,\sigma}(T+\sigma)\int_{\sigma}^{T+\sigma}\Omega^{-1}_{\frac{1}{\lambda}K,\sigma}(s)(s-\sigma)\,\mathrm{d}s\right]c = 0. \tag{5.39}$$

Solution (5.38) of system (5.39) is non-trivial because, in the contrary case, (5.36) and (5.34) imply the triviality of the eigenfunction u of the operator $H_\sigma K$. Thus, for the given λ, the matrix $B_{\frac{1}{\lambda}K,\sigma}$, where

$$B_{P,\sigma} := \lambda \mathbb{1}_n - \frac{1}{T}\Omega_{P,\sigma}(T+\sigma)\int_{\sigma}^{T+\sigma}\Omega^{-1}_{P,\sigma}(s)(s-\sigma)\,\mathrm{d}s, \tag{5.40}$$

is singular. Consequently,

$$R_{K,T} \leqslant \max\left\{|\lambda| \mid \det B_{\frac{1}{\lambda}K,\sigma} = 0\right\}. \tag{5.41}$$

On the other hand, if we assume that

$$\det B_{\frac{1}{\lambda}K,\sigma} = 0 \tag{5.42}$$

for a certain non-zero λ, then there exists a non-trivial solution c of the linear algebraic system (5.39). Then, as can be shown by using (5.42), the unique solution y of the initial-value problem

$$y'(t) = \frac{1}{\lambda}\left(1 - 2\frac{t-\sigma}{T}\right)K(t)y(t) + \frac{t-\sigma}{\lambda T}K(t)c, \quad t \in [\sigma, T+\sigma], \tag{5.43}$$

$$y(\sigma) = 0 \tag{5.44}$$

satisfies Eq. (5.37), whence the validity of equality (5.38) follows. Therefore, the function y is also a solution of problem (5.36).

Let us set

$$u(t) := \frac{1}{\lambda}\left(1 - 2\frac{t-\sigma}{T}\right)y(t) + \frac{t-\sigma}{\lambda T}c, \quad t \in [\sigma, T+\sigma]. \tag{5.45}$$

It is obvious from (5.43) and (5.45) that

$$y'(t) = K(t)u(t), \quad t \in [\sigma, T+\sigma],$$

and, hence, by virtue of (5.44), equality (5.34) is true.

We have thus established the existence of a non-trivial function $u \in C_T^n$ whose corresponding function (5.34) satisfies Eq. (5.35) for all t from $[\sigma, T+\sigma]$. Integrating the expressions on both sides of (5.35) from σ to $t \in [\sigma, T+\sigma]$ and taking (5.44) into account, we obtain that u satisfies relation (5.33), i.e., u is an eigenfunction of the operator $H_\sigma K$ with the eigenvalue λ. Consequently,

$$R_{K,T} \geqslant \max\Big\{|\lambda| \;\Big|\; \det B_{\frac{1}{\lambda}K,\sigma} = 0\Big\}. \tag{5.46}$$

Finally, combining (5.41) and (5.46), we arrive at the equality

$$R_{K,T} = \max\Big\{|\lambda| \;\Big|\; \det B_{\frac{1}{\lambda}K,\sigma} = 0\Big\}. \tag{5.47}$$

Let us now show that the value of $R_{K,T}$ is independent of σ, which fact will justify notation (5.30). Indeed, since, by the well-known property of the fundamental matrix, we have

$$\Omega_{P,\sigma}(t) = \Omega_{P,0}(t)\Omega_{P,\sigma}(0),$$

it follows that

$$\begin{aligned} B_{P,\sigma} &= \lambda\mathbb{1}_n - \frac{1}{T}\Omega_{P,\sigma}(T+\sigma)\int_\sigma^{T+\sigma} \Omega_{P,\sigma}^{-1}(s)(s-\sigma)\,\mathrm{d}s \\ &= \lambda\mathbb{1}_n - \frac{1}{T}\Omega_{P,0}(T+\sigma)\Omega_{P,\sigma}(0)\int_0^T s\Omega_{P,\sigma}^{-1}(0)\Omega_{P,0}^{-1}(s+\sigma)\,\mathrm{d}s \\ &= \lambda\mathbb{1}_n - \frac{1}{T}\Omega_{P,0}(T+\sigma)\int_0^T s\Omega_{P,0}^{-1}(s+\sigma)\,\mathrm{d}s. \end{aligned} \tag{5.48}$$

However,

$$\Omega_0(s+\sigma) = \Omega_0(s)\Omega_0(\sigma),$$

and, therefore, (5.48) yields

$$\begin{aligned} B_{P,\sigma} &= \lambda\mathbb{1}_n - \frac{1}{T}\Omega_{P,0}(T)\Omega_0(\sigma)\int_0^T s\Omega_0^{-1}(\sigma)\Omega_{P,0}^{-1}(s)\,\mathrm{d}s \\ &= \lambda\mathbb{1}_n - \frac{1}{T}\Omega_{P,0}(T)\int_0^T s\Omega_{P,0}^{-1}(s)\,\mathrm{d}s. \end{aligned}$$

The last equality implies, in particular, that the matrix $B_{K\lambda^{-1},\sigma}$, in fact, does not depend on σ. This, by virtue of equality (5.47), means that number (5.30) has the same property.

Finally, setting $\sigma = 0$ in (5.47), we arrive at equality (5.32), which completes the proof of our lemma. □

COROLLARY 5.23. *Let the matrix-valued function* $K:(-\infty,\infty)\to \mathrm{GL}_n(\mathbb{R})$ *have the form*

$$K(t)=\operatorname{diag}\bigl(k_1(t),k_2(t),\dots,k_n(t)\bigr),\quad t\in(-\infty,\infty),$$

where $k_1, k_2, \dots, k_n$ *are certain continuous scalar functions on* $(-\infty,\infty)$. *Then the number* $1/R_{K,T}$ *is equal to the least of the absolute values of the numbers* μ *satisfying the relation*

$$\frac{T}{\mu}=\int_0^T s\exp\biggl[\mu\int_s^T\biggl(1-\frac{2\xi}{T}\biggr)k_\nu(\xi)\,\mathrm{d}\xi\biggr]\,\mathrm{d}s$$

for a certain $\nu\in\{1,2,\dots,n\}$.

PROOF. It remains to notice that, in the case considered, $\Omega_{\mu K,0}(t)$ is a diagonal matrix with the components

$$\exp\biggl[\mu\int_0^t\biggl(1-\frac{2\xi}{T}\biggr)k_\nu(\xi)\,\mathrm{d}\xi\biggr], \tag{5.49}$$

where $\nu\in\{1,2,\dots,n\}$. Applying Lemma 5.22 and substituting functions (5.49) into the expressions for the components of $\Omega_{\mu K,0}(t)$ in (5.32), we obtain the conclusion desired. □

REMARK 5.24. By using the well-known Krein–Rutman theorem on the maximal eigenvalue of a completely continuous positive operator (see Krein and Rutman [38], Theorem 6.2), one can show that, in formula (5.32), one may consider the positive real μ only.

COROLLARY 5.25. *For an arbitrary constant square matrix* K *of dimension* n, *the relation*

$$R_{K,T}\approx\frac{Tr(K)}{3.4161}$$

is true.

The assertion of Corollary 5.25 is known, e.g., from [64].

5.4. *Successive approximations and their convergence*

Having fixed some $\sigma\in(-\infty,\infty)$ and $z\in D_{\frac{T}{2}M}$, we introduce the sequence of functions $\{x_m(\cdot,z)\mid m\geqslant 0\}$ by the recurrence formula

$$x_{m+1}(t,z)=z+\int_\sigma^t f\bigl(s,x_m(s,z)\bigr)\,\mathrm{d}s-\frac{t-\sigma}{T}\int_\sigma^{T+\sigma} f\bigl(s,x_m(s,z)\bigr)\,\mathrm{d}s, \tag{5.50}$$

where $x_0(t,z)=z$ for all $t\in[\sigma,T+\sigma]$. Clearly, all the members of thus defined sequence belong to the space C_T^n.

LEMMA 5.26. *For all $x\in C_T^n$ and $t\in[\sigma,T+\sigma]$, the estimate*

$$\begin{aligned}&\left|\int_\sigma^t\left[x(s)-\frac{1}{T}\int_\sigma^{T+\sigma}x(\xi)\,\mathrm{d}\xi\right]\mathrm{d}s\right|\\&\quad\leqslant\frac{1}{2}\alpha_{1,T}(t-\sigma)\Big(\max_{s\in(-\infty,\infty)}x(s)-\min_{s\in(-\infty,\infty)}x(s)\Big)\end{aligned}\tag{5.51}$$

is true, where $\alpha_{1,T}$ is the function given by formula (3.34).

PROOF. Consider the expression from the left-hand side of the desired inequality (5.51). Taking into account the fact that, for y from C_T^n, the relation

$$\int_\sigma^{T+\sigma}y(s)\,\mathrm{d}s=\int_0^T y(s)\,\mathrm{d}s$$

holds, and carrying out the change of variables $\eta=s-\sigma$ in each of the integrals, we obtain

$$\int_\sigma^t\left[x(s)-\frac{1}{T}\int_0^T x(\xi)\,\mathrm{d}\xi\right]\mathrm{d}s=\int_0^{t-\sigma}\left[x(\xi+\sigma)-\frac{1}{T}\int_0^T x(\eta+\sigma)\,\mathrm{d}\eta\right]\mathrm{d}\xi.$$

In view of Lemma 2.3 from [70], the estimate

$$\left|\int_0^t\left[x(s)-\frac{1}{T}\int_0^T x(\xi)\,\mathrm{d}\xi\right]\mathrm{d}s\right|\leqslant\frac{1}{2}\alpha(t)\Big[\max_{s\in(-\infty,\infty)}x(s)-\min_{s\in(-\infty,\infty)}x(s)\Big]$$

holds for all $x\in C_T^n$ and $t\in(-\infty,\infty)$, where the scalar function α is given by (3.34). This yields that, for $t\in[\sigma,T+\sigma]$,

$$\begin{aligned}&\left|\int_\sigma^t\left[x(s)-\frac{1}{T}\int_0^T x(\sigma)d\xi\right]\mathrm{d}s\right|\\&\quad\leqslant\frac{1}{2}\alpha(t-\sigma)\Big[\max_{\xi\in(-\infty,\infty)}x(\xi+\sigma)-\min_{\xi\in(-\infty,\infty)}x(\xi+\sigma)\Big]\\&\quad=\frac{1}{2}\alpha(t-\sigma)\Big[\max_{\xi\in(-\infty,\infty)}x(\xi)-\min_{\xi\in(-\infty,\infty)}x(\xi)\Big],\end{aligned}$$

which coincides with the required estimate (5.51). The lemma is proved. □

Recall that, throughout all the paper, the inequalities and the symbols max and min are understood in the sense of Remark 5.20.

LEMMA 5.27. *For an arbitrary z from $D_{\frac{T}{2}M}$, the values of all the functions of sequence* (5.50) *belong to D.*

PROOF. The assertion of Lemma 5.27 is established similarly to the argument in Section 2, Chapter 1 of [70]. For this purpose, one should use assumptions 1, 3, and 4. □

Having fixed a certain arbitrary $\sigma \in (-\infty, \infty)$, for all u from C_T^n and $t \in (-\infty, \infty)$, we put

$$(\Lambda_\sigma u)(t) := \int_\sigma^{t - T k_\sigma(t)} u(s)\,\mathrm{d}s - \left(\frac{t-\sigma}{T} + k_\sigma(t)\right) \int_0^T u(s)\,\mathrm{d}s, \tag{5.52}$$

where $k_\sigma : (-\infty, \infty) \to \mathbb{Z}$ is the function defined by formula (5.29). By analogy with Lemma 5.21, it is not difficult to verify that $(\Lambda_\sigma u)(t+T) = (\Lambda_\sigma u)(t)$ for all $u \in C_T^n$ and $t \in (-\infty, \infty)$. Therefore, formula (5.52) defines a (bounded) linear operator $\Lambda_\sigma : C_T^n \to C_T^n$.

LEMMA 5.28. *For arbitrary non-negative u from C_T^n and all $t \in (-\infty, \infty)$, the componentwise inequality*

$$\left|(\Lambda_\sigma u)(t)\right| \leqslant (H_\sigma u)(t) \tag{5.53}$$

is true, where H_σ is the operator in C_T^n defined by formula (5.28).

PROOF. Let u belong to C_T^n, $u \geqslant 0$. Due to the T-periodicity of the functions Λ_σ and H_σ, it suffices to verify the validity of estimate (5.53) for $t \in [\sigma, T+\sigma]$ only.

According to (5.52) and (5.29), for $t \in [\sigma, T+\sigma]$, we have

$$(\Lambda_\sigma u)(t) = \left(1 - \frac{t-\sigma}{T}\right) \int_\sigma^t u(s)\,\mathrm{d}s + \frac{t-\sigma}{T} \int_t^{T+\sigma} u(s)\,\mathrm{d}s,$$

whence

$$\left|(\Lambda_\sigma u)(t)\right| \leqslant \left(1 - \frac{t-\sigma}{T}\right) \int_\sigma^t u(s)\,\mathrm{d}s + \frac{t-\sigma}{T} \int_t^{T+\sigma} u(s)\,\mathrm{d}s.$$

Recalling now the definition (5.28) of the operator H_σ, we arrive at (5.53). □

LEMMA 5.29. *For all $z \in D_{\frac{T}{2}M}$, $t \in (-\infty, \infty)$, and $m = 1, 2, \ldots$, the estimate*

$$\varrho_m(t, z) \leqslant \big(H_\sigma K \varrho_{m-1}(\cdot, z)\big)(t) \tag{5.54}$$

is true, where

$$\varrho_m(t,z) := \left|x_m(t,z) - x_{m-1}(t,z)\right| \tag{5.55}$$

and the functions $x_m(\cdot,z)$ $(m \geqslant 1)$ *are given by the recurrence formula* (5.50).

PROOF. According to Lemma 5.27, for all $z \in D_{\frac{T}{2}M}$, $t \in (-\infty,\infty)$, and $m = 1,2,\ldots,$ we have $x_m(t,z) \in D$. Therefore, in the estimate

$$\begin{aligned}\varrho_{m+1}(t) \leqslant{}& \left(1 - \frac{t-\sigma}{T}\right)\int_\sigma^t \left[f\left(s, x_m(s,z)\right) - f\left(s, x_{m-1}(s,z)\right)\right] \mathrm{d}s \\ &+ \frac{t-\sigma}{T}\int_t^{T+\sigma} \left[f\left(s, x_m(s,z)\right) - f\left(s, x_{m-1}(s,z)\right)\right] \mathrm{d}s, \\ t \in [\sigma, T+\sigma],\end{aligned}$$

which is an easy consequence of Corollary 5.50, we can use inequality (5.27) from assumption 2:

$$\begin{aligned}\varrho_{m+1}(t) \leqslant{}& \left(1 - \frac{t-\sigma}{T}\right)\int_\sigma^t K(s)\left|x_m(s,z) - x_{m-1}(s,z)\right| \mathrm{d}s \\ &+ \frac{t-\sigma}{T}\int_t^{T+\sigma} K(s)\left|x_m(s,z) - x_{m-1}(s,z)\right| \mathrm{d}s, \quad t \in [\sigma, T+\sigma].\end{aligned}$$

Taking notation (5.55) and definition (5.28) of the operator H_σ into account, we arrive immediately at (5.54). Moreover, due to the T-periodicity of the functions considered, relation (5.54) holds for all $t \in (-\infty,\infty)$. The lemma is proved. □

LEMMA 5.30. *Under the assumption that*

$$R_{K,T} < 1, \tag{5.56}$$

for all $z \in D_{\frac{T}{2}M}$, *the sequence* $\{x_m(\cdot,z) \mid m \geqslant 0\}$ *given by formula* (5.50) *converges uniformly to a certain function* $x(\cdot,z) \in C_T^n$. *Furthermore, for all* $[\sigma, T+\sigma]$, *the pointwise and componentwise estimate*

$$\begin{aligned}\left|x_m(\cdot,z) - x(\cdot,z)\right| \leqslant{}& \frac{1}{2}(H_\sigma K)^m (I - H_\sigma K)^{-1} \alpha_{1,T}(\cdot - \sigma) \\ &\times \left[\max_{s\in(-\infty,\infty)} f(s,z) - \min_{s\in(-\infty,\infty)} f(s,z)\right]\end{aligned} \tag{5.57}$$

is true, in which the function $\alpha_{1,T} : \mathbb{R} \to \mathbb{R}$ *is given by formula* (3.34).

For all z indicated, the function $x(\cdot, z)$ is the unique solution of the non-linear integral equation

$$x(t) = z + \int_\sigma^t f\bigl(s, x(s)\bigr)\,\mathrm{d}s - \frac{t-\sigma}{T}\int_0^T f\bigl(s, x(s)\bigr)\,\mathrm{d}s,$$
$$t \in [\sigma, T+\sigma]. \tag{5.58}$$

This solution, moreover, satisfies the differential system (5.1) *if, and only if*

$$\int_0^T f\bigl(s, x(s, z)\bigr)\,\mathrm{d}s = 0. \tag{5.59}$$

The symbol I in inequality (5.57) stands for the identity operator in C_T^n.

PROOF. It is easy to see that the linear continuous operator $H_\sigma K : C_T^n \to C_T^n$ is positive in the sense that the (pointwise and componentwise) inequality $x \geqslant 0$ always yields $H_\sigma K x \geqslant 0$. Then, according to condition (5.56), the inverse operator $(I - H_\sigma K)^{-1}$ exists and, by virtue of the equality

$$(I - H_\sigma K)^{-1} = \sum_{m=0}^{+\infty} (H_\sigma K)^m, \tag{5.60}$$

is positive in the same sense. By virtue of the positivity of the operator $H_\sigma K$, Lemma 5.29 yields the pointwise inequality

$$\varrho_{m+1}(\cdot, z) \leqslant (H_\sigma K)^m \varrho_1(\cdot, z)$$

valid for all $m = 1, 2, 3, \ldots$. Therefore, according to (5.60), we have

$$\begin{aligned}\bigl|x_{m+k}(\cdot, z) - x_m(\cdot, z)\bigr| &\leqslant (H_\sigma K)^m \sum_{\nu=1}^{k} (H_\sigma K)^\nu \varrho_1(\cdot, z)\\ &\leqslant (H_\sigma K)^m \sum_{\nu=1}^{+\infty} (H_\sigma K)^\nu \varrho_1(\cdot, z)\\ &= (H_\sigma K)^m (I - H_\sigma K)^{-1} \varrho_1(\cdot, z).\end{aligned}$$

Thus, the sequence of continuous T-periodic vector-functions on $(-\infty, \infty)$ given by the recurrence formula (5.50) is fundamental in the uniform norm and, hence, is uniformly convergent.

On the other hand, it is clear from (5.55) that, for all (t, z) from $[\sigma, T+\sigma] \times D$,

$$\varrho_1(t,z) \leqslant \frac{1}{2}\alpha_{1,T}(t-\sigma)\Big[\max_{s\in(-\infty,\infty)} f(s,z) - \min_{s\in(-\infty,\infty)} f(s,z)\Big],$$

whence estimate (5.57) immediately follows.

Passing to the limit as $m \to +\infty$ in (5.50), we easily show that the relation

$$\lim_{m\to+\infty} \max_{t\in(-\infty,\infty)} \big|x_m(t,z) - x(t,z)\big| = 0$$

yields (5.58). The continuous differentiability of the function $x(\cdot,z)$ is obvious from the structure of equality (5.58). Differentiating both sides of the relation mentioned, we arrive at the system of non-linear integro-functional equations

$$x'(t,z) = f\big(t, x(t,z)\big) - \frac{1}{T}\int_0^T f\big(s, x(s,z)\big)\,\mathrm{d}s, \quad t \in (-\infty,\infty). \tag{5.61}$$

It is clear from the form of system (5.61) that every solution $x(\cdot,z)$, of system (5.61) satisfying equality (5.59) is also a solution of system (5.1), and *vice versa*. Thus, all the assertions of the lemma are proved. □

REMARK 5.31. It follows from assumption 5.19 that, under the conditions of Lemma 5.30, for all $m \geqslant 1$, the estimate

$$\big|x_m(\cdot,z) - x(\cdot,z)\big| \leqslant \frac{1}{2}(H_\sigma K)^m (I - H_\sigma K)^{-1}\alpha_{1,T}(\cdot-\sigma)M \tag{5.62}$$

holds on $[\sigma, T+\sigma]$. This estimate rougher but simpler than (5.57).

Note that, in the right-hand side of inequalities (5.57) and (5.62), the value of the operator $(H_\sigma K)^m(I - H_\sigma K)^{-1}$ is calculated on the function of the form

$$[\sigma, T+\sigma] \ni t \longmapsto \alpha_{1,T}(t-\sigma)c,$$

where $\alpha_{1,T}$ is the scalar function defined by formula (3.34) and $c \in \mathbb{R}^n$ is a constant vector independent of z.

REMARK 5.32. In the case where

$$K(t) \equiv K, \quad t \in (-\infty,\infty),$$

in assumption 2 from Section 5.2, condition (5.56), according to Corollary 5.25, takes the form

$$r(K) < \frac{3.4161\ldots}{T}, \tag{5.63}$$

where $r(K)$ is the greatest positive eigenvalue of the constant matrix K.

REMARK 5.33. The constant $r = 1/3.41613062\ldots$ has also appeared in [85] in the form $r = \inf_{t\in[0,1]} r(t)$, where $r(t)$, $t \in [0,1]$, is the convergence radius of the Poisson–Abel functional series

$$\sum_{k=0}^{\infty} \alpha_{k,1}(t)\lambda^k, \quad |\lambda| < r, \tag{5.64}$$

and $\alpha_{k,1}$, $k = 1, 2, \ldots$, are the functions given by formulae (3.35) for $T = 1$.

5.5. *Case of "large" Lipschitz constants*

Assuming the global Lipschitz condition 2 ((5.27) in Assumption 5.19), which, in other words, means the fulfilment of inequality (5.27) for arbitrary x and y from $\mathbb{R}^n$, one can drop condition (5.56). For this purpose, one can modify the recurrence formula (5.50) by using certain considerations from the paper of Ronto [61].

Let us consider problem (3.3), (3.4) under the following assumptions.

ASSUMPTION 5.34.

(1) $f : [0,T] \times D \to \mathbb{R}^n$ satisfies the Carathéodory conditions, i.e.,
  (a) The mapping $D \ni x \mapsto f(t,x)$ is continuous for almost every t from $[0,T]$;
  (b) The mapping $[0,T] \ni t \mapsto f(t,x)$ is Lebesgue measurable for all $x \in D$ and, moreover, the function

$$[0,T] \ni t \longmapsto M(t) := \sup_{x\in D} \bigl|f(t,x)\bigr| \tag{5.65}$$

  is Lebesgue integrable.

(2) There exists a matrix-valued function $K : [0,T] \to \mathrm{GL}_n(\mathbb{R})$ with Lebesgue integrable elements such that relation (5.27) holds for a.e. $t \in [0,T]$ and all $\{x_1, x_2\} \subset D$.

Let $\Phi : [0,T] \to \mathrm{GL}_n(\mathbb{R})$ be an arbitrary matrix-valued function with absolutely continuous elements possessing the property

$$\Phi(T) = \Phi(0) + \mathbb{1}_n. \tag{5.66}$$

It is obvious that such functions exist. One can always put, e.g.,

$$\Phi(t) := \operatorname{diag}\left(\frac{t^{\alpha_1} - \tau_1^{\alpha_1}}{T^{\alpha_1}}, \frac{t^{\alpha_2} - \tau_2^{\alpha_2}}{T^{\alpha_2}}, \ldots, \frac{t^{\alpha_n} - \tau_n^{\alpha_n}}{T^{\alpha_n}}\right), \quad t \in [0,T],$$

where $\{\tau_k \mid k = 1, 2, \ldots, n\} \subset \mathbb{R}$ and $\{\alpha_k \mid k = 1, 2, \ldots, n\} \subset (0, +\infty)$.

For any integrable function $x : [0,T] \to \mathbb{R}^n$, we put

$$(W_\Phi x)(t) := \bigl|\mathbb{1}_n - \Phi(t)\bigr| \int_0^t x(s)\,\mathrm{d}s + \bigl|\Phi(t)\bigr| \int_t^T x(s)\,\mathrm{d}s$$

at a.e. $t \in [0, T]$. It is clear that W_Φ can be regarded as a linear mapping from $L_1([0, T], \mathbb{R}^n)$ to $C([0, T], \mathbb{R}^n)$.

Let

$$\beta_\Phi := \max_{t \in [0,T]} (W_\Phi M)(t),$$

where $M : [0, T] \to \mathbb{R}^n$ is the integrable function (5.65) from Assumption 5.34.

ASSUMPTION 5.35. The set D_{β_Φ} is non-empty.

We recall that the set D_{β_Φ} is defined according to formula (1.3) in notation 1, Section 1. Assumption 5.35 means that the set D where f satisfies the Carathéodory and Lipschitz conditions is "large enough."

PROPOSITION 5.36. *Under Assumptions* 5.34 *and* 5.35:

(1) *For any solution x of the periodic problem* (3.3), (3.4), *there exists some $\xi \in \mathbb{R}^n$ such that*

$$x(t) = \xi + \int_0^t f\big(s, x(s)\big)\,\mathrm{d}s - \Phi(t) \int_0^T f\big(s, x(s)\big)\,\mathrm{d}s, \quad t \in [0, T]; \tag{5.67}$$

(2) *For any $\xi \in \mathbb{R}^n$, every solution x of Eq.* (5.67) *satisfies the T-periodic boundary condition* (3.4);

(3) *If $x(\cdot, \xi)$ is a solution of Eq.* (5.67) *and the value of ξ is such that*

$$\int_0^T f\big(s, x(s, \xi)\big)\,\mathrm{d}s = 0, \tag{5.68}$$

then $x(\cdot, \xi)$ also satisfies the differential system (3.3).

For the sake of simplicity, we now assume that

$$\Phi(0) = 0 \tag{5.69}$$

and, therefore, by (5.66), it follows that

$$\Phi(T) = \mathbb{1}_n. \tag{5.70}$$

THEOREM 5.37. (See [61, Theorem 2].) *Under Assumptions* 5.34 *and* 5.35, *there exists an absolutely continuous matrix-valued function $\Phi : [0, T] \to \mathrm{GL}_n(\mathbb{R})$ with properties* (5.69) *and* (5.70) *such that, for any $\xi \in D_{\beta_\Phi}$, Eq.* (5.67) *has a unique absolutely continuous solution $x(\cdot, \xi)$. For all such ξ, the function $x(\cdot, \xi)$ satisfies the relation*

$$\lim_{m \to \infty} \max_{t \in [0,T]} \big| x(t, \xi) - x_m(t, \xi) \big| = 0,$$

where, for any $t \in [0, T]$,

$$x_m(t,\xi) := \xi + \int_0^t f\bigl(s, x_{m-1}(s,\xi)\bigr)\,\mathrm{d}s \\ - \Phi(t)\int_0^T f\bigl(s, x_{m-1}(s,\xi)\bigr)\,\mathrm{d}s, \tag{5.71}$$

and $x_0(\cdot,\xi) \in C([0,T], D)$ *is arbitrary. Furthermore, for all* $m = 1, 2, \ldots$, *the error estimate*

$$|x - x_m| \leqslant \phi_m, \tag{5.72}$$

where

$$\phi_m := (I - W_\Phi K)^{-1}(W_\Phi K)^m |x_1 - x_0|, \tag{5.73}$$

holds coordinatewise and pointwise on $[0, T]$.

In (5.73), the symbol K stands for the operator of multiplication by the matrix-valued function $K : [0,T] \to \mathrm{GL}_n(\mathbb{R})$.

By analogy with formula (3.91), Section 3.2, the properties of sequence (5.71) described in Theorem 5.37 allow us to introduce the functions

$$D_{\beta_\Phi} \ni z \longmapsto \Delta_m(z) := \int_0^T f\bigl(t, x_m(t,z)\bigr)\,\mathrm{d}t \tag{5.74}$$

and consider the equation

$$\Delta_m(\xi) = 0$$

as an approximation of Eq. (5.68).

NOTATION 5.38. Let $G \subset \mathbb{R}^n$, $\Omega \subseteq G$, $\Omega \neq \varnothing$, and $X = (X_k)_{k=1}^n$ and $Y = (Y_k)_{k=1}^n$ be two functions from G to $\mathbb{R}^n$. We write $X \sqsupset_\Omega Y$ if and only if there exists a mapping $k : G \to \{1, 2, \ldots, n\}$ such that $X_{k(\xi)} > Y_{k(\xi)}$ for any $\xi \in \Omega$.

THEOREM 5.39. (See [61, Corollary 3].) *Let* $f : [0,T] \times D \to \mathbb{R}^n$ *satisfy Assumptions* 5.34 *and* 5.35. *Let* $\Phi : [0,T] \to \mathrm{GL}_n(\mathbb{R})$ *be an absolutely continuous matrix-valued function possessing properties* (5.69), (5.70) *and such that*

$$r(W_\Phi K) < 1. \tag{5.75}$$

Assume that the set D_{β_Φ} *contains a certain closed subdomain* Ω *such that, for a certain natural* m, *the relation*

$$|\Delta_m| \sqsupset_{\partial\Omega} \int_0^T K(t)\phi_m(t)\,\mathrm{d}t, \tag{5.76}$$

holds, where $\phi_m : [0, T] \to \mathbb{R}^n$ *is the function given by formula* (5.73). *If moreover,* $\deg(\Delta_m, \Omega, 0) \neq 0$, *then the boundary value problem* (3.3), (3.4) *has a solution* x *with* $x(0) \in \Omega$.

Theorems 5.37 is established with the aid of Theorem 5.39. The proof of Theorem 5.39, in turn, is based upon the following statement [61, Lemma 1].

LEMMA 5.40. *For an arbitrary* $\varepsilon \in (0, \infty)$ *one can specify some* $\delta_\varepsilon \in (0, \infty)$ *such that* $r(W_\Phi K) < \varepsilon$, *provided that* $\max_{k,j=1,2,\dots,n} \int_0^T |\Phi_{kj}(t)|\, \mathrm{d}t < \delta_\varepsilon$.

Here, Φ_{kj}, $k = 1, 2, \dots, n$, stand for the corresponding elements of the matrix $\Phi = (\Phi_{kj})_{k,j=1}^n$. It follows from Lemma 5.40 that an absolutely continuous function $\Phi : [0, T] \to \mathrm{GL}_n(\mathbb{R})$ having properties (5.69), (5.70), and (5.75) exists even in the cases where the integrals $\int_0^T |K_{kj}(t)|\, \mathrm{d}t$, $k, j = 1, 2, \dots, n$, of the elements of the variable matrix $K = (K_{kj})_{k,j=1}^n : [0, T] \to \mathrm{GL}_n(\mathbb{R})$ are arbitrarily large.

Thus, we can always suppose that condition (5.75) of Theorem 5.39 (unlike condition (5.56) of Lemma 5.30) is satisfied.

5.6. *Symmetries* (τ, E) *of solutions of the integral equation* (5.58)

Here, we describe certain conditions under which a certain (generally speaking, infinite) set of solutions of the n-parametric family of Eqs. (5.58) possess property (τ, E). For this purpose, we first state an important property of the linear operator $\Lambda_\sigma : C_T^n \to C_T^n$ given by formula (5.52).

LEMMA 5.41. *For an arbitrary point* $\sigma \in (-\infty, \infty)$ *and an arbitrary non-singular* n-*dimensional matrix* E, *the inclusion*

$$\Lambda_\sigma C_T^n(-2\sigma, -E^{-1}) \subset C_T^n(-2\sigma, E) \tag{5.77}$$

is true.

Inclusion (5.77) of Lemma 5.41 is proved by using [67, Lemma 5]. Recall that the linear set $C_T^n(-2\sigma, E)$ is defined by formula (5.3).

LEMMA 5.42. *Assume that the continuous function* $f : (-\infty, \infty) \times D \to \mathbb{R}^n$, *point* $\tau \in (-\infty, \infty)$, *and the non-singular* n-*dimensional matrix* E *are such that*

$$-f(-t - \tau, z) = Ef(t, Ez) \tag{5.78}$$

for all $z \in \mathbb{R}^n$ *and* $t \in (-\infty, \infty)$. *Then, for an arbitrary* (τ, E)-*proper function* $x \in C_T^n$, *the corresponding function*

$$(-\infty, \infty) \ni t \longmapsto f\big(t, x(t)\big)$$

possesses property $(\tau, -E^{-1})$.

Recall that the symmetry properties (τ, E) and $(\tau, -E^{-1})$ are described by Definition 5.1.

PROOF OF LEMMA 5.42. Let a function $x \in C_T^n$ have property (τ, E). According to Definition 5.1, this means that, for all t from $(-\infty, \infty)$, equality (5.2) is satisfied, whence

$$x(-t-\tau) = E^{-1}x(t), \quad t \in (-\infty, \infty). \tag{5.79}$$

Setting

$$u(t) := f\big(t, x(t)\big), \quad t \in (-\infty, \infty), \tag{5.80}$$

and taking relation (5.79) into account, we get

$$\begin{aligned} -E^{-1}u(-t-\tau) &= -E^{-1}f\big(-t-\tau, x(-t-\tau)\big) \\ &= -E^{-1}f\big(-t-\tau, E^{-1}x(t)\big). \end{aligned} \tag{5.81}$$

Using now condition (5.78) and definition (5.80), from equality (5.81) we find

$$-E^{-1}u(-t-\tau) = u(t), \quad t \in (-\infty, \infty).$$

The last equality means that u has property $(\tau, -E^{-1})$. Since the (τ, E)-proper function u is chosen arbitrarily, we have thus proved our lemma. □

REMARK 5.43. Condition (5.78) resembles condition (6.17) from Hale [28] that, in our notation, has form (5.78) with $\tau = 0$ and the matrix E possessing the properties

$$E^2 = \mathbb{1}_n, \qquad EP_0 = P_0E, \tag{5.82}$$

where P_0 is a certain projection operator defined on the set of continuous T-periodic functions. The former condition is, obviously, more general, because it involves the shift of the argument by an arbitrary τ. Moreover, we do not assume that the matrix E should satisfy (5.82) or any other related conditions used in the work cited.

Lemmata 5.30, 5.41, and 5.42 allow us to obtain the following statement.

THEOREM 5.44. *Assume that, for all real t, the function $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ satisfies condition* (5.78) *with certain number τ and non-singular n-dimensional matrix E. Let Assump-*

tion 5.19 *be true and, moreover, let inequality* (5.56) *hold. Then, for all* z *from* $D_{\frac{T}{2}M}$ *such that*[2]

$$Ez = z, \tag{5.83}$$

the integro-functional equation

$$\begin{aligned} x(t) = z &+ \int_{-\frac{\tau}{2}}^{t - T\lfloor T^{-1}(t+\frac{\tau}{2})\rfloor} f\big(s, x(s)\big)\,\mathrm{d}s \\ &- \frac{2(t - T\lfloor T^{-1}(t + \frac{\tau}{2})\rfloor) + \tau}{2T} \int_{-\frac{\tau}{2}}^{T-\frac{\tau}{2}} f\big(s, x(s)\big)\,\mathrm{d}s, \\ & t \in (-\infty, \infty), \end{aligned} \tag{5.84}$$

has a unique T*-periodic solution* $x(\cdot, z) : (-\infty, \infty) \to \mathbb{R}^n$*, and this solution possesses property* (τ, E)*.*

Moreover, the fulfilment of equality (5.59) *is necessary and sufficient for the function* $x(\cdot, z)$ *to be a solution of the differential system* (5.1).

Note that the restriction of Eq. (5.84) to the interval $[-\tau/2, T - \tau/2]$ has the form

$$\begin{aligned} x(t) = z &+ \int_{-\frac{\tau}{2}}^{t} f\big(s, x(s)\big)\,\mathrm{d}s \\ &- \frac{2t + \tau}{2T} \int_{-\frac{\tau}{2}}^{T-\frac{\tau}{2}} f\big(s, x(s)\big)\,\mathrm{d}s, \quad t \in \left[-\frac{\tau}{2}, T - \frac{\tau}{2}\right], \end{aligned} \tag{5.85}$$

and that every solution of (5.85) satisfies the condition $x(-\tau/2) = x(T - \tau/2)$.

PROOF. It follows from (5.29) that

$$k_{-\frac{\tau}{2}}(t) = \left\lfloor \frac{1}{T}\left(t + \frac{\tau}{2}\right)\right\rfloor, \quad t \in (-\infty, \infty).$$

Therefore, as can be readily verified, Eq. (5.84) can be rewritten as

$$x = z + \Lambda_{-\frac{\tau}{2}} \mathfrak{f} x,$$

where $\mathfrak{f}$ is the Nemytsky operator generated by the function $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$,

$$u(\cdot) \longmapsto f\big(\cdot, u(\cdot)\big). \tag{5.86}$$

---

[2] A solution of Eq. (5.84) is unique for all $z \in D_{\frac{T}{2}M}$, not only for those satisfying condition (5.83). However, the fulfilment of the property (τ, E) is guaranteed only for those z indicated.

The T-periodicity of the function $f(\cdot, z)$ for z fixed ensures that the operator $\mathfrak{f}$ acts in the space C_T^n. We recall that the operator $\Lambda_{-\frac{\tau}{2}} : C_T^n \to C_T^n$ is defined by formula (5.52) for

$$\sigma = -\frac{\tau}{2}. \tag{5.87}$$

Let us set

$$x_{m+1} = z + \Lambda_{-\frac{\tau}{2}} \mathfrak{f} x_m, \quad m = 0, 1, 2, \ldots, \tag{5.88}$$

where $x_0 \equiv z$. It is obvious from (5.85) that the restriction of (5.88) to $[-\tau/2, T - \tau/2]$ coincides with (5.50) for σ given by (5.87).

Equality (5.83) that, by assumption, is satisfied by z, guarantees that the constant function x_0 belongs to $C_T^n(\tau, E)$ (see Definition 5.1). In view of Lemma 5.42, it follows from (5.78) and (5.87) that the function $\mathfrak{f} x_0$ possesses property $(\tau, -E^{-1})$, i.e., $u \in C_T^n(\tau, -E^{-1})$. By Lemma 5.41, this implies that, together with x_0, the function $\Lambda_{-\frac{\tau}{2}} \mathfrak{f} x_0$, also belongs to $C_T^n(\tau, E)$. Therefore, taking equality (5.88) and Lemma 5.2 into account, we conclude that $x_1 \in C_T^n(\tau, E)$. Arguing similarly, we show that all the members of the sequence of functions (5.88) possess property (τ, E).

Finally, applying Lemma 5.30 with σ given by (5.87) and taking condition (5.56) into account, we conclude that sequence (5.88) converges uniformly to a unique solution of the integro-functional Eq. (5.84). Since, by the consideration above, $x_m \in C_T^n(\tau, E)$ for all $m \geqslant 0$, we see that the corresponding limit function is also an element of the space $C_T^n(\tau, E)$, i.e., has property (τ, E).

It remains to note that the last assertion of Theorem 5.44 is also a consequence of Lemma 5.30. □

5.7. *Periodic solutions with property* (τ, E)

Now we are ready to establish the main theorem of the section. To prove it, besides the results of the preceding subsections, we need the following

LEMMA 5.45. *For an arbitrary function $x \in C_T^n$ having property (τ, G) with certain real τ and non-singular matrix G satisfying the condition*

$$\det(\mathbb{1}_n - G) \neq 0, \tag{5.89}$$

the equality

$$\int_0^T x(s)\,\mathrm{d}s = 0 \tag{5.90}$$

is true.

PROOF. Let us carry out the change of variables $s = -\xi - \tau$ in the integral from the left-hand side of (5.90). Taking into account the T-periodicity of the function $x \in C_T^n(\tau, G)$ and its property (τ, G), we get

$$\begin{aligned} G\int_0^T x(s)\,\mathrm{d}s &= -G\int_{-\tau}^{-T-\tau} x(-\xi-\tau)\,\mathrm{d}\xi \\ &= -G\int_0^{-T} x(-\xi-\tau)\,\mathrm{d}\xi \\ &= G\int_0^{T} x(-\xi-\tau)\,\mathrm{d}\xi \\ &= \int_0^T x(\xi)\,\mathrm{d}\xi, \end{aligned}$$

i.e.,

$$(\mathbb{1}_n - G)\int_0^T x(s)\,\mathrm{d}s = 0. \tag{5.91}$$

In view of (5.89), relation (5.91) yields the required equality (5.90). The lemma is proved. □

The statements obtained above lead us to the following theorem.

THEOREM 5.46. *Let us suppose that Assumption* 5.19[3] *hols for the functions* $f:\mathbb{R}^{n+1}\to\mathbb{R}^n$*, inequality* (5.56) *is true, and, moreover,* f *satisfies condition* (5.78) *with certain real* τ *and non-singular* n*-dimensional matrix* E *such that*

$$\det(\mathbb{1}_n + E) \neq 0. \tag{5.92}$$

Then every solution x *of the system of differential Eqs.* (5.1) *for which the value* $x(-\frac{\tau}{2})$ *belongs to the set* $D_{\frac{\tau}{2}M}$ *and satisfies the equality*

$$x\left(-\frac{\tau}{2}\right) = Ex\left(-\frac{\tau}{2}\right), \tag{5.93}$$

is a T*-periodic solution possessing property* (τ, E).

PROOF. According to Theorem 5.44, under our conditions, the integro-differential equation (5.84) has a unique solution $x(\cdot, z)$ for an arbitrary $z \in \mathbb{R}^n$. For all vectors z satisfying equality (5.83), this solution, moreover, has property (τ, E).

---

[3] Namely, conditions (5.19)–(5.19) from Section 5.2, p. 517.

By Lemma 5.42, condition (5.78) imposed on the function f allows us to claim that the function

$$(-\infty, \infty) \ni t \longmapsto f\bigl(t, x(t, z)\bigr) \tag{5.94}$$

has property $(\tau, -E^{-1})$. Since, obviously,

$$\mathbb{1}_n + E^{-1} = E^{-1}(\mathbb{1}_n + E),$$

this, in view of (5.92), implies that the matrix $G := -E^{-1}$ and function (5.94) belonging to the subspace $C_T^n(\tau, G)$ satisfy the assumptions of Lemma 5.45. The lemma mentioned yields the relation

$$\int_{-\frac{\tau}{2}}^{T-\frac{\tau}{2}} f\bigl(s, x(s, z)\bigr)\, \mathrm{d}s = 0$$

which, by virtue of the T-periodicity of the functions considered, coincides with equality (5.59).

Thus, the solution $x(\cdot, z)$ of Eq. (5.84) for all real t satisfies the relation

$$x(t, z) = z + \int_{-\frac{\tau}{2}}^{t} f\bigl(s, x(s, z)\bigr)\, \mathrm{d}s,$$

whence it follows that this function is a solution of the initial-value problem

$$x'(t) = f\bigl(t, x(t)\bigr), \quad -\infty < t < \infty,$$
$$x\left(-\frac{\tau}{2}\right) = z.$$

On the other hand, we have proved above that $x(\cdot, z)$ lies in $C_T^n(\tau, E)$ and, in particular, is periodic with period T. It remains only to note that all these facts are true for an arbitrary z with property (5.83), which means that every solution of system (5.1) satisfying equality (5.93) belongs to the subspace $C_T^n(\tau, E)$. The theorem is proved. □

A review of the proof of Theorem 5.46 given above allows us to complement it by the following remark.

REMARK 5.47. The derivative of each of the (τ, E)-proper solutions of the differential system (5.1) that are dealt with in Theorem 5.46 has property $(\tau, -E^{-1})$.

In the case where the Lipschitz condition (5.27) for the function f in the right-hand side of system (5.1) is satisfied in the entire space and with a constant matrix, i.e., when (3.5) holds for all $t \in (-\infty, \infty)$ and $\{x_1, x_2\} \subset \mathbb{R}^n$, where $K \in \mathrm{GL}_n(\mathbb{R})$, the conditions of Theorem 5.46 are simplified. Namely, the following statement holds.

COROLLARY 5.48. *Let the continuous function* $f:\mathbb{R}^{n+1}\to\mathbb{R}^n$ *be* T*-periodic in the first variable and satisfy the Lipschitz condition* (3.5) *with a certain* n*-dimensional matrix* K*, whose spectral radius fulfils inequality* (5.63). *Assume also that, for certain real number* τ *and non-singular* n*-dimensional matrix* E *with property* (5.92)*, the function* f *satisfies condition* (5.78).

Then every solution x *of the system of differential Eqs.* (5.1)*, whose value at the point* $-\frac{\tau}{2}$ *satisfies relation* (5.93)*, is a solution periodic with period* T *and possessing property* (τ, E).

A model example of a non-singular matrix E satisfying condition (5.92) of Theorem 5.46 is provided by the unit matrix $E = \mathbb{1}_n$, for which the validity of equality (5.2) means the τ-evenness of the function x. Recall that the notion of a τ-even function is described by Definition 5.4.

COROLLARY 5.49. *If the matrix* K *in the Lipschitz condition* (5.27) *is chosen to be constant, and the continuous function* $f:\mathbb{R}^{n+1}\to\mathbb{R}^n$ *is* T*-periodic with respect to the first variable and satisfies the condition*

$$-f(-t-\tau, z) = f(t, z), \quad -\infty < t < \infty, \tag{5.95}$$

then the fulfilment of inequality (5.63) *guarantees that all the solutions of the differential system* (5.1) *whose values at the point* $-\frac{\tau}{2}$ *lie in the set* $D_{\frac{T}{2}M}$*, are* τ*-even and periodic with period* T.

PROOF. According to Remark 5.32, inequality (5.63) guarantees the fulfilment of condition (5.56). It is also clear that (5.95) is a particular case of (5.78) with $E = \mathbb{1}_n$. Thus, the required assertion is a consequence of Theorem 5.46 for $E = \mathbb{1}_n$. □

The following statement is an immediate consequence of Corollary 5.49.

COROLLARY 5.50. *Let the continuous function* $f:\mathbb{R}^{n+1}\to\mathbb{R}^n$ *be* T*-periodic and odd in the first variable. If, moreover,* f *satisfies the Lipschitz condition* (3.5) *with a matrix* K *whose spectral radius* $r(K)$ *satisfies inequality* (5.63)*, then all the solutions of the differential system* (5.1) *are even and periodic with period* T.

REMARK 5.51. Corollary 5.50 implies, in particular, Theorem 6 from [70, p. 326].

COROLLARY 5.52. *If the function* f *in Eq.* (5.1) *satisfies condition* (5.78) *with a non-singular matrix* E*, then the non-linear differential system* (5.1) *admits the trivial solution.*

PROOF. It is easy to see that, by virtue of (5.78), the function

$$(-\infty, \infty) \ni t \longmapsto f(t, 0)$$

has property $(\tau, -E)$. Proposition 5.14 implies that

$$E^2 f(t,0) = 0, \quad t \in (-\infty, \infty). \tag{5.96}$$

Since E^2 is non-singular, relation (5.96) yields $f(t,0) = 0$ for all $t \in (-\infty, \infty)$. Therefore, system (5.1) admits the trivial solution $x \equiv 0$. □

6. Two-point boundary value problem with non-linear non-separated conditions

In Samoilenko and Ronto [91], Samoilenko and Taĭ [92], the methodology of the successive approximation method is extended in order to make possible studying the non-linear two-point boundary problem of the form

$$y'(t) = f\bigl(t, y(t)\bigr), \quad t \in [0, T], \tag{6.1}$$

$$g\bigl(y(0), y(T)\bigr) = 0, \tag{6.2}$$

for which purpose a general non-linear change of variable was applied to Eq. (6.1). Here, we use a simpler substitution, which, as is shown, essentially facilitates the subsequent application of the successive approximation method. In particular, all the assumptions are formulated in terms of the original problem, and not the transformed one.

6.1. *Reduction to a problem with linear conditions*

Consider the boundary value problem (6.1), (6.2), where the functions $f : [0, T] \times G \to \mathbb{R}^n$ and $g : G \times G \to \mathbb{R}^n$ are continuous, $G \subset \mathbb{R}^n$ being the closure of a bounded domain. Assume that, for $t \in [0, T]$ fixed, f satisfies the Lipschitz condition

$$\bigl|f(t, y_1) - f(t, y_2)\bigr| \leqslant K|y_1 - y_2|, \quad \{y_1, y_2\} \subset G, \tag{6.3}$$

where the square matrix K is supposed to have non-negative components. In (6.3), the absolute value sign and inequality are understood component-wise.

Let us introduce the traditional substitution

$$y(t) = x(t) + w, \tag{6.4}$$

where $w \in \Omega \subset \mathbb{R}^n$ is an unknown parameter. The domain, Ω, of w is chosen so that $D + \Omega \subset G$, whereas the new variable, x, is supposed to have range in D, the closure of a bounded subdomain of G. Substitution (6.4) allows one to rewrite (6.1), (6.2) as

$$x'(t) = f\bigl(t, x(t) + w\bigr), \quad t \in [0, T], \tag{6.5}$$

$$g\bigl(x(0) + w, x(T) + w\bigr) = 0. \tag{6.6}$$

Let us rewrite (6.6) in the form

$$Ax(0) + Bx(T) + g\bigl(x(0) + w, x(T) + w\bigr) = Ax(0) + Bx(T), \tag{6.7}$$

where A and B are some square matrices of dimension n such that $\det B \neq 0$.

The parameter w is natural to be determined from the additional equation

$$Ax(0) + Bx(T) + g\bigl(x(0) + w, x(T) + w\bigr) = 0. \tag{6.8}$$

Due to (6.7), condition (6.8) obviously implies that

$$Ax(0) + Bx(T) = 0. \tag{6.9}$$

Thus, the essentially non-linear problem (6.1), (6.2) is transformed into the equivalent problem (6.5), (6.6), and (6.8).

On the other hand, system (6.5), (6.6), (6.8) can be regarded as a collection of problems (6.9) for the differential system

$$x'(t) = f\bigl(t, x(t) + w\bigr), \tag{6.10}$$

parametrised by the vector parameter w belonging to a certain given set Ω. The essential advantage obtained thereby is that the boundary condition (6.9) is linear. Therefore, problems of the family (6.10), (6.9) can be studied by using the successive approximation method developed in [90,91].

Assume that

$$D_\beta := \bigl\{x \in \mathbb{R}^n \mid B\bigl(x, \beta(x)\bigr) \subset D\bigr\} \neq \varnothing, \tag{6.11}$$

where

$$\beta(x) := \frac{T}{4}\delta_G(f) + \bigl|(B^{-1}A + \mathbb{1}_n)x\bigr|. \tag{6.12}$$

Here, by definition,

$$\delta_G(f) := \max_{(t,y)\in[0,T]\times G} f(t, y) - \min_{(t,y)\in[0,T]\times G} f(t, y). \tag{6.13}$$

Moreover, we suppose that

$$r(K) < \frac{10}{3T}. \tag{6.14}$$

Introduce the sequence of functions

$$\begin{aligned} x_{m+1}(t, w, z) := z &+ \int_0^t f\bigl(s, x_m(s, w, z) + w\bigr)\,\mathrm{d}s \\ &- \frac{t}{T}\int_0^T f\bigl(s, x_m(s, w, z) + w\bigr)\,\mathrm{d}s - \frac{t}{T}(B^{-1}A + \mathbb{1}_n)z, \end{aligned} \tag{6.15}$$

where $m = 0, 1, \ldots$ and $x_0(t, w, z) \equiv z$. It is easily seen that all the members of sequence (6.15) satisfy condition (6.9) for every $z \in D_\beta$ and $w \in \Omega$. Furthermore, $x_m(0, w, z) = z$ for all $m \in \mathbb{N}$.

By virtue of (6.9), a solution x of (6.10), (6.9) satisfies the condition

$$x(T) = -B^{-1}Ax(0)$$

and, therefore, Eq. (6.8) can be rewritten as

$$g\bigl(x(0) + w, -B^{-1}Ax(0) + w\bigr) = 0. \tag{6.16}$$

Summarising the above-said, we conclude that problem (6.10), (6.9), (6.8) is equivalent to system (6.10), (6.9), (6.16). We suggest to solve it sequentially: first solve (6.10), (6.9), and then try to determine whether (6.16) can be fulfilled.

THEOREM 6.1. *Assume conditions* (6.3), (6.11), *and* (6.14). *Then*:

(1) *Sequence* (6.15) *has a limit* $x^*(\cdot, w, z)$ *uniform in* $(t, w, z) \in [0, T] \times D \times D_\beta$;

(2) *The limit function* $x^*(\cdot, w, z)$ *is a solution of problem* (6.9) *for the "perturbed" differential system*

$$x'(t) = f\bigl(t, x(t) + w\bigr) + \Delta(w, z), \quad t \in [0, T],$$

where, by definition,

$$\Delta(w, z) := -\frac{1}{T}(B^{-1}A + \mathbb{1}_n) - \frac{1}{T}\int_0^T f\bigl(s, x^*(s, w, z)\bigr)\, \mathrm{d}s.$$

Furthermore, $x^*(0, w, z) = z$.

(3) *The error estimate*

$$\bigl|x^*(t, w, z) - x_m(t, w, z)\bigr| \leqslant \frac{20}{9}t\left(1 - \frac{t}{T}\right)Q^{m-1}(\mathbb{1}_n - Q)^{-1}\left[\frac{1}{2}\delta_G(f) + K\bigl|(B^{-1}A + \mathbb{1}_n)z\bigr|\right] \tag{6.17}$$

holds, where $Q := \frac{3T}{10}K$ *and* $\delta_G(f)$ *is as in* (6.13).

PROOF. It is carried out similarly to that of Theorem 1 from [68] and Theorem 2.1 from [91, p. 32]. □

The following statement shows the relation of the function $x^*(\cdot, w, z)$ to problem (6.10), (6.9).

THEOREM 6.2. *Under assumptions* (6.3), (6.11), *and* (6.14), *the function* $x^*(\cdot, w, z^*)$ *is a solution of the boundary value problem* (6.10), (6.9) *if, and only if* $z = z^*$ *satisfies the "determining equation,"*

$$\bigl(B^{-1}A + \mathbb{1}_n\bigr)z + \int_0^T f\bigl(s, x^*(s, w, z)\bigr)\,\mathrm{d}s = 0,$$

where w *is considered as a parameter.*

PROOF. Analogous to that of Theorem 2.4 from the book of Samoilenko and Ronto [91, p. 36]. ☐

THEOREM 6.3. *Assume conditions* (6.3), (6.11), *and* (6.14). *Then, for the function* $x^*(\cdot, w, z^*)$ *to be a solution of the boundary value problem* (6.1), (6.2), *it is sufficient that the parameters* $z = z^*$, $w = w^*$ *satisfy the system of determining equations*

$$(B^{-1}A + \mathbb{1}_n)z + \int_0^T f\bigl(s, x^*(s, w, z)\bigr)\,\mathrm{d}s = 0, \tag{6.18}$$

$$g(z + w, -B^{-1}Az + w) = 0. \tag{6.19}$$

In this case,

$$y^*(t) = x^*\bigl(t, w^*, z^*\bigr) + w^* \tag{6.20}$$

is a solution of the boundary value problem (6.1), (6.2).

PROOF. It is obvious from the form of substitution (6.4) that Eqs. (6.18) and (6.19) hold whenever the transformed boundary value problem (6.10), (6.9), (6.16) coincides with (6.1), (6.2). ☐

REMARK 6.4. In [92,91], a change of variable more general than (6.4) was used,

$$y(t) = x(t) + h(t, w). \tag{6.21}$$

However, when $\partial h(t, w)/\partial t \not\equiv 0$, substitution (6.21) leads one to a more complicated transformed boundary value problem, as well as more complicated recursion sequence and determining equations. Furthermore, the basic assumptions should then be made on the *transformed* differential equation.

REMARK 6.5. Considering (6.20), on can treat the function

$$y_m(t) = x_m(t, w_m, z_m) + w_m \tag{6.22}$$

as the mth approximation to the exact solution, $y^*(\cdot, w^*, z^*)$, of the boundary value problem (6.1), (6.2). In (6.22), w_m and z_m are solutions of the "approximate determining equations,"

$$(B^{-1}A + \mathbb{1}_n)z + \int_0^T f\bigl(s, x_m(s, w, z)\bigr)\,\mathrm{d}s = 0, \tag{6.23}$$

$$g(z + w, -B^{-1}Az + w) = 0, \tag{6.24}$$

and $x_m(\cdot, z, w)$ is given by (6.15). We do not consider the strict substantiation, referring the reader to [91] where similar techniques are described.

REMARK 6.6. It follows from the consideration in Section 7.2 that the convergence of the recursion sequence (6.15) can be proved under the condition $r(K) < 3.4161\ldots \cdot T^{-1}$, which is weaker than (6.14). However, estimate (6.17) does not hold in this case.

REMARK 6.7. By using an idea from [61], one can obtain similar statements based on replacing (6.15) by the sequence

$$\begin{aligned} x_{m+1}(t, w, z) := z - (A + B)z + \int_0^t f\bigl(s, x_m(s, w, z) + w\bigr)\,\mathrm{d}s \\ - \Phi(t)\int_0^T f\bigl(s, x_m(s, w, z) + w\bigr)\,\mathrm{d}s, \end{aligned}$$

where Φ is an arbitrary continuous matrix-valued function such that

$$A\Phi(0) + B\Phi(T) = \mathbb{1}_n.$$

In this case, the restrictive conditions of type (6.14) can be avoided. However, an analogue of condition (6.11) is more difficult to verify unless f in Eq. (6.1) is globally Lipschitzian.

EXAMPLE 6.8. Consider the differential equation

$$y''(t) + \frac{t}{16}y(t) + \frac{t}{4}\bigl(y'(t)\bigr)^2 = \frac{t^2}{128}, \quad t \in [0, 1], \tag{6.25}$$

under the non-linear boundary conditions

$$\begin{aligned} &y'(1)y'(0) + \frac{1}{4}y(0) = 0, \\ &y(1)y'(0) - \frac{1}{2}y'(1) = -\frac{7}{128}. \end{aligned} \tag{6.26}$$

We reduce Eq. (6.25) to the first order system

$$\begin{aligned} y_1'(t) &= y_2(t), \\ y_2'(t) &= \frac{t^2}{128} - \frac{t}{16} y_1(t) - \frac{t}{4} y_2^2(t); \end{aligned} \tag{6.27}$$

conditions (6.26) then rewrite as

$$\begin{aligned} y_2(1) y_2(0) + \frac{1}{4} y_1(0) &= 0, \\ y_1(1) y_2(0) - \frac{1}{2} y_2(1) &= -\frac{7}{128}. \end{aligned}$$

Put $A = B = I_2$ and $T = 1$. It is not difficult to verify that the assumptions of the theorems above hold in the domain $G := \{(y_1, y_2) \mid |y_1| \leqslant 1, |y_2| \leqslant \frac{1}{2}\}$.

Namely, relation (6.11) holds whenever β from (6.12) satisfies the component-wise inequality

$$\beta(x_1, x_2) \leqslant \begin{pmatrix} \frac{1}{4} \\ \frac{17}{256} \end{pmatrix} + 2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The vector function in the right-hand side of (6.27) satisfies (6.3) in G with $K = \left(\begin{smallmatrix} 0 & 1 \\ \frac{1}{16} & \frac{1}{4} \end{smallmatrix}\right)$. Since $r(K) = (1 + \sqrt{2})/4$, it is obvious that (6.14) holds.

Substitution (6.4) in (6.27) leads us to the system

$$\begin{aligned} x_1'(t) &= x_2(t) + w_2, \\ x_2'(t) &= \frac{t^2}{128} - \frac{t}{16}\big[x_1(t) + w_1\big] - \frac{t}{4}\big[x_2(t) + w_2\big]^2. \end{aligned}$$

In this case, the determining equation (6.24) has the form

$$\begin{aligned} w_2^2 - z_2^2 + \frac{1}{4}(z_1 + w_1) &= 0, \\ (w_1 - z_1)(w_2 + z_2) - \frac{1}{2}(w_2 - z_2) &= -\frac{7}{128}. \end{aligned} \tag{6.28}$$

Let us find the first approximation, $x_1 = \left(\begin{smallmatrix} x_{1,1} \\ x_{1,2} \end{smallmatrix}\right)$, in the sense of Remark 6.5. According to (6.15), we have

$$\begin{aligned} x_{1,1}(t, w_1, w_2, z_1, z_2) &= z_1 - 2t z_1, \\ x_{1,2}(t, w_1, w_2, z_1, z_2) &= z_2 + \frac{t^3}{384} - \frac{t^2}{32}(z_1 + w_1) - \frac{t^2}{8}(z_1 + w_2)^2 - \frac{t}{384} \\ &\quad + \frac{t}{32}(z_1 + w_1) + \frac{t}{8}(z_2 + w_2)^2 - 2t z_2. \end{aligned}$$

The first approximate determining equation (6.23) (i.e., that corresponding to $m = 1$), is equivalent to $\Delta_1(w, z) = 0$, where the function $\Delta_1 = \binom{\Delta_{1,1}}{\Delta_{1,2}}$ is given by the formula

$$\Delta_{1,1}(w_1, w_2, z_1, z_2) = \frac{385}{192} z_1 - \frac{1}{1536} + \frac{1}{192} w_1 + \frac{1}{48} (z_2 + w_2)^2 + w_2 \tag{6.29}$$

and

$$\begin{aligned}
\Delta_{1,2}(w_1, w_2, z_1, z_2) = {} & -\frac{1}{15360} z_1 z_2 w_2 + \frac{36863}{14155776} + \frac{107531}{10321920} z_1 \\
& + \frac{46079}{23040} z_2 - \frac{322549}{10321920} w_1 + \frac{1}{5760} w_2 \\
& - \frac{1}{15360} w_1 z_2 w_2 - \frac{1}{15360} {z_2}^4 - \frac{1}{15360} {w_2}^4 \\
& - \frac{107509}{2580480} {z_2}^2 + \frac{107531}{1290240} z_2 w_2 - \frac{322549}{2580480} {w_2}^2 \\
& - \frac{1}{245760} {z_1}^2 - \frac{1}{122880} z_1 w_1 - \frac{1}{30720} z_1 {z_2}^2 \\
& - \frac{1}{30720} z_1 {w_2}^2 - \frac{1}{245760} {w_1}^2 - \frac{1}{30720} w_1 {z_2}^2 \\
& - \frac{1}{30720} w_1 {w_2}^2 - \frac{1}{3840} {z_2}^3 w_2 - \frac{1}{2560} {z_2}^2 {w_2}^2 \\
& - \frac{1}{3840} z_2 {w_2}^3 + \frac{1}{3840} z_2 z_1 + \frac{1}{3840} z_2 w_1 \\
& - \frac{1}{320} {z_2}^2 w_2 - \frac{3}{320} z_2 {w_2}^2 - \frac{1}{768} w_2 z_1 \\
& - \frac{1}{768} w_2 w_1 + \frac{1}{960} {z_2}^3 - \frac{1}{192} {w_2}^3 .
\end{aligned} \tag{6.30}$$

Solving system (6.28) together with the approximate determining equations

$$\Delta_{1,1}(w_1, w_2, z_1, z_2) = 0, \qquad \Delta_{1,2}(w_1, w_2, z_1, z_2) = 0,$$

where $\Delta_{1,1}$ and $\Delta_{1,2}$ are given by (6.29) and (6.30), we obtain the following values for $z_1 = \binom{z_{1,1}}{z_{1,2}}$ and $w_1 = \binom{w_{1,1}}{w_{1,2}}$:

$$z_{1,1} \approx -0.06206890391, \qquad z_{1,2} \approx -0.00001967788609,$$
$$w_{1,1} \approx -0.0002202244703, \qquad w_{1,2} \approx 0.1247889518.$$

With these values of parameters, the first approximation, x_1, has the form

$$x_{1,1}(t) \approx -0.06228912838 + 0.1241378078t,$$

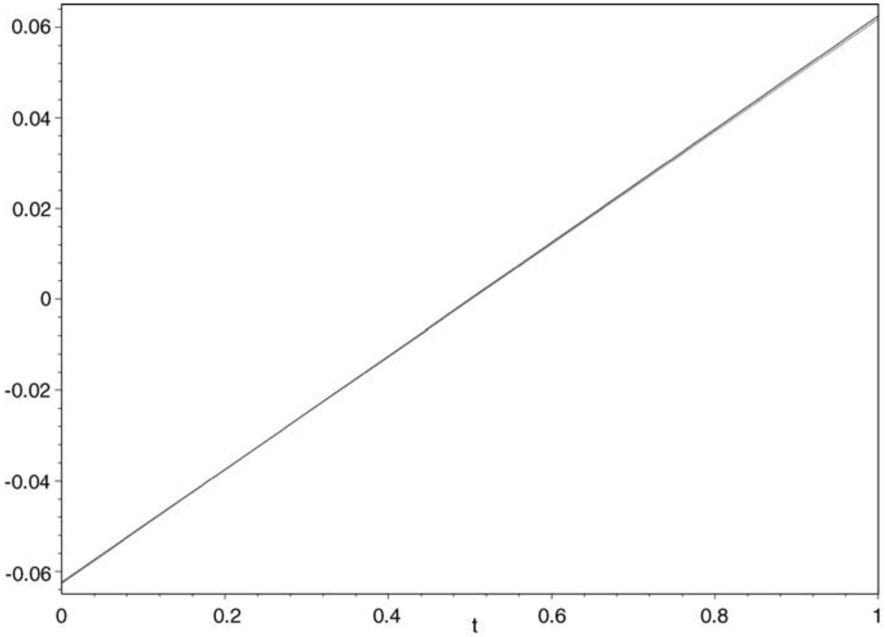


Fig. 4. The exact solution and the first approximation (Example 6.8, problem (6.25), (6.26)). The curves graphically almost coincide with one another.

$$
\begin{aligned}
x_{1,2}(t) \approx{}& -0.196778861 \times 10^{-4} + \frac{1}{384}t^3 + 0.613798 \times 10^{-6}t^2 \\
&- 0.2565424693 \times 10^{-2}t.
\end{aligned}
$$

Consequently, according to substitution (6.4), the solution, y, of problem (6.25), (6.26), up to the first approximation $y_1(t) = x_1(t) + w_1$, has the form

$$
\begin{aligned}
y_{1,1}(t) &\approx -0.06250935285 + 0.1241378078t, \\
y_{1,2}(t) &\approx 0.1247692739 + \frac{1}{384}t^3 + 0.613798 \times 10^{-6}t^2 \\
&\quad - 0.2565424693 \times 10^{-2}t.
\end{aligned}
$$

Note that the function $y^*(t) = \frac{t}{8} - \frac{1}{16}$ is an exact solution of problem (6.25), (6.26). As is seen from Figs. 4 and 5, the graphs of the exact solution and the first approximation almost coincide, whereas the deviation of their derivatives does not exceed 0.001. Thus, even the first iteration of the method gives a satisfactory approximation for the solution of the problem considered.

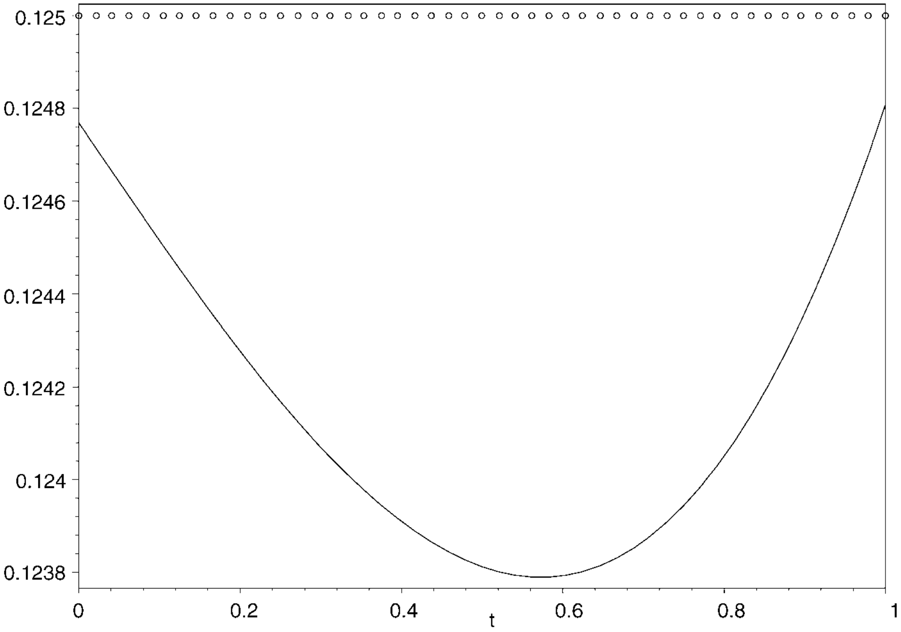


Fig. 5. The derivatives of the exact solution and the first approximation (Example 6.8, problem (6.25), (6.26)).

6.2. *A separated non-linear boundary condition*

It turns out that, in the case of separated, in a sense, boundary conditions, the successive approximation method can be applied without any change of variable.

Consider the two-point boundary value problem

$$x'(t) = f\big(t, x(t)\big), \quad t \in [0, T], \tag{6.31}$$

$$x(T) = a\big(x(0)\big), \tag{6.32}$$

where $a : D \to \mathbb{R}^n$ is continuous and f satisfies conditions (6.3) and (6.14) in the closure, D, of a bounded domain from $\mathbb{R}^n$.

Put $\gamma(x) := \frac{T}{4}\delta_D(f) + |a(x) - x|$, $x \in D$, and assume that

$$D_\gamma \neq \varnothing, \tag{6.33}$$

where the set D_γ is defined similarly to (6.11), whereas the number $\delta_D(f)$ is given by (6.13). Introduce the sequence of functions

$$\begin{aligned} x_{m+1}(t, z) := {} & z + \int_0^t f\big(s, x_m(s, z)\big)\,\mathrm{d}s \\ & - \frac{t}{T}\int_0^T f\big(s, x_m(s, z)\big)\,\mathrm{d}s + \frac{t}{T}\big[a(z) - z\big], \end{aligned} \tag{6.34}$$

where $m = 0, 1, \dots$ and $x_0(t, z) \equiv z \in D_\gamma$. Obviously, all the members of sequence (6.34) satisfy condition (6.32) for every $z \in D_\gamma$.

THEOREM 6.9. *Assume conditions* (6.3), (6.14), *and* (6.33). *Then*:

(1) *Sequence* (6.34) *converges uniformly in* $[0, T] \times D_\gamma$,

$$\lim_{n\to+\infty} \sup_{(t,z)\in[0,T]\times D_\gamma} \left\| x_n(t, z) - x^*(t, z)\right\| = 0.$$

1. *The function* $x^*(\cdot, z)$ *is a solution of problem* (6.32) *for the "perturbed" system*

$$x'(t) = f\big(t, x(t)\big) + \Delta(z), \quad t \in [0, T],$$

where

$$\Delta(z) := \frac{1}{T}\left(a(z) - z - \int_0^T f\big(s, x^*(s, z)\big)\, \mathrm{d}s\right).$$

Moreover, $x^*(\cdot, z)$ *satisfies* $x^*(0, z) = x_m(0, z) = z$ $(\forall m \in \mathbb{N})$.

(3) *The error estimate*

$$\begin{aligned} \left|x^*(t, w, z) - x_m(t, w, z)\right| \leqslant \frac{20}{9} t\left(1 - \frac{t}{T}\right) Q^{m-1}(\mathbb{1}_n - Q)^{-1} \\ \times \left[\frac{1}{2}\delta_D(f) + K\left|a(z) - z\right|\right] \end{aligned} \tag{6.35}$$

holds, where $Q := \frac{3T}{10}K$ *and* $\delta_D(f)$ *is defined by* (6.13).

PROOF. Analogous to that of Theorem 2.1 from [91, p. 32]. □

THEOREM 6.10. *Under conditions of Theorem* 6.9, *the limit function of sequence* (6.34) *is a solution of problem* (6.31), (6.32) *if, and only if the parameter* $z = z^*$ (*which stands for the initial value of the solution at* $t = 0$) *satisfies the determining equation*

$$a(z) - z = \int_0^T f\big(s, x^*(s, z)\big)\, \mathrm{d}s.$$

PROOF. Similar to that of Theorem 2.3 from [91, p. 36]. □

REMARK 6.11. If $a(z) \equiv z$ in (6.32), then (6.31), (6.32) is nothing but the T-periodic boundary value problem. In this case, sequence (6.34) is reduced to the well-known T-periodic successive approximations [89],

$$x_{m+1}(t, z) := z + \int_0^t f\big(s, x_m(s, z)\big)\, \mathrm{d}s - \frac{t}{T}\int_0^T f\big(s, x_m(s, z)\big)\, \mathrm{d}s, \quad t \in [0, T].$$

REMARK 6.12. Estimate (6.35) implies that the function $t \mapsto x_m(t, z_m)$ is natural to be taken as the mth approximation of the exact solution of problem (6.31), (6.32), when $z = z_m$ is a root of the mth approximate determining equation,

$$a(z) - z = \int_0^T f\big(s, x_m(s, z)\big)\,\mathrm{d}s. \tag{6.36}$$

EXAMPLE 6.13. Let us apply the technique described above to the problem

$$\begin{aligned} &x''(t) - \frac{t}{8}x'(t) + \frac{1}{2}\big(x'(t)\big)^2 = \frac{1}{4}, \quad t \in [0, 1],\\ &x(1) = \frac{1}{2}\big(x'(0)\big)^2 + \frac{1}{16},\\ &x'(1) = -x(0) + \frac{3}{16}. \end{aligned} \tag{6.37}$$

As usual, we rewrite the second order problem (6.37) in the form

$$\begin{aligned} &x_1'(t) = x_2,\\ &x_2'(t) = \frac{1}{4} + \frac{t}{8}x_2(t) - \frac{1}{2}x_2^2(t), \quad t \in [0, 1],\\ &x_1(1) = \frac{1}{2}x_2^2(0) + \frac{1}{16},\\ &x_2(1) = -x_1(0) + \frac{3}{16}. \end{aligned} \tag{6.38}$$

Consider (6.38) in the domain

$$(t, x_1, x_2) \in [0, 1] \times D, \quad D = \left\{(x_1, x_2) \;\middle|\; |x_1| \leqslant 1, |x_2| \leqslant \frac{3}{4}\right\}. \tag{6.39}$$

It is not difficult to verify that the conditions of Theorem 6.9 are fulfilled with

$$K = \begin{pmatrix} 0 & 1 \\ 0 & \frac{7}{8} \end{pmatrix}, \qquad \gamma(x) \leqslant \begin{pmatrix} \frac{7}{16} + \frac{1}{2}x_2^2 - x_1 \\ \frac{1}{2} - x_1 - x_2 \end{pmatrix}.$$

The corresponding quantity (6.13) is estimated as

$$\delta_D(f) \leqslant \begin{pmatrix} \frac{3}{4} \\ \frac{5}{8} \end{pmatrix},$$

and K satisfies (6.14) because $r(K) = \frac{7}{8}$.

According to (6.34), we have the following formulae for $x_1 = \binom{x_{1,1}}{x_{1,2}}$:

$$x_{1,1}(t,\zeta_1,\zeta_2) = \zeta_1 + t\left(\frac{1}{2}\zeta_2^2 - \zeta_1 + \frac{1}{16}\right), \tag{6.40}$$

$$x_{1,2}(t,\zeta_1,\zeta_2) = \zeta_2 + \frac{t^2}{16}\zeta_2 - \frac{1}{16}t\zeta_2 - t\left(\zeta_1 + \zeta_2 - \frac{3}{16}\right). \tag{6.41}$$

The first approximate determining equations (6.36) have the form

$$\Delta_{1,1}(\zeta_1,\zeta_2) = 0, \qquad \Delta_{1,2}(\zeta_1,\zeta_2) = 0, \tag{6.42}$$

where

$$\Delta_{1,1}(\zeta_1,\zeta_2) = \frac{1}{2}\zeta_2^2 - \frac{47}{96}\zeta_2 - \frac{1}{2}\zeta_1 - \frac{1}{32},$$
$$\Delta_{1,2}(\zeta_1,\zeta_2) = -\frac{49}{48}\zeta_1 - \frac{33}{512} - \frac{88165}{89088}\zeta_2 + \frac{827}{5120}\zeta_2{}^2 - \frac{31}{192}\zeta_2\zeta_1 + \frac{1}{6}\zeta_1{}^2.$$

The solution, $\zeta = \binom{\zeta_1}{\zeta_2}$, of system (6.42) has the form

$$\zeta_1 = -\frac{1}{16}, \qquad \zeta_2 = 0. \tag{6.43}$$

Inserting (6.43) into (6.40) and (6.41), we obtain the first approximation

$$x_{1,1}(t) = \frac{t}{8} - \frac{1}{16}, \qquad x_{1,2}(t) = \frac{t}{4}. \tag{6.44}$$

Note that the vector function with components

$$x_1^*(t) = \frac{t^2}{8} - \frac{1}{16}, \qquad x_2^*(t) = \frac{t}{4} \tag{6.45}$$

is an exact solution of problem (6.38).

Let us construct the second approximation. By (6.34), we have

$$x_{2,1}(t,\zeta_1,\zeta_2) = \zeta_1 + \frac{1}{48}\zeta_2 t^3 - \frac{1}{2}t^2\zeta_1 + \frac{3}{32}t^2 - \frac{17}{32}t^2\zeta_2$$
$$+ \frac{49}{96}\zeta_2 t - \frac{1}{2}t\zeta_1 - \frac{1}{32}t + \frac{1}{2}t\zeta_2^2, \tag{6.46}$$

$$
\begin{aligned}
x_{2,2}(t,\zeta_1,\zeta_2) = {} & \frac{95}{512}t - \frac{31}{192}t\zeta_2\zeta_1 + z_2 + \frac{17}{768}\zeta_2 t^3 - \frac{1}{32}t^2\zeta_2 - \frac{49}{48}t\zeta_1 \\
& - \frac{1733}{5120}t\zeta_2^2 + \frac{1}{512}t^3 - \frac{3041}{3072}\zeta_2 t - \frac{1}{2560}\zeta_2^2 t^5 + \frac{1}{64}t^4\zeta_2\zeta_1 \\
& - \frac{17}{48}t^3\zeta_2\zeta_1 + \frac{1}{2}t^2\zeta_2\zeta_1 - \frac{1}{1024}t^4\zeta_2 + \frac{17}{1024}t^4\zeta_2^2 + \frac{1}{48}t^3\zeta_1 \\
& - \frac{107}{512}t^3\zeta_2^2 - \frac{1}{6}t^3\zeta_1^2 + \frac{17}{32}t^2\zeta_2^2 + \frac{1}{6}t\zeta_1^2.
\end{aligned}
\tag{6.47}
$$

It is not difficult to verify that inserting into (6.46) and (6.47) even the roots (6.43) of the *first* determining system (6.42), we arrive at the exact solution (6.45) of problem (6.38).

7. A non-linear problem with another type of separated boundary conditions

Here, we suggest a constructive scheme for the investigation of non-linear differential systems (3.3) with non-linear two-point boundary conditions of the form (2.4).

The right-hand side of the equation is supposed to be Lipschitzian in a suitable bounded domain, Ω, and the spectrum of the corresponding Lipschitz matrix is assumed to be contained inside the circle of radius less than $r \approx 3.4161/T$. Under the latter assumption, the convergence of the method is proved by using some techniques developed in [70,64] and their appropriate modifications.

It is worth mentioning that we impose the only technical restriction (7.21) on the continuous mapping $b\colon \mathbb{R}^n \to \mathbb{R}^n$, and this condition is automatically satisfied whenever the domain Ω is large enough.

7.1. *Subsidiary lemmata*

The estimates related to the successive approximation method under consideration essentially use properties of the sequence of functions $\{\alpha_{k,T}\colon\ k \geqslant 0\} \subset C([0,T])$ given by the recurrence formula (3.35):

$$
\alpha_{k+1,T}(t) = \left(1 - \frac{t}{T}\right)\int_0^t \alpha_{k,T}(s)\,\mathrm{d}s + \frac{t}{T}\int_t^T \alpha_{k,T}(s)\,\mathrm{d}s, \qquad k = 0,1,\ldots,
\tag{7.1}
$$

and $\alpha_{0,T} \equiv 1$. Let us obtain some further properties of this sequence.

LEMMA 7.1.

(1) *The functions of sequence* (7.1) *are symmetric with respect to the point $T/2$, i.e., for all $k \geqslant 0$,*

$$
\alpha_{k,T}(t) = \alpha_{k,T}(T - t), \qquad t \in [0,T];
\tag{7.2}
$$

$$
\alpha_{k,T}(T/2 - t) = \alpha_{k,T}(T/2 + t), \qquad t \in [0,T/2];
\tag{7.3}
$$

(2) *Sequence* (7.1) *can be represented alternatively as*

$$\alpha_{k+1,T}(t) = \int_0^t \alpha_{k,T}(s)\,\mathrm{d}s + \frac{t}{T}\int_t^{T-t} \alpha_{k,T}(s)\,\mathrm{d}s \tag{7.4}$$

$$= \frac{t}{T}\int_0^{T-t} \alpha_{k,T}(s)\,\mathrm{d}s + \bigl(1-\frac{t}{T}\bigr)\int_{T-t}^{T} \alpha_{k,T}(s)\,\mathrm{d}s \tag{7.5}$$

$$= \int_0^{T-t} \alpha_{k,T}(s)\,\mathrm{d}s + \left(1-\frac{t}{T}\right)\int_{T-t}^{t} \alpha_{k,T}(s)\,\mathrm{d}s; \tag{7.6}$$

(3) *The relations*

$$\alpha_{k,T}(0) = \alpha_{k,T}(T) = 0; \qquad \alpha_{k,T}(t) > 0 \quad \textit{for all } t \in (0,T),$$

hold for every $k \geqslant 0$;

(4) *For every* $k \geqslant 1$, *the maximal value of* $\alpha_{k,T}(t)$ *is achieved at* $t = T/2$, *namely,*

$$\max_{t\in[0,T]} \alpha_{k,T}(t) = \alpha_{k,T}(T/2) = T^k \alpha_{k,1}(1/2). \tag{7.7}$$

PROOF. *Assertion* 1. Obviously, (7.2) holds for $k = 0$. Let us assume that (7.2) is true for some k and prove that this relation remains valid when k is replaced by $k+1$. Indeed, we have

$$\int_0^{T-t} \alpha_{k,T}(s)\,\mathrm{d}s = \int_0^{T-t} \alpha_{k,T}(T-s)\,\mathrm{d}s = \int_t^{T} \alpha_{k,T}(s)\,\mathrm{d}s \tag{7.8}$$

and, by (7.8),

$$\int_{T-t}^{T} \alpha_{k,T}(s)\,\mathrm{d}s = \int_0^{T} \alpha_{k,T}(s)\,\mathrm{d}s - \int_0^{T-t} \alpha_{k,T}(s)\,\mathrm{d}s = \int_0^{t} \alpha_{k,T}(s)\,\mathrm{d}s. \tag{7.9}$$

Considering (7.1), (7.8), and (7.9), we obtain

$$\alpha_{k,T}(T-t) = \left(1-\frac{T-t}{T}\right)\int_0^{T-t} \alpha_{k,T}(s)\,\mathrm{d}s + \frac{T-t}{T}\int_{T-t}^{T} \alpha_{k,T}(s)\,\mathrm{d}s$$
$$= \left(1-\frac{t}{T}\right)\int_0^{t} \alpha_{k,T}(s)\,\mathrm{d}s + \frac{t}{T}\int_t^{T} \alpha_{k,T}(s)\,\mathrm{d}s = \alpha_{k,T}(t),$$

i.e., (7.2) holds. Equality (7.3) is an immediate consequence of (7.2) with t replaced by $t + T/2$.

Assertion 2. By virtue of (7.1) and (7.9), we have

$$\alpha_{k+1,T}(t) = \int_0^t \alpha_{k,T}(s)\,\mathrm{d}s - \frac{t}{T}\int_0^t \alpha_{k,T}(s)\,\mathrm{d}s$$

$$+\frac{t}{T}\int_t^{T-t}\alpha_{k,T}(s)\,\mathrm{d}s+\frac{t}{T}\int_{T-t}^{T}\alpha_{k,T}(s)\,\mathrm{d}s$$
$$=\int_0^t\alpha_{k,T}(s)\,\mathrm{d}s+\frac{t}{T}\int_t^{T-t}\alpha_{k,T}(s)\,\mathrm{d}s,$$

i.e. (7.4) holds. To establish relation (7.5), we note that (7.5) with t replaced by $T-t$ can be rewritten as

$$\alpha_{k+1,T}(t)=\left(1-\frac{t}{T}\right)\int_0^t\alpha_{k,T}(s)\,\mathrm{d}s+\frac{t}{T}\int_t^{T}\alpha_{k,T}(s)\,\mathrm{d}s,$$

which coincides with the left-hand side of equality (7.1), i.e., with the $(k+1)$th member of the sequence defined by formula (7.1). Taking assertion 1 into account and using the fact that the change of variable $t\mapsto T-t$ is invertible, we show that representation (7.5) holds. Finally, (7.6) follows immediately when, in (7.4), we replace t by $T-t$ and take (7.2) into account.

Assertion 3. The equalities $\alpha_{k,T}(0)=\alpha_{k,T}(T)=0$ $(k=0,1,\ldots)$ follow immediately from (7.1) and (7.2). Arguing by induction and using representation (7.4) of sequence (7.1), we show that all the functions $\alpha_0,\alpha_1,\ldots$ take positive values inside the interval $(0,T)$.

Assertion 4. Differentiating (7.4) with respect to t, we obtain

$$\alpha'_{k,T}(t)=\left(1-\frac{2}{T}t\right)\alpha_{k-1,T}(t)+\frac{1}{T}\int_t^{T-t}\alpha_{k-1,T}(s)\,\mathrm{d}s, \tag{7.10}$$

whence it follows by induction that, for every k, the function $\alpha_{k,T}$ is increasing on $(0,T/2)$ and, by (7.2), is decreasing on $(T/2,T)$. Since $\alpha'_{k,T}(T/2)=0$, we see that $T/2$ is the global maximum point of $\alpha_{k,T}$, and assertion 4 is thus proved.

It is obvious that, in order to prove the second part of equality (7.7), it will suffice to establish the following relation:

$$\alpha_{m,T}(\sigma T)=T^m\alpha_{m,1}(\sigma)\quad\text{for all }\sigma\in[0,1]. \tag{7.11}$$

By (7.1), we have

$$\alpha_{m,T}(\sigma T)=(1-\sigma)\int_0^{\sigma T}\alpha_{m-1,T}(u)\,\mathrm{d}s+\sigma\int_{\sigma T}^{T}\alpha_{m-1,T}(u)\,\mathrm{d}u$$
$$=T(1-\sigma)\int_0^{\sigma}\alpha_{m-1,T}(u)\,\mathrm{d}s+T\sigma\int_{\sigma}^{1}\alpha_{m-1,T}(u)\,\mathrm{d}u$$

for all $\sigma\in[0,1]$. Introducing the mappings $A:C([0,1],\mathbb{R}^n)\to C([0,1],\mathbb{R}^n)$ and $\mathcal{S}_T:C([0,1],\mathbb{R}^n)\to C([0,T],\mathbb{R}^n)$ by the formulae

$$(Ax)(\sigma):=(1-\sigma)\int_0^{\sigma}x(u)\,\mathrm{d}u+\sigma\int_{\sigma}^{1}x(u)\,\mathrm{d}u,\quad \sigma\in[0,1], \tag{7.12}$$

and

$$(\mathcal{S}_T x)(\sigma) := x(\sigma T), \quad \sigma \in [0,1],$$

we arrive at the relations

$$\mathcal{S}_T \alpha_{m,T} = A \mathcal{S}_T \alpha_{m-1,T}, \quad m = 1, 2, \ldots,$$

whence it follows that

$$\mathcal{S}_T \alpha_{m,T} = T^m \cdot A^m \mathcal{S}_T \alpha_0 \equiv T^m \cdot A^m \mathcal{S}_T 1 \equiv T^m \cdot A^m 1. \tag{7.13}$$

On the other hand, equality (7.1) with $T = 1$ yields

$$\alpha_{m,1}(\sigma) = (1-\sigma) \int_0^\sigma \alpha_{m-1,1}(u)\,\mathrm{d}u + \sigma \int_\sigma^1 \alpha_{m-1,1}(u)\,\mathrm{d}u, \quad \sigma \in [0,1],$$

or, which is the same, $\alpha_{m,1} = A\alpha_{m-1,1}$, $m = 1, 2, \ldots$, whence

$$\alpha_{m,1} = A^m \alpha_{0,1} \equiv A^m 1, \quad m = 1, 2, \ldots. \tag{7.14}$$

Combining (7.13) and (7.14), we arrive at (7.11). The second part of (7.7) is an immediate consequence of relation (7.11). The lemma is proved. □

REMARK 7.2. Assertion 4 of Lemma 7.1 can also be derived from Corollary 10 of Rontó, Rontó, Samoilenko, and Trofimchuk [64]. Properties 1, 3, and 4 of sequence (7.1), established by Lemma 7.1, are also given by Samoilenko [85] for $T = 1$.

LEMMA 7.3. *The relations*

$$\lim_{m\to+\infty} \max_{t\in[0,T]} \frac{\alpha_{m,T}(t)}{\alpha_{m+1,T}(t)} = \lim_{m\to+\infty} \frac{\max_{t\in[0,T]} \alpha_{m,T}(t)}{\max_{t\in[0,T]} \alpha_{m+1,T}(t)}$$
$$= \lim_{m\to+\infty} \frac{\alpha_{m,T}(T/2)}{\alpha_{m+1,T}(T/2)} = \frac{1}{Tr(A)}$$

hold, where $r(A)$ is the spectral radius of operator (7.12).

PROOF. In view of (7.11), for all $m = 0, 1, 2, \ldots$,

$$\frac{\alpha_{m,T}(T/2)}{\alpha_{m+1,T}(T/2)} = \frac{\alpha_{m,1}(1/2)}{T\alpha_{m+1,1}(1/2)}.$$

By [64, Lemma 10], one has $\lim_{m\to+\infty} \sqrt[m]{\alpha_{m,1}(1/2)} = r(A)$, whereas [64, Corollary 12] implies that $\lim_{m\to+\infty} \sqrt[m]{\alpha_{m,1}(\sigma)\,/\,\alpha_{1,1}(\sigma)} = r(A)$ uniformly in $\sigma \in [0,1]$. Since

$\lim_{m\to\infty} c_m = c$ implies $\lim_{m\to\infty} \sqrt[m]{c_m} = \lim_{m\to\infty} c_{m+1}c_m^{-1} = c$, we arrive at the desired equality. □

REMARK 7.4. The value of the spectral radius $r(A)$ can be found numerically (see, e.g., [64,85]). Note that $r(A) \approx 0.292728853011 \approx 1\,/\,3.416130626393$.

Lemma 7.5 below extends Lemma 3.16.

LEMMA 7.5. *For the sequence of functions defined by relation* (7.1),[4] *given an arbitrarily small* $\varepsilon \in (0, +\infty)$, *one can specify an* $m_\varepsilon \in \mathbb{N}$ *such that*

$$\alpha_{m+1,T}(t) \leqslant T\bigl(r(A) + \varepsilon\bigr)\alpha_{m,T}(t), \quad m \geqslant m_\varepsilon, \tag{7.15}$$

and, consequently,

$$\alpha_{m+1,T}(t) \leqslant T^{m-m_\varepsilon+1}\bigl(r(A) + \varepsilon\bigr)^{m-m_\varepsilon+1}\alpha_{m_\varepsilon,T}(t), \quad m \geqslant m_\varepsilon. \tag{7.16}$$

PROOF. Estimate (7.15) is easily proved with the help of Lemma 7.3, whereas (7.16) is obtained immediately from (7.15) by the induction argument. □

The latter lemma will be used to establish error estimates for the successive approximations to be constructed below. However, when obtaining the error estimates on practice, one may want the above logic somewhat reversed. Namely, having the mth iteration (whose error, as will be shown below, is estimated with the help of the function $\alpha_{m-1,T}$), specify a (possibly negative) number ε_m such that the estimate $\alpha_{m+1,T}(t) \leqslant T(r(A) + \varepsilon_m)\alpha_{m,T}(t)$ holds for all $t \in [0, T]$. This aim can obviously be achieved by putting $\varepsilon_m := \alpha_{m+1,1}(1/2)/\alpha_{m,1}(1/2) - r(A)$, a few values of ε_m being shown in Table 1.

REMARK 7.6. Lemma 7.5 improves the result of Lemma 4 from Rontó and Mészáros [69] because the former, together with the table below, shows, in particular, that

$$\alpha_{m+1}(t) \leqslant T\bigl(r(A) + \frac{1}{1000}\bigr)\alpha_{m,T}(t)$$

for all $t \in [0, T]$ and $m \geqslant 3$, whereas $r(A) + \frac{1}{1000} \approx 0.29372885 < \frac{3}{10}$.

We also need a generalisation of Lemma 3.13.

LEMMA 7.7. *For every continuous* $x : [0, T] \to \mathbb{R}$, *the inequality*

$$\begin{aligned}&\left|\int_0^{\theta(t)} \left(x(\tau) - \frac{1}{T}\int_0^T x(s)\,\mathrm{d}s\right)\mathrm{d}\tau\right| \\ &\quad \leqslant \frac{1}{2}\alpha_{1,T}(t)\Bigl[\max_{\tau\in[0,T]} x(\tau) - \min_{\tau\in[0,T]} x(\tau)\Bigr]\end{aligned} \tag{7.17}$$

---

[4] Or, which is the same, by (7.4), (7.5), or (7.6)

Table 1

m	approximate ε_m	$r(A)+\varepsilon_m$
1	$0.406044803222 \times 10^{-1}$	$\frac{1}{3}$
2	$0.727114698891 \times 10^{-2}$	$\frac{3}{10}$
3	$0.921940639708 \times 10^{-3}$	$\frac{37}{126} \approx 0.293650793650$
4	$0.639397817078 \times 10^{-4}$	$\frac{65}{222} \approx 0.292792792792$
5	$-0.158028381223 \times 10^{-5}$	$\frac{161}{550} \approx 0.292727272727$
6	$-0.956604258782 \times 10^{-6}$	$\frac{6827}{23322} \approx 0.292727896407$
7	$-0.101923715775 \times 10^{-6}$	$\frac{195849}{669046} \approx 0.292728751087$
8	$-0.981324932500 \times 10^{-9}$	$\frac{3248737}{11098110} \approx 0.292728852030$
9	$0.114183386672 \times 10^{-8}$	$\frac{975725035}{3333204162} \approx 0.292728854153$
10	$0.153987313839 \times 10^{-9}$	$\frac{571245741}{1951450070} \approx 0.292728853165$
11	$0.488928250983 \times 10^{-11}$	$\frac{2961468157897}{10116762073110} \approx .292728853016$
12	$-0.130772159250 \times 10^{-11}$	$\frac{5201443062517}{17768808947382} \approx 0.292728853010$
$\cdots$	$\cdots$	$\cdots$
$+\infty$	0	$r(A) \approx 0.292728853011$

holds for all $t \in [0, T]$, *where* $\theta : [0, T] \to [0, T]$ *is defined either with* $\theta(t) := t$ *or* $\theta(t) := T - t$.

PROOF. When $\theta(t) \equiv t$, the statement constitutes inequality (4) of Lemma 3 from [69]. When $\theta(t) \equiv T - t$, by the above-said, we have

$$\begin{aligned}&\left|\int_0^{T-t}\left(x(\tau) - \frac{1}{T}\int_0^T x(s)\,\mathrm{d}s\right)\mathrm{d}\tau\right| \\ &\quad \leqslant \frac{1}{2}\alpha_{1,T}(T-t)\Bigl[\max_{\tau\in[0,T]} x(\tau) - \min_{\tau\in[0,T]} x(\tau)\Bigr].\end{aligned}$$

Since, in view of Lemma 7.1 (assertion 1), we have $\alpha_{m,T}(t) = \alpha_{m,T}(T-t)$ for all $m \geqslant 0$ and $t \in [0, T]$, the latter inequality yields (7.17). □

7.2. *Successive approximations for a problem with separated boundary conditions*

We consider the two-point boundary value problem

$$x'(t) = f\bigl(t, x(t)\bigr), \quad t \in [0, T], \tag{7.18}$$

$$x(0) = b\bigl(x(T)\bigr), \tag{7.19}$$

where $f : [0, T] \times \Omega \to \mathbb{R}^n$ and $b : \Omega \to \mathbb{R}^n$ are continuous and Ω is the closure of a bounded domain. We suppose that, for every $t \in [0, T]$ fixed, f satisfies the componentwise Lipschitz condition

$$\bigl|f(t, y_1) - f(t, y_2)\bigr| \leqslant K|y_1 - y_2| \text{ for all } \{y_1, y_2\} \subset \Omega, \tag{7.20}$$

where K is a square matrix of dimension n having non-negative components. We recall that, in (7.20) and similar relations below, the signs $|\cdot|$ and $\leqslant$ are understood componentwise.

Assume that

$$\Omega(\beta) := \bigl\{x \in \mathbb{R}^n \mid H_b(x, \beta) \subset \Omega\bigr\} \neq \varnothing, \tag{7.21}$$

where

$$H_b(x, \beta) := \bigl\{u \in \mathbb{R}^n \mid \max\bigl(|u - x|, |u - b(x)|\bigr) \leqslant \beta\bigr\} \tag{7.22}$$

with $\beta := \frac{T}{4}\delta_\Omega(f)$,

$$\delta_\Omega(f) := \max_{(t,x)\in[0,T]\times\Omega} f(t, x) - \min_{(t,x)\in[0,T]\times\Omega} f(t, x), \tag{7.23}$$

and

$$r(K) < \frac{1}{Tr(A)}. \tag{7.24}$$

Recall that A is the linear operator on the space $C([0, T], \mathbb{R}^n)$ defined by formula (7.12), and $1/r(A) \approx 3.4161306$.

In connection with the study of problem (7.18), (7.19), we introduce the sequence of functions

$$\begin{aligned} x_{m+1}(t, z) := {} & \frac{t}{T}z + \left(1 - \frac{t}{T}\right)b(z) + \int_0^t f\bigl(s, x_m(s, z)\bigr)\,\mathrm{d}s \\ & - \frac{t}{T}\int_0^T f\bigl(s, x_m(s, z)\bigr)\,\mathrm{d}s, \quad m = 0, 1, 2, \ldots, \end{aligned} \tag{7.25}$$

or, which is the same,

$$\begin{aligned} x_{m+1}(t, z) := {} & \frac{t}{T}z + \left(1 - \frac{t}{T}\right)b(z) - \int_t^T f\bigl(s, x_m(s, z)\bigr)\,\mathrm{d}s \\ & + \left(1 - \frac{t}{T}\right)\int_0^T f\bigl(s, x_m(s, z)\bigr)\,\mathrm{d}s, \quad m = 0, 1, 2, \ldots, \end{aligned}$$

where

$$x_0(t,z) := \frac{t}{T}z + \left(1 - \frac{t}{T}\right)b(z) \quad \text{for all } t \in [0,T], \tag{7.26}$$

and $z = \operatorname{col}(z_1, z_2, \dots, z_n) \in \Omega(\beta)$ is a parameter.

It can be easily verified that, for every $m \in \mathbb{N} \cup \{0\}$,

$$x_m(T,z) = z \quad \text{and} \quad x_m(0,z) = b(z), \tag{7.27}$$

i.e., functions (7.25) satisfy the boundary condition (7.19) for all $z \in \Omega(\beta)$.

THEOREM 7.8. *Assume that conditions* (7.20), (7.21), *and* (7.24) *hold for the boundary value problem* (7.18), (7.19). *Then*:

(1) *Sequence* (7.25) *has a limit $x^*(\cdot,z)$ uniform in $(t,z) \in [0,T] \times \Omega(\beta)$.*

(2) *The limit function $x^*(\cdot,z)$ is a solution of the "perturbed" problem*

$$\begin{aligned} &x'(t) = f\big(t, x, (t)\big) + \frac{1}{T}\Delta(z), \quad t \in [0,T], \\ &x(0) = b\big(x(T)\big) \end{aligned} \tag{7.28}$$

such that $x^*(T,z) = z$, where

$$\Delta(z) := z - b(z) - \int_0^T f\big(s, x^*(s,z)\big)\,\mathrm{d}s. \tag{7.29}$$

(3) *Given an arbitrarily small positive ε, one can choose a number $m_\varepsilon \in \mathbb{N}$ such that, for all $m \geqslant m_\varepsilon$, the following error estimate holds*:

$$\big|x_m(t,z) - x^*(t,z)\big| \leqslant \frac{1}{2}\alpha_{m_\varepsilon,T}(t)Q_\varepsilon^{m-m_\varepsilon}(\mathbb{1}_n - Q_\varepsilon)^{-1}Q_\varepsilon\delta_\Omega(f), \tag{7.30}$$

where $\delta_\Omega(f)$ is given by the component-wise relation (7.23) *and*

$$Q_\varepsilon := T\big(r(A) + \varepsilon\big) \cdot K. \tag{7.31}$$

REMARK 7.9. The value of $\alpha_{m_\varepsilon,T}(t)$ in (7.30) can be estimated from above by using formula (7.7).

PROOF OF THEOREM 7.8. We shall prove that, under the conditions assumed, (7.25) is a Cauchy sequence in the Banach space $C([0,T],\mathbb{R}^n)$ equipped with the usual uniform norm. First, we show that $x_m(t,z) \in \Omega$ for all $(t,z) \in [0,T] \times \Omega(\beta)$ and $m \in \mathbb{N}$.

Indeed, in view of Lemma 7.7, we have

$$x_1(t,z) - \frac{t}{T}z - \left(1 - \frac{t}{T}\right)b(z) = \int_0^t \left[f(s,z) - \int_0^T f(\tau,z)\,\mathrm{d}\tau\right]\mathrm{d}s$$

and, hence,

$$\begin{aligned}\bigl|x_1(t,z) - tT^{-1}z - \bigl(1 - tT^{-1}\bigr)b(z)\bigr| &\leqslant \frac{1}{2}\alpha_{1,T}(t)\delta_\Omega(f)\\ &\leqslant \frac{T}{4}\delta_\Omega(f) \equiv \beta. \qquad (7.32)\end{aligned}$$

It is easy to see from (7.21) and (7.22) that $z \in \Omega(\beta)$ if, and only if, for every $u \in \mathbb{R}^n$, the relation $\max_{t\in[0,T]} |tT^{-1}z + (1 - tT^{-1})b(z) - u| \leqslant \beta$ implies $u \in \Omega$. Therefore, by virtue of (7.32), we conclude that $x_1(t,z) \in \Omega$ whenever $(t,z) \in [0,T] \in \Omega(\beta)$. The fact that $x_m(t,z) \in \Omega$ for all $(t,z) \in [0,T] \in \Omega(\beta)$ and $m \in \mathbb{N}$ is proved quite similarly.

Now, consider the difference

$$\begin{aligned}x_{m+1}(t,z) - x_m(t,z) = &\int_0^t \bigl[f\bigl(s, x_m(s,z)\bigr) - f\bigl(s, x_{m-1}(s,z)\bigr)\bigr]\,\mathrm{d}s\\ &- \frac{t}{T}\int_0^T \bigl[f\bigl(s, x_m(s,z)\bigr) - f\bigl(s, x_{m-1}(s,z)\bigr)\bigr]\,\mathrm{d}s \qquad (7.33)\end{aligned}$$

and introduce the notation

$$d_m(t,z) := \bigl|x_m(t,z) - x_{m-1}(t,z)\bigr|, \quad m = 1, 2, \ldots. \qquad (7.34)$$

By virtue of identity (7.33) and the Lipschitz condition (7.20), we have

$$d_{m+1}(t,z) \leqslant K\left[\left(1 - \frac{t}{T}\right)\int_0^t d_m(s,z)\,\mathrm{d}s + \frac{t}{T}\int_t^T d_m(s,z)\,\mathrm{d}s\right] \qquad (7.35)$$

for every $m \geqslant 0$. According to (7.32),

$$d_1(t,z) \equiv \bigl|x_1(t,z) - tT^{-1}z - \bigl(1 - tT^{-1}\bigr)b(z)\bigr| \leqslant \frac{1}{2}\alpha_{1,T}(t)\delta_\Omega(f). \qquad (7.36)$$

Iterating (7.35) and taking into account (7.1) and (7.36), we obtain that, for all $m \geqslant 0$, $t \in [0,T]$, and $z \in \Omega_\beta$,

$$d_{m+1}(t,z) \leqslant \frac{1}{2}K^m\alpha_{m+1,T}(t)\delta_\Omega(f). \qquad (7.37)$$

By Lemma 7.5, for an arbitrary $\varepsilon \in (0,+\infty)$, estimate (7.16) holds for the function α_{m+1} [see (7.1)] whenever m is not less than some m_ε. Fixing the number ε and combining (7.37) and (7.16), we obtain

$$d_{m+1}(t,z) \leqslant \frac{1}{2}\alpha_{m_\varepsilon,T}(t)Q_\varepsilon^{m-m_\varepsilon+1}\delta_\Omega(f) \qquad (7.38)$$

for all $t \in [0, T]$, $z \in \Omega(\beta)$, and $m \geqslant m_\varepsilon$. Therefore,

$$\begin{aligned}\left|x_{m+j}(t,z) - x_m(t,z)\right| &\leqslant \sum_{i=0}^{j-1} \left|x_{m+j-i}(t,z) - x_{m+j-i-1}(t,z)\right| \\ &= \sum_{i=1}^{j} d_{m+i}(t,z) \quad \text{by (7.38)} \\ &\leqslant \frac{1}{2}\alpha_{m_\varepsilon,T}(t) \sum_{i=0}^{j-1} Q_\varepsilon^{m-m_\varepsilon+i+1} \delta_\Omega(f) \\ &= \frac{1}{2}\alpha_{m_\varepsilon,T}(t) Q_\varepsilon^{m-m_\varepsilon} Q_\varepsilon \sum_{i=0}^{j-1} Q_\varepsilon^i \delta_\Omega(f). \qquad (7.39)\end{aligned}$$

Since ε in (7.31) can be chosen arbitrarily small, relation (7.39) together with assumption (7.24) and definition (7.31) imply $\sum_{k=0}^{+\infty} Q_\varepsilon^k = (\mathbb{1}_n - Q_\varepsilon)^{-1}$ and $\lim_{m\to\infty} Q_\varepsilon^m = 0$. This means that (7.25) is a Cauchy sequence and, hence, assertion 7.8 holds.

Passing to the limit as $m \to +\infty$ in equality (7.25), we show that the function $x^*(\cdot, z) := \lim_{m\to+\infty} x_m(\cdot, z)$ satisfies the following integral equation:

$$\begin{aligned}x(t) = \frac{t}{T}z + \left(1 - \frac{t}{T}\right)b(z) + \int_0^t f\big(\tau, x(\tau)\big)\,\mathrm{d}\tau \\ - \frac{t}{T}\int_0^T f\big(\tau, x(\tau)\big)\,\mathrm{d}\tau. \qquad (7.40)\end{aligned}$$

It follows from (7.40) and (7.27) that $x^*(\cdot, z)$ satisfies the two-point boundary condition (7.19). It is also obvious from (7.40) that $x^*(T, z) = z$, which means that $x^*(\cdot, z)$ is a solution of problem (7.28), i.e., assertion 7.8 holds. Estimate (7.30) of assertion 7.8 is an immediate consequence of (7.39). ☐

REMARK 7.10. In a number of earlier works in this area (see, e.g., [70] for a survey), conditions of the form $r(K) < cT^{-1}$ with $2 < c < 1/r(A)$ were assumed instead of (7.24).

REMARK 7.11. Depending on the form of a particular boundary value problem, one may want to choose a different starting approximation, x_0, for the iteration process (7.25). For instance, if a "slowly varying" solution is expected, one may put $x_0(t, z) := z$, which was often done in the earlier works on the numerical-analytic method. In the latter case, on the assumption that

$$z \in \left\{u \in \mathbb{R}^n \;\middle|\; B\left(u, \frac{T}{4}\delta_\Omega(f)\right) \subset \Omega\right\} \neq \varnothing,$$

the convergence of sequence (7.25) as well as estimate (7.30) of Theorem 7.8 can easily be proved, e.g., by using Lemma 4 from [70, p. 384]. For this purpose, it suffices to replace (7.38) by the inequality

$$d_{m+1}(t,z) \leqslant \alpha_{m_\varepsilon,T}(t)\left[\frac{1}{2}Q_\varepsilon^{m-m_\varepsilon+1}\delta_\Omega(f) + K\,Q_\varepsilon^{m-m_\varepsilon}\big|b(z)+z\big|\right],$$

put $a_m := T\alpha_{m,1}(\frac{1}{2})K^{m-1}\delta_\Omega(f)$, $b_m := T\alpha_{m-1,1}(\frac{1}{2})K^{m-1}|b(z)+z|$, $u_m := \operatorname{col}(a_m, b_m)$ and $\phi_1(t) := \phi_2(t) := 1$, and notice that, for an arbitrary fixed positive ε and all m greater than some $m_\varepsilon \in \mathbb{N}$, estimate (7.15) of Lemma 7.5 implies the component-wise relation $u_{m+1} \leqslant T[r(A)+\varepsilon]\operatorname{diag}(K,K)u_m$. Then, similarly to [70, p. 385], assumption (7.24) implies that $\lim_{m\to+\infty} d_m(t,z) = 0$ uniformly in $(t,z) \in [0,T]\times\Omega(\beta)$, which, by (7.34), means that (7.25) is a Cauchy sequence.

Now we show that, in view of Theorem 7.8, the boundary value problem (7.18), (7.19) can be formally interpreted as a family of initial value problems for differential equations with "additively forced" right-hand side members. Namely, consider the Cauchy problem

$$x(T) = z \tag{7.41}$$

for the system

$$x'(t) = f\big(t, x(t)\big) + \mu, \quad t \in [0,T], \tag{7.42}$$

where $\mu \in \mathbb{R}^n$ and $z \in \Omega(\beta)$ are parameters.

THEOREM 7.12. *Under the conditions of Theorem* 7.8, *the solution of the initial value problem* (7.42) *satisfies the two-point condition* (7.19) *if, and only if* $\mu = \frac{1}{T}\Delta(z)$, *where* $\Delta : \Omega(\beta) \to \mathbb{R}^n$ *is the mapping defined by* (7.29).

REMARK 7.13. It is easy to show that the Lipschitz condition (7.20) implies that (7.42) has a unique solution for all $(\mu, z) \in \mathbb{R}^n \times \Omega(\beta)$.

PROOF OF THEOREM 7.12. It follows from the proof of Theorem 7.8 that, for every fixed $z \in \Omega(\beta)$, the function

$$x = x^*(\cdot, z) = \lim_{m\to+\infty} x_m(\cdot, z) \tag{7.43}$$

[see (7.25)] satisfies the integral equation (7.40). This implies immediately that (7.43) is the unique (see Remark 7.13) solution of the initial value problem (7.41) for the system

$$x'(t) = f\big(t, x(t)\big) + \frac{1}{T}\Delta(z), \quad t \in [0,T], \tag{7.44}$$

and, in addition, (7.43) satisfies the two-point condition (7.19). Since (7.44) coincides with problem (7.42) corresponding to $\mu = \frac{1}{T}\Delta(z)$. The fact that function (7.43) is not solution of (7.42) for any other value of μ, is obvious, e.g., from Eq. (7.40). □

COROLLARY 7.14. *Under the conditions of Theorem* 7.8, *the limit function* (7.43) *of sequence* (7.25) *is a solution of the two-point boundary value problem* (7.18), (7.19) *if, and only if the parameter* $z \in \Omega(\beta)$ *involved in* (7.25) *and* (7.43) *satisfies the following "determining equation"*:

$$z - b(z) = \int_0^T f\bigl(s, x^*(s, z)\bigr)\,\mathrm{d}s. \tag{7.45}$$

In Corollary 7.14, the parameter z can be interpreted as the value at $t = T$ of a possible solution of problem (7.18), (7.19).

PROOF. It suffices to apply Theorem 7.12 and notice that the differential equation in (7.44) coincides with (7.18) if, and only if z satisfies $\Delta(z) = 0$, i.e., when relation (7.45) holds. □

REMARK 7.15. On practise, it is natural to fix some natural m and consider the "approximate determining equation"

$$z - b(z) = \int_0^T f\bigl(s, x_m(s, z)\bigr)\,\mathrm{d}s. \tag{7.46}$$

In case Eq. (7.46) has an isolated root, say $z = z_m$, in some open subdomain of $\Omega(\beta)$, one can prove that, under certain additional conditions, the exact determining equation (7.45) is also solvable and, hence, the two-point problem (7.18), (7.19) has a solution x satisfying $x(T) \in \Omega(\beta)$; in addition, the function

$$\xi_m(t) := x_m(t, z_m), \quad t \in [0, T], \tag{7.47}$$

can be regarded as the "mth approximation" to the solution mentioned [see estimate (7.30)]. For that purpose, one can use some topological degree techniques (cf. Theorem 3.1 in [70, p. 43]) or the methods for the approximate solution of non-linear equations in Banach spaces extensively developed in the book of Krasnoselskii et al. [37] (see, e.g., Theorem 19.2 in [37, p. 281]). Here, we do not consider this problem in more detail.

We apply the above techniques to the following

EXAMPLE 7.16. Consider the two-point problem

$$\begin{aligned} & x''(t) - \frac{t}{8}x'(t) + \frac{1}{2}\bigl(x'(t)\bigr)^2 + \frac{1}{2}x(t) = \frac{9}{32} + \frac{t^2}{16}, \quad t \in [0, 1], \\ & x(0) = \bigl(x'(1)\bigr)^2, \\ & x'(0) = x'(1) - x(1) - \frac{1}{16}. \end{aligned} \tag{7.48}$$

By setting $x_1 := x$ and $x_2 := x'$, system (7.48) is rewritten as

$$\begin{aligned} x_1'(t) &= x_2(t), \\ x_2'(t) &= \frac{9}{32} + \frac{t^2}{16} + \frac{t}{8} x_2(t) - \frac{1}{2}\bigl(x_2(t)\bigr)^2 - \frac{1}{2} x_1(t), \\ x_1(0) &= \bigl(x_2(1)\bigr)^2, \\ x_2(0) &= x_2(1) - x_1(1) - \frac{1}{16}. \end{aligned} \tag{7.49}$$

Let us consider the first order problem (7.49) in the domain $(t, x_1, x_2) \in [0,1] \times \Omega$ with $\Omega := \{(x_1, x_2) \mid |x_1| \leqslant 1, |x_2| \leqslant \frac{3}{4}\}$. It is not difficult to verify that the corresponding quantity (7.23) is estimated as $\delta_\Omega(f) \leqslant \operatorname{col}(\frac{3}{2}, \frac{5}{4})$, and (7.20) holds with $\beta = \frac{1}{4}\delta_\Omega(f) \leqslant \operatorname{col}(\frac{3}{8}, \frac{5}{16})$, because $x \in \Omega(\beta)$ whenever $|x|$ is small enough. Condition (7.21) is satisfied with $K := \bigl(\begin{smallmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{8} \end{smallmatrix}\bigr)$ and, by Remark 7.4, the spectral radius of the matrix K satisfies inequality (7.24): $r(K) = \frac{7+\sqrt{177}}{16} \approx 1.269 < 1/r(A)$.

Let us see how the “zero approximation” (in the sense of Remark 7.15), behaves. According to (7.26), we have

$$\begin{aligned} x_{0,1}(t,z) &= \bigl(z_1 - {z_2}^2\bigr)t + {z_2}^2, \\ x_{0,2}(t,z) &= \left(z_1 + \frac{1}{16}\right)t + z_2 - z_1 - \frac{1}{16}, \end{aligned} \tag{7.50}$$

where $x_k := \operatorname{col}(x_{k,1},\ x_{k,2})$ for $k = 0, 1, \ldots$. The corresponding system of approximate determining equations is written as

$$\begin{aligned} &\frac{3}{2} z_1 - {z_2}^2 + \frac{1}{32} - z_2 = 0, \\ &-\frac{3}{32} z_2 + \frac{31}{24} z_1 - \frac{365}{1536} + \frac{1}{6} {z_1}^2 - \frac{1}{2} z_2 z_1 + \frac{3}{4} {z_2}^2 = 0. \end{aligned}$$

It has a positive solution

$$z_1 \approx 0.1801357504, \qquad z_2 \approx 0.2425992362, \tag{7.51}$$

which, substituted into (7.50) according to (7.47), yields the zeroth approximation to the solution of problem (7.48):

$$\xi_0(t) \approx 0.1212813610 \cdot t + 0.05885438940, \quad t \in [0,1]. \tag{7.52}$$

Let us construct the first and the second approximations. According to formula (7.25), we have

$$x_{1,1}(t,z) = \frac{1}{2}\left(z_1 + \frac{1}{16}\right)t^2 + \left(\frac{1}{2} z_1 - {z_2}^2 - \frac{1}{32}\right)t + {z_2}^2, \tag{7.53}$$

and

$$
\begin{aligned}
x_{1,2}(t,z) = {} & \left(\frac{3}{128} - \frac{1}{6}\left(z_1 + \frac{1}{16}\right)^2 + \frac{1}{24}z_1\right)t^3 \\
& + \left(\frac{1}{16}z_2 - \frac{5}{16}z_1 - \frac{1}{256} - \frac{1}{2}\left(z_2 - z_1 - \frac{1}{16}\right)\left(z_1 + \frac{1}{16}\right)\right. \\
& \left. + \frac{1}{4}{z_2}^2\right)t^2 + \left[-\frac{1}{16}z_2 + \frac{61}{48}z_1 + \frac{11}{256} - \frac{1}{4}{z_2}^2 + \frac{1}{6}\left(z_1 + \frac{1}{16}\right)^2\right. \\
& \left. + \frac{1}{2}\left(z_2 - z_1 - \frac{1}{16}\right)\left(z_1 + \frac{1}{16}\right)\right]t + z_2 - z_1 - \frac{1}{16}.
\end{aligned}
$$

The corresponding two approximate determining equations (7.46) are

$$
\frac{281}{192}z_1 - \frac{23}{24}{z_2}^2 + \frac{75}{2048} + \frac{1}{24}{z_1}^2 - \frac{191}{192}z_2 - \frac{1}{12}z_2 z_1 = 0 \tag{7.54}
$$

and

$$
\begin{aligned}
& \frac{5524531}{4423680}z_1 - \frac{32309}{327680}z_2 + \frac{18407}{24576}{z_2}^2 + \frac{98797}{645120}{z_1}^2 - \frac{28847}{61440}z_2 z_1 \\
& \quad + \frac{1}{945}{z_1}^4 - \frac{1}{240}{z_2}^3 z_1 + \frac{1}{960}{z_2}^4 - \frac{53}{1280}{z_2}^3 + \frac{2467}{120960}{z_1}^3 + \frac{1}{160}{z_2}^2{z_1}^2 \\
& \quad - \frac{1}{240}{z_1}^3 z_2 - \frac{61}{768}z_2{z_1}^2 + \frac{391}{3840}{z_2}^2 z_1 = \frac{16965793}{70778880},
\end{aligned} \tag{7.55}
$$

respectively. Computing a non-negative solution of (7.54), (7.55), we obtain

$$
z_1 \approx 0.1875156187, \qquad z_2 \approx 0.2500152909, \tag{7.56}
$$

and, inserting (7.56) into (7.53), we arrive at the first approximation:

$$
\xi_1(t) \approx 0.06250764568 + 0.163710^{-6}t + 0.1250078094t^2, \quad t \in [0,1]. \tag{7.57}
$$

In a similar manner, one can establish that the second approximate system of determining equations has a positive solution

$$
z_1 \approx 0.1874363008, \qquad z_2 \approx 0.2499449264 \tag{7.58}
$$

and, by (7.47) and (7.58), the corresponding function of the second approximation has the form

$$
\begin{aligned}
\xi_2(t) \approx {} & 0.06247246623 - 0.2235410^{-5}t + 0.005208996698t^4 \\
& - 0.01041383249t^3 + 0.1301709057t^2, \quad t \in [0,1].
\end{aligned} \tag{7.59}
$$

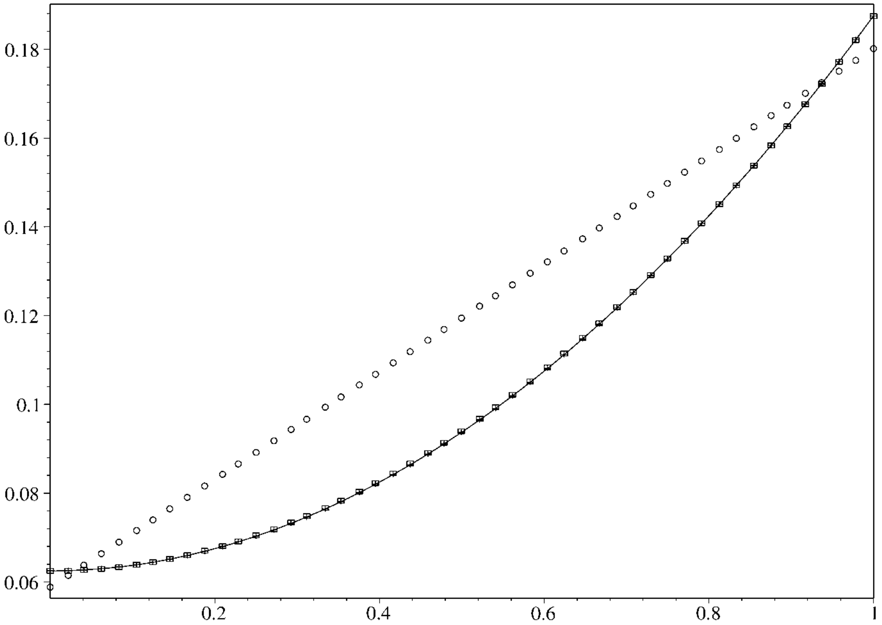


Fig. 6. The exact solution of problem (7.48) and functions (7.52), (7.57), and (7.59).

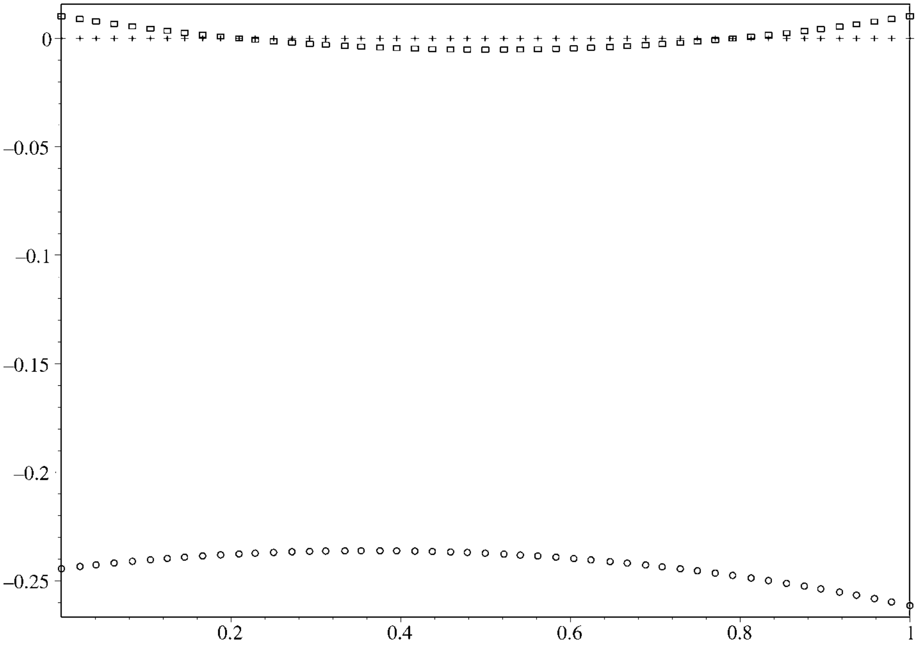


Fig. 7. Functions (7.61) for $k = 0, 1, 2$.

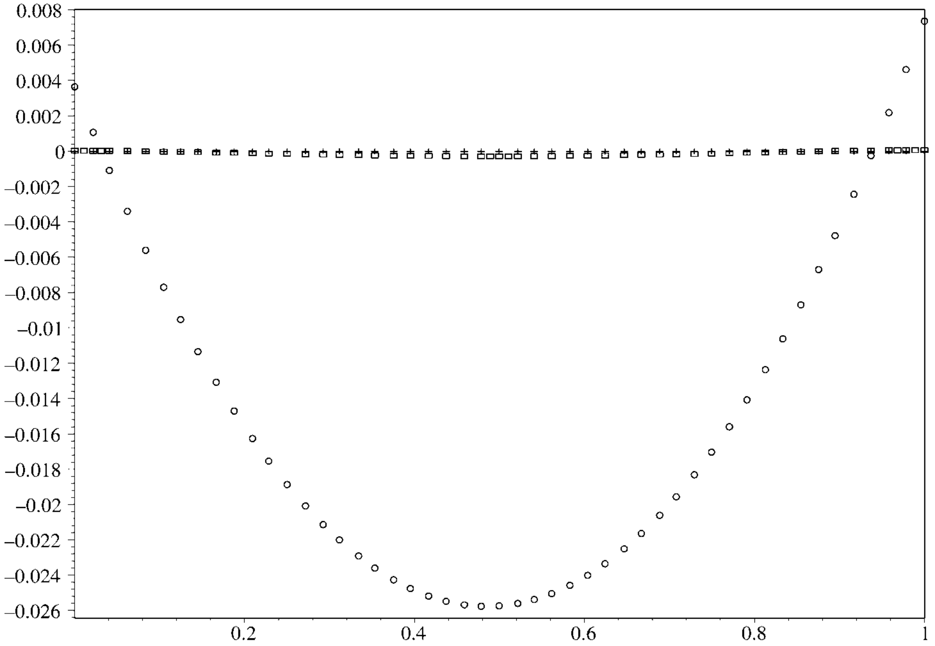


Fig. 8. The difference between the exact solution (7.60) of problem (7.48) and functions (7.52), (7.57), and (7.59).

Note that the two-point problem (7.48) has a solution

$$x(t) = \frac{t^2}{8} + \frac{1}{16}, \quad t \in [0, 1], \tag{7.60}$$

which, as is seen from Fig. 6, is well enough approximated by functions (7.52), (7.57), and (7.59) constructed above according to formula (7.47).

The functions

$$t \mapsto \xi_k''(t) - \frac{t}{8}\xi_k'(t) + \frac{1}{2}\bigl(\xi_k'(t)\bigr)^2 + \frac{1}{2}x\xi_k(t) - \frac{t^2}{16} - \frac{9}{32}, \quad k = 0, 1, 2, \tag{7.61}$$

obtained as a result of substitution of functions (7.52), (7.57), and (7.59) into the given differential equation, are plotted on Fig. 7, and the error of the approximate solutions with respect to function (7.60) is illustrated by Fig. 8. On all the figures, the symbols “circle”, “cross”, and “box” indicate the curves related to the zero, first, and second approximations, respectively. The graph of the exact solution (7.60) of problem (7.48) is drawn with the solid line [which, in fact, almost coincides with the graphs of functions (7.57) and (7.59)].

The computation shows that, e.g., function (7.59) satisfies the differential equation in (7.48) “up to 0.01” and differs from the exact solution (7.60) at most by 0.001.

8. Parametrisation method for three-point boundary value problems

Here, a three-point boundary-value problem for a system of non-linear differential equations is transformed to certain family of two-point problems the solutions of which are investigated by using successive approximation techniques.

Numerous approaches to investigation of boundary value problems with unknown parameters are available. Various iteration methods are used for such problems in the papers of Goma [26,25,27], Hosabekov [30], Kurpel′ and Marusjak [41]. In the works of Akhmedov, Svarichevskaya, and Yagubov [3], Luchka [45], Luchka and Zakhari͡ĭchenko [46], Samoilenko, Luchka, and Listopadova [93,94], the method of averaging of functional corrections and other projection-iterative methods are developed. Estimates of the number of solutions of certain singular two-point problems with parameters were obtained by Fečkan [19,20] by using the Nielsen fixed point theory. Various versions of the continuation method are studied, e.g., in the book of Gaines and Mawhin [24]. The are many works where different kinds of two-point and multipoint boundary value problems are studied by using numerical methods (see, e.g., the books of Ascher, Mattheij, and Russell [6], Keller [35] and the bibliography given therein for a survey).

It is important to note that the majority of works dealing with the numerical finding of solutions of boundary value problems are based upon the shooting method (see, e.g., [35, 6,10,1]). In particular, in the book of Keller [35], the two-point problem

$$y'(x) = f(x, y(x), \lambda_1, \lambda_2, \dots, \lambda_m), \quad x \in [a, b], \tag{8.1}$$

$$Ay(a) + By(b) = \gamma, \tag{8.2}$$

is considered, where $y : [a, b] \to \mathbb{R}^n$, A and B are given matrices of dimension $(m+n) \times n$, and λ_1, λ_2, …, λ_m are unknown parameters. Recall that the idea of the shooting method for the two-point problem (8.1), (8.2) is to replace it by the Cauchy problem for the same equation subjected to the initial conditions

$$y(a) = s. \tag{8.3}$$

In this case, the solution $y(\cdot, s, \lambda)$ of problem (8.1), (8.3) contains the unknown parameters $\lambda = \operatorname{col}(\lambda_1, \lambda_2, \dots, \lambda_m)$ and $s = \operatorname{col}(s_1, s_2, \dots, s_n)$ whose values are to be determined from the system of $n + m$ equations

$$Ay(a, s, \lambda) + By(b, s, \lambda) = \gamma. \tag{8.4}$$

It is then stated that the two-point problem (8.1), (8.2) has the same number of solutions as the system of numerical equations (8.4).

The main complication in this approach obviously lies in the numerical construction and investigation of system (8.4), which is possible only if the initial value problems (8.1), (8.3) for arbitrary values of λ and s have a unique solution *well-defined on the entire interval* $[a, b]$. In this relation, it is usually assumed that the function f involved in the right-hand side of (8.1) is continuous on the entire space and satisfies the Lipschitz condition with respect to the second argument on the entire $\mathbb{R}^n$.

The rigorous investigation of boundary value problems by using numerical methods tends to avoid the case where the range of solutions in question is restricted to a certain given (possibly, bounded) closed $D \subset \mathbb{R}^n$ and, correspondingly, one or another regularity assumption is imposed on the function f on the set $[a, b] \times D \times \mathbb{R}^m$ only. For example, in the case of the scalar two-point boundary value problem (8.2) for the equation

$$x'(t) = \big(x(t)\big)^2, \quad t \in [a, b], \tag{8.5}$$

the global Lipschitz condition for the right-hand side of (8.5) is obviously not satisfied. The local solution of the corresponding Cauchy problem (8.5), (8.3) has the form

$$x(t, s) = -\frac{1}{t - s^{-1} - a}, \quad t \in [a, b], \tag{8.6}$$

where s is a real parameter. It is clear from (8.6) that if $0 \leqslant s^{-1} \leqslant b - a$ then the solution of the initial value problem (8.5), (8.3) is undefined on the entire interval $[a, b]$ and, therefore, it does not make sense even to construct Eq. (8.4), not talking of its solution.

In many cases, the complications indicated can be overcome by using the approach of [89,70,60,66].

8.1. *Problem setting*

Let us consider the problem on finding a solution of the differential system (2.1) satisfying the three-point non-separated linear boundary conditions

$$Ax(0) + Bx(\xi) + x(T) = d, \tag{8.7}$$

where $f : [0, T] \times D \to \mathbb{R}^n$, D is the closure of a bounded domain in $\mathbb{R}^n$, $\{A, B\} \subset \mathrm{GL}_n(\mathbb{R})$, and $\xi \in (0, T)$ is given. It is easy to see that the tree-point boundary condition (2.5) with a non-singular matrix C can be brought to form (8.7) (it suffices to replace A, B, and d in (8.7) by $C^{-1}A$, $C^{-1}B$, and $C^{-1}d$, respectively). Conditions (8.7) are non-separated unless the matrices A and B are both equal to zero.

Here we show an approach to the investigation of problem (2.1), (8.7), which is based upon artificial introduction of a suitable number of parameters into the boundary conditions and allows one to construct approximate solutions in an analytic form.

We assume that the function $f : [0, T] \times D \to \mathbb{R}^n$ is continuous and, in the second variable, satisfies the Lipschitz condition (3.5) with a certain n-dimensional square matrix K with non-negative elements.

We restrict our consideration by the class of differential systems (2.1) for which the spectral radius $r(K)$ of the non-negative matrix K appearing in the Lipschitz condition (3.5) satisfies inequality (6.14).

8.2. *Transformation to a two-point problem with a parameter in boundary conditions*

Let us replace formally the value of the unknown solution of problem (2.1), (8.7) at the point ξ by the n-dimensional vector of parameters $(\lambda_k)_{k=1}^n$ that varies in a certain set $\Lambda \subset D$:

$$x(\xi) = \lambda. \tag{8.8}$$

Every function x satisfying relations (8.8) and (8.7) obviously has the property

$$Ax(0) + x(T) = d - B\lambda. \tag{8.9}$$

Let us consider problem (2.1), (8.9), in which the two-point boundary condition (8.9) contains an unknown parameter $\lambda \in \Lambda$.

REMARK 8.1. The set of solutions of the original three-point problem (2.1), (8.7), clearly, coincides with the set of those solutions of the two-point problem (2.1), (8.9) for which the parameter λ satisfies the additional condition (8.8).

For any $\{z, \lambda\} \subset \mathbb{R}^n$, we put

$$\beta(z, \lambda) := \frac{T}{4}\delta_D(f) + \left|d - B\lambda - (A + \mathbb{1}_n)z\right|, \tag{8.10}$$

where $\mathbb{1}_n$ is the unit matrix of dimension n and $\delta_D(f)$ is defined by equality (6.13). We shall assume in the sequel that the set $\Gamma \subseteq D$ given by the formula

$$\Gamma := \left\{z \in D\colon\ B\big(z, \beta(z, \lambda)\big) \subset D \text{ for all } \lambda \in \Lambda\right\} \tag{8.11}$$

is non-empty: $\Gamma \neq \varnothing$. Recall that, here and above, notation (12) of Section 1 is used.

With the parametrised problem (2.1), (8.9), we associate the sequence of functions

$$\begin{aligned} x_{m+1}(t, z, \lambda) := z &+ \int_0^t f\big(s, x_m(s, z, \lambda)\big)\,\mathrm{d}s - tT^{-1}\int_0^T f\big(s, x_m(s, z, \lambda)\big)\,\mathrm{d}s \\ &+ tT^{-1}[d - B\lambda - Az - z], \end{aligned} \tag{8.12}$$

where $\lambda \in \Lambda$, $z \in \Gamma$, $m = 0, 1, 2, \ldots$, $x_0(t, z, \lambda) \equiv z$. It is easy to see that $x_m(0, z, \lambda) = z$ for any $m \geqslant 0$, $\lambda \in \Lambda$, and $z \in \Gamma$ and, moreover,

$$x_m(T, z, \lambda) = d - B\lambda - Az, \tag{8.13}$$

i.e., function (8.12) satisfies the two-point condition (8.9).

REMARK 8.2. The two-point condition (8.9) obviously explicitly only p parameters among λ_1, λ_2, $\ldots$, λ_n, where $p = \operatorname{rank} B$. Therefore, relation (8.8), according to which the three-point condition (8.7) is replaced by the two-point boundary condition (8.9), can be replaced by the relation

$$Bx(\xi) = B\lambda. \tag{8.14}$$

In other words, it is sufficient to consider only those solutions of the two-point problem (2.1), (8.9), for which the additional condition (8.14) is satisfied. In this case, one completely excludes from consideration the $n - p$ parameters the values of which one need not to determine. If $p = n$ (i.e., $\det B \neq 0$), then conditions (8.14) and (8.8) have the same meaning.

One could also construct a similar scheme by replacing the three-point condition (8.7) by the condition

$$Ax(0) + x(T) = d - \mu \tag{8.15}$$

and seeking for those solutions x of the two-point problem (2.1), (8.15), for which $Bx(\xi) = \mu$.

8.3. *Convergence of successive approximations*

Let us present conditions sufficient for the uniform convergence of the recurrence sequence (8.12) and establish the relation between its limit and the sets of solutions of problems (2.1), (8.7) and (2.1), (8.9).

THEOREM 8.3. *Assume that the continuous function* $f : [0, T] \times D \to \mathbb{R}^n$ *satisfies condition* (3.5) *with the matrix* K *possessing property* (6.14). *Moreover, let the matrices* A, B *and the vector* d *be such that* $\Gamma \neq \varnothing$.

Then:

(1) *Sequence* (8.12) *converges uniformly in* $t \in [0, T]$ *for all* $(z, \lambda) \in \Gamma \times \Lambda$;

(2) *For any* $(z, \lambda) \in \Gamma \times \Lambda$, *the limit function*

$$x^*(t, z, \lambda) := \lim_{m\to+\infty} x_m(t, z, \lambda) \tag{8.16}$$

of sequence (8.12) *is the unique solution of the integral equation*

$$\begin{aligned} x(t) := z + \int_0^t f\bigl(s, x(s)\bigr)\,\mathrm{d}s - tT^{-1}\int_0^T f\bigl(s, x(s)\bigr)\,\mathrm{d}s \\ + tT^{-1}[d - B\lambda - Az - z], \quad t \in [0, T], \end{aligned} \tag{8.17}$$

or, which is the same, of the integro-differential equation

$$x'(t) = f\bigl(t, x(t)\bigr) + T^{-1}\Delta(z, \lambda), \quad t \in [0, T], \tag{8.18}$$

where

$$\Delta(z,\lambda) := d - B\lambda - Az - z - \int_0^T f\bigl(s, x^*(s,z,\lambda)\bigr)\,\mathrm{d}s. \tag{8.19}$$

For $(z,\lambda) \in \Gamma \times \Lambda$, *this solution satisfies the boundary conditions* (8.9);

(3) *For all* $(t,z,\lambda) \in [0,T] \times \Gamma \times \Lambda$, *the estimate*

$$\bigl|x^*(t,z,\lambda) - x_m(t,z,\lambda)\bigr| \leqslant h(t,z,\lambda), \tag{8.20}$$

is true, where

$$h(t,z,\lambda) := \frac{20t}{9}\left(1 - \frac{t}{T}\right) Q^m (\mathbb{1}_n - Q)^{-1}\bigl[\delta_D(f) + K|d - B\lambda - Az - z|\bigr]$$

and $Q := \frac{3T}{10}K$.

Proof. Let us show that sequence (8.12) is a Cauchy sequence in the Banach space $C([0,T],\mathbb{R}^n)$. Let us first prove that the values of each of the functions (8.12) for all $(z,\lambda) \in \Gamma \times \Lambda$ are contained in D. By Lemma 7.7, for any continuous $x : [0,T] \to \mathbb{R}^n$, the estimate

$$\left|\int_0^t \left(x(\tau) - \frac{1}{T}\int_0^T x(s)\,\mathrm{d}s\right)\mathrm{d}\tau\right| \leqslant \frac{1}{2}\alpha_{1,T}(t)\Bigl[\max_{\tau\in[0,T]} x(\tau) - \min_{\tau\in[0,T]} x(\tau)\Bigr]$$

is true, where the function $\alpha_{1,T}$ has form (3.34). It follows from Lemma 7.1 that $\max_{t\in[0,T]}\alpha_{1,T}(t) = T/2$. Therefore, for $m = 0$, relation (8.12), in view of (8.10), implies that

$$\bigl|x_1(t,z,\lambda) - z\bigr| \leqslant \frac{1}{2}\alpha_{1,T}(t)\delta_D(f) + |d - B\lambda - z - Az| \leqslant \beta(z,\lambda),$$

for any $(t,z,\lambda) \in [0,T] \times \Gamma \times \Lambda$. Here the value $\delta_D(f)$ is given by formula (6.13). Therefore, in view of (8.11), it follows that $x_m(t,z,\lambda) \in D$ for all such (t,z,λ) and any $m = 0,1,2,\ldots$. Arguing similarly to the proof of [71, Theorem 1], one can show that, for all $m \geqslant 0$, $j \geqslant 0$ and $(t,z,\lambda) \in [0,T] \times \Gamma \times \Lambda$

$$\begin{aligned}\bigl|x_{m+j}(t,z,\lambda) - x_m(t,z,\lambda)\bigr| \leqslant \frac{10t}{9}\alpha_{1,T}(t)Q^m\Biggl[\frac{1}{2}\sum_{i=0}^{j-1} Q^i\delta_D(f) \\ + K\sum_{i=0}^{j-1} Q^{i-1}|d - B\lambda - z - Az|\Biggr].\end{aligned} \tag{8.21}$$

Since, by (6.14), $\lim_{m\to+\infty} Q^m = 0$ and $\sum_{i=0}^{+\infty} Q^i = (\mathbb{1}_n - Q)^{-1}$, it is obvious from (8.21) that (8.12) is a Cauchy sequence in the uniform norm. Passing to the limit when j

tends to $+\infty$ in (8.21), we obtain that the uniform limit (8.16) of sequence (8.12) satisfies inequality (8.20). Passing to the limit as $m \to +\infty$ in (8.12) and (8.13), we conclude that function (8.16) satisfies Eqs. (8.17), (8.18) and condition (8.9). ☐

REMARK 8.4. It is clear from the form of Eq. (8.17) that, under assumptions of Theorem 8.3, its unique solution (8.16) for all $(z, \lambda) \in \Gamma \times \Lambda$ satisfies condition

$$x^*(0, z, \lambda) = z. \tag{8.22}$$

REMARK 8.5. In some papers (e.g., in [41,100,101]), parametrised problems of type (2.1), (8.9) are studied by means of construction of two separate iteration processes for the unknown function and the vector of parameters.

The form of the "perturbed" equation (8.18) suggests the idea that the parametrised boundary value problem (2.1), (8.9) might be interpreted as the family of the Cauchy problems

$$x(0) = z \tag{8.23}$$

for the differential equations of the form

$$x'(t) = f\bigl(t, x(t)\bigr) + \mu, \quad t \in [0, T], \tag{8.24}$$

where the vector of parameters μ takes values in a suitable set. More precisely, the following statement holds.

THEOREM 8.6. *Let conditions of Theorem* 8.3 *be satisfied and let* $(z, \lambda) \in \Gamma \times \Lambda$ *and* $\mu \in \mathbb{R}^n$. *Then a solution of the Cauchy problem* (8.24), (8.23) *satisfies the two-point boundary conditions* (8.9) *if and only if the parameter* μ *in* (8.24) *is given by the equation*

$$\mu = \frac{1}{T}\Delta(z, \lambda), \tag{8.25}$$

where $\Delta : \Gamma \times \Lambda \to \mathbb{R}^n$ *is the function defined by formula* (8.19).

REMARK 8.7. It follows from Theorem 8.3 that, under the conditions indicated above, for any $(z, \lambda) \in \Gamma \times \Lambda$, there exists a limit function (8.16) of the recurrence sequence (8.12) and, therefore, mapping (8.19) is well-defined.

PROOF OF THEOREM 8.6. *Sufficiency.* Let us put $\mu = T^{-1}\Delta(z, \lambda)$, in (8.24), where $z \in \Gamma$ and $\lambda \in \Lambda$ are some fixed vectors, and Δ is given b formula (8.19). According to Remark 8.7, the expression $\Delta(z, \lambda)$ makes sense for all $(z, \lambda) \in \Gamma \times \Lambda$. Theorem 8.3 implies that, for the given z and λ, the uniform limit (8.16) of the corresponding sequence (8.12) coincides with the unique solution of the two-point problem (8.18), (8.9). As indicated in Remark 8.4, this function also satisfies the initial condition (8.22), i.e., is a solution of the Cauchy problem (8.24), (8.23) corresponding to the value (8.25) of the parameter μ.

Necessity. Let $z \in \Gamma$, $\lambda \in \Lambda$, and $\bar{\mu} \in \mathbb{R}^n$ be arbitrary vectors, and $\bar{x} : [0, T] \to \mathbb{R}^n$ be a solution of the Cauchy problem (8.23) for the equation

$$x'(t) = f\big(t, x(t)\big) + \bar{\mu}, \quad t \in [0, T], \tag{8.26}$$

satisfying, for the given value of λ, the two-point condition (8.9). Assume that there exists another value $\bar{\bar{\mu}}$, of the parameter μ, for which a certain solution $\bar{\bar{x}}$ of problem (8.23) for the equation

$$x'(t) = f\big(t, x(t)\big) + \bar{\bar{\mu}}, \quad t \in [0, T], \tag{8.27}$$

satisfies the two-point condition (8.9). It is obvious from (8.23), (8.26), and (8.27) that the functions $\bar{x}$ and $\bar{\bar{x}}$ satisfy the integral equations

$$\bar{x}(t) = z + \int_0^t f\big(s, \bar{x}(s)\big)\,\mathrm{d}s + \bar{\mu}t, \quad t \in [0, T], \tag{8.28}$$

and

$$\bar{\bar{x}}(t) = z + \int_0^t f\big(s, \bar{\bar{x}}(s)\big)\,\mathrm{d}s + \bar{\bar{\mu}}t, \quad t \in [0, T], \tag{8.29}$$

respectively. For $t = T$, relations (8.28), (8.29) yield

$$T\bar{\mu} = \bar{x}(T) - z - \int_0^T f\big(s, \bar{x}(s)\big)\,\mathrm{d}s \tag{8.30}$$

and

$$T\bar{\bar{\mu}} = \bar{\bar{x}}(T) - z - \int_0^T f\big(s, \bar{\bar{x}}(s)\big)\,\mathrm{d}s. \tag{8.31}$$

The function $\bar{x}$, by assumption, satisfies the two-point condition (8.9),

$$A\bar{x}(0) + \bar{x}(T) = d - B\lambda,$$

and the initial condition $\bar{x}(0) = z$, whence

$$\bar{x}(T) = d - B\lambda - Az.$$

Similarly, on can verify that

$$\bar{\bar{x}}(T) = d - B\lambda - Az.$$

Therefore, (8.30) and (8.31) yield

$$\bar{\mu} = \frac{1}{T}\left(d - B\lambda - Az - z - \int_0^T f\big(s, \bar{x}(s)\big)\,\mathrm{d}s\right) \tag{8.32}$$

and

$$\bar{\bar{\mu}} = \frac{1}{T}\left(d - B\lambda - Az - z - \int_0^T f\big(s, \bar{\bar{x}}(s)\big)\,\mathrm{d}s\right). \tag{8.33}$$

Substituting (8.32) and (8.33) into (8.28) and (8.29), respectively, we obtain that

$$\begin{aligned} \bar{x}(t) = z &+ \int_0^t f\big(s, \bar{x}(s)\big)\,\mathrm{d}s \\ &+ \frac{t}{T}\left[d - B\lambda - Az - z - \int_0^T f\big(s, \bar{x}(s)\big)\,\mathrm{d}s\right] \end{aligned} \tag{8.34}$$

and

$$\begin{aligned} \bar{\bar{x}}(t) = z &+ \int_0^t f\big(s, \bar{x}(s)\big)\,\mathrm{d}s \\ &+ \frac{t}{T}\left[d - B\lambda - Az - z - \int_0^T f(s, \bar{\bar{x}}(s))\,\mathrm{d}s\right] \end{aligned} \tag{8.35}$$

for any $t \in [0, T]$. We recall that here $z \in \Gamma$ and $\lambda \in \Lambda$. Therefore, similarly to the proof of Theorem 8.3, by using the form of Eqs. (8.34), (8.35) and definition (8.11) of the set Γ, we can show that all the values of the functions $\bar{\bar{x}}$ and $\bar{x}$ are contained in D:

$$\bar{x}\big([0, T]\big) \cup \bar{\bar{x}}\big([0, T]\big) \subset D. \tag{8.36}$$

It is obvious from (8.34) and (8.35) that

$$\begin{aligned} \bar{\bar{x}}(t) - \bar{x}(t) = &\int_0^t \big[f\big(s, \bar{\bar{x}}(s)\big) - f\big(s, \bar{x}(s)\big)\big]\,\mathrm{d}s \\ &- \frac{t}{T}\int_0^T \big[f\big(s, \bar{\bar{x}}(s)\big) - f\big(s, \bar{x}(s)\big)\big]\,\mathrm{d}s, \quad t \in [0, T], \end{aligned}$$

and, hence, by virtue of relation (8.36) and the Lipschitz condition (3.5), the function

$$r(t) := \big|\bar{\bar{x}}(t) - \bar{x}(t)\big|, \quad t \in [0, T],$$

satisfies the integral inequality

$$\begin{aligned} r(t) &\leqslant K\left[\int_0^t r(s)\,\mathrm{d}s + \frac{t}{T}\int_0^T r(s)\,\mathrm{d}s\right] \\ &\leqslant K\alpha_{1,T}(t) \max_{s\in[a,b]} r(s), \quad t \in [0, T], \end{aligned} \tag{8.37}$$

where the function $\alpha_{1,T}$ is given by formula (3.34). Using (8.37) sequentially, we arrive at the inequality

$$r(t) \leqslant K^m \alpha_{m,T}(t) \max_{s\in[a,b]} r(s), \quad t \in [0,T], \tag{8.38}$$

where the natural m is arbitrary and $\alpha_{m,T}$ is defined by relation (3.35). According to Lemma 2.4 from [70], functions (3.35) for all $t \in [0,T]$ and $k = 1, 2, \ldots$ admit estimate (3.36). Consequently, it follows from (8.38) that, for any natural m,

$$r(t) \leqslant \frac{10}{9} K\alpha_{1,T}(t)\left(\frac{3T}{10}K\right)^{m-1} \max_{s\in[a,b]} r(s), \quad t \in [0,T]. \tag{8.39}$$

Letting in (8.39) m tend to $+\infty$ and taking condition (6.14) into account, we conclude that $r \equiv 0$ on $[0,T]$, i.e., the functions $\bar{x}$ and $\bar{\bar{x}}$ coincide with one another, and hence $\bar{\mu} = \bar{\bar{\mu}}$. The contradiction obtained proves that $\mu = T^{-1}\Delta(z,\lambda)$ is the unique value of the parameter μ in Eq. (8.24), for which the solution of the initial value problem (8.24), (8.23) satisfies the two-point condition (8.9). □

REMARK 8.8. It follows from Lemma 7.5 that for an arbitrarily small positive ε there always exists a number $k_\varepsilon \in \mathbb{N}$ such that functions (3.35), starting with the k_εth one, admit the estimate

$$\alpha_{k,T}(t) \leqslant \left(\frac{T}{q} + \varepsilon\right)^{k-k_\varepsilon} \alpha_{k_\varepsilon,T}(t), \quad t \in [0,T],\ k \geqslant k_\varepsilon, \tag{8.40}$$

where $q \approx 3.416131$. Using (8.40) instead of estimate (3.36) and modifying appropriately the proofs of Theorems 8.3 and 8.6, one ca show that inequality (6.14) can be replaced by the weaker condition $r(K) < qT^{-1}$. All the related statements are preserved in this case at the expense of certain technical modifications of the argument (e.g., the form of the function h in estimate (8.20)).

Let us make clear how the function (8.16) defined in Theorems 8.3 and 8.6 is related to the set of solutions of the two-point problem (2.1), (8.9) containing a free parameter $\lambda \in \Lambda$.

PROPOSITION 8.9. *Under conditions of Theorem* 8.3, *the function* $x^*(\cdot, z, \lambda)$ *given by formula* (8.16) *for* $(z,\lambda) \in \Gamma \times \Lambda$, *is a solution of the two-point problem* (2.1), (8.9) *with the parameter* λ *if and only if* z *and* λ *satisfy the relation*

$$\int_0^T f\big(s, x^*(s,z,\lambda)\big)\,\mathrm{d}s = d - B\lambda - Az - z. \tag{8.41}$$

Function (8.16) *is a solution of the original three-point problem* (2.1), (8.7) *if and only if the pair* (z,λ) *satisfies condition* (8.41), *and, moreover,*

$$x^*(\xi, z, \lambda) = \lambda. \tag{8.42}$$

PROOF. It follows from Theorem 8.3 that for all $(z, \lambda) \in \Gamma \times \Lambda$ the function $x^*(\cdot, z, \lambda)$ is a solution of the two-point problem (8.18), (8.9). Equation (8.18) coincides with Eq. (2.1) if and only if z and λ satisfy condition $\Delta(z, \lambda) = 0$, i.e., equality (8.41) holds. The second assertion follows from Remark 8.1. □

8.4. *Practical realisation*

When putting the suggested scheme to practise, it is natural to fix a certain iteration number, m, in (8.12) and regard $x_m(\cdot, z, \lambda)$ as an approximation to the unknown function $x^*(\cdot, z, \lambda)$, the existence of which is stated in Theorem 8.3. In this case, instead of Eqs. (8.41), (8.42), the "approximate determining equations"

$$\int_0^T f\bigl(s, x_m(s, z, \lambda)\bigr)\,\mathrm{d}s = d - B\lambda - Az - z \tag{8.43}$$

and

$$\begin{aligned} z + \int_0^t f\bigl(s, x_m(s, z, \lambda)\bigr)\,\mathrm{d}s - tT^{-1}\int_0^T f\bigl(s, x_m(s, z, \lambda)\bigr)\,\mathrm{d}s \\ + tT^{-1}[d - B\lambda - Az - z] = \lambda. \end{aligned} \tag{8.44}$$

arise, from which we can find suitable values of the pair of parameters $(z, \lambda) \in \Gamma \times \Lambda$. If the system of $2n$ Eqs. (8.43), (8.44) has an isolated solution $(\bar{z}, \bar{\lambda}) \in \Gamma \times \Lambda$, then, by using topological and functional-analytic methods (see, e.g., Theorem 3.1 in [70] and Theorem 19.2 in [37]), under the appropriate additional assumptions, one can prove that system (8.41), (8.42) is also solvable and, thus establishing the existence of a solution of the original three-point problem. The function

$$\bar{x}_m(t) := x_m(\cdot, \bar{z}, \bar{\lambda}), \quad t \in [0, T], \tag{8.45}$$

can then be regarded as an approximation to a solution of problem (2.1), (8.7).

8.5. *Example of a two-dimensional three-point problem*

Let us consider the three-point boundary value problem

$$x_1'(t) = x_2(t), \tag{8.46a}$$

$$x_2'(t) = \frac{1}{8}tx_2 - \frac{1}{2}{x_2}^2 - \frac{1}{2}x_1 + \frac{9}{32} + \frac{1}{16}t^2, \quad t \in [0, 1], \tag{8.46b}$$

$$x_2(0) + x_1\left(\frac{1}{2}\right) + x_1(1) = \frac{9}{32}, \tag{8.46c}$$

$$x_1(0) + x_2(1) = \frac{5}{16}. \tag{8.46d}$$

Conditions (8.46c), (8.46d), clearly, can be written in form (8.7) with $T = 1$ and

$$A := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad d := \begin{bmatrix} \frac{9}{32} \\ \frac{5}{16} \end{bmatrix}, \quad \xi := \frac{1}{2}. \tag{8.47}$$

Let us restrict our consideration by those solutions of system (8.46a), (8.46b), all the values of which are contained in the set

$$D := \left\{ (x_1, x_2) \colon \ |x_1| \leqslant 1, |x_2| \leqslant \frac{3}{4} \right\}. \tag{8.48}$$

It is not difficult to verify that, for all $t \in [0, 1]$ and $\{u, v\} \subset D$, the function $f : [0, 1] \times D \to \mathbb{R}^2$ given by the formula

$$f(t, x_1, x_2) := \left[\ x_2 \tfrac{1}{8} t x_2 - \tfrac{1}{2} {x_2}^2 - \tfrac{1}{2} x_1 + \tfrac{9}{32} + \tfrac{t^2}{16} \ \right], \quad (t, x_1, x_2) \in [0, 1] \times D,$$

satisfies the Lipschitz condition (3.5) with the matrix

$$K := \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{7}{8} \end{bmatrix}. \tag{8.49}$$

The greatest eigenvalue of matrix (8.49) is equal to $(7 + \sqrt{177})/16 \approx 1.27$, which is less than $\frac{10}{3}$, and, hence, inequality (6.14) holds.

Let us seek for the solutions of the three-point problem (8.46) among the solutions of system (8.46a), (8.46b) satisfying the parametrised two-point conditions

$$x_2(0) + x_1(1) = \frac{9}{32} - \lambda_1, \tag{8.50}$$

$$x_1(0) + x_2(1) = \frac{5}{16}. \tag{8.51}$$

Conditions (8.50), (8.51), obviously, coincide with (8.9) for A, B, and d given by (8.47). For the domain where the two-dimensional vector of parameters $\lambda = (\lambda_1, \lambda_2)$ varies, we take, e.g., the rectangle

$$\Lambda := \left\{ (\lambda_1, \lambda_2) \colon \ |\lambda_1| \leqslant \frac{1}{4}, |\lambda_2| \leqslant \frac{1}{2} \right\} \tag{8.52}$$

contained in set (8.48). A solution (x_1, x_2) of problem (8.46a), (8.46b), (8.50), (8.51) is a solution of the original three-point problem (8.46) if and only if the additional conditions $x_1(\frac{1}{2}) = \lambda_1$ and $x_2(\frac{1}{2}) = \lambda_2$ are satisfied.

Let us use the approach based on Theorem 8.3, for which purpose verify the fulfilment of its assumptions. It is easy to see that the value $\delta_D(f)$ in the case considered is given by the equality

$$\delta_D(f) = \frac{1}{2}\begin{bmatrix} 3 \\ \frac{355}{128} \end{bmatrix}. \tag{8.53}$$

Furthermore, it is clear from (8.48), (8.52), (8.47), and (8.53) that, for our problem, the components of the function $\binom{\beta_1}{\beta_2} : D \times \Lambda \to \mathbb{R}^2$ defined by equality (8.10) have the form

$$\beta_1(z_1, z_2, \lambda_1, \lambda_2) = \frac{3}{8} + \left|\frac{9}{32} - \lambda_1 - z_1 - z_2\right|,$$

$$\beta_2(z_1, z_2, \lambda_1, \lambda_2) = \frac{355}{1024} + \left|\frac{5}{15} - z_1 - z_2\right|.$$

Therefore, according to the definition (8.11) of the set Γ, a point (z_1, z_2) from D belongs to Γ if and only if

$$\left|\frac{9}{32} - \lambda_1 - z_1 - z_2\right| \leqslant \frac{5}{8}, \tag{8.54}$$

$$\left|\frac{5}{15} - z_1 - z_2\right| \leqslant \frac{413}{1024} \tag{8.55}$$

for all $\lambda_1 \in [-\frac{1}{4}, \frac{1}{4}]$. The set Γ of such pairs $(z_1, z_2) \in D$ is obviously non-empty. Thus, all the conditions of Theorem 8.3 are satisfied in this case and, therefore, for any $(z, \lambda) \in \Gamma \times \Lambda$, on [0, 1], the function (8.16) discussed in Theorem 8.6 and Proposition 8.9 is well-defined.

Let us construct approximate solutions of problem (8.46) by using the approach indicated in Section 8.4. For this purpose, we compute analytically several members of the recurrence function sequence (8.12) with the help of Maple [29]. Let us set the zeroth iteration to be equal identically to z, $z \in \Gamma$. Then for $m = 1$ relation (8.12) yields

$$x_{11}(t, z_1, z_2, \lambda_1, \lambda_2) = z_1 + \frac{9}{32}t - t\lambda_1 - z_1 t - z_2 t, \tag{8.56}$$

$$x_{12}(t, z_1, z_2, \lambda_1, \lambda_2) = z_2 + \frac{7}{24}t + \frac{1}{48}t^3 + \frac{1}{16}z_2 t^2 - \frac{17}{16}z_2 t - z_1 t \tag{8.57}$$

for any $t \in [0, 1]$, $(z_1, z_2) \in \Gamma$, and λ_1, λ_2 satisfying the inequalities

$$|\lambda_1| \leqslant \frac{1}{4}, \qquad |\lambda_2| \leqslant \frac{1}{2}.$$

Here and below, the symbols x_{k1} and x_{k2} designate, respectively, the first and the second components of the vector x_k.

For $m = 1$, the system approximate determining equations (8.43), (8.44) corresponding to problem (8.46) consists of the equations

$$\frac{25}{192} - \frac{143}{96}z_2 - \frac{1}{2}z_1 - \lambda_1 = 0, \tag{8.58}$$

$$\frac{40391}{483840} - \frac{28157}{23040}z_2 - \frac{583}{720}z_1 + \frac{827}{5120}z_2^2 - \frac{31}{192}z_2z_1 + \frac{1}{6}z_1^2 - \frac{1}{4}\lambda_1 = 0, \tag{8.59}$$

$$\frac{5}{8}z_1 + \frac{313}{3072} - \frac{3}{8}z_2 - \frac{3}{2}\lambda_1 = 0, \tag{8.60}$$

and

$$\begin{aligned}&\frac{129067}{294912}z_2 + \frac{98569}{589824} - \frac{899}{1536}z_1 - \frac{63}{1024}z_2^2\\ &\quad + \frac{1}{1024}z_2z_1 + \frac{1}{16}z_1^2 - \frac{1}{16}\lambda_1 - \lambda_2 = 0.\end{aligned} \tag{8.61}$$

REMARK 8.10. The rank of the matrix B corresponding to the two-point conditions (8.46c), (8.46d) rewritten in form (8.7) is equal to 1 (the conditions do not contain explicitly $x_2(\frac{1}{2})$). Therefore, in view of Remark 8.2, the determining equations — both those given above and the other ones — need not be solved with respect to the parameter λ_2 because its value does not play any role in for problem considered. In this example we, however, write down all the equations completely in order to illustrate the approach.

Solving the system of Eqs. (8.58)–(8.60), (8.61) numerically, we obtain

$$z_1 \approx 0.06775109879, \qquad z_2 \approx 0.006168107627, \tag{8.62}$$

$$\lambda_1 \approx 0.08714487359, \qquad \lambda_2 \approx 0.125. \tag{8.63}$$

Substituting (8.62), (8.63) into (8.56) and (8.57), we find a function of form (8.45) – the "first approximation" to a solution of the three-point problem (8.46) corresponding to the values (8.62), (8.63) of the parameters z_1, z_2 and λ_1, λ_2 computed above:

$$\begin{aligned}\bar{x}_{11}(t) &= 0.06775109879 + 0.1201859200t,\\ \bar{x}_{12}(t) &= 0.006168107627 + 0.2173619535t\\ &\qquad + \frac{t^3}{48} + 0.0003855067267t^2.\end{aligned} \tag{8.64}$$

Note that the given three-point problem (8.46) has the solution

$$x_1(t) = \frac{t^2}{8} + \frac{1}{16}, \quad x_2(t) = \frac{t}{4}, \quad t \in [0, 1], \tag{8.65}$$

passing through the point $(\frac{1}{16}, 0)$ of the set Γ at the moment of time 0.

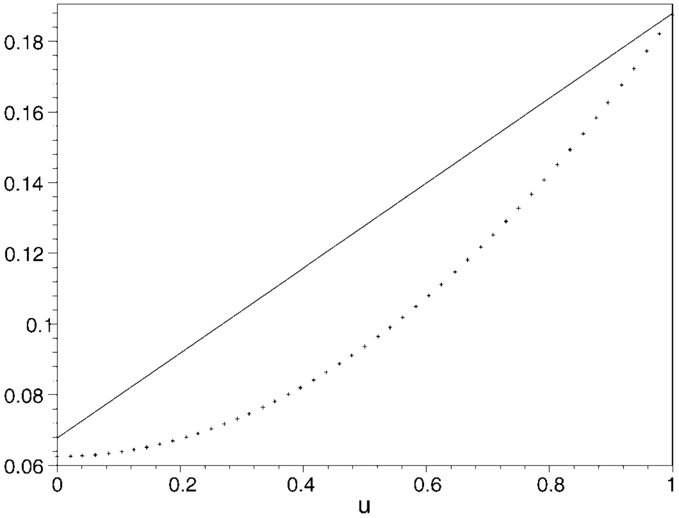


Fig. 9. The first component of the exact solution (8.65) and its “first approximation” (8.64).

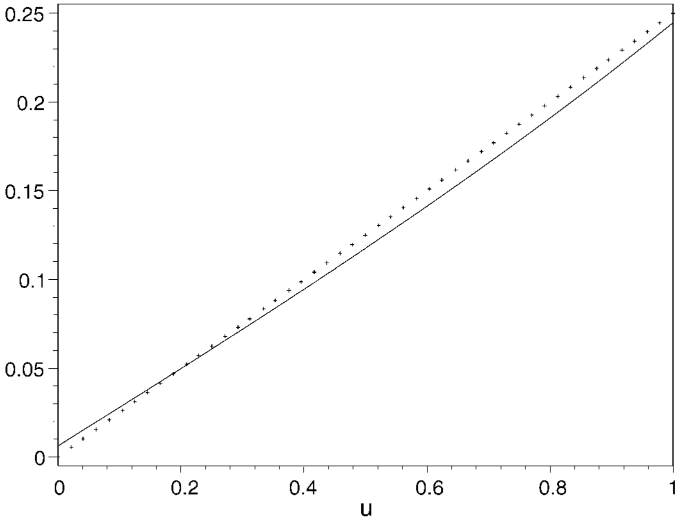


Fig. 10. The second component of the exact solution (8.65) and its “first approximation” (8.64).

Figs. 9 and 10 show the graphs of the first and the second components of the exact solution (8.65) of problem (8.46) (dotted line) and its “first approximation” (8.64) (solid line). The components of the deviation of the “first approximation” (8.64) from solution (8.65), i.e., the functions $x_1 - \bar{x}_{11}$ and $x_2 - \bar{x}_{12}$, are shown, respectively, on Figs. 13 and 14.

In a similar way, by using formulae (8.56) and (8.57) that have been derived above for the first iteration, one can find the second iteration $x_2(\cdot, z_1, z_2, \lambda_1, \lambda_2)$ (i.e., $m = 2$ in formula

(8.12)), the components of which have the form

$$\begin{aligned} x_{21}(t, z_1, z_2, \lambda_1, \lambda_2) = z_1 &+ \frac{t^4}{192} + \frac{t^3}{48} z_2 + \frac{7t^2}{48} - \frac{17t^2}{32} z_2 - \frac{t^2}{2} z_1 \\ &- \frac{47t}{96} z_2 + \frac{25t}{192} - \frac{t}{2} z_1 - t\lambda_1, \quad t \in [0, 1], \end{aligned} \tag{8.66}$$

and

$$\begin{aligned} x_{22}(t, z_1, z_2, \lambda_1, \lambda_2) = z_2 &+ \frac{176471}{483840} t - \frac{28157}{23040} z_2 t + \frac{17t^3}{288} z_2 + \frac{t^2}{6} z_2 + \frac{t^2}{4} z_1 \\ &- \frac{t^6 z_2}{4608} - \frac{943t}{720} z_1 - \frac{1733}{5120} z_2^2 t + \frac{17t^5 z_2}{3840} + \frac{t^5 z_1}{240} - \frac{t^5 z_2^2}{2560} \\ &- \frac{t^4 z_2}{128} + \frac{17t^4 z_2^2}{1024} - \frac{107t^3 z_2^2}{512} + \frac{t^3 z_1}{18} - \frac{t^3 z_1^2}{6} + \frac{65t^3}{3456} \\ &- \frac{9t^2}{128} - \frac{t^5}{1440} - \frac{t^7}{32256} + \frac{t^2}{2} z_2 z_1 + \frac{17t^2}{32} z_2^2 + \frac{t^2}{4} \lambda_1 \\ &+ \frac{t}{6} z_1^2 - \frac{t}{4} \lambda_1 + \frac{t^4}{64} z_2 z_1 - \frac{31}{192} t z_2 z_1 \\ &- \frac{17}{48} t^3 z_2 z_1, \quad t \in [0, 1], \end{aligned} \tag{8.67}$$

and construct the corresponding system of Eqs. (8.43), (8.44) for determining the values of z_1, z_2, λ_1, and λ_2. The system mentioned, as the computation shows, has the approximate solution

$$z_1 \approx 0.06242777432, \qquad z_2 \approx 0.0001215436768, \tag{8.68}$$

$$\lambda_2 \approx 0.125, \qquad \lambda_1 \approx 0.09364329365. \tag{8.69}$$

Substituting (8.68), (8.69) in (8.66) and (8.67), we obtain the "second approximation," the components of which are given by the formulae

$$\begin{aligned} \bar{x}_{21}(t) = 0.06242777432 &+ \frac{t^4}{192} + 0.2532159933 \times 10^{-5} t^3 \\ &+ 0.1145548760 t^2 + 0.0052916467 t, \quad t \in [0, 1], \end{aligned} \tag{8.70}$$

and

$$\begin{aligned} \bar{x}_{22}(t) = &-0.2637666597 \times 10^{-7} t^6 + 0.2600559795 t + 0.0001215436768 \\ &- 0.8307568908 \times 10^{-6} t^4 + 0.02163102627 t^3 - 0.03127067403 t^2 \\ &- 0.0004337906399 t^5 - \frac{t^7}{32256}, \quad t \in [0, 1], \end{aligned} \tag{8.71}$$

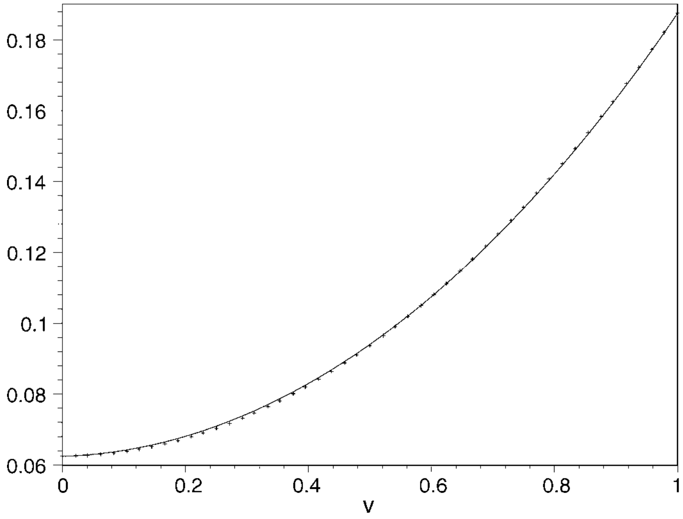


Fig. 11. The first component of solution (8.65) and its “second approximation” (8.70).

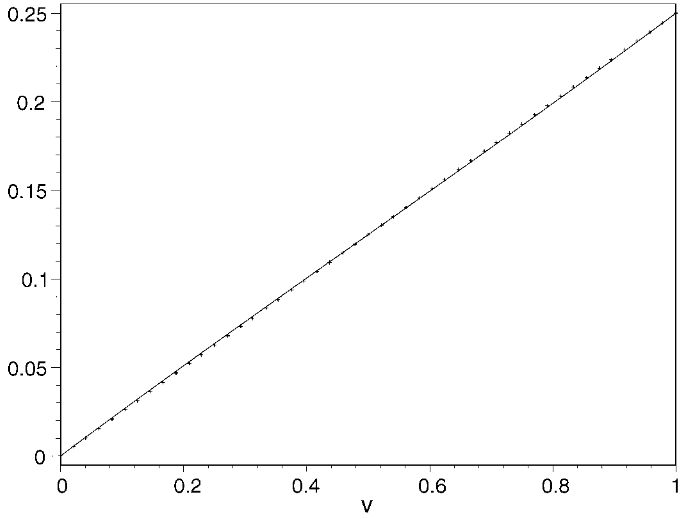


Fig. 12. The second component of solution (8.65) and its “second approximation” (8.71).

respectively.

Figure 11 shows the graph of the first component of the solution (8.65) of problem (8.46) (dotted line) and its “second approximation” (8.70) (solid line). On Fig. 12, the graphs of the second component of solution (8.65) (dotted line) and its “second approximation” (8.71) (solid line) are presented. The curves almost coincide with one another. The error of the second approximation is drawn on Figs. 15 and 16.

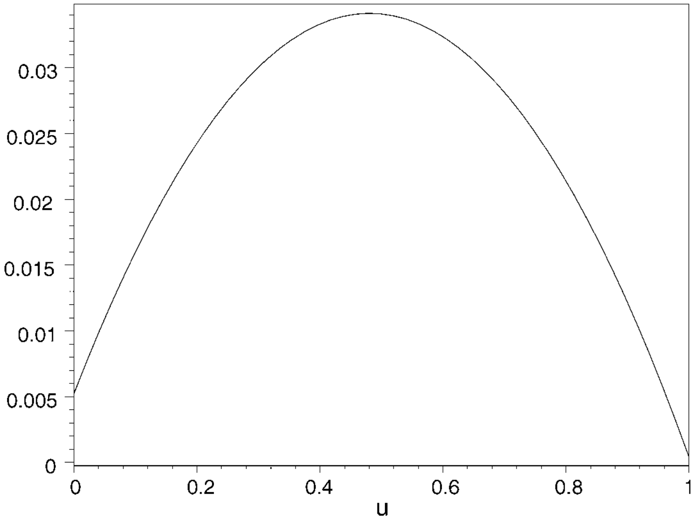


Fig. 13. Error of the first component of the “first approximation” (8.64) with respect to solution (8.65).

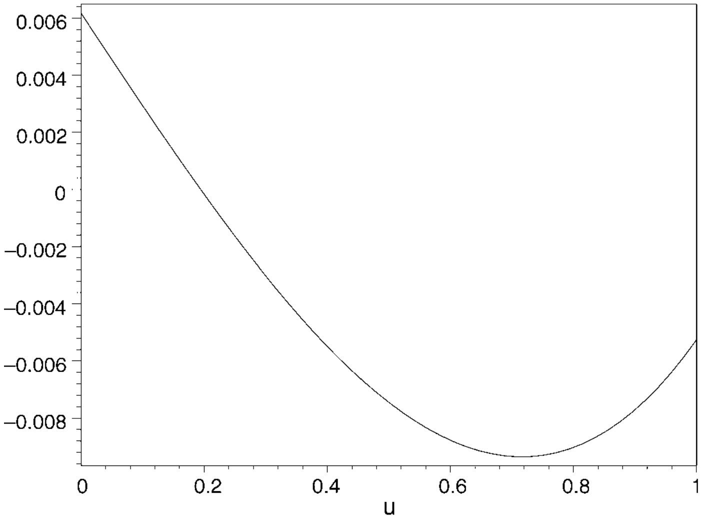


Fig. 14. Error of the second component of the “first approximation” (8.64) with respect to solution (8.65).

The computation shows that the third approximation constructed according to the scheme indicated above gives the absolute error which does not exceed 0.00035 for the first component and 0.0001 for the second one.

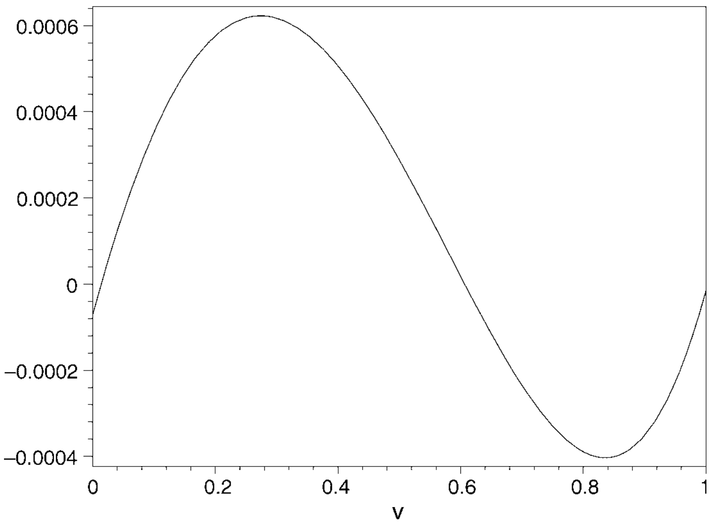


Fig. 15. Error of the “second approximation” (8.70) to the first component of solution (8.65).

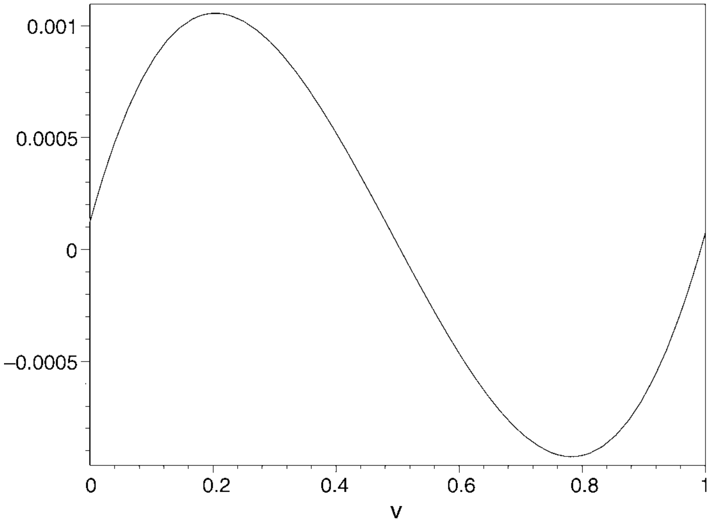


Fig. 16. Error of the “second approximation” (8.71) to the second component of solution (8.65).

9. Historical remarks

The basic concepts of the successive approximation method considered here had been first introduced by Samoilenko [82,83] in connection with the study of the periodic boundary value problem (2.1), (2.2) and was referred to as the “numerical-analytic method” and “periodic successive approximations method”. Some its aspects, as was noticed afterwards, have common points with other methods oriented to the investigation of periodic solutions of non-autonomous systems of differential equations.

The survey papers [74,76,77,75,78–80] analyse the history of the method and many interesting works inspired by it. Among the other works concerning the successive approximation method for boundary value problems, we mention the paper [74] and the Appendix of the book [70] which analyse the relation of the method to the Lyapunov–Schmidt equation [48], with the Cesari–Hale method for weakly non-linear differential equations [12], with the Cesari scheme for essentially non-linear systems of ordinary differential equations [11], and with the integral operator M_J introduced by Mawhin [81, pp. 98–107] as a mapping in C_T whose fixed points coincide with the solutions of a T-periodic boundary value problem for system (2.1). For example, the result on the existence of T-periodic solutions established in [81] (see Theorem 2.3, Corollary 4.4, and Remark 4.5 in [81]) can also be obtained with the help of the successive approximations method discussed here. In [70, Appendix] by Rontó, Samoilenko, and Trofimchuk, such a result is presented as Theorem 10, in the formulation borrowed from [72].

Note that a result close to the above-mentioned Theorem 10 from [70] had been first proved by Mawhin in 1969 for a locally Lipschitzian equation by using the Cesari method [81]. The statement indicated can be regarded as a consequence of the coincidence degree theory (see [81, p. 225] and [24]).

The improvement of the applicability conditions of the successive approximations method under consideration and estimates appearing in Theorems 3.17, 3.22, 3.23, 3.28, 3.36, 4.2, 6.1, and 8.3, Lemma 3.26, and inequality (4.25), involving the matrix Q of form (3.9), is closely related to the refinement of the constant q. Clearly, the more q is, the wider the domain of application of the iteration scheme will be. The following list describes, in the chronological order, the constants used in various estimates related to the successive approximation method:

1965 Samoilenko [82,83]; $q = 3.1$.

1976 Samoilenko and Ronto [88]; $q = \pi$.

1982 Samoilenko and Laptinskii [86]; $q = 3.416\ldots$. The proof has a gap which can be eliminated by using the Krein–Rutman theorem [38].

1985 Evhuta and Zabreiko [15,16]; $q = 3.416\ldots$.

1990 Trofimchuk [104]; $q = 3.416\ldots$. It is also proved that $q \leqslant 2\pi$. The estimates obtained are rather complicated and difficult to apply due to the need to compute the values of some integral operators.

1992 Kwapisz [42]; $q = \sqrt{10} = 3.1622\ldots$.

1996 Rontó and Mészáros [69]; $q = \frac{10}{3}$.

1996 Rontó, Rontó, Samoilenko, and Trofimchuk [64]; $q = 3.416\ldots$. Convenient techniques for the computation of constants in estimates are suggested.

2001 Perov, Dikareva, Oleinikova, and Portnov [57]; $q = \pi$. Three methods of proof are presented. In the first one, certain inaccuracies in some auxiliary statements used in [82,83,88] are corrected, whereas the second one uses the Steklov inequality [55, 56,14]. The third approach suggested allowed the authors to improve the a priori estimates by determining the maximal eigenvalues of certain integral operators.

2003 Samoilenko [85]; $q = 3.416\ldots$. The value of the constant was determined as the convergence radius of the Poisson–Abel series (5.64).

Using the idea of the successive approximations method first introduced for the study of the periodic solutions, many authors developed its versions adopted to the investigation of other kinds of boundary value problems for various types of ordinary differential, integro-differential, and more general functional differential equations and their systems. Many such papers are discussed in [74,76,77,75,78–80] and in the Appendix of [70].

Section A3.1 in [70, Appendix] analyses, in particular, the analogues of the successive approximations (3.29) constructed directly for the scalar second-order non-linear Lipschitzian differential equation

$$x''(t) + \omega^2 x(t) = f\big(t, x(t), x'(t)\big), \quad t \in [0, T], \tag{9.1}$$

without reduction of (9.1) to an equivalent system of first order equations. The convergence corresponding conditions turned out to be significantly worse compared to their counterparts for first order systems [70, p. 339]. The scheme of the successive approximations was adjusted for finding the periodic solutions of second order equations by Strizhak [102] and Samoilenko and Ronto [89]. Some modifications of the method were applied by Mitropolskii, Khoma, and Gromyak [50] in order to prove existence theorems to the T-periodic problem

$$x(0) = x(T), \qquad x'(0) = x'(T), \tag{9.2}$$

for Eq. (9.1) with $\omega \neq 0$, both in the non-resonance and resonance cases. The conditions of convergence of the scheme from [50] in the non-resonance case are rather restrictive.

Existence and uniqueness results for the non-resonance T-periodic boundary value problem (9.1), (9.2) were established by Chornyĭ [13]. Another scheme, which is also applicable in the cases where the Lipschitz constants for f in (9.1) are large, was suggested by Perestyuk and Ronto [53]. Problem (9.2) for some types of second order differential equations was studied by Shlapak [97,98], Sobkovich [99], Ronto and Martynyuk [73], and many others (see, e.g., [70, Appendix] for references).

Modified versions of the successive approximation scheme of form (3.29) for the study of T-periodic solutions of the delay differential system

$$x'(t) = f\big(t, x(t), x(t-\delta)\big), \quad t \in [0, T], \tag{9.3}$$

was first introduced by Martynyuk and Samoilenko [47], Mitropolʹskiĭ, Samoilenko, and Martynyuk [51] and the role of the delay, δ, pointed out.

Versions of the successive approximations obtained from (3.29) by a certain change of variables were used by Augustynowicz and Kwapisz [7], Kwapisz [42,44,43] for the investigation of solutions of the neutral type functional differential equation

$$x'(t) = f\big(t, x\big(\alpha(t)\big), x'\big(\beta(t)\big)\big), \quad t \in [0, T], \tag{9.4}$$

subjected to the two-point boundary conditions of the form

$$Ax(0) + Bx(T) = d,$$

where $\{A, B\} \subset \mathrm{GL}_n(\mathbb{R})$, $d \in \mathbb{R}^n$, and α, β are functions mapping the interval $[0, T]$ into itself (see Section A3.2.2 in [70, p. 369]).

Techniques based upon successive approximations of the kind indicated were used by Baĭnov and Sarafova [8] for integro-differential equations and Sarafova [95,96,9] Ronto [62] for differential equations with “maxima.” Zavalykut [107], Zavalykut and Nurzhanov [106] extended the techniques discussed to the more general differential operator equations of the form

$$x'(t) = (Ax)(t), \quad t \in (-\infty, \infty),$$

where A is a continuous mapping defined on the space C_T (see also [70, Section A3.3]).

Successive approximation techniques were applied for impulsive systems by Samoilenko [84], Samoilenko and Perestyuk [87], Kurbanbaev [40], Perestyuk and Shovkoplyas [54].

In [78], the application of the successive approximations method to difference equations is analysed. The paper [79] deals with its application to various non-linear boundary value problems (see also [60,39]). The paper [80] gives a survey of works devoted to successive approximation techniques for differential equations in Banach spaces, certain types of implicit differential equations, and to boundary value problems containing a control parameter (see also Akhmet and Zafer [4], Jankowski [34,32,33,31], Ronto, Ronto, and Shchobak [65]).

10. Exercises

Here, we suggest several exercises related to the techniques discussed.

EXERCISE 10.1. Study in which sense the family of sets $\{\Omega_{m,N}\}$ tends to the set Ω^* in Remark 3.37.

EXERCISE 10.2. For Eq. (4.96) in Example 4.9, by using the scheme (4.96), construct the functions $y_1(\theta, c, \mu)$ and $y_2(\theta, c, \mu)$. Write down the corresponding determining equations (4.45) and construct the 1st and the 2nd approximate periodic solutions (4.56) of Eq. (4.96). Find the corresponding values of c and T, compare them to ε and 2π (cf. (4.111)).

EXERCISE 10.3. Investigate the properties of the sequence $\alpha_{m,T}$, $m \geqslant 0$, given by formula (3.35). Prove the equality

$$\begin{aligned}\lim_{m\to\infty} \max_{t\in[0,T]} \frac{\alpha_{m,T}(t)}{\alpha_{m+1,T}(t)} &= \lim_{m\to\infty} \frac{\max_{t\in[0,T]} \alpha_{m,T}(t)}{\max_{t\in[0,T]} \alpha_{m+1,T}(t)} \\ &= \lim_{m\to\infty} \frac{\alpha_{m,T}(T/2)}{\alpha_{m+1,T}(T/2)} = \frac{1}{Tr(\mathcal{A})},\end{aligned}$$

where $\mathcal{A}$ is the linear operator in the space $C([0,1])$ given by the formula

$$(\mathcal{A}x)(s) := (1-s)\int_0^s x(\xi)\,\mathrm{d}\xi + s\int_s^1 x(\xi)\,\mathrm{d}\xi, \quad s \in [0,1]. \tag{10.1}$$

EXERCISE 10.4. Compute approximately the spectral radius $r(\mathcal{A})$ of operator (10.1). Hint: $r(\mathcal{A}) \approx 0.29$, see [60].

EXERCISE 10.5. Show how is it possible to relax condition (3.7). Hint: use Problems 10.3 and 10.4 and see [60] to show that the constant $3/10$ in (3.7) can be replaced by $r(\mathcal{A})$.

EXERCISE 10.6. For the periodic boundary value problem for the differential equation of the form

$$x'(t) = f\big(t, x, x'(t)\big), \quad t \in [0,T],$$

construct a suitable successive approximation method instead of that based upon the iteration scheme (3.29). Hint: see Section 3.10 in [80].

References

[1] A.A. Abramov, V.I. Ul'yanova, and L.F. Yukhno, *A method for solving the multiparameter eigenvalue problem for certain systems of differential equations*, Comput. Math. Math. Phys. **40** (1) (2000), 18–26.

[2] R.P. Agarwal and Donal O'Regan, *A survey of recent results for initial and boundary value problems singular in the dependent variable*, Handbook of Differential Equations, Elsevier/North-Holland, Amsterdam (2004), 1–68.

[3] K.T. Akhmedov, N.A. Svarichevskaya and M.A. Yagubov, *Approximate solution of a two-point boundary value problem with a parameter by the method of averaging functional corrections*, Dokl. Akad. Nauk Azerb. SSR **29** (8) (1973), 3–7.

[4] M. Akhmet and A. Zafer, *Controllability of two-point nonlinear boundary-value problems by the numerical-analytic method*, Appl. Math. Comput. **151** (3) (2004), 729–744. ISSN 0096-3003.

[5] C. Arzelà, *Works. Volume I, II (Opere. Volume I, II)*, Edizioni Cremonese, Bologna (1992).

[6] U.M. Ascher, R.M.M. Mattheij and R.D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Classics in Applied Mathematics, Vol. 13, SIAM, Philadelphia (1995).

[7] A. Augustynowicz and M. Kwapisz, *On a numerical-analytic method of solving of boundary value problem for functional-differential equation of neutral type*, Math. Nachr. **145** (1990), 255–269. ISSN 0025-584X.

[8] D.D. Baĭnov and G.H. Sarafova, *An application of the numerical-analytic method of A.M. Samoĭlenko for investigation of periodic systems of integro-differential equations*, Arch. Math. (Brno) **15** (2) (1979), 67–80, . ISSN 0044-8753.

[9] D.D. Baĭnov and G.H. Sarafova, *Application of the numerical-analytic method to the investigation of periodic systems of ordinary differential equations with maxima*, Rev. Roumaine Sci. Tech. Sér. Méc. Appl. **26** (3) (1981), 371–382. ISSN 0035-4074.

[10] T. Bhattacharyya, P.A. Binding and K. Seddighi. *Multiparameter Sturm–Liouville problems with eigenparameter dependent boundary conditions*, J. Math. Anal. Appl. **264** (2) (2001) 560–576.

[11] L. Cesari, *Functional Analysis And Periodic Solutions Of Non-Linear Differential Equations*, Vol. I, Interscience Publishers, John Wiley & Sons, Inc., New York (1963), 149–167.

[12] L. Cesari and J.K. Hale, *A new sufficient condition for periodic solutions of weakly nonlinear differential systems*, Proc. Amer. Math. Soc. **8** (1957), 757–764. ISSN 0002-9939.

[13] V.Z. Chornyĭ, *Periodic solutions of second-order wave integro-differential and difference equations*, Akad. Nauk Ukrainy Inst. Mat. Preprint (46) (1991), 31.

[14] L.Yu. Dikareva, *On A.M. Samoilenko's successive approximations method*, Trudy Mol. Uchen.: Sb. Nauchn. Trudov, Vol. 1, Voronezh. Gos. Univ., Voronezh (1999), 28–32.

[15] N.A. Evhuta and P.P. Zabreiko, *On the convergence of A.M. Samoilenko's successive approximation method of finding periodic solutions*, Dokl. Akad. Nauk BSSR **29** (1) (1985), 15–18 (in Russian).

[16] N.A. Evhuta and P.P. Zabreiko, *On A.M. Samoilenko's method of finding periodic solutions to quasilinear differential equations in Banach spaces*, Ukrain. Mat. Zh. **37** (2) (1985), 162–168.

[17] M. Farkas, *Periodic Motions*, Applied Mathematical Sciences, Vol. 104, Springer-Verlag, New York–London (1994).

[18] M. Fečkan, *A symmetry theorem for variational problems*, Nonlinear Anal. **16** (6) (1991), 499–506. ISSN 0362-546X.

[19] M. Fečkan, *Parametrized singular boundary value problems*, J. Math. Anal. Appl. **188** (2) (1994), 417–425. ISSN 0022-247X.

[20] M. Fečkan, *Parametrized singularly perturbed boundary value problems*, J. Math. Anal. Appl. *188* (2) (1994), 426–435. ISSN 0022-247X.

[21] M. Fečkan, *A symmetry theorem for ordinary differential equations*, Nonlinear Anal. **23** (11) (1994), 1437–1452. ISSN 0362-546X.

[22] M. Fečkan, *A symmetry theorem for dynamical systems*, Nonlinear Anal. **25** (6) (1995), 591–605. ISSN 0362-546X.

[23] M. Fečkan, *Minimal periods of periodic solutions*, Miskolc Math. Notes **7** (2) (2006), 123–141. ISSN 1787-2405.

[24] R.E. Gaines and J.L. Mawhin, *Coincidence degree, and nonlinear differential equations*, Lecture Notes in Mathematics, Vol. 568, Springer-Verlag, Berlin–Heidelberg–New York (1977).

[25] I.A. Goma, *On the theory of the solutions of a boundary value problem with a parameter*, Azerbaĭdžan. Gos. Univ. Učen. Zap. Ser. Fiz.-Mat. Nauk (1) (1976), 11–16.

[26] I.A. Goma, *Method of successive approximations in a two-point boundary problem with parameter*, Ukrain. Math. J. **29** (6) (1977), 594–599. ISSN 0041-6053.

[27] I.A. Goma, *A boundary value problem with a parameter for differential equations with retarded argument*, Azerbaĭdžan. Gos. Univ. Učen. Zap. Ser. Fiz.-Mat. Nauk (4) (1975) 25–33.

[28] J.K. Hale, *Oscillations in Nonlinear Systems*, McGraw-Hill, New York (1963).

[29] A. Heck, *Introduction to Maple*, third ed. Springer-Verlag, New York (2003). ISBN 0-387-00230-8.

[30] O. Hosabekov, *Sufficient conditions for the convergence of the Newton–Kantorovič method for a boundary value problem with a parameter*, Dokl. Akad. Nauk Tadžik. SSR **16** (8) (1973), 14–17. ISSN 0002-3469.

[31] T. Jankowski, *Numerical-analytic methods for differential-algebraic systems*, Acta Math. Hungar. **95** (3) (2002), 243–252. ISSN 0236-5294.

[32] T. Jankowski, *The application of numerical-analytic method for systems of differential equations with a parameter*, Ukrain. Math. J. **54** (4) (2002), 671–683. ISSN 0041-5995.

[33] T. Jankowski, *Monotone and numerical-analytic methods for differential equations*, Comput. Math. Appl. **45** (12) (2003), 1823–1828. ISSN 0898-1221.

[34] T. Jankowski, *Numerical-analytic method for implicit differential equations*, Math. Notes (Miskolc) **2** (2) (2001), 137–144. ISSN 1586-8850.

[35] H.B. Keller, *Numerical Methods for Two-Point Boundary-Value Problems*, Dover Publications, Inc., New York (1992).

[36] I.T. Kiguradze, *Some Singular Boundary Value Problems for Ordinary Differential Equations*, Izdat. Tbilis. Univ., Tbilisi (1975).

[37] M.A. Krasnoselskii, G.M. Vainikko, P.P. Zabreiko, Ya.B. Rutitskii and V.Ya. Stetsenko, *Approximate Solution of Operator Equations*, Wolters–Noordhoff Scientific Publications Ltd., Groningen (1972).

[38] M.G. Krein and M.A. Rutman, *Linear Operators Leaving Invariant a Cone in a Banach Space*, Amer. Math. Soc. Transl., Vol. 26, Amer. Math. Soc., Providence, RI (1950).

[39] O.O. Kurbanbaev, *A numerical-analytic method for boundary value problems with separated boundary conditions*, Uzbek. Mat. Zh. (3–4) (2001), 18–23.

[40] O.O. Kurbanbaev, *A numerical-analytic method for nonlinear boundary value problems with impulse action*, Uzbek. Mat. Zh. (4) (2004), 14–20.

[41] N.S. Kurpel′ and A.G. Marusjak, *A multipoint boundary value problem for differential equations with parameters*, Ukrain. Mat. Zh. **32** (2) (1980), 223–226, 285. ISSN 0041-6053.

[42] M. Kwapisz, *On modification of the integral equation of A.M.Samoilenko's numerical-analytic method*, Math. Nachr. **157** (1992), 125–135.

[43] M. Kwapisz, *Some remarks on an integral equation arising in applications of numerical-analytic method of solving of boundary value problems*, Ukrain. Math. J. **44** (1) (1992), 115–119. ISSN 0041-5995.

[44] M. Kwapisz, *On integral equations arising in numerical-analytic method of solving boundary value problems for differential-functional equations*, International Conference on Differential Equations, Vols. 1, 2 (Barcelona, 1991), World Sci. Publ., River Edge, NJ (1993), 671–677.

[45] A.Yu. Luchka, *The Method of Averaging Functional Corrections. Theory and Applications*. Translated from the Russian by Scripta Technica, Inc. Academic Press, New York (1965).

[46] A.Yu. Luchka and Yu.O. Zakharı̄ı̆chenko, *Investigation of systems of differential equations with parameters under impulse conditions and with restrictions*, Nelı̄nı̄ı̆nı̄ Koliv. **3** (2) (2000), 218–225. ISSN 1562-3076.

[47] D.I. Martynyuk and A.M. Samoilenko, *On periodic solutions of non-linear systems with delay*, Mat. Fiz. (3) (1967), 128–145.

[48] J. Mawhin, *Topological degree methods in nonlinear boundary value problems*, CBMS Regional Conference Series in Mathematics, Vol. 40, American Mathematical Society, Providence, RI (1979). ISBN 0-8218-1690-*. Expository lectures from the CBMS Regional Conference held at Harvey Mudd College, Claremont, CA, June 9–15, 1977.

[49] J. Mawhin and W. Walter, *A general symmetry principle and some implications*, J. Math. Anal. Appl. *186* (3) (1994), 778–798. ISSN 0022-247X.

[50] Yu. Mitropolskii, G. Khoma and M. Gromyak, *Asymptotic Methods for Investigating Quasiwave Equations of Hyperbolic Type*, Mathematics and its Applications, Vol. 402, Kluwer Academic Publishers Group, Dordrecht (1997). ISBN 0-7923-4529-0. Translated from the 1991 Russian original by Andrei Khruzin and revised by the authors.

[51] Yu.A. Mitropol′skiĭ, A.M. Samoilenko and D.I. Martynyuk, *Systems of Evolution Equations with Periodic and Quasiperiodic Coefficients*, Kluwer Academic Publishers, Dordrecht (1993). ISBN 0792320549.

[52] S.K. Ntouyas, *Nonlocal initial and boundary value problems: a survey*, Handbook of Differential Equations: Ordinary Differential Equations, Vol. II, Elsevier B.V., Amsterdam (2005), 461–557.

[53] N.A. Perestyuk and A.N. Ronto, *Numerical-analytic method for the equation of non-linear oscillator*, Publ. Univ. Miskolc Ser. D Nat. Sci. Math. 36 (2) (1996), 115–124. ISSN 1419-7006.

[54] N.A. Perestyuk and V.N. Shovkoplyas, *Periodic solutions of nonlinear differential equations with impulsive action*, Ukrain. Math. J. **31** (5) (1979), 408–414. ISSN 0041-5995.

[55] A.I. Perov, *Variational Methods in the Theory of Non-Linear Oscillations*, Izd. Voronezh. Gos. Univ., Voronezh (1981).

[56] A.I. Perov and A.A. Tananika, *A generalization of Wirtinger's inequality*, Differentsial′nye Uravneniya **22** (6) (1986), 1074–1076, 1103. ISSN 0374-0641.

[57] A.I. Perov, L.Yu. Dikareva, S.A. Oleinikova and M.M. Portnov, *On A.M. Samoilenko's method convergence condition*, Vestnik Voronezh. Gos. Univ. Ser. Fiz. Mat. (1) (2001), 111–119.

[58] I.G. Petrovski, *Ordinary Differential Equations*, revised English ed., Dover Publications, Inc., New York (1973). Translated and edited by Richard A. Silverman.

[59] I. Rachůnková, S. Staněk and M. Tvrdý, *Singularities and Laplacians in boundary value problems for nonlinear ordinary differential equations*, Handbook of Differential Equations, Elsevier/North-Holland, Amsterdam (2006), 607–723.

[60] A. Ronto and M. Rontó, *A note on the numerical-analytic method for nonlinear two-point boundary-value problems*, Nonlinear Oscilations **4** (1) (2001), 112–128. ISSN 1562-3076.

[61] A.N. Ronto, *On some boundary value problems for Lipschitz differential equations*, Nelı̄nı̄ı̆nı̄ Koliv. (1) (1998), 74–94. ISSN 1562-3076.

[62] A.N. Ronto, *On periodic solutions of systems with "maxima"*, Dopov. Nats. Akad. Nauk Ukr. Mat. Prirodozn. Tekh. Nauki (12) (1999), 27–31. ISSN 1025-6415.

[63] A.N. Ronto and N.I. Ronto, *On some symmetric properties of periodic solutions*, Nonlinear Oscillations **6** (1) (2003), 82–107. ISSN 1562-3076.

[64] A.N. Rontó, M. Rontó, A.M. Samoilenko and S.I. Trofimchuk, *On periodic solutions of autonomous difference equations*, Georgian Math. J. **8** (1) (2001), 135–164. ISSN 1072-947X.

[65] A.N. Ronto, M. Ronto and N.M. Shchobak, *On the parametrization of three-point nonlinear boundary value problems*, Nonlinear Oscillations **7** (3) (2004), 384–402. ISSN 1562-3076.

[66] A. Ronto and M. Rontó, *On the investigation of some boundary value problems with non-linear conditions*, Math. Notes (Miskolc) **1** (1) (2000), 43–55. ISSN 1586-8850.

[67] A. Ronto and M. Rontó, *On the* (τ, E) *property of periodic solutions*, Colloquium on Differential and Difference Equations, CDDE 2002 (Brno), Folia Fac. Sci. Natur. Univ. Masaryk. Brun. Math., Vol. 13, Masaryk Univ., Brno (2003), 247–267.

[68] M. Rontó, *On numerical-analytic method for BVPs with parameters*, Publ. Univ. Miskolc Ser. D Nat. Sci. Math. **36** (2) (1996), 125–132. ISSN 1419-7006.

[69] M. Rontó and J. Mészáros, *Some remarks on the convergence analysis of the numerical-analytic method based on successive approximations*, Ukrain. Math. J. 48 (1) (1996), 101–107.

[70] M. Ronto and A.M. Samoilenko, *Numerical-Analytic Methods in the Theory of Boundary-Value Problems*, World Scientific Publishing Co. Inc., River Edge, NJ (2000). ISBN 981-02-3676-X. With a preface by Yu.A. Mitropolsky and an appendix by the authors and S.I. Trofimchuk.

[71] M. Rontó and N. Shchobak, *On the numerical-analytic investigation of parametrized problems with nonlinear boundary conditions*, Nonlinear Oscillations **6** (4) (2003), 482–510.

[72] M. Rontó and S. Trofimchuk, *Numerical-analytic method for non-linear differential equations*, Publ. Univ. Miskolc Ser. D Nat. Sci. Math. **38** (1998), 97–116. ISSN 1419-7006.

[73] N.I. Ronto and O.M. Martynyuk, *Investigation of periodic solutions of countable second-order systems*, Ukrain. Math. J. **44** (1) (1992), 74–83. ISSN 0041-5995.

[74] N.I. Ronto, A.M. Samoilenko and S.I. Trofimchuk, *The theory of the numerical-analytic method: achievements and new directions of development. I*, Ukrain. Math. J. **50** (1) (1998), 116–135. ISSN 0041-6053.

[75] N.I. Ronto, A.M. Samoilenko and S.I. Trofimchuk, *The theory of the numerical-analytic method: achievements and new directions of development. IV*, Ukrain. Math. J. **50** (12) (1998), 1888–1907.

[76] N.I. Ronto, A.M. Samoilenko and S.I. Trofimchuk, *The theory of the numerical-analytic method: achievements and new directions of development. II*, Ukrain. Math. J. **50** (2) (1998), 255–277.

[77] N.I. Ronto, A.M. Samoilenko and S.I. Trofimchuk, *The theory of the numerical-analytic method: achievements and new directions of development. III*, Ukrain. Math. J. **50** (7) (1998), 1091–1114.

[78] N.I. Ronto, A.M. Samoilenko and S.I. Trofimchuk, *The theory of the numerical-analytic method: achievements and new directions of development. V*, Ukrain. Math. J. **51** (5) (1999), 735–747.

[79] N.I. Ronto, A.M. Samoilenko and S.I. Trofimchuk, *The theory of the numerical-analytic method: achievements and new directions of development. VI*, Ukrain. Math. J. **51** (7) (1999), 1079–1094.

[80] N.I. Ronto, A.M. Samoilenko and S.I. Trofimchuk, *The theory of the numerical-analytic method: achievements and new directions of development. VII*, Ukrain. Math. J. **51** (9) (1999), 1399–1418.

[81] N. Rouche and J. Mawhin, *Ordinary Differential Equations*, Surveys and Reference Works in Mathematics, Vol. 5, Pitman (Advanced Publishing Program), Boston, MA, 1980. ISBN 0-273-08419-4. Stability and periodic solutions. Translated from the French and with a preface by R.E. Gaines.

[82] A.M. Samoilenko, *A numerical-analytic method for investigation of periodic systems of ordinary differential equations. I*, Ukrain. Mat. Zh. **17** (4) (1965), 82–93. ISSN 0041-6053.

[83] A.M. Samoilenko, *A numerical-analytic method for investigation of periodic systems of ordinary differential equations. II*, Ukrain. Mat. Zh. **18** (2) (1966), 50–59. ISSN 0041-6053.

[84] A.M. Samoilenko, *On the justification of the averaging method for investigation of oscillations in impulsive systems*, Ukrain. Mat. Zh. **19** (5) (1967), 96–104. ISSN 0041-6053.

[85] A.M. Samoilenko, *On a sequence of polynomials and the radius of convergence of its Abel–Poisson sum*, Ukrain. Math. J. **55** (7) (2003), 1119–1130.

[86] A.M. Samoilenko and V.N. Laptinskii, *On estimates of periodic solutions of differential equations*, Dokl. Akad. Nauk Ukr. SSR (1) (1982), 30–32.

[87] A.M. Samoilenko and N.A. Perestyuk, *Impulsive Differential Equations*, World Scientific Series on Nonlinear Science. Series A: Monographs and Treatises, Vol. 14, World Scientific Publishing Co. Inc., River Edge, NJ (1995). ISBN 981-02-2416-8. With a preface by Yu.A. Mitropolʹskiĭ and a supplement by S.I. Trofimchuk, Translated from the Russian by Y. Chapovsky.

[88] A.M. Samoilenko and N.I. Ronto, *Numerical-Analytic Methods of Investigation of Periodic Solutions*. Višča Škola, Kiev (1976).

[89] A.M. Samoilenko and N.I. Ronto, *Numerical-Analytic Methods of Investigating Periodic Solutions*, Mir, Moscow (1979). Translated from the Russian by Vladimir Shokurov. With a foreword by Yu.A. Mitropolʹskiĭ.

[90] A.M. Samoilenko and N.I. Ronto, *Numerical-Analytic Methods for Investigation of Solutions of Boundary-Value Problems*, Naukova Dumka, Kiev (1986). With an English summary, Edited and with a preface by Yu.A. Mitropolʹskiĭ.

[91] A.M. Samoilenko and N.I. Ronto, *Numerical-Analytic Methods in the Theory of Boundary Value Problems for Ordinary Differential Equations*, Naukova Dumka, Kiev (1992).

[92] A.M. Samoilenko and Le Lyong Taĭ. *A method for studying boundary value problems with nonlinear boundary conditions*, Ukrain. Math. J. 42 (7) (1991), 844–850 1990. ISSN 0041-6053.

[93] A.M. Samoilenko, A.Yu. Luchka and V.V. Listopadova, *A boundary value problem for a system of differential equations with impulse action and parameters, and its solution by the projection method*, Dopov./Dokl. Akad. Nauk Ukraïni (1) (1994), 21–25. ISSN 0868-8044.

[94] A.M. Samoilenko, A.Yu. Luchka and V.V. Listopadova, *Application of iterative processes to a boundary value problem for a system of differential equations with impulse action and with parameters*, Dopov./Dokl. Akad. Nauk Ukraïni (2) (1994), 15–20. ISSN 0868-8044.

[95] G. Sarafova, *Periodic solutions of nonlinear systems of differential equations with maxima*, Plovdiv. Univ. Nauchn. Trud. **22** (1) (1985), 155–167, 1984.

[96] G.Kh. Sarafova and D.D. Baĭnov, *Application of A.M. Samoĭlenko's numerical-analytic method to the investigation of periodic linear differential equations with maxima*, Studia Sci. Math. Hungar. **17** (1–4) (1982), 221–228. ISSN 0081-6906.

[97] Ju.D. Shlapak, *Periodic solutions of nonlinear second-order equations which are not solved for the highest derivative*, Ukrain. Math. J. **26** (1974), 702–706. ISSN 0041-5995.

[98] Ju.D. Shlapak, *Periodic solutions of a second order linear system that is not solved for the derivatives*, Analytic Methods for the Study of the Solutions of Nonlinear Differential Equations (Russian), Izdanie Inst. Mat. Akad. Nauk Ukrain. SSR, Kiev (1975), 183–188.

[99] R.I. Sobkovich, *Periodic control problem for systems of second-order differential equations*, Analytic Methods of Nonlinear Mechanics, Akad. Nauk Ukrain. SSR Inst. Mat., Kiev (1981), 125–133, 171.

[100] R.I. Sobkovich, *Periodic solutions of systems of differential equations of first order with a parameter*, Ukrain. Math. J. **33** (6) (1981), 627–632. ISSN 0041-5995.

[101] R.I. Sobkovich, *A certain boundary-value problem for a first-order differential equation with several parameters*, Ukrain. Math. J. **34** (6) (1982), 650–656. ISSN 0041-5995.

[102] T.G. Strizhak, *On periodic solutions of systems of second order non-linear equations*, Asymptotic and Qualitative Methods in the Theory of Non-Linear Oscillations, Institute of Mathematics, Ukrainian Academy of Sciences, Kiev (1971), 35–46 (in Russian).

[103] G.P. Tolstov, *Fourier Series*, Dover Publications Inc., New York (1976), Second English translation, Translated from the Russian and with a preface by Richard A. Silverman.

[104] E.P. Trofimchuk, *Integral operators of the method of periodic successive approximations*, Matematicheskaya Fizika i Nelineinaya Mekhanika, **13** (1990), 31–36.

[105] A. Vanderbauwhede, *Local Bifurcation and Symmetry*, Research Notes in Mathematics, Vol. 75, Pitman (Advanced Publishing Program), Boston, MA (1982). ISBN 0-273-08569-7.

[106] G.D. Zavalykut and O.D. Nurzhanov, *A periodic boundary value problem for a class of differential-operator equations*, Ukrain. Math. J. **39** (3) (1987), 228–232. ISSN 0041-5995.

[107] G.D. Zavalykut, *A numerical-analytic method for investigating periodic solutions of a class of operator-differential equations*, Differentsialʹnye Uravneniya **19** (4) (1983), 569–575. ISSN 0374-0641.

CHAPTER 6

Analytic Ordinary Differential Equations and Their Local Classification

Henryk Żołądek*

Institute of Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland
E-mail: *zoladek@mimuw.edu.pl*

Contents

1. Introduction . . . 595
2. Linear differential equations . . . 595
    2.1. Regular singularities . . . 597
    2.2. Irregular singularities . . . 603
    2.3. Appendix. Proof of the normalization theorems . . . 612
3. Holomorphic vector fields . . . 616
    3.1. Resolution for planar vector fields and the center–focus problem . . . 616
    3.2. Poincaré–Dulac normal form . . . 629
    3.3. Ecalle–Voronin moduli . . . 632
    3.4. Martinet–Ramis moduli . . . 642
    3.5. Holonomy maps and normal forms for resonant saddles . . . 647
    3.6. Geometric realization of Martinet–Ramis moduli . . . 651
    3.7. Linearization . . . 658
    3.8. Appendix 1. Complex structures and almost complex structures . . . 665
    3.9. Appendix 2. Proof of the Hukuhara–Kimura–Matuda theorem . . . 669
4. Bogdanov–Takens singularity . . . 671
    4.1. Formal orbital normal form . . . 671
    4.2. Analytic Takens prenormal form . . . 681
References . . . 684

Abstract

The article is devoted to the local theory of analytic differential equations. We describe classification of meromorphic systems $\dot{x} = A(t)x$ near regular and irregular singular point. We present a local theory of non-linear holomorphic equations $\dot{x} = V(x)$, with the proof of

---

[*] Supported by Polish MNiSzW Grant No. 1 P03A 015 29.

HANDBOOK OF DIFFERENTIAL EQUATIONS
Ordinary Differential Equations, volume 4
Edited by F. Battelli and M. Fečkan
© 2008 Elsevier B.V. All rights reserved

the resolution theorem and its application to the center–focus problem, with Ecalle–Voronin–Martinet–Ramis moduli and with Briuno–Yoccoz classification of non-resonant saddles. Finally, we present formal classification of nilpotent singularities and prove analyticity of the Takens prenormal form.

MSC: primary 05C38, 15A15; secondary 05A15, 15A18

1. Introduction

The origins of the theory of Ordinary Differential Equations lie in the classical mechanics. The second Newton's principle of dynamics states that the further configurations of a mechanical system are determined by the initial positions and velocities via a second order differential equation. Therefore the theory is essentially real, because the time is real. But in the conservative mechanical systems the direction of time is not that important, the corresponding Newton's equations are time–reversible. Even in the non-conservative case the knowledge of the initial positions and velocities at $t = 0$ allows to reconstruct the history of the system.

Therefore, in the philosophical sense, the theory of Ordinary Differential Equations resembles other mathematical theories: the Riemannian geometry, the real algebraic geometry, the Morse theory of functions, etc. In all these theories the extension to the complex domain turned out very effective. For example, the Morse theory describes the change of topology of level surfaces of a function during passage through a critical value; but in the case of holomorphic function one can pass along a loop around the critical value and the resulting change in the topology of level surface is described by the Picard–Lefschetz formula.

It is not surprising that also the theory of Ordinary Differential Equations has its holomorphic analogue. It is the analytic theory of Ordinary Differential Equations, which does not concentrate on solutions in the form of power series, but rather treats the time as complex and represents solutions as embeddings of Riemann surfaces into a complex phase space. This approach was used already in the works of K. Gauss (on the hypergeometric equations) and of B. Riemann (in the so-called Riemann–Hilbert problem).

The aim of this article is to present the local theory of analytic Ordinary Differential Equations. It is an analogue of the singularity theory of holomorphic functions (generalization of the Morse Lemma). It turns out that the classification of singularities of differential equations is more complicated than in the geometrical case. It is related with the Stokes phenomenon for linear differential systems, when some asymptotic expansions of solutions are valid only in some sectors (in the complex time plane) and the 'differences' between the expansions in adjacent sectors constitute functional moduli of the singularity. There exists also a non-linear version of the Stokes phenomenon in the orbital classification of holomorphic planar vector fields near elementary singular point. In the case of non-elementary singularity the situation is even more complicated, e.g. in the nilpotent case only the formal orbital classification is known.

2. Linear differential equations

The subject in this chapter are the non-autonomous linear differential systems

$$\dot{z} = A(t)z, \quad z \in \mathbb{C}^n, \tag{2.1}$$

and the linear higher order differential equations

$$x^{(n)} + a_1(t)x^{(n-1)} + \cdots + a_n(t)x = 0, \quad x \in \mathbb{C}. \tag{2.2}$$

In the above two equations the ‘time’ t usually takes value in the complex plane $\mathbb{C}$. Often we will consider Eqs. (2.1) and (2.2) locally, then $t \in (\mathbb{C}, 0)$.

The entries of the matrix $A(t)$ and the coefficients $a_i(t)$ are meromorphic functions.

The equations of this type are met very often. For example, the *Gauss hypergeometric equation*

$$t(t-1)\ddot{x} + \big[(\alpha+\beta+1)t - \gamma\big]\dot{x} + \alpha\beta x = 0 \tag{2.3}$$

belongs to this class. Probably it is the most important linear differential equation. Most of the differential equations appearing in the theory of special functions, e.g. the Bessel equation, the Weber equation, the Legendre equation, the Hermite equation, the Jacobi equation, are obtained from the hypergeometric equation by application of some limit process.

Algebraic functions are also solutions of such differential equations. It follows from the Riemann theorem about multivalued holomorphic functions with regular singularities (see [89]). For example, the function $x(t) = \sqrt{t} + \sqrt[3]{t-1}$ satisfies the equation

$$(t-3)\frac{\mathrm{d}}{\mathrm{d}t}\big(6t(t-1)\dot{x}\big) = 3(3t-1)\big[3(t-1)\dot{x} - x\big] + 2(4t-3)[2t\dot{x} - x].$$

DEFINITION 2.1. A point t_* at which the matrix $A(t)$ in (2.1) or one of the coefficients $a_i(t)$ in (2.2) has a pole is the *singular* point for the equation. In the case of local situation we shall assume that $t_* = 0$.

The singular point $t = 0$ of Eq. (2.1) or (2.2) is called *regular* if any its solution $\phi(t)$ has at most polynomial growth in any sector with vertex at $t = 0$. It means that

$$|\phi(t)| < C_N/|t|^N, \quad |t| \to 0,\ \alpha < \arg t < \beta.$$

Otherwise the singular point is called *irregular*.

The division of the singular points of systems (2.1) into regular and irregular is very rough. Our aim is to classify such systems with respect to equivalence relations defined below.

DEFINITION 2.2. Two local differential systems $\dot{z}_1 = A_1(t)z_1$ and $\dot{z}_2 = A_2(t)z_2$ are called *holomorphically equivalent* near the singular point $t = 0$ if there exists a holomorphic matrix $H(t)$ such that, after the change $z_2 = H(t)z_1$, the first system is transformed to the other system. In other words,

$$A_2 = \dot{H}H^{-1} + HA_1H^{-1}.$$

These systems are called *meromorphically equivalent* (respectively, *formally equivalent* or *formally meromorphically equivalent)* if there exists a meromorphic matrix $H(t)$ (respectively a formal series $H(t) \sim \sum_{j=0}^{\infty} H_j t^j$ or $\sum_{j=-m}^{\infty} H_j t^j$) which transforms one system to another system. In the formal cases this holds at the level of formal expansions of these systems in powers of t.

2.1. *Regular singularities*

We begin the study of singularities of the meromorphic differential equations with the one-dimensional phase space, $n = 1$. It means that we have the equation

$$t^r \dot{z} = a(t)z$$

where $a(t)$ is an analytic function near 0.

If $r = 0$ then the solutions are analytic; in fact, the point $t = 0$ is not singular for this equation.

If $r > 0$ then we assume that the coefficient $a_0 \neq 0$ in the expansion $a(t) = a_0 + a_1 t + \cdots$. The general solution is written in the form

$$z(t) = C \cdot \exp\left[\int^t a(s)s^{-r}\,\mathrm{d}s\right].$$

If $r = 1$ then we get $z(t) = Ct^{a_0} \times \Psi(t)$, where Ψ is an analytic function. This means the regularity of the singular point.

If $r > 1$ then $z(t) = C\mathrm{e}^{R(1/t)}t^b \times \Psi(t)$, Ψ analytic, where

$$R(x) = \frac{a_0}{1-r}x^{r-1} + \frac{a_1}{2-r}x^{r-2} + \cdots - a_{r-2}x,$$

$b = a_{r-1}$. Here the singularity is irregular, because for $\mathrm{Re}(a_0 t^{1-r}) < 0$ the solution diverge faster than any power of t. We proved the following

PROPOSITION 2.3. *If $n = 1$ then the singular point of Eq.* (2.1) *is regular if and only if the function $A(t)$ has pole of order exactly* 1.

In general, Eq. (2.1) *is analytically equivalent to the equation $t^r \dot{z} = (\sum_0^{r-1} a_j t^j)z$.*

Consider now the case of n-dimensional system with the following singularity

$$t\dot{z} = Az,$$

where A is a constant matrix. The fundamental matrix of solutions of this system is

$$t^A = \mathrm{e}^{A\ln t},$$

which grows at most polynomially (regularity). This fact is generalized in the following way.

PROPOSITION 2.4. *If $A(t) = A_0/t + A_1(t)$, where $A_0 = \mathrm{const} \neq 0$ and $A_1(t)$ is a holomorphic matrix-valued function, then the point $t = 0$ is regular.*

PROOF. We put $t = \rho\mathrm{e}^{\mathrm{i}\theta}$, where $\theta = \mathrm{const}$ and ρ tends to 0. Let $\phi(t) \not\equiv 0$ be a solution, and $r(\rho) = |\phi(t)|_{\theta=\mathrm{const}}|$.

From the identity $\frac{\partial \phi}{\partial \rho} = \mathrm{e}^{\mathrm{i}\theta}\dot{\phi} = (\tilde{A}(t)/\rho)\phi$ (where $\tilde{A}$ is holomorphic) we get the inequality $\frac{\mathrm{d}r}{\mathrm{d}\rho} > -\frac{Kr}{\rho}$ for $0 < \rho < \rho_0$, with some constant K and ρ_0.

From this we see that the graph of the function $r(\rho)$ must lie in the domain $\{r < \rho^{-K}\}$ for ρ near 0. □

DEFINITION 2.5. System (2.1), i.e. $\dot{z} = A(t)z$, has singularity of the *Fuchs type* at $t = 0$ if $A(t)$ has simple pole at 0. Equation (2.2), i.e. $x^{(n)} + \cdots + a_n(t)x = 0$, has singularity of the *Fuchs type* at $t = 0$ if all the functions $t^j a_j(t)$ are holomorphic near 0.

Proposition 2.4 says that, if the system (2.1) has singularity of the Fuchs type, then this singularity is regular. The following example shows that the converse is not true.

EXAMPLE 2.6. The *Euler equation*

$$t^n x^{(n)} + c_1 t^{n-1} x^{(n-1)} + \cdots + c_n x = 0,$$

where c_i are constants, by means of the change $t = \mathrm{e}^u$, $\frac{\mathrm{d}}{\mathrm{d}t} = \mathrm{e}^{-u}\frac{\mathrm{d}}{\mathrm{d}u}$, $\frac{\mathrm{d}^2}{\mathrm{d}t^2} = \mathrm{e}^{-2u}(\frac{\mathrm{d}^2}{\mathrm{d}u^2} - \frac{\mathrm{d}}{\mathrm{d}u}), \ldots$ is reduced to the equation

$$\frac{\mathrm{d}^n x}{\mathrm{d}u^n} + d_1 \frac{\mathrm{d}^{n-1} x}{\mathrm{d}u^{n-1}} + \cdots + d_0 x = 0$$

with constant coefficients. Its general solution has the form $x = \sum a_{\alpha,k} \mathrm{e}^{\alpha u} u^k = \sum a_{\alpha,k} t^\alpha (\ln t)^k$. Thus the point $t = 0$ is regular. On the other hand, if we rewrite the Euler equation in form of the linear system

$$\dot{z} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -c_n/t^n & -c_{n-1}/t^{n-1} & \ldots & \ldots & -c_1/t \end{pmatrix} z$$

(where $z_1 = x, z_2 = \dot{x}, \ldots$) then the corresponding matrix $A(t)$ has non-simple pole.

In the case of linear equation of n-th order the situation is much more clear.

THEOREM 2.7. *The singular point $t = 0$ for Eq.* (1.2) *is regular if and only if it is of Fuchs type.*

Before proof of this theorem we need one important definition.

DEFINITION 2.8. Let t_0 be a non-singular point of Eq. (2.1) or (2.2), defined near the singularity $t = 0$. Let $\mathcal{V}$ be the space of its solutions defined near t_0. As t turns around the point 0, with the beginning and end at t_0, the solutions are prolonged analytically and define the *monodromy operator* $\mathcal{M} : \mathcal{V} \to \mathcal{V}$: if $\phi(t)$ is a (local) solution then $\mathcal{M}\phi(t) = \phi(\mathrm{e}^{2\pi \mathrm{i}} t)$.

Let $\phi_1(t), \dots, \phi_n(t)$ be a basis of $\mathcal{V}$. This basis defines the *fundamental matrix* $\mathcal{F} = (\phi_1, \dots, \phi_n)$ in the case (2.1) and

$$\mathcal{F} = \begin{pmatrix} \phi_1 & \dots & \phi_n \\ \vdots & \ddots & \vdots \\ \phi_1^{(n-1)} & \dots & \phi_n^{(n-1)} \end{pmatrix}$$

in the case (2.2). Then the monodromy operator, written in this basis, is the *monodromy matrix* M satisfying $\mathcal{F}(t)M = \mathcal{F}(\mathrm{e}^{2\pi \mathrm{i}}t)$. Here the matrix M depends on the choice of $\mathcal{F}$ but locally does not depend on the point t_0.

PROOF OF THEOREM 2.7. (We follow the proof of Theorem 5.2 in the book of E. Coddington and N. Levinson [24].)

Assume that the Fuchs conditions, i.e. all $t^j a_j(t)$ are holomorphic, hold. We are going to show that the singularity $t = 0$ is regular.

Following the latter example we perform a transformation reducing equation (2.2) to a system (1.1) with the Fuchs type singularity. Then Proposition 2.4 will give us the regularity property.

The change is following

$$z_1 = x, \quad z_2 = (t\,\mathrm{d}/\mathrm{d}t)x, \quad z_3 = (t\,\mathrm{d}/\mathrm{d}t)^2 x, \dots, z_n = (t\,\mathrm{d}/\mathrm{d}t)^{n-1}x.$$

It gives $\dot{z}_1 = (1/t)z_2$, $\dot{z}_2 = (1/t)z_3, \dots, \dot{z}_{n-1} = (1/t)z_n$ and $z_j = t^j x^{(j)}$ plus a combination of $t^k x^{(k)}$, $k < j$. Thus $t^j x^{(j)}$ can be expressed as linear combination of z_k, $k \leqslant j$.

The last equation from the promised system is $\dot{z}_n = \frac{\mathrm{d}}{\mathrm{d}t}(t^{n-1}x^{(n-1)}) +$ plus combination of $\dot{z}_i$, $i < n$. The term $\frac{\mathrm{d}}{\mathrm{d}t}(t^{n-1}x^{(n-1)})$ contains $(n-1)z_n/t$ and $t^{n-1}x^{(n)} = -t^{n-1}\sum a_j x^{(n-j)} = -t^{-1}\sum (a_j t^j)(x^{(n-j)}t^{n-j})$. The functions $a_j t^j$ are holomorphic by the assumption and $x^{(n-j)}t^{n-j}$ are expressed by means of z_i's. All this shows that $t\dot{z}_n$ is analytic in t and z.

Assume that $t = 0$ is a regular point. We want to show that the functions $b_j(t) = a_j(t)t^j$ are analytic.

We use the monodromy operator. If $\mathcal{F}$ is the fundamental matrix and $\mathcal{F} \to \mathcal{F}M$ is the monodromy operation and $G(t) = t^{-\ln M/2\pi \mathrm{i}}$, then the matrix-function $\mathcal{F}G$ is univalent. By the regularity of $\mathcal{F}$ it is a meromorphic function. Thus $\mathcal{F} =$ (meromorphic matrix) $\times$ $t^{\ln M/2\pi \mathrm{i}}$.

The monodromy matrix M has always some (left) eigenvector $v = (v_1, \dots, v_n)$, $vM = \lambda v$. Using it we can find a solution $\phi(t) = \sum v_i \phi_i$ of the form $t^\mu \times$ (analytic function), where the exponent μ satisfies $\mathrm{e}^{2\pi \mathrm{i}\mu} = \lambda$. Note that the function $\phi(t)$ does not contain logarithms in its expansion.

Other solutions are searched for in the form $x = \phi \cdot y$. For $u = \dot{y}$ we obtain an equation of order $n - 1$.

Indeed, if we define the differential operator $P = (\mathrm{d}/\mathrm{d}t)^n + \sum a_j(\mathrm{d}/\mathrm{d}t)^{n-j}$ then $P(\phi y) = P(\phi) \cdot y$ plus expressions depending on $\dot{y}$. The original equation for x gives an equation is of the form

$$\sum c_j(t)u^{(j)} = 0,$$

where $c_0 = a_0 = 1$ and

$$c_j = a_j + \text{const} \cdot a_{j-1} \cdot \dot{\phi}/\phi + \text{const} \cdot a_{j-2} \cdot \ddot{\phi}/\phi + \cdots, \quad \phi^{(l)}/\phi \sim \text{const} \cdot t^{-l}.$$

Next we use induction with respect to the order n. In the case $n = 1$ the statement was proved in Proposition 2.3; the regularity implies Fuchs property.

Assume that we know that any regular equation of order $n - 1$ satisfies the Fuchs conditions and consider our system.

Let ψ_i be the system of independent solutions of the equation for u. Then $\psi_i = \frac{\mathrm{d}}{\mathrm{d}t}(\phi_i/\phi)$, where ϕ_i are independent solutions of the equation for x. Because all ϕ_i are regular (by assumption) and $\phi \sim t^{\mu}$ then also $\psi_i(t)$ behave regularly. Therefore the equation for u has regular singularity at $t = 0$. By the induction assumption the coefficients $c_j(t)$ have good behavior, $c_j t^j$ are analytic. Because of the above relation between a_j's and c_k's we find that also the functions $a_j t^j$ are holomorphic. □

PROPOSITION 2.9. *The point $t = 0$ is a regular singular point for the system $\dot{z} = A(t)z$ if and only if this system is meromorphically equivalent to the system $\dot{z} = (C/t)z$ where the matrix C is the logarithm of the monodromy matrix M, $M = \mathrm{e}^{2\pi \mathrm{i} C}$.*

PROOF. Let $\mathcal{F}(t)$ be the fundamental matrix of the system with the matrix A. Define $H(t) = \mathcal{F}(t)t^{-C}$. It is univalent and meromorphic matrix-valued function (by the regularity). It is easy to check that $H(t)$ is the conjugating transformation.

The reverse implication is obvious. □

Our next aim is to improve the latter result. Let us rewrite system (2.1) in the form

$$t^r \dot{z} = B(t)z, \quad B(0) \neq 0, \tag{2.4}$$

with an analytic matrix $B(t)$. This system is associated with the following autonomous system in the extended phase space

$$z' = B(t)z, \quad t' = t^r, \tag{2.5}$$

where $' = \mathrm{d}/\mathrm{d}\tau$. Note that the graphs of solutions to Eq. (2.4) form the phase curves of Eq. (2.5). Assume also that $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the matrix $B(0)$.

The point $z = 0$, $t = 0$ is an equilibrium point of system (2.5). If $r = 1$, i.e. in the Fuchs case, then the eigenvalues of the linear part at $(0, 0)$ of system (2.5) are $\lambda_1, \ldots, \lambda_n, 1$. If $r > 1$, then the eigenvalues of system (2.5) are $\lambda_1, \ldots, \lambda_n, 0$.

DEFINITION 2.10. The system $\lambda_1, \dots, \lambda_n$ of eigenvalues of $B(0)$ is *resonant* if:
(a) $\lambda_i - \lambda_j \in \mathbb{Z}$ for some $i \neq j$ in the case $r = 1$;
(b) $\lambda_i - \lambda_j = 0$ for some $i \neq j$ in the case $r > 1$.

Now we use some results from the theory of normal forms for germs of analytic vector fields $(\mathbb{C}^n, 0)$

$$\dot{x} = V(x) = Ax + \cdots,$$

where A is a constant matrix of the linearization of the vector field at 0 (with the eigenvalues $\mu_1, \dots \mu_n$). The *Poincaré–Dulac theorem* (Theorem 3.17 from the next section) says that:

there exists a change $x \to y = G(x)$, where $G(x)$ is given by formal power series, which transforms the above vector field to a formal vector field $\dot{y} = W(y)$, such that the expansion of each component $W_j(y)$ of W is of the form

$$W_i(y) = \mu_i y_i + \sum a_{i,k} y^k$$

and the summation runs over multiindices $k = (k_1, \dots, k_n)$ which satisfy the resonance relations

$$\mu_i = \mu_1 k_1 + \cdots + \mu_n k_n.$$

Moreover, if the components $V_1(x), \dots, V_k(x)$ are linear with respect to the variables $x_1, \dots, x_k$ then also the components $G_1(x), \dots, G_k(x)$ and $W_1(y), \dots, W_k(y)$ are linear with respect to suitable variables.

We apply this result to vector field (2.5), where we use changes of the form $(z, t) \to (H(t)z, t)$. (We note that the idea to apply the Poincaré–Dulac theorem to linear systems was used firstly in the review article by Arnold and Il'yashenko [5].) The reader can notice that the resonant relations from the Poincaré–Dulac theorem correspond to the relations between the eigenvalues of the matrix $B(0)$ defined in Definition 2.10. The corresponding resonant terms in the normal form are

$$t^l z_j e_i,$$

i.e. if $\lambda_i = \lambda_j + l$, where $(e_1, \dots, e_n)$ is the eigenvector basis for the matrix $B(0)$.

COROLLARY 2.11. *If the singular point $t = 0$, with $r = 1$ or $r > 1$, is non-resonant then the Poincaré–Dulac theorem says that system* (2.4) *is formally equivalent to the system $t^r \dot{z} = B(0)z$.*

On the other hand, Proposition 2.9 says about the meromorphic equivalence. But the formal equivalence means that, in the meromorphic matrix $H(t) = H_{-d} t^{-d} + \cdots$, all the terms with negative power of t vanish. This gives the following result.

THEOREM 2.12. *The system $t\dot{z} = B(t)z$ with non-resonant Fuchsian singular point is analytically equivalent to $t\dot{z} = B(0)z$.*

Of course, in the above theorem we can assume that the matrix $B(0)$ is diagonal. Hence the normal form variables are completely separated. In the resonant case we can divide the eigenvalues into groups such that:

- the differences $\lambda_i - \lambda_j \notin \mathbb{Z}$ for eigenvalues from different groups,
- the differences $\lambda_i - \lambda_j \in \mathbb{Z}$ for eigenvalues from the same group.

We order the eigenvalues in each group such that the differences $\lambda_i - \lambda_j$ are non-negative integers for $j < i$, $\lambda_i - \lambda_j \in \mathbb{Z}_+$. We have the following result whose proof is postponed to Subsection 2.3 below.

THEOREM 2.13. *If a Fuchsian singular point is resonant then it is analytically equivalent to a system of the form*

$$t\dot{z}_i = \lambda_i z_i + \sum_{\lambda_i - \lambda_j \in \mathbb{Z}_+} a_{ij} t^{\lambda_i - \lambda_j} z_j, \quad i = 1, \dots, n, \ a_{ij} = \text{const.}$$

REMARK 2.14. The system from the thesis of Theorem 2.13 can be easily integrated. Its general solution has the form $z_i(t) = t^{\lambda_i} \times$ (polynomial of $\log t$). For more details we refer the reader to [24].

REMARK 2.15. If a singular point $t = 0$ is regular for Eq. (2.2) and $a_i(t) = \nu_i t^{-i} + \cdots$ then we can seek solutions ϕ such that $\phi(t) \sim Ct^{\lambda}$. We arrive to the equation

$$P(\lambda) = \lambda(\lambda - 1) \cdots (\lambda - n + 1) + \lambda \cdots (\lambda - n + 2) \cdot \nu_1 + \cdots + \nu_n = 0$$

which is called the *defining equation* for (2.2).

The defining equation corresponds to the characteristic equation for the matrix $B(0)$ from the system (2.4). Namely, we apply the transformation of Eq. (2.2) to a Fuchsian system from the proof of Theorem 2.7. This Fuchsian system has the form (2.4) with $r = 1$.

The roots of the defining equation allow to determine the first terms of the asymptotic of solutions of the differential equation.

If the system $(\lambda_i, \dots, \lambda_n)$ of roots is not resonant in the sense of Definition 2.10 (for $r = 1$) then the general solution is a linear combination of functions of the form $\phi_i(t) = t^{\lambda_i} \times$ (analytic function). Otherwise we have also functions of the form $t^{\lambda_i} \times (\log t)^j \times$ (analytic function).

EXAMPLE 2.16. Consider the Gauss hypergeometric equation (2.3), i.e. $t(t-1)\ddot{x} + [(\alpha + \beta + 1)t - \gamma]\dot{x} + \alpha\beta x = 0$. It has singular points at $t = 0$, $t = 1$ and also at $t = \infty$, all Fuchsian.

The defining equation at $t = 0$ is $\lambda(\lambda - 1 + \gamma) = 0$ and the two linearly independent solutions are $\phi_1 = F(\alpha, \beta; \gamma; t)$ and $\phi_2 = t^{1-\gamma} F(\alpha - \gamma + 1, \beta - \gamma + 1; 2 - \gamma; t)$, where $F(\alpha, \beta; \gamma; t) = \sum_{n=0}^{\infty} \frac{(\alpha)_n (\beta)_n}{(\gamma)_n n!} t^n$, $(a)_n = a(a+1) \cdots (a + n - 1)$, is the so-called the *hypergeometric series*. This holds when γ is not integer and for integer γ one of these functions is not defined. For example, when $\gamma = 0$ the solution ϕ_2 is the same but $\phi_1 = tF(\alpha, \beta; 1; t) \cdot \log t + G(t)$, where $G(t)$ is some uniquely defined series.

For $t = 1$ the defining equation is $\lambda(\lambda - \gamma + \alpha + \beta) = 0$ and the (generically) independent solutions are $F(\alpha, \beta; \alpha + \beta + 1 - \gamma; 1 - t)$ and $t^{\gamma - \alpha - \beta} F(\gamma - \beta, \gamma - \alpha; 1 - \alpha - \beta + \gamma; 1 - t)$.

Near $t = \infty$ one rewrites the Gauss equation in the local variable $\tau = 1/t$. The defining equation is $(\lambda - \alpha)(\lambda - \beta) = 0$ and the functions $\tau^\alpha F(\alpha, 1 - \gamma + \alpha; 1 + \alpha - \beta; \tau)$, $\tau^{-\beta} F(\beta, 1 - \gamma + \beta; 1 + \beta - \alpha; \tau)$ constitute the basis of the space of solutions.

2.2. *Irregular singularities*

Firstly we find a formal normal form for linear systems (2.4) with an irregular singularity.

THEOREM 2.17. *If $r > 1$ and $B(0)$ is non-resonant then the system $t^r \dot{z} = B(t)z$ is formally equivalent to the system*

$$t^r \dot{w} = D(t)w, \tag{2.6}$$

where the matrix $D(t)$ is diagonal, $D = \operatorname{diag}(d_1(t), \dots, d_n(t))$, and $d_i(t)$ are polynomials of degree at most $r - 1$.

PROOF. By the Poincaré–Dulac theorem all the terms $t^k z_j e_i$, $i \neq j$, can be canceled (because they are non-resonant). There remain only the terms $t^k z_i e_i$ which are diagonal. Next one applies Proposition 2.3 to each component. □

If we write $D(t)t^{-r} = D_r t^{-r} + D_{r-1} t^{-r+1} + \dots + D_2 t^{-2} + E t^{-1}$, where D_j and E are diagonal matrices, then the formal fundamental matrix of the system from Theorem 2.17 takes the form $\mathcal{F}(t) = \widehat{G}(t) t^E \exp[\sum D_j t^{j+1}/(j+1)]$, where $\widehat{G}$ is a formal power series.

In the resonant case there exists also a diagonal formal normal form, but in another category. One should change the local ring $\mathbb{C}[[t]]$ (of formal power series in t) by the ring $\mathbb{C}[[t^{1/b}]]$, where b is some positive integer.

EXAMPLE 2.18. (See [35,82].) The system

$$\dot{z}_1 = (1/t) z_3, \quad \dot{z}_2 = (1/t^2) z_1, \quad \dot{z}_3 = (1/t^2) z_2$$

is of the form $t^2 \dot{z} = B(t)z$, where the matrix $B(0)$ is nilpotent. The substitution $y_1 = t^{-1/3} z_1$, $y_2 = z_2$, $y_3 = t^{1/3} z_3$ gives

$$\dot{y}_1 = t^{-5/3} y_3 - (1/3t) y_1, \quad \dot{y}_2 = t^{-5/3} y_1, \quad \dot{y}_3 = t^{-5/3} y_2 + (1/3t) y_3.$$

Thus we get $\dot{y} = (t^{-5/3} F + t^{-1} E) y$, where

$$F = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad E = \operatorname{diag}(-1/3, 0, 1/3).$$

We see that the leading matrix F has three distinct eigenvalues, the cubic roots of -1. Transforming it to the diagonal form and applying the series of transformations from the proof of Theorem 2.17, which are power series in $t^{1/3}$, we can diagonalize also the remaining part of the system for y.

A general result concerning normal forms in the case of resonant irregular singular point belongs to M. Hukuhara H. Turrittin and A. Levelt and is formulated in the next theorem.

THEOREM 2.19. (See [39,81,46].) *If $t = 0$ is an irregular singular point of the system $\dot{z} = A(t)z$ then there exist a positive integer b and a change $y = H(t^{1/b})z$ in the class of formal series in $s = t^{1/b}$, such that H transforms the system for z to a system of the form*

$$\dot{y} = (D_m t^{-m/b} + D_{m-1} t^{-(m-1)/b} + \cdots + D_{b+1} t^{-(b+1)/b} + E t^{-1}) y \tag{2.7}$$

where the matrices D_j and E are diagonal.

Two forms (2.7) *with D_j, E and D'_j, E' are formally equivalent* (*over* $\mathbb{C}[[t^{1/b}]]$) *iff the matrices D_j and D'_j and E and E' are mutually spontaneously conjugated by means of some matrix from* $\mathrm{GL}(n, \mathbb{C})$.

REMARK 2.20. In the variable $s = t^{1/b}$ system (2.7) takes the form

$$\mathrm{d}y/\mathrm{d}s = b(D_m s^{b-m-1} + \cdots + D_{b+1} s^{-2} + E s^{-1}) y.$$

The number $r_1 = \frac{m}{b}$ is called the *Katz invariant* and the smallest integer b is called the *ramification index*.

PROOF OF THEOREM 2.19. (We follow the papers [72] of Y. Sibuya and [82] of V. Varadarajan and the book [89].)

Assume that we have $t^r \dot{z} = B(t)z$.

1. Repeating the proof of Theorem 2.17 we show that the division of the matrix $B(0)$ into diagonal blocks corresponding to different eigenvalues λ_i can be prolonged to a formal splitting of the system into several independent systems, which are characterized by having only one eigenvalue.

2. Let $\lambda_1 = \cdots = \lambda_n = \lambda$. The change $z = \mathrm{e}^{\lambda/(1-r)t^{r-1}} y$ leads to a system $t^r \dot{y} = B(t)y$ (with new $B(t)$) such that

$$B(0) = \begin{pmatrix} 0 & d_1 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & d_{n-1} \\ 0 & 0 & \ldots & 0 \end{pmatrix},$$

where $d_i = 0, 1$

3. Assume that all $d_j = 1$. We introduce the variables $u_1 = y_1$, $u_2 = t^r \dot{u}_1, \ldots, u_n = t^r \dot{u}_{n-1}$. Then we obtain a system $t^r \dot{u} = B(t)u$ such that

$$B(t) = \begin{pmatrix} 0 & 1 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \\ b_1(t) & b_2(t) & \ldots & b_n(t) \end{pmatrix},$$

where each $b_j(t) = \mathrm{O}(t)$.

Now the substitution $u_j = t^{a_j} v_j$, with suitable rational exponents a_j, leads to the equation $\dot{v} = (t^{-\alpha} C + \cdots)v$ where the matrix C has the form like the matrix $B(0)$ but with at least one constant $b_j(0) \neq 0$. This means that C has at least one non-zero eigenvalue and we can apply the induction with respect to the dimension of the eigenspace.

4. If some of the d_i's in the matrix $B(0)$ (from the point 2) are equal to zero then, using the change $y \to u$ analogous as in the point 3, we reduce the matrix $B(t)$ to a matrix

$$\begin{pmatrix} B_1 & E_{12} & \ldots \\ E_{21} & B_2 & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix},$$

such that $B_j = B_j(t)$ are of the form from the point 3 and

$$E_{ij} = \begin{pmatrix} 0 & \ldots & 0 \\ e_1(t) & \ldots & e_s(t) \end{pmatrix},$$

where $e_k(t)$ are functions. It means that most of the rows of $B(t)$ consist of zeroes and the other rows are composed of functions $b_k(t)$ and $e_l(t)$.

One can show that, after applying a suitable change $u \to v$ of the same form as in the point 3, we obtain $\dot{v} = (t^{-\alpha} C + \cdots)v$ where either:

(i) the matrix C has a non-zero eigenvalue (if some $b_k(0)$ becomes $\neq 0$ or some $e_l(0) \neq 0$ in some E_{ij}, $i > j$), or

(ii) the matrix C is nilpotent but its Jordan normal form contains more non-zero d_j's than in the matrix $B(0)$ from the point 2 (here some $e_l(0) \neq 0$ in a $E_{i,j}$, $i < j$), or

(iii) none of the above.

In the cases (i) and (ii) we apply the induction with respect to the dimension of the eigenspace and to the number of non-zero d_j's. (See also [82].)

5. The case (iii) occurs when some supposedly non-trivial row of $B(t)$ from the point 4 vanishes identically; (such phenomenon has infinite codimension). Note that all the changes used in the above point are polynomial in $s = t^{1/b}$. Therefore the system becomes analytically separated, although the first diagonal matrix D_m from the thesis of Theorem 2.19 has coinciding eigenvalues. □

Having solved the problem of formal classification of meromorphic systems with irregular singularity we pass to the analytic classification. It turns out that formal solutions of such systems (or higher order equations) can be divergent. The series from the formal normal forms are only asymptotic.

EXAMPLE 2.21. (See [34].) The series $\sum k!t^k$ formally satisfies the equation $t^2\ddot{x} + (3t-1)\dot{x} + x = 0$ but is divergent.

In what follows we assume that the eigenvalues of the matrix D_m from the thesis of Theorem 2.19 are different. From the point 5 of the above proof it follows that it is sufficient for the analytic classification.

We apply the change of time $t \to t^{1/b}$, and the new time we again denote by t. We apply a polynomial change $z \to \sum_{j=0}^{N} H_j(t)z$, which reduces the polar part of the matrix $A(t) = B/t^r$ to the diagonal form. Thus we get a system $t^r\dot{z} = B(t)z$, where $B(t) = D(t) + O(t^r)$ (we do not change notation of B) and whose formal normal form is (2.6), i.e. $t^r\dot{y} = D(t)y$ with diagonal and polynomial $D(t)$.

Recall that the eigenvalues of $B(0) = D(0)$ satisfy the condition (see the remark above)

$$\lambda_i \neq \lambda_j, \quad i \neq j.$$

If $i \neq j$ then the rays in the complex t plane defined by

$$\operatorname{Re}\bigl[(\lambda_i - \lambda_j)t^{1-r}\bigr] = 0$$

are called the *rays of division corresponding to the pair* (λ_i, λ_j). The following theorem of H. Poincaré is principal in this subsection. Its proof is sketched in Appendix.

THEOREM 2.22. (See [67].) *Let* $S = \{|t| < \epsilon, \alpha < \arg t < \beta\}$ *be a sector in the* t*-plane not containing two rays of division* (*corresponding to any pair* (λ_i, λ_j)). *Then there exists a unique matrix function* $H(t) = I + O(t)$ *analytic in* S *which transforms the system* $t^r\dot{z} = B(t)z$ *to the formal normal form system* (2.6).

Let S_1 and S_2 be two adjacent sectors (i.e. with non-empty intersection) satisfying the assumptions of Theorem 2.22. Let H_{S_1} and H_{S_2} be the corresponding operators guaranteed by this theorem. From Theorem 2.17 it follows that these two matrix-functions have the same Taylor expansions. Because these Taylor expansions are only asymptotic, then H_{S_i} are generally different in the intersection of the two sectors.

One defines the matrix-function

$$G_S(t) = H_{S_1}H_{S_2}^{-1}, \quad t \in S = S_1 \cap S_2.$$

Let $\mathcal{L}_S$ be the space of solutions of the normalized system. The transformation induced by G_S transforms solutions from $\mathcal{L}_S$ to solutions from the same space. Therefore it defines certain automorphism $\mathcal{C}_S : \mathcal{L}_S \to \mathcal{L}_S$.

If we choose a basis of $\mathcal{L}_S$ in the form

$$\phi_j = a_j(t)e_j, \quad a_j(t) = \exp\bigl[\lambda_j t^{1-r}/(1-r) + \cdots\bigr],$$

where (e_j) is the standard basis of $\mathbb{C}^n$, then the operator $\mathcal{C}_S$ expressed in this basis is a constant matrix $C_S = (c_{ij})$.

DEFINITION 2.23. The operator $\mathcal{C}_S$ which is called the *Stokes operator* and the matrix C_S is called the *Stokes matrix*.

If $\psi_i = G_S(t)\phi_i(t) = \mathcal{C}_S\phi_i = \sum_j c_{ij}\phi_j$ then the matrix G_S acts as follows: $G_S e_i = \psi_i/a_i(t)$. We have

$$(G_S)_{ij} = a_i^{-1}(t)\cdot c_{ij}\cdot a_j(t) = c_{ij}\cdot \exp\bigl[(\lambda_j-\lambda_i)t^{1-r}/(1-r)+\cdots\bigr].$$

PROPOSITION 2.24. *The matrix $G_S(t)$ preserves the normalized system* (2.6) *and has the property $G_S(t) = I + O(t^N)$ as $t\to 0$ for any N.*

This implies that: $c_{ii} = 1$ *and* $c_{ij} = 0$ *when* $\mathrm{Re}(\lambda_i - \lambda_j)t^{1-r} \to -\infty$ *as* $t \to 0$. *Therefore, after suitable ordering of the basis (e_i), we obtain that the matrix C_S is unipotent upper triangular.*

The system of Stokes operators can be naturally described using the cohomology language. Take the circle S^1, which we treat as the circle $\{r=0\}$ in $\mathbb{R}_+\times S^1 = \{(r,\theta)\}$ where r,θ are the polar coordinates of the $t = r\mathrm{e}^{\mathrm{i}\theta}\in\mathbb{C}$. If $U\subset S^1$ is an open arc then we associate with a germ of sector $S = S_U$ in $(\mathbb{C},0)$ with base at U.

DEFINITION 2.25. The *Stokes sheaf St* is defined by the presheaf of groups $\mathrm{St}(U)$, $U\subset S^1$ (as above), where $\mathrm{St}(U)$ consists of matrix-functions $S = S_U \ni t \to G_U(t)$ satisfying the following properties:

- $G_U(t)\sim I$,
- $G_U(t)$ preserve the normalized system.

REMARK 2.26. St is not a standard sheaf with which we are acquainted. The groups $\mathrm{St}(U)$ are (generally) non-Abelian groups and the Čech coboundary operator, Čech cocycles and Čech cohomologies must take proper meaning. Here we recall the definition of the first cohomology group of sheaves $\mathcal{F}$ of non-Abelian groups.

The cochains are the same as in the Abelian theory. In particular, a 0-cochain associated with a covering $\mathcal{U} = \{U\}$ is a system $(G_U)_{U\in\mathcal{U}}$ with $G_U\in\mathcal{F}(U)$. 1-cochains are systems (G_{UV}) with $G_{UV}\in\mathcal{F}(U\cap V)$ such that $G_{VU} = G_{UV}^{-1}$. This 1-cochain is a cocycle iff $G_{UV}G_{VW}G_{WU} = e$. Two 1-cochains (G_{UV}) and (G'_{UV}) are equivalent iff there is 0-cochain (K_U) such that

$$G'_{UV} = K_U G_{UV} K_V^{-1}.$$

The first cohomology group $H^1(\mathcal{F},\mathcal{U})$ is defined as the set of classes of 1-cocycles with respect to the this equivalence. Such cocycles with $G_{UV}\in\mathrm{St}(U\cap V)$ are called the *Stokes cochains*.

Next one should takes the direct limit with respect to the coverings, but usually fine finite coverings are sufficient.

A Stokes cochain can be interpreted as a cocycle of Stokes matrices. If $\mathcal{F}_0$ is the fundamental matrix of the normalized system, then

$$C_{UV} = \mathcal{F}_0 G_{UV}\mathcal{F}_0^{-1}$$

is a cocycle of constant matrices. Two Stokes cocycles (C_{UV}) and (C'_{UV}) are called equivalent iff there exists a system (C_U) of constant matrices such that $C'_{UV} = C_U^{-1} C_{UV} C_V$.

If we calculate the cohomology group of the Stokes sheaf using C_{UV} then we see that it is finite-dimensional.

Now we pass to the description of moduli of analytic classification of linear systems with irregular singular point. We fix the normalized system (2.6), which we denote by $\mathcal{S}_0$, and consider the space $\mathcal{M} = \mathcal{M}_D$ of systems $\mathcal{S}: t^r \dot{z} = B(t)$ which are formally equivalent to $\mathcal{S}_0$. On the space $\mathcal{M}$ we introduce the following equivalence relation (compare Definition 2.2):

$\mathcal{S} \sim \mathcal{S}'$ iff there is a matrix function $H(t)$ holomorphic in a (*whole*) *neighborhood of $t = 0$ and realizing equivalence between $\mathcal{S}$ and $\mathcal{S}'$.*

The following fundamental in this theory result was proved by B. Malgrange and Y. Sibuya and is called also the *Malgrange–Sibuya Theorem.*

THEOREM 2.27. (See [52,73].) *The space of equivalence classes defined above coincides with the first cohomology group of the circle with coefficients in the Stokes sheaf. In other words, the group $H^1(S^1, St)$ parameterizes the orbits of the action of the group of analytic equivalences on the space of germs of meromorphic linear systems with fixed formal normal form.*

This property holds in the non-resonant case as well as in the resonant case with the formal normal form the thesis of Theorem 2.19.

PROOF. Let $\mathcal{M}$ be the above space of the equivalence classes. We define a natural map from $\mathcal{M}$ to $H^1(S^1, St)$.

Take a system $\mathcal{S} \in \mathcal{M}$. By the sectorial normalization theorem there exists a covering of a neighborhood of the point $t = 0$ by sectors S_i such that $\mathcal{S}$ is transformed to $\mathcal{S}_0$ by means of holomorphic maps $H_i(t)$ in S_i. Let U_i be the bases of the sectors S_i. Then the family of operators $H_i H_j^{-1}$ defines a Stokes cocycle associated with the covering $\{U_i\}$.

If two systems $\mathcal{S}$ (defining a cocycle $H_i H_j^{-1}$) and $\mathcal{S}'$ (defining a cocycle $H'_i (H'_i)^{-1}$) are equivalent by means of a holomorphic matrix $H(t)$, then $H'_i = H_i H$ (by the uniqueness in Theorem 2.22) and the cocycles define the same cohomology class. This gives the map $\Phi : \mathcal{M} \to H^1(S^1, St)$.

The injectivity of the map Φ is simple. If $H'_i (H'_j)^{-1} = G_i^{-1} H_i H_j^{-1} G_j$ then $H_i^{-1} G_i H'_i = H_j^{-1} G_j H'_j$. The latter matrix is thus univalent and defines a holomorphic conjugation between the systems $\mathcal{S}'$ and $\mathcal{S}$.

The surjectivity of the map Φ is the most difficult part of the proof. Here we follow [6], but we omit the details. Probably the first proof of this surjectivity was given by G.D. Birkhoff [11].

Let $(G_{ij}) = (G_{U_i U_j})$ be a cocycle with values in the Stokes sheaf. If we could find a system $H_i = I + O(t)$ of holomorphic matrices in sectors S_i (with bases U_i) and such that $G_{ij} = H_j^{-1} H_i$, then this would give us a corresponding meromorphic linear equation $\dot{z} = Az$.

Indeed, let $\mathcal{F}_0$ be a fundamental matrix for the normalized system $\dot{z} = A_0 z$, $A_0 = t^{-r}C(t)$. The matrices $\mathcal{F}_i = H_i\mathcal{F}_0$ are the fundamental matrices for systems with matrices $A_i = \dot{\mathcal{F}}_i\mathcal{F}_i^{-1}$ in the sectors S_i. We have $A_i = \dot{H}_i H_i^{-1} + H_i A_0 H_i^{-1}$, $A_0 = \dot{\mathcal{F}}_0\mathcal{F}_0^{-1}$. Calculations which use the fact that G_{ij} preserve A_0 show that $A_i = A_j$ at the intersections of sectors. Thus A_j's define one univalent matrix $A(t)$. Because $H_i = I + O(t)$ the matrix $A(t)$ has the same order of pole as A_0, $A(t) = t^{-r}B(t)$ with analytic B.

To prove the existence of such H_j's we use apparatus of the sheaf theory. It turns out that the fact that the Stokes sheaf (and the sheaf $\mathcal{E}$ defined below) is not Abelian is not a serious obstacle to apply the results from the Abelian theory to our purposes.

Firstly we solve the system of equations $G_{ij} = G_j^{-1}G_i$ in the class of smooth (C^∞) matrix-valued functions $G_i = I + O(t)$ in sectors S_i. To do this one introduces the sheaf $\mathcal{E}$ (on S^1) of germs of smooth matrix-functions in sectors, $\mathcal{E}(U) = C^\infty(S_U, GL(n))$. It is a flabby sheaf, admitting partitions of unity, and therefore has trivial higher cohomology groups, $H^i(S^1, \mathcal{E}) = 0$, $i > 0$. In the Abelian case it is given in [36], in the non-Abelian case it is proved in [6].

Assume that we have such G_i's. Because $\bar{\partial}G_{ij} = 0$, we have $\bar{\partial}G_i \cdot G_i^{-1} = \bar{\partial}G_j \cdot G_j^{-1}$. The latter expression defines a univalent in a punctured neighborhood of $t = 0$ smooth function F.

Now, applying a variant of the Poincaré $\bar{\partial}$-lemma, we find a matrix function $G = I + O(t)$ such that $\bar{\partial}G \cdot G^{-1} = F$.

The functions $H_i = G^{-1}G_i$ are holomorphic and satisfy the relation $H_j^{-1}H_i = G_{ij}$. □

Theorem 2.27, together with Theorems 2.17 and 2.19, gives a complete solution of the problem of analytic classification of systems with irregular singularity. The formal normal form is determined in analytic way (using finite number of polynomial transformations). This formal normal form gives the moduli space. A separate problem constitutes the task of computations of the Malgrange–Sibuya moduli, i.e. how to calculate the elements from $H^1(S^1, St)$ from data of the initial analytic system. In the below examples we show how it is done in special cases.

EXAMPLE 2.28. This example comes from the work [88] (see also [89]).

Consider the following function

$$F(t) = \int_{0^+}^{t} s^a \mathrm{e}^{-1/s^k}\varphi(s)\,\mathrm{d}s = \int f(s)\,\mathrm{d}s$$

where φ is a germ of analytic function. We ask when this function is of the local Darboux type

$$F(t) = t^{a+k+1}\mathrm{e}^{-1/t^k}\psi(t), \tag{2.8}$$

with analytic ψ.

To get the answer we introduce the loops σ_j (called also *asymptotic cycles*) as in Fig. 1. The punctured neighborhood of $t = 0$ is divided into sectors of fall and of jump of e^{-1/s^k}

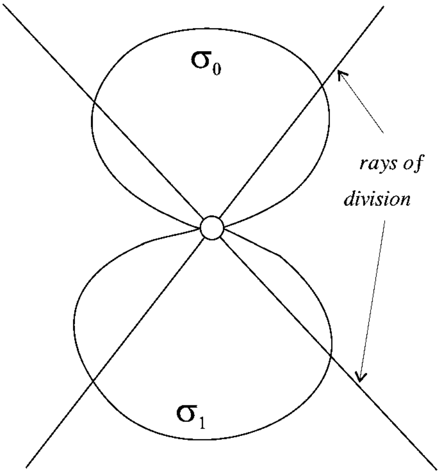


Fig. 1.

as $|t| \to 0$, $\arg t = \text{const}$. Each loop starts and ends at successive sectors of fall (at $t = 0$) and 'surrounds' one sector of jump.

We define the quantities

$$A_j = \int_{\sigma_j} f(s)\,\mathrm{d}s.$$

We have the following property:

The values of $F(t)$ *in neighboring sectors* (*from Fig.* 1) *differ by the constants* A_j. *Therefore the function* F *has the form* (2.8) *iff all* $A_j = 0$.

Consider the linear system

$$t^{k+1}\dot{z} = \begin{pmatrix} -k & t^{k+1}\varphi(t) \\ 0 & at^k \end{pmatrix} z. \tag{2.9}$$

Its general solution is $z_1(t) = D\mathrm{e}^{1/t^k} + C\mathrm{e}^{1/t^k} \int^t f(s)$, $z_2(t) = Ct^a$. Because the singular part of the matrix $A(t)$ is in the diagonal form $A_0(t) = \operatorname{diag}(-kt^{-k-1}, at^{-1})$ then such is also the normalized system.

We choose the system of fundamental solutions of the normalized system in the form $w_1 = \mathrm{e}^{1/t^k}$, $w_2 = t^a$. Therefore $z_1 = w_1 + \eta(t)w_2$, $z_2 = w_2$ where $\eta(t) = t^{-a}\mathrm{e}^{1/t^k}F(t)$.

In different sectors S_j (see Fig. 1) we have different branches of the function F and different branches η_j of η. We have $\eta_{j+1} = \eta_j + A_j t^{-a}\mathrm{e}^{1/t^k}$.

Therefore the normalizing matrices take the form $H_j = \left(\begin{smallmatrix} 1 & \eta_j \\ 0 & 1 \end{smallmatrix}\right)$ and the Stokes cocycle $C_{j,j+1}$ is given by the matrices $\left(\begin{smallmatrix} 1 & A_j \\ 0 & 1 \end{smallmatrix}\right)$.

We have $H^1(S^1, St) \simeq \mathbb{C}^k$ and the Stokes cocycle associated with the system (2.9) is equal to $G_{j,j+1} = \left(\begin{smallmatrix} 1 & A_j t^{-a}\exp(1/t^k) \\ 0 & 1 \end{smallmatrix}\right)$.

EXAMPLE 2.29. (We follow [6] and [37].) Recall the *Bessel equation*

$$t^2\ddot{x} + t\dot{x} + (t^2 - \nu^2)x = 0.$$

We see that the point $t = 0$ is regular (see Definition 2.1 and Theorem 2.7). The point $t = \infty$ turns out irregular.

We introduce the notations $\tau = 1/t$, $u(\tau) = (x(1/\tau), \dot{x}(1/\tau))^\top$ and below by the dot we denote the derivative with respect to τ. Then we get the system $\tau^2\dot{u} = B(\tau)u$, where

$$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\tau + \begin{pmatrix} 0 & 0 \\ -\nu^2 & 0 \end{pmatrix}\tau^2.$$

In order to get the formal normal form we have to transform only the first two terms of $B(\tau)$ to diagonal form. The matrix B_0 has distinct eigenvalues ±i, i.e. we have the non-resonant case. The transformation matrix is chosen in the form $H(t) \sim H_0(I + H_1\tau + \cdots)$, where only the first two terms are important.

We have $H_0 = \left(\begin{smallmatrix} -1 & -\mathrm{i} \\ -\mathrm{i} & -1 \end{smallmatrix}\right)$ and $H_0B_0H_0^{-1} = D_0 = \operatorname{diag}(\mathrm{i}, -\mathrm{i})$. Next $D_1 = H_0(B_1 + [H_1, B_0])H_0^{-1} = H_0B_1H_0^{-1} + [\tilde{H}_1, D_0]$, $\tilde{H}_1 = H_0H_1H_0^{-1}$. The commutator in the latter expression acts onto the non-diagonal entries (because D_0 is diagonal) and, choosing suitable $\tilde{H}_1$, we can obtain D_1 in the diagonal form. Because $H_0B_1H_0^{-1} = \frac{1}{2}\left(\begin{smallmatrix} 1 & * \\ * & 1 \end{smallmatrix}\right)$, then $D_1 = \frac{1}{2}I$. Thus the basis of solutions of the normalized system is $u_{1,2} = \tau^{1/2}\mathrm{e}^{\pm\mathrm{i}/\tau}e_{1,2} = t^{-1/2}\mathrm{e}^{\pm\mathrm{i}t}e_{1,2}$.

There are two rays of division, both in the real axis, and we can take the covering of $\mathbb{C} \setminus 0$ by two sectors with angles $\frac{3\pi}{2}$: S_1 around the real half-line $\{t > 0\}$ and S_2 around $\{t < 0\}$. The intersection $S_1 \cap S_2$ consists of two sectors around $\{\operatorname{Im} t > 0\}$ and $\{\operatorname{Im} t < 0\}$ respectively.

Recall that the Stokes sheaf consists of such matrix functions $G(t)$ that $G \sim I$ and $GDG^{-1} + \tau^2\dot{G}G^{-1} = D$ (preservation of the normalized system). In view of the fact that $2D_1 = I$ this is equivalent to the following condition:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(\chi^{-1}G\chi) = 0 \quad \text{where } \chi(\tau) = \operatorname{diag}(\mathrm{e}^{-\mathrm{i}/\tau}, \mathrm{e}^{\mathrm{i}/\tau}).$$

Therefore the map $G \to G = \chi^{-1}G\chi$ defines an isomorphism of the sheaf St with some sheaf of subgroups of $\mathrm{GL}(2, \mathbb{C})$. The first condition for the Stokes sheaf $\chi F\chi^{-1} \sim I$ implies that the matrix elements of F fulfill the conditions $F_{11} = F_{22} = 1$, $F_{21}\mathrm{e}^{-2\mathrm{i}/\tau} \sim 0$, $F_{12}\mathrm{e}^{2\mathrm{i}/\tau} \sim 0$. Therefore, depending in which part of the τ-plane we are, the matrix F takes one of the two forms

$$\begin{pmatrix} 1 & f_+ \\ 0 & 1 \end{pmatrix}, \quad \operatorname{Im} t > 0; \qquad \begin{pmatrix} 1 & 0 \\ f_- & 1 \end{pmatrix}, \quad \operatorname{Im} t < 0. \tag{2.10}$$

The parameters $f_\pm$ parameterize the space $H^1(S^1, St)$ and constitute the moduli of analytic classification of equations which are formally equivalent to the Bessel equation.

We calculate the constants f_+, f_- for the very Bessel equation following the book of J. Heading [37]. One uses the basis of solutions of the Bessel equation consisting of the *Bessel functions*

$$J_{\pm\nu}(t) = \sum_{j=0}^{\infty} \frac{(-1)^j (t/2)^{2j\pm\nu}}{j!\Gamma(\pm\nu + j + 1)} = t^{\pm\nu} P_\pm(t^2). \tag{2.11}$$

Note that the function $P_\pm$ are integer. Assume that a solution $AJ_\nu(t) + BJ_{-\nu}(t)$ has the asymptotic expansion $\sim t^{-1/2}\mathrm{e}^{\mathrm{i}t}$ for $\arg t = 0$ and $t \to \infty$. After analytic prolongation of this solution to the rays $\arg t = \pi$ and $\arg t = 2\pi$, with the use of formulas (2.10) and (2.11), we obtain

$$\begin{aligned} &A\mathrm{e}^{\mathrm{i}\pi\nu} J_\nu(t) + B\mathrm{e}^{-\mathrm{i}\pi\nu} J_{-\nu}(t) \sim \mathrm{e}^{-\mathrm{i}\pi/2} t^{-1/2} \mathrm{e}^{-\mathrm{i}t}, \\ &A\mathrm{e}^{2\mathrm{i}\pi\nu} J_\nu(t) + B\mathrm{e}^{-2\mathrm{i}\pi\nu} J_{-\nu}(t) \sim \mathrm{e}^{-\mathrm{i}\pi} t^{-1/2} \mathrm{e}^{\mathrm{i}t} + f_- \mathrm{e}^{-\mathrm{i}\pi} t^{-1/2} \mathrm{e}^{-\mathrm{i}t}, \end{aligned}$$

as $t \to +\infty$. Together with $AJ_\nu(t) + BJ_{-\nu}(t) \sim t^{-1/2}\mathrm{e}^{\mathrm{i}t}$ this gives the value

$$f_- = 2\mathrm{i}\cos(\pi\nu).$$

Analogous calculations give $f_+ = 2\mathrm{i}\cos(\pi\nu)$.

2.3. *Appendix. Proof of the normalization theorems*

Here we prove Theorems 2.13 and 2.22. These proof are conceptually similar, since they are reduced to the contraction principle for a suitable map in a Banach space. We begin with the proof of Theorem 2.22, where we follow the book of W. Wasow [85]. Another proof of Theorem 2.22 uses the Gevrey expansions and multi-summability (see next Subsection 3.6 below) and is given in the paper of B. Braaksma [14].

PROOF OF THEOREM 2.22. Recall that this theorem states that:

there exists a unique matrix function $H(t) = I + O(t)$ analytic in a sector S, not containing two rays of division, which transforms the system $t^r \dot{z} = B(t)z$ to the formal normal form system (2.6).

The essential fact which has to be proved is the following

PROPOSITION. *Let the eigenvalues of $B(0)$ be divided into two groups $\lambda_1, \dots, \lambda_p$ and $\lambda_{p+1}, \dots, \lambda_n$ such that $\lambda_i \neq \lambda_j$ for $i \leqslant p < j$. Then there exists a matrix $H(t)$, holomorphic in S, which transforms the system to an analogous system with the matrix $C(t)$ of the block-diagonal form*

$$\begin{pmatrix} C^{11}(t) & 0 \\ 0 & C^{22}(t) \end{pmatrix}.$$

Moreover, we can assume that the sector S is symmetric with respect to the real positive semi-axis and has the magnitude $< \pi/(r-1)$.

PROOF OF PROPOSITION. If $z = H(t)y$ then y satisfies the equation $t^r H\dot{y} = (BH - t^r \dot{H})y$ and we have

$$HC = BH - t^r \dot{H}.$$

So we assume that $B(0)$ is in the block-diagonal form $B(0) = \operatorname{diag}(B^{11}(0), B^{22}(0))$ and we seek $H(t)$ in the form

$$\begin{pmatrix} I & H^{12}(t) \\ H^{21}(t) & I \end{pmatrix}.$$

In what follows we will deal only with H^{12}. The analysis for H^{21} is quite analogous.

The matrix $X(t) = H^{12}(t)$ satisfies the following non-linear differential equation

$$t^r \dot{X} = B^{11}X - XB^{22} + B^{12} - XB^{21}X$$

where $B^{ij}(t)$ are the corresponding block-matrices.

The latter equation is a particular case of the following non-linear equation

$$t^r \dot{u} = \Lambda u + p(t,u), \quad u \in \mathbb{C}^m, \tag{2.12}$$

where Λ has non-zero eigenvalues Λ_k (equal to $\lambda_i - \lambda_j$ in our case) and the term p is non-linear.

Equation (2.12) is equivalent to the following integral equation

$$u(t) = V(t)D + \int_{t_0}^{t} V(t)V^{-1}(s)s^{-r} p\bigl(s, u(s)\bigr)\,\mathrm{d}s \tag{2.13}$$

where $V(t) = \mathrm{e}^{\Lambda/(1-r)t^{r-1}}$ and U is some constant vector.

In fact, the constant vector D and the limit t_0 of integration are in some way connected one with another. One can replace the integration in (2.13) by the integration along some contour $\Gamma(t)$ in the complex s-plane in such a way that it ends at t. In this case the constant vector can be put equal to zero. Moreover, for each component of Eq. (2.13) we can choose the integration path independently.

The main trick of the proof is to choose the path $\Gamma(t)$ in such a way that the expression $V(t)V^{-1}(s)$ is not too big along the path of integration; then we will be able to apply the principle of contracting maps. In particular, we want that the expressions

$$\exp\bigl[\operatorname{Re}\bigl(\Lambda_k(t^{1-r} - s^{1-r})/(1-r)\bigr)\bigr],$$

associated with the k-th components, are not big.

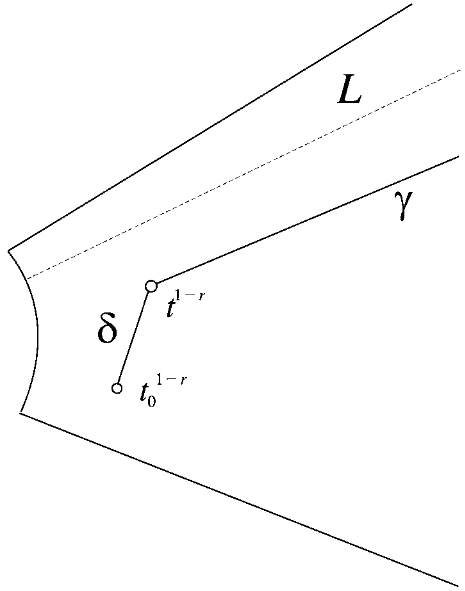


Fig. 2.

We fix for a while the index k and some point $t_0 \in \mathbb{R} \cap S$. We shall consider only such t's that $|t| << t_0$. Moreover it is useful to pass to the chart $\zeta = s^{1-r}$; then the sector S is replaced by a sector with vertex at infinity.

Assume that S does not contain any ray of division corresponding to Λ_k. It means that $\operatorname{Re} \Lambda_k \zeta/(1-r)$ has definite sign in the whole S. Depending on this sign we choose either:

- the straight semi-line γ from $\zeta = \infty$ to $\zeta = t^{1-r}$ with constant $\arg \zeta$, if the sign is positive, or
- the interval δ joining t_0^{1-r} with t^{1-r}, otherwise; (see Fig. 2).

In both cases we have $\operatorname{Re} \Lambda_k (t^{1-r} - \zeta)/(1-r) < 0$.

If S contains a ray of division L (only one by the assumption) then the choice of the path δ remains unchanged but the path γ passes along a straight half-line parallel to L (from infinity to t^{1-r}, see Fig. 2). Here we have $\operatorname{Re} \Lambda_k (t^{1-r} - \zeta)/(1-r) = 0$ along γ.

The further proof uses the standard analytic methods and we describe it only shortly, without detailed estimates.

We solve the fixed point equation

$$u = T(u),$$

where $T(u)$ is the non-linear integral operator $\int V(t) V^{-1}(s) s^{-r} p(s, u(s))\, \mathrm{d}s$ and the integration runs along the paths described above.

In order to apply the method of contracting maps, we must define some Banach space of holomorphic functions u and get an estimate for $\|T(u_1) - T(u_2)\|$ (i.e. the Lipschitz continuity of $p(t, u)$ with respect to u).

The class of u's consists of those which satisfy the estimate $|u(t)| < c|t|^m$ with certain fixed m. This m is defined as the order of $p(t,0) = \mathrm{const}\cdot t^m + \cdots$. One shows that, if u satisfies this estimate, then $|T(u)(t)| < Kc|t|^m$ with a constant K not depending on u (see Lemma 14.2 in [85]).

The Lipschitz estimate $|p(t,u_1) - p(t,u_2)| < \mu|u_1 - u_2|$, with μ arbitrarily small, follows from the fact that p contains only terms of the form $\mathrm{O}(t^m)u^0$, $\mathrm{O}(t)u$ and $\mathrm{O}(|u|^2)$.

Next, one chooses the sequence $u_0 = 0$, $u_1 = T(u_0)$, $u_2 = T(u_1), \ldots$ of successive approximations. One gets $|u_1| < c|t|^m$ and $|u_{n+1} - u_n| < \mu K|u_n - u_{n-1}|$. If $\mu < K^{-1}$ then this series converges to a function holomorphic in some sector S_0 of small radius. □

PROOF OF THEOREM 2.13. Recall that this theorem states that:

for a system $t\dot z = B(t)z$ with Fuchsian and resonant singularity there exists an analytic matrix $H(t)$ which transforms it to a system of the form,

$$t\dot z_i = \lambda_i z_i + \sum_{\lambda_i-\lambda_j\in\mathbb{Z}_+} a_{ij}t^{\lambda_i-\lambda_j}z_j, \quad i = 1,\ldots,n,$$

with constant coefficients a_{ij}.

Note also that the method of the proof of Theorem 2.12 does not work in this situation.

There are finitely many resonant relations $\lambda_i - \lambda_j = k_{ij} \in \mathbb{Z}_+$. Let an integer $m > 0$ be such that $m > \mathrm{Re}(\lambda_i - \lambda_j)$ for all i, j. We can assume that the system takes the form

$$t\dot z = Dz + tB_0(t)z + t^m B_1(t)z,$$

where the matrix D is constant in the Jordan form and $tB_0(t)z$ contains only the resonant terms (like in thesis of the theorem).

We seek the matrix H in the form

$$H(t) = I + t^m G(t)$$

and obtain the equation

$$t\dot G = \bigl([D,G] - mG\bigr) + \bigl(\bigl[tB_0(t), G\bigr] + B_1(t) + t^m B_1(t)G\bigr).$$

It has the same type as Eq. (2.13) but with $r = 1$ and with the eigenvalues of Λ equal $\lambda_i - \lambda_j - m$, which are all non-zero. The fundamental matrix $V(t)$ has the eigenvalues $t^{\lambda_i-\lambda_j-m}$.

The corresponding integral equation is following

$$u(t) = \bigl(T(u)\bigr)(t) = V(t)\int_{t_0}^{t} V(s)^{-1}s^{-1}p\bigl(s,u(s)\bigr)\,\mathrm{d}s,$$

where the initial point t_0 can be chosen equal to 0. Indeed, since $p(s,u(s))$ is bounded and the entries of the matrix $V(s)^{-1}s^{-1}$ vanish at $s = 0$, the above integral is well defined.

It is easy to check that the Lipschitz condition for $p(t,\cdot)$ is satisfied and we can apply the contraction principle. Therefore there exists a unique fixed point for the map T which is an analytic solution to our equation. $\square$

3. Holomorphic vector fields

The previous section was devoted to linear analytic ODEs. In this section we develop the theory of normal forms of non-linear analytic differential equations. These are the equations of the form

$$\dot{x} = V(x), \quad x \in \mathbb{C}^n, \tag{3.1}$$

where $V(x)$ is a holomorphic vector-function, i.e. *holomorphic vector fields*. Here the time t is also complex. Therefore the phase curves become Riemann surfaces and the phase portrait of such a vector field defines a foliation $\mathcal{F}$ into complex phase curves. Such foliations are called *holomorphic foliations*.

The zeroes x_0 of the vector field (3.1) are the *singular points* of the vector field $V(x)$ and of the foliation $\mathcal{F}$. In this (and next) section we study the holomorphic foliations near a singular point, which we assume $x_0 = 0$.

DEFINITION 3.1. Two germs of holomorphic vector fields V and W are called (analytically) *equivalent* if there exists a local analytic diffeomorphism $H:(\mathbb{C}^n,0)\to(\mathbb{C}^n,0)$ transforming V to W, i.e.

$$W = H^*V := (H_*^{-1}V)\circ H.$$

These germs are *formally equivalent* if $H(x)$ is defined by means a formal invertible series and the above equality holds at the level of formal power series.

The germs V and W are called (analytically) *orbitally equivalent* if there exists a local analytic diffeomorphism $H:(\mathbb{C}^n,0)\to(\mathbb{C}^n,0)$ transforming the phase curves of V to phase curves of V. This means that

$$W = F\cdot H^*V$$

where $F \neq 0$ is some analytic function. These germs are *formally orbitally equivalent* if H and F are formal series and the above equality holds at the formal level.

3.1. *Resolution for planar vector fields and the center–focus problem*

This section is not devoted to classification of vector fields with respect to equivalence or orbital equivalence. We present here a result about reduction of a complex singular point of a vector field to a collection of more simple singularities. Moreover we restrict the analysis to the case $n = 2$, i.e. the plane vector fields.

DEFINITION 3.2. A singular point of a germ of planar holomorphic planar vector field is called *elementary* (or *reduced*) iff at least one of its eigenvalues is non-zero.

Thus we can assume that we have

$$\dot{x} = \lambda_1 x + \cdots, \quad \dot{y} = \lambda_2 y + \cdots, \quad \lambda_1 \neq 0.$$

Important invariant of the singular point is the *ratio of eigenvalues*

$$\lambda = \lambda_2 / \lambda_1.$$

If $\lambda \in \mathbb{C} \setminus \mathbb{R}$ then we say that the singular point is a *focus*; if $\lambda > 0$ then it is a *node*; if $\lambda < 0$ then it is a *saddle*; if $\lambda = 0$ then it is *saddle–node*.

The singular point is called *resonant* if λ is a rational number. The resonant node is called *dicritical* if it is formally equivalent to its linear part.

DEFINITION 3.3. The *elementary blowing-up* is the map

$$\pi : (M, E) \to (\mathbb{C}^2, 0),$$

where M is a 2-dimensional complex manifold equal (as a set) to $(\mathbb{C}^2 \setminus 0) \sqcup E$, where $E \simeq \mathbb{C}P^1$ is the set of lines passing through the origin. E is called the *exceptional divisor*. If $(u_1 : u_2)$ are the homogeneous coordinates in $\mathbb{C}P^1$ then the blowing-up map is locally given by the formulas

$$(x, u) \to (x, xu), \quad u = u_2/u_1, \quad (u_1 \neq 0),$$
$$(y, z) \to (yv, y), \quad v = u_1/u_2, \quad (u_2 \neq 0).$$

If $V : (\mathbb{C}^2, 0) \to (\mathbb{C}^2, 0)$ is a germ of holomorphic vector field, then after the elementary blowing-up we obtain usually not a vector field in M but rather a field of directions $\tilde{V}$ near the distinguished divisor E. In the local (x, u)-charts (or (y, v)-charts) in M this field of directions is defined by means of an analytic vector field.

Let $V : (\mathbb{C}^2, 0) \to (\mathbb{C}^2, 0)$ be a germ of holomorphic vector field. We say that it admits a *good resolution* if after a *finite* number of elementary blowing-ups we obtain a field of directions in a 2-dimensional complex manifold with only elementary and non-dicritical singular points.

EXAMPLE 3.4. *A general homogeneous vector field of degree n*

$$\dot{x} = P_n(x, y), \qquad \dot{y} = Q_n(x, y)$$

has $n + 1$ *invariant lines*.

Indeed, in the blowing-up coordinates $x, u = y/x$ we get the system $\dot{x} = x^n P_n(1, u)$, $\dot{u} = x^{n-1}[Q_n(1, u) - u P_n(1, u)] = x^{n-1} H_{n-1}(u)$, where the polynomial $H_{n+1}(u)$ in the square brackets is (generally) non-zero. In order to obtain a field of directions one has

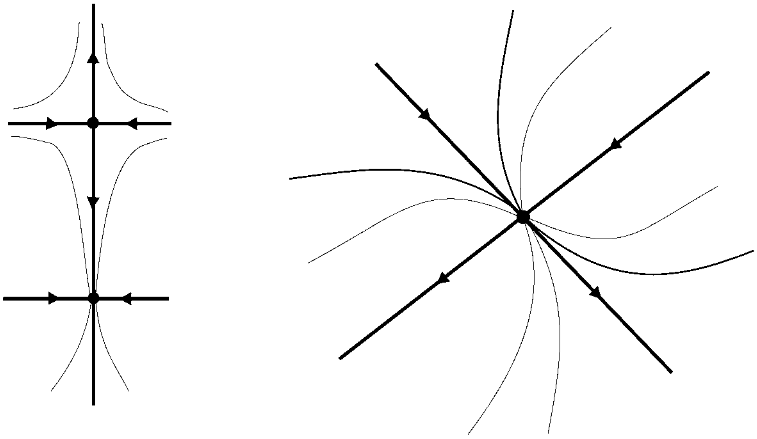


Fig. 3.

to divide the latter system by x^{n-1}. The zeroes u_i of $H_{n+1}(u)$ define the invariant lines $y = u_i x$, see Fig. 3.

If the zeroes u_i are isolated then we can perturb this vector field by adding terms of degree $> n$ and the above blowing-up gives the resolution of the perturbed vector field; the singular points $(u_i, 0)$ remain elementary.

If $P_n = xR_{n-1}$, $Q_n = yR_{n-1}$ then $H_{n+1} \equiv 0$ and the foliation in the (x,u)-coordinates is the flow-box foliation. If additionally we consider perturbation, with terms of higher degree of general type, then in the resolution we must divide the blown-up vector field by x^n. The blowing-up defines the resolution of the singularity but the exceptional divisor $x = 0$ is not invariant for the corresponding foliation. In such a case we say that the divisor E is *dicritical*.

EXAMPLE 3.5. Consider the *Bogdanov–Takens singularity*

$$\dot{x} = 2y, \quad \dot{y} = ax^2 + bxy + cy^2 + \cdots,$$

where we assume that $a \neq 0$ (and we can put $a = 3$). This system can be treated as a small perturbation of the Hamiltonian system X_H with the Hamilton function $H = y^2 - x^3$ with the cusp singularity. Indeed, in the quasi-homogeneous filtration with the exponents $d(x) = 2$, $d(y) = 3$, the Hamiltonian part has degree 1 and the remaining terms have greater degrees.

The blowing-up of the cusp curve is presented at Fig. 4. It is easy to see that the same elementary blowing-up maps can be applied to the Bogdanov–Takens singularity. The resolved field of directions has the same singularities as X_H which are also hyperbolic saddles. The curve Γ is an invariant curve of the vector field (separatrix) and has the cusp singularity. We shall return to this singularity in the next section.

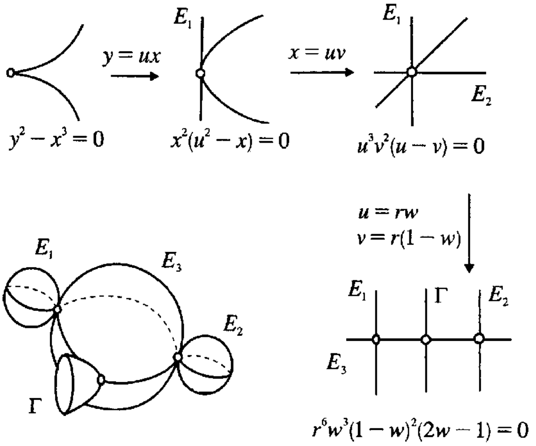


Fig. 4.

The following theorem is principal in this section. It was firstly formulated by I. Bendixson under the assumption that the vector field is real and analytic. F. Dumortier also considered the real case with only elementary singularities at the end. Moreover, he assumed that the initial vector field is only smooth. The assumption that the singular point is isolated is replaced by the following *Łojasiewicz condition* $|V(x)| > C|x|^{\alpha}$. Our version of the desingularization theorem was firstly formulated by A. Seidenberg. However its final proof, which is the just proof, was given by A. van den Essen.

THEOREM 3.6. (See [10,70,27,32].) *Any germ of analytic planar vector field with isolated singularity admits a good resolution.*

REMARK 3.7. There exists an analogous theorem (about desingularization) for germs of holomorphic foliations of codimension 1 in $(\mathbb{C}^3, 0)$ (see [22]). But no analogous result about the codimension 1 foliations is known in dimensions >3. Also there is no desingularization theorem for holomorphic vector fields in dimensions >2.

PROOF OF THEOREM 3.6. (We follow the article of J.-F. Mattei and R. Moussu [56] and [89].)

1. If we have a $p : q$ resonant node $\dot{x} = px + \cdots$, $\dot{y} = qy + \cdots$, $p < q$, then the resolved vector field has two singular points in the divisor $E = \{(u_1 : u_2)\}$: the $p : (q - p)$ resonant node $(1 : 0)$ (i.e. $\dot{x} \approx px$, $\dot{u} \approx (q - p)u$) and $p : (p - q)$ resonant saddle $(0 : 1)$. If the node $x = y = 0$ is dicritical then the node $(1 : 0)$ is also dicritical.

Repeating this several times, we arrive to a $1 : 1$ resonant node which after elementary blowing-up is transformed to the flow-box (see Example 3.4).

2. We work with the vector fields $V = b(x, y)\partial_x - a(x, y)\partial_y$ and with the equivalent to them 1-forms

$$\omega = a\,\mathrm{d}x + b\,\mathrm{d}y.$$

We recall the definition of *index* of the vector field (or of the 1-form) $i_0(V) = i_0(\omega) = i(a, b; 0)$. Let $b(x, y) = 0$ be an irreducible (analytic) curve passing through the origin. It can be locally parametrized by an analytic map $(\mathbb{C}, 0) \ni t \to \gamma(t) \in (\mathbb{C}^2, 0)$ such that $b \circ \gamma \equiv 0$. There is the so called *primitive parametrization* such that any other parametrization of the same curve is induced from the primitive parametrization by a change of parameters. (Example: the primitive parametrization of the cusp $y^2 = x^3$ is $x = t^2$, $y = t^3$.) Then we have $a \circ \gamma(t) = \alpha t^d + \cdots$, $\alpha \neq 0$. We get $i(a, b; 0) = d$. If the function b has the representation $b = b_1^{l_1} \cdots b_r^{l_r}$ with irreducible b_j's then $i(a, b; 0) = \sum l_j \cdot i(a, b_j; 0)$.

Note that $i_0(V) = 1$ iff 0 is a non-degenerate elementary singular point, i.e. not a saddle-node.

The *rank* of the vector field (or of the 1-form) $\nu(V) = \nu(\omega) = \nu_0(\omega)$ equals to the order of the first non-zero term in the Taylor expansion of V at 0. Analogously one defines the rank of a function.

3. LEMMA. *Let $\pi : (x, u) \to (x, xu)$ be the blowing-up map with the exceptional divisor $E = \pi^{-1}(0)$. We define the functions $\tilde{a}, \tilde{b}$ by the formulas*

$$a \circ \pi = x^{\nu(a)} \tilde{a}, \qquad b \circ \pi = x^{\nu(b)} \tilde{b}.$$

Then we have

$$i(a, b; 0) = \nu(a)\nu(b) + \sum_{z \in E} i(\tilde{a}, \tilde{b}; z).$$

PROOF. When the polynomial $\tilde{b}(0, u)$ has only simple zeroes $z \in E$ (there is $\nu(b)$ of them) then each such zero gives the contribution $\nu(a) + i(\tilde{a}, \tilde{b}; z)$ to the index $i(a, b; 0)$. The general case is a consequence of the additivity of the index; for a family V_ε of vector fields with singular points $x_i(\varepsilon)$, $i = 1, \ldots, k$, degenerating to 0 as $\varepsilon \to 0$ we have $i_0 V_0 = \sum i_{x_j}(V_\varepsilon)$. □

4. PROPOSITION. *There exists a sequence of elementary blowing-ups such that at the end we obtain only singularities with rank $\leqslant 1$.*

PROOF. Let $\omega_\nu = a_\nu\, \mathrm{d}x + b_\nu\, \mathrm{d}y$ be the homogeneous part of ω of order $\nu = \nu(\omega) = \min(\nu(a), \nu(b))$. We put

$$P_{\nu+1} = x a_\nu + y b_\nu.$$

Then we have

$$\begin{aligned}
\pi^* \omega &= a(x, ux)\, \mathrm{d}x + b(x, ux)\, \mathrm{d}(ux) \\
&= (a + ub)\, \mathrm{d}x + xb\, \mathrm{d}u \\
&= x^\nu \bigl\{ \bigl[P_{\nu+1}(1, u) + \cdots \bigr] \mathrm{d}x + x \bigl[b_\nu(1, u) + \cdots \bigr] \mathrm{d}u \bigr\}.
\end{aligned}$$

Here the case with $P_{\nu+1} \equiv 0$ is dicritical.

For any $z \in E$ we define the germ $\tilde{\omega} = \tilde{\omega}_z$ as the germ of $\pi^*\omega$ at z, divided by the greatest common multiplier of the germs its components. Thus if $\pi^*\omega = a_1\,\mathrm{d}x + b_1\,\mathrm{d}u$, then $\tilde{\omega} = \pi^*\omega/\gcd(a_1, b_1)$. The Pfaff equation $\tilde{\omega} = 0$ defines the blown-up foliation near z.

Proposition 4 follows from the following result. □

5. LEMMA. *If $\nu > 1$ then for any $z \in E$ we have*

$$i_z(\tilde{\omega}) < i_0(\omega).$$

PROOF. Consider firstly the dicritical case, i.e. $P_{\nu+1} \equiv 0$. We shall show the identity

$$i_0(\omega) = \nu^2 + \nu - 1 + \sum i_z(\tilde{\omega}).$$

We can choose a local system of coordinates $(x, y) \to (x + \cdots, y + \cdots)$ such that the following conditions hold:

(i) $P(x, y) = xa + yb$ has rank exactly $\nu + 2$;
(ii) $b(0, y) = y^{\nu+1}$ (then $P(0, y) = y^{\nu+2}$).

We have

$$i(P, b; 0) = i(x, b; 0) + i(a, b; 0)$$

where

$$i(P, b; 0) = \nu(\nu + 2) + \sum i(\tilde{P}, \tilde{b}; z)$$

(by Lemma 3) and $i(x, b; 0) = \nu + 1$ (by (ii)). Now one must only notice that $\tilde{\omega} = \tilde{P}\,\mathrm{d}x + \tilde{b}\,\mathrm{d}u$ and thus $i(\tilde{P}, \tilde{b}; z) = i_z(\tilde{\omega})$.

Consider now the non-dicritical case $P_{\nu+1} \not\equiv 0$. We shall prove that

$$i_0(\omega) = \nu^2 - \nu - 1 + \sum i_z(\tilde{\omega}).$$

We have (by Lemma 3)

$$i(a, b; 0) = \nu^2 + \sum i(\tilde{a}, \tilde{b}; z). \tag{3.2}$$

Next

$$i_z(\tilde{\omega}) = i(\tilde{a} + u\tilde{b}, x\tilde{b}; z) = i(\tilde{a} + u\tilde{b}, x; z) + i(\tilde{a}, \tilde{b}; z)$$

where $i(\tilde{a} + u\tilde{b}, x; z)$ is the multiplicity of z as zero of the polynomial $P_{\nu+1}$; the sum of these multiplicities is $\nu + 1$.

Summing up the latter equality over z's and taking into account (3.2) we get the result. □

6. Let $\nu = 1$. The only case with non-elementary singular point of rank one is the case with nilpotent linear part $\dot{x} = y + \cdots$, $\dot{y} = \cdots$. The corresponding 1-form is following

$$\omega = \left[y + A_1(x, y)\right] \mathrm{d}y + B_1(x, y)\, \mathrm{d}x,$$

with $\nu(A_1), \nu(B_1) \geqslant 2$.

After blowing-up in the x-direction, i.e. with $y = xu$, we get $\tilde{\omega} = x(u + xA_2)\, \mathrm{d}u + (u^2 + xB_2)\, \mathrm{d}x$ where A_2, B_2 are analytic functions (without restrictions onto ranks). The latter form can be included into the following series of 1-forms

$$\eta = x(y + xA)\, \mathrm{d}y + (ny^2 + xB)\, \mathrm{d}x \tag{3.3}$$

where n is a natural number. In what follows we deal with (3.3).

Note also that $\nu(\eta) = 2$.

7. Let $B(0) = b_0 \neq 0$. Here we blow-up η in the y-direction by putting $x = vy$. We obtain

$$\tilde{\eta} = v\left[(n + 1)y + b_0 v + A'\right] \mathrm{d}y + y[ny + b_0 v + B']\, \mathrm{d}v.$$

This form has good quadratic part. Note that the polynomial

$$P_3(\tilde{\eta}) = vy\left[(2n + 1)y + 2b_0 v\right]$$

has distinct factors. After the next blowing up we obtain three elementary singular points (see Example 3.4).

8. Let $B(0) = 0$. From (3.3) we get

$$P_3(\eta) = x\left[(n + 1)y^2 + \alpha xy + \beta y^2\right] = (n + 1)x(y - c_1 x)(y - c_2 x).$$

If $c_1 \neq c_2$ then the next blowing-up solves the problem (like in the point 7).

Let $c_1 = c_2$. Eventually applying the change $y \to y - cx$ we can assume that $c_1 = c_2 = 0$, i.e. $P_3(\eta) = (n + 1)xy^2$. The singular point $(0 : 1) \in E$, corresponding to the line $x = 0$ in $P_3 = 0$, is elementary; the point $(1 : 0)$ is non-elementary. Then calculations of the blowing up with $y = ux$ give

$$\tilde{\eta} = x(u + A')\, \mathrm{d}u + \left[(n + 1)u^2 + xB'\right] \mathrm{d}x$$

where $A' = A(x, ux)$ and $B' = B(x, ux)/x + uA(x, ux)$.

If $A(0) \neq 0$ then the point $x = u = 0$ is elementary.

9. Let $A(0) = 0$. Then $A' = xA''$ and $\tilde{\eta}$ has the same form as η in (3.3) but with n replaced by $n + 1$. So we can repeat the analysis from the points 7 and 8. We have two possibilities: either after finitely many steps we obtain elementary singular point or infinitely many times we obtain the form (3.3).

However the second possibility contradicts Proposition 4, because we would get infinitely many times $\nu > 1$. This contradiction proves Theorem 3.6. □

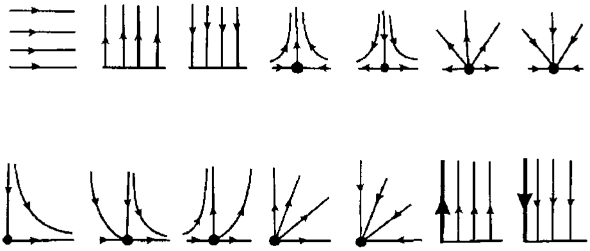


Fig. 5.

Theorem 3.6 has important consequences in the complex domain and in the real domain. Below we present its application in the study of phase portraits of real analytic vector field about a singular point.

Theorem 3.6 implies that after blowing-ups the phase portrait near the distinguished divisors can be composed from the finite collection of standard portraits (presented at Fig. 5).

COROLLARY 3.8. *The resolution of singular point of a real vector field allows to determine its topological type modulo solution of the center–focus problem. If there is a characteristic trajectory* (*see Definition* 3.9) *of the resolved field* (*tending to a point in the distinguished divisors*) *then the topological type of the vector field is completely determined*; *if there is no characteristic trajectory then the topological type is determined by the stability type of the singularity.*

Therefore we arrive to the center–focus problem for germs of analytic vector fields.

DEFINITION 3.9. The *characteristic trajectory* of a real planar vector field V at the singular point $x = 0$ is a phase curve which approaches the singularity with definite limit direction. It means that after resolution some local phase portrait different from the first and eighth portraits in Fig. 5 appears.

If there is no characteristic trajectories then the singular point is called *monodromic*.

The *center–focus problem* is the problem of describing conditions for stability or unstability of a monodromic singular point for analytic vector field.

V. Arnold [3] formulated this problem as follows. The space J^k of k–jets j^kV at $x = 0$ of germs $V : (\mathbb{R}^2, 0) \to (\mathbb{R}^2, 0)$ of analytic vector fields is divided into three subspaces J_s^k, J_u^k and J_n^k. Here the jets j^kV from J_s^k (respectively from J_u^k) are such that any germ W with $j^kW = j^kV$ has 0 as an asymptotically stable point (respectively unstable point). The subspace J_n^k consists of 'neutral' jets. Arnold asked whether the sets $J_{s,u,n}^k$ are semi-algebraic, or semi-analytic, and conjectured that the codimension of J_n^k grows to infinity with k.

In practice one divides the center–focus problem into two problems. Firstly one asks whether the singular point is monodromic; then one can define the Poincaré return map from an arc ending at the singularity to itself. Next one calculates the asymptotic expansion of the Poincaré map.

Theorem 3.6 implies that the monodromicity problem is algebraically solvable, i.e. that the division of the space J^k into 'stably' monodromic, non-monodromic and neutral jets is defined by algebraic equations and inequalities onto the coefficients of the Taylor expansion of V.

Therefore there appears the problem of calculating the leading term in the asymptotic expansion of the Poincaré map for germs from a given class. Such class may consists, for example, of germs with a fixed resolution of the singularity.

The first case with elementary singular point, i.e. with imaginary eigenvalues, was solved by H. Poincaré and A. Lyapunov.

THEOREM 3.10. (See [67,51].) *The problem of distinguishing between center and focus for germs of the form*

$$\dot{x} = -y + \cdots, \qquad \dot{y} = x + \cdots$$

is algebraically solvable. More precisely, the Poincaré map $\mathcal{P}$: $\{x = r \geqslant 0, y = 0\} \to \{x \geqslant 0, y = 0\}$ *has the expansion*

$$\mathcal{P}(r) = r + g_3 r^3(1 + \cdots) + g_5 r^5(1 + \cdots) + \cdots$$

where the Poincaré–Lyapunov focus quantities g_{2j+1} *are polynomials with respect to* (*finite number*) *of coefficients of the Taylor expansion of the vector field.*

Moreover, if all $g_j = 0$ *for a given germ* V *then there exists a local analytic first integral*

$$F = x^2 + y^2 + \cdots$$

and the singular point $x = y = 0$ *is a center for* V.

PROOF. It is useful to pass to the complex variables $z = x + \mathrm{i}y$, $\bar{z} = x - \mathrm{i}y$. Then we have

$$\dot{z} = \mathrm{i}z + \sum a_{kl} z^k \bar{z}^l.$$

We look for a Lyapunov function (in the stable or unstable case) or for a first integral $F = z\bar{z} + \sum f_{kl} z^k \bar{z}^l$. It should satisfy the equation

$$\dot{F} = \frac{1}{\pi}\left(g_3|z|^4 + g_5|z|^6 + \cdots\right),$$

where the coefficients f_{kl} and $\frac{1}{\pi} g_j$ are defined recursively. We get a system of equations $\sqrt{-1}\,(k-l) f_{kl} = -(a_{k,l-1} + \bar{a}_{k-1,l}) +$ (terms depending on previous f_{ij}'s). We see that we can calculate f_{kl} always when $k \neq l$, when $k = l$ we find the obstacle which equals $\frac{1}{\pi} g_k$.

To relate the expansion of $\dot{F}$ with the expansion of the Poincaré map we parametrize the section by $F = r^2 + \cdots$. Then $\mathcal{P}(r) - r = \int_0^T \dot{r}\, \mathrm{d}t \approx \frac{1}{2r} \int_0^{2\pi} \dot{F}\, \mathrm{d}t \approx \frac{1}{2r} \cdot 2\pi \cdot \frac{g_{2j-1}}{\pi} r^{2j}$, where $T \approx 2\pi$ is the return time to the section and g_{2j-1} is the first non-zero coefficients in $\dot{F}$.

The proof of the analyticity of the first integral in the center case can be obtained by direct estimations of the coefficients (like in [67] and [51]). There exists a geometrical proof, due to R. Moussu [63], which uses the monodromy transformation Δ_γ associated with a loop γ in one of the separatrices $\Gamma_+ = \{x + \mathrm{i}y + \cdots = 0\} \subset (\mathbb{C}^2, 0)$ (see Theorem 3.20 and Definition 3.42 below). Δ_γ is a holomorphic map from a small disc $D \simeq (\mathbb{C}, 0)$ transversal to Γ_+ at the base point of the loop γ (which surrounds the point $x = y = 0$ in Γ_+) to itself. Due to the vanishing of the Poincaré–Lyapunov quantities it is an identity map. So, we can define the first integral F; firstly on $(\mathbb{C}^2, 0)\backslash$separatrices by the intersection of the leaves of the complex foliation $\mathcal{F}$ with the disc D (parametrized by $F|_D : D \to (\mathbb{C}, 0)$), next on the separatrices by putting $F = 0$. This function F turns out holomorphic in $(\mathbb{C}^2, 0)$. For more details we refer the reader to [63]. □

The next is the case of nilpotent singular points:

$$\dot{x} = y + \cdots, \qquad \dot{y} = \cdots.$$

Here we have so-called Takens–Bogdanov normal form

$$\dot{x} = y + a(x), \qquad \dot{y} = \varepsilon x^{s-1}, \tag{3.4}$$

where $a(x) = a_r x^r + a_{r+1} x^{r+1} + \cdots$, $a_r \neq 0$ and $\varepsilon = \pm 1$ (see Theorem 4.2 and Proposition 4.5 in Section 4). The coefficients a_j in the expansion of $a(x)$ are algebraic functions of the coefficients in the Taylor expansion of the vector field V and this normal form is analytic. The following result belongs to A. Sadovski and R. Moussu.

THEOREM 3.11. (See [68,62].) *The singularity $x = y = 0$ of a vector field V with the Takens–Bogdanov form* (3.4) *is monodromic if and only if*: *s is even, $\varepsilon = -1$ and either $r > \frac{s}{2}$ or $r = \frac{s}{2} = 2\tilde{r}$ is even and $ra_r^2 < 4$.*

In this case the singular point is a center iff all the coefficients a_{2j} with odd indices vanish. If a_{2j+1} is the first non-zero odd coefficient, then the singular point is stable (*respectively unstable*) *if $a_{2j+1} < 0$* (*respectively $a_{2j+1} > 0$*).

PROOF. The conditions: s even, $\varepsilon = -1$, $s \leqslant 2r$ and $s = 2r$ but r even follow from qualitative analysis of the phase portrait of the system (3.4).

If $s = 2m$ is even then the leading part of this system, with respect to the quasi-homogeneous gradation such that degree of y equals m times degree of x, is

$$\dot{x} = y + a_{2\tilde{m}} x^{2\tilde{m}}, \qquad \dot{y} = -x^{2m-1} \tag{3.5}$$

(compare Section 4.1). Here the term $a_{2\tilde{m}} x^{2\tilde{m}} = a_m x^m$, $m = 2\tilde{m}$, vanishes when $s < 2r$ and when $s = 2r$ but r is odd. This system is invariant with respect to the reflection $x \to -x$ and reversion of time; we say that the system is *time–reversible*. Hence system (3.4) is monodromic iff system (3.5) has center at the origin. Moreover, the change $z = x^m$ and division by x^{m-1} gives a linear system with eigenvalues $\lambda_{1,2} = \frac{1}{2}(ma_m \pm \sqrt{m^2 a_m^2 - 4m})$.

In the case with real eigenvalues the linear system has an invariant line (eigenspace) which gives a characteristic trajectory for system (3.4). Therefore $m^2a_m^2 - 4m < 0$.

To prove the second statement we assume that a_{2j+1} is the first non-zero odd coefficient. Then the initial part

$$\dot{x} = y + a_{2\tilde{m}}x^{2\tilde{m}} + \cdots + a_{2j}x^{2j}, \qquad \dot{y} = -x^{2m-1}$$

of the (3.4) is time-reversible and hence with center at the origin. It has first integral $F(x, y)$, which can be not analytic at the origin, but its first approximation is a first integral $F_0(x, y)$ of the quasi-homogeneous vector field (3.5). F_0 is of the form $(y - \mu_1x^m)^{\lambda_1}(y - \mu_2x^m)^{\lambda_2}$. The term $a_{2j+1}x^{2j+1}\partial_x$ is the first perturbation of the latter system and we can calculate the increments of the first integral F along trajectories of the perturbed system. Near the origin it is approximately the increment of F_0 and can be expressed in terms of the Melnikov type integral $\int_{F_0=h} \frac{\partial F_0}{\partial x}x^{2j+1}\,\mathrm{d}t$, where $\mathrm{d}t = -\mathrm{d}y/x^{2m+1}$.

The result is that the latter integral is positive. □

The following example demonstrates that in the case of more degenerate non-elementary monodromic singular point the center focus problem is not analytically solvable. (Another example of similar type was given by Y. Il'yashenko in [41].)

EXAMPLE 3.12. Consider the following family of vector fields

$$\begin{aligned}\dot{x} &= -y\big[(2\mu + 1)x^2 + y^2\big] + x^2\big[\lambda_1x^2 + \lambda_2(x^2 + y^2)\big],\\ \dot{y} &= x\big[x^2 + (1 - 2\mu)y^2\big] + xy\big[\lambda_1x^2 + \lambda_2(x^2 + y^2)\big].\end{aligned}$$

Each vector field from this family is of the type

$$V = V_d + f \cdot E,$$

where V_d is a homogeneous vector field of degree d, $E = x\partial_x + y\partial_y$ is the *Euler vector field* and $f(x, y)$ is a homogeneous polynomial of degree $d + k$. It has first integral of the so called *Darboux–Schwarz–Christoffel type*

$$H = x^{-k}M(u) + k\int^u M(v)P(v)B(v)^{-1}\,\mathrm{d}v, \quad u = y/x,$$

where $M(u) = \mathrm{e}^{g(u)}\prod(u - u_j)^{\alpha_j}$ is of generalized Darboux type and $B(u) = \prod(u - u_j)^{m_j}$ and $P(u)$ are polynomials. Indeed, in the variables x, u we get the Bernoulli equation $\frac{\mathrm{d}x}{\mathrm{d}u} = \frac{A(u)}{B(u)}x + \frac{P(u)}{B(u)}x^{k+1}$, $P(u) = f(1, u)$, with the first integral as above, where $M = \exp\int A/B$. The center conditions mean that the function $H(x, y)$ is single–valued in $\mathbb{R}^2\setminus 0$. This implies that the homogeneous function $x^{-k}M(y/x)$ is single valued and that $\int_{-\infty}^{\infty} M(v)P(v)B(v)^{-1}\,\mathrm{d}v = 0$.

In our case the first integral equals $H = (x^2 + y^2)^{-1/2} e^{\mu x^2/(x^2+y^2)} + F(y/x)$, where

$$F(u) = \int_{-\infty}^{u} e^{\mu/(v^2+1)} (v^2 + 1)^{-3/2} \bigl[\lambda_1 + \lambda_2 (v^2 + 1)\bigr] \, dv.$$

The center condition means that $F(-\infty) = F(\infty) = 0$. It can be rewritten in the form

$$\lambda_1 \Psi'(\mu) + \lambda_2 \Psi(\mu) = 0,$$

where

$$\begin{aligned}\Psi(\mu) &= \int_{-\infty}^{\infty} \frac{e^{\mu/(v^2+1)}}{\sqrt{v^2+1}} \, dv = \frac{2}{\sqrt{\mu}} e^{\mu} \int_0^{\sqrt{\mu}} e^{-z^2} \, dz \\ &= \frac{2}{\sqrt{\mu}} e^{\mu} \operatorname{Erf}(\sqrt{\mu}) = 2 e^{\mu} \Phi(1/2, 3/2; -\mu)\end{aligned}$$

where Erf is the probability integral and $\Phi(a, c; x) = 1 + \frac{a}{c}\frac{x}{1!} + \frac{a(a+1)}{c(c+1)}\frac{x^2}{2!} + \cdots$ is the confluent hypergeometric function. By the Kummer identity $\Phi(a, c; x) = e^x \Phi(c - a, c; -x)$ (see [9]) we have $\Psi(\mu) = \Phi(1, 3/2; \mu)$.

The function Ψ has the following properties (see [9]):

- $\Psi(\mu)$ is an entire function,
- $\Psi(\mu) \sim \Gamma(\frac{3}{2}) e^{\mu} \mu^{-1/2}$ as $\operatorname{Re} \mu \to \infty$,
- $\Psi(\mu) \sim -\frac{1}{2\mu}$ as $\operatorname{Re} \mu \to -\infty$.

Therefore Ψ is not of the form polynomial× exp(polynomial). The same holds for the function $\psi(\mu) = \lambda_1 \Psi'(\mu) + \lambda_2 \Psi(\mu)$. Therefore the function $\psi(\mu)$ has infinitely many zeroes in $\mathbb{C}$ and hence the set $\{(\lambda_1, \lambda_2, \mu)\colon \lambda_1 \Psi'(\mu) + \lambda_2 \Psi(\mu) = 0\}$ is not semi-algebraic.

The above shows that the problem of center is not algebraically solvable.

On the other hand Y. Il'yashenko proved the following almost algebraic solvability of the center–focus problem.

THEOREM 3.13. (See [41].) *Consider a class of germs of vector fields of the form*

$$\dot{x} = P_n + \cdots, \qquad \dot{y} = Q_n + \cdots \tag{3.6}$$

where $P_n(x, y)$ *and* $Q_n(x, y)$ *are fixed homogeneous polynomials such that the polynomial*

$$R_{n+1}(x, y) = x Q_n - y P_n \tag{3.7}$$

is non-zero in $\mathbb{R}^2 \setminus 0$. *Then the problem of stability of the singularity* 0 *is algebraically solvable with respect to the coefficients of higher order terms in* (3.6).

PROOF. The phase curves of the system (3.6) in the polar coordinates r, φ satisfy the equation

$$\frac{\mathrm{d}r}{\mathrm{d}\varphi} = \frac{A_{n+1}(\varphi)r + A_{n+2}(\varphi)r^2 + \cdots}{B_{n+1}(\varphi) + B_{n+2}(\varphi)r + \cdots}$$

where

$$A_{n+1}(\varphi) = S_{n+1}(\cos\varphi, \sin\varphi),\ A_{n+2}(\varphi), \dots ,$$
$$B_{n+1}(\varphi) = R_{n+1}(\cos\varphi, \sin\varphi),\ B_{n+2}(\varphi), \dots, S_{n+1}(x, y) = xP_n + yQ_n, \dots ,$$

are trigonometric polynomials. Moreover, $B_{n+1}(\varphi) \neq 0$.

Therefore, the solutions $r = r(\varphi; r_0)$ which satisfy the initial condition $r(0; r_0)$ can be expanded into power series $\sum a_j(\varphi)r_0^j$, where the coefficients are calculated recursively. Hence the Poincaré map is also expanded into power series $\mathcal{P}(r_0) = c_1 r_0 + c_2 r_0^2 + \cdots$. The first coefficient equals

$$c_1 = \exp\left(\int_0^{2\pi} \frac{A_{n+1}(\psi)}{B_{n+1}(\psi)}\,\mathrm{d}\psi\right) = \exp\left(2\int_{-\infty}^{\infty} \frac{P_n(1,u)}{R_{n+1}(1,u)}\,\mathrm{d}u\right) \tag{3.8}$$

and the sequel coefficients depend polynomially on the coefficients in the polynomials $A_{n+2}, B_{n+2}, A_{n+3}, \dots$. □

The quantity

$$g_1 = \ln c_1$$

is called the *first focus quantity*.

In formula (3.8) for the first focus quantity only one blowing-up of the singularity was used. N. Medvedeva generalized this formula to a more general case.

THEOREM 3.14. (See [57].) *Assume that we have a germ of the form* (3.6) *with monodromic singularity and such that the resolution of this singularity involves two blowing-ups* (*so the polynomial* $R_{n+1}(x, y)$ *has a zero*). *Then the first focus value is defined, i.e.* $c_1 = \mathcal{P}'(0) \neq 0$, *and equals*

$$g_1 = 2 \cdot \left(\text{V.P.} \int_{-\infty}^{\infty} \frac{P_n(1,u)}{R_{n+1}(1,u)}\,\mathrm{d}u\right),$$

where V.P. denotes the principal value of an integral from a rational function.

The proof of this theorem relies on decomposing the Poincaré map into composition of correspondence maps (defined by phase curves of the resolved foliation) near the saddle point (which is the intersection point of two exceptional divisors) and near the remaining parts of the divisor. We do not present this proof.

We finish this section by presenting (without proof) the following recent result of N. Medvedeva which solves the problem of analytical solvability of the Arnold's problem about center and focus. She restricts the problem to a fixed *class of germs with monodromic singular point*; such class consists of germs with fixed scheme of resolution of the singularity.

THEOREM 3.15. (See [58,59].) *The center–focus problem is analytically solvable in any class of germs with monodromic singularity. More precisely, the Poincaré map admits asymptotic expansion*

$$\mathcal{P}(\rho) = c_1\rho + \sum P_k(\ln\rho)\rho^{\nu_k},$$

where $c_1 > 0$, *the exponents* ν_j *grow to infinity and the coefficients* $P_k(\ln\rho)$ *are polynomials whose coefficients depend analytically on the coefficients of the jets* $j^{s_k}V$ *and* $s_k \to \infty$ *with* k.

3.2. *Poincaré–Dulac normal form*

Consider a germ of holomorphic vector field in $(\mathbb{C}^n, 0)$ of the form

$$\dot{x} = V(x) = Ax + \cdots \tag{3.9}$$

where A is the linearization matrix at the singular point $x = 0$. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of the matrix A.

DEFINITION 3.16. We say that the eigenvalues satisfy the *resonant relation of the type* $(i; k)$ if

$$\lambda_i = \lambda_1 k_1 + \cdots + \lambda_n k_n$$

where $k = (k_1, \dots, k_n)$ and k_j are non-negative integers.

We say that the germ (3.9) is of the *Poincaré type* if the convex hull of the set $\{\lambda_1, \dots, \lambda_n\}$ in the complex plane does not contain the origin.

The following result is known as the Poincaré–Dulac normal form.

THEOREM 3.17. (See [4,67,26].) *There exists a change* $x \to y = G(x)$ *given by formal power series which transforms the system* (3.9) *to the system*

$$\dot{y}_i = \lambda_i y_i + \sum_{(i;k)\ \text{resonant}} c_{i;k} y^k, \quad i = 1, \dots, n.$$

PROOF. We apply a series of changes of the type $x \to x' = x + \phi(x)$, where $\phi = (\phi_1, \dots, \phi_n)$ is homogeneous transformation of degree m,

$$\phi_i(x) = \sum_{|k|=m} b_{i,k} x^k.$$

The inverse map has the form $x' \to x = x' - \phi(x') + \cdots$. We strive to cancel all possible terms in the vector field of homogeneous degree m.

We have $\dot{x}_i' = (\text{old part of degree} < m) + (\text{old part of degree } m) + (L_A\phi)_i(x') + \cdots$ where

$$(L_A\phi)_i = \partial\phi_i(x)/\partial x \cdot Ax' - (A\phi)_i$$

is called the *homological operator*.

If the matrix A is diagonal then L is also diagonal:

$$L_A(x^k e_i) = \Bigl(\sum k_j\lambda_j - \lambda_i\Bigr)(x^k e_i)$$

where e_i form the standard basis in $\mathbb{C}^n$. Because the eigenvalues of L_A vanish only at the resonant terms all the non-resonant terms in the vector field of degree m can be canceled in this way.

If A is upper-triangular then, taking the suitable ordering of the basis $x^k e_i$, we get the triangular form of the operator L_A

$$L_A(x^k e_i) = \Lambda_{i,k} x^k e_i + \sum_{j<i,l} \text{const} \cdot x^l e_j.$$

□

EXAMPLE 3.18. Consider the case $n = 2$. The resonance means

(i) $\lambda_1 = k_1\lambda_1 + k_2\lambda_2$ and/or (ii) $\lambda_2 = k_1\lambda_1 + k_2\lambda_2$.

Next we divide the problem into several subcases.

(a) If $\lambda_1 = \lambda_2 = 0$ then all terms in the expansion of the vector field are resonant and nothing can be reduced (by means of Theorem 3.17).

(b) If $\lambda_1 = 0 \neq \lambda_2$ (the s*addle–node* case) of then in (i) we must have $k_2 = 0$ and in (ii) we must have $k_2 = 1$. Thus the Poincaré–Dulac formal normal form is

$$\dot{x} = x^2 f(x), \qquad \dot{y} = \lambda_2 y\bigl(1 + g(x)\bigr).$$

(c) If $\lambda_2/\lambda_1 \geqslant 1$ is a rational number (the case of *resonant node*) then (i) cannot hold and (ii) can hold with $k_1 = \lambda_2/\lambda_1$, $k_2 = 0$ provided that $\lambda_2/\lambda_1 = m$ is integer. The normal form is either linear system or

$$\dot{x} = \lambda_1 x, \qquad \dot{y} = m\lambda_1 y + ax^m.$$

(d) If $\lambda_2/\lambda_1 = -p/q < 0$ is rational (the case of *resonant saddle*) then the normal form is

$$\dot{x} = \lambda_1 x\bigl(1 + f(x^q y^p)\bigr), \qquad \dot{y} = \lambda_2 y\bigl(1 + g(x^q y^p)\bigr).$$

(e) If λ_2/λ_1 is irrational (the non-resonant case) then the normal form is linear.

In some cases the Poincaré–Dulac normal form is analytic.

THEOREM 3.19. (See [67,17].) *If the germ* $V(x)$ *is of the Poincaré type* (*see Definition* 3.16) *then the change* $y = G(x)$ *which reduces it to the polynomial Poincaré–Dulac normal form can be chosen analytic.*

PROOF. (We follow the paper [17] by N. Brushlinskaya.) After changing (if necessary) the time $t \to ct$ we can assume that the eigenvalues lie in the right half-plane, $\operatorname{Re}\lambda_j < 0$.

Note also that there can exist only finitely many resonant relations; for instance in the case (c) of Example 3.18 there is only one such relation. Let $0 < m < -\operatorname{Re}\lambda_j < M$ and let an integer r be such that $\frac{M}{n} < r$. Then all resonant terms $x^k e_i$ have degree $|k| = k_1 + \cdots + k_n < r$. We assume that the part $W(x) = Ax + W_2(x)$ of order $< r$ of the vector field $V(x)$ is in the normal form, $V_1 = V - W = \mathrm{O}(|x|^r)$.

Choose so-called Lyapunov metric $\|\cdot\|$ in $\mathbb{C}^n$, such that for $t > 0$ we have

$$\left\|g^t(x_0)\right\| \leqslant \mathrm{e}^{-mt}\|x_0\|, \qquad \left\|(g^t_*)^{-1}\right\| \leqslant \mathrm{e}^{Mt},$$

where $g^t(x_0)$ is the phase flow map generated by the vector field W (solution $x(t)$ after the time t with initial condition $x(0) = x_0$) and $g^t_* = \frac{\partial g^t}{\partial x_0} = \Psi(t, x_0)$ is the solution of the corresponding variation equation $\dot{\Psi} = \frac{\partial W}{\partial x}\Psi$, $\Psi(0) = I$. If we transform the matrix A to the form $\operatorname{diag}(\lambda_1, \ldots, \lambda_n) + \epsilon A_1$, where A_1 is nilpotent, then the Lyapunov metric is the Hermitian metric associated with this new basis of $\mathbb{C}^n$.

As in other similar situations we reduce the problem to a fixed point problem for some operator which is contracting in suitable Banach space. We consider changes of the form $x = y + G(y)$, where $\|G(y)\| = \mathrm{O}(\|y\|^r)$. As the Banach space we choose the space $E_{r,\rho}$ of vector fields $G(y)$, analytic in a ball $\|y\| \leqslant \rho$ and having zero at $y = 0$ of order r, and with the norm $\|G\|_{r,\rho} = \sup_{\|y\|\leqslant\rho} \|G(y)\|/\|y\|^r$.

Let $L_W G = \frac{\partial G}{\partial y}W - \frac{\partial W}{\partial y}G$ be the Poisson bracket of the vector fields V and G. Then the equation for conjugation of V and W_0 by means of the diffeomorphism $y + G(y)$ leads to the equation

$$L_V G = \Phi(G), \tag{3.10}$$

where $\Phi(G) = V_1(y + G(y)) + W_2(y + G(y)) - W_2(y) - \frac{\partial W_2}{\partial y}G(y)$.

The operator L_W has the right inverse defined by means of the method of characteristics for the non-homogeneous PDE $L_W G = H$. If we put $g(t) = G(g^t(y_0))$ and $h(t) = H(g^t(y_0))$, then we get the equation $\dot{g} - a(t)g = h$, $a(t) = \frac{\partial W}{\partial y}(g^t(y_0))$. Its solution is $g(t) = -\Psi(t, y_0)\int_t^\infty \Psi^{-1}(s, y_0)h(s)\,\mathrm{d}s$, where the latter integral is absolutely convergent. Putting $t = 0$ and $y_0 = y = g^0(y)$ we find

$$G(y) = (L_W^{-1}H)(y) = -\int_0^\infty \Psi^{-1}(t, y)H\bigl(g^t(y)\bigr)\,\mathrm{d}t.$$

Due to the properties of the Lyapunov metric we have $\|H(g^t(y))\|_{r,\rho} \leqslant \|H\|_{r,\rho}\|y\|^r \mathrm{e}^{-rmt}$ and $\|\Psi^{-1}\| \leqslant \mathrm{e}^{Mt}$ and hence the operator L_W^{-1} is bounded.

Finally, standard estimated for the non-linear part $\Phi(G)$ in (3.10) imply that the operator $L_W^{-1} \circ \Phi$ is a contraction on some ball $\{\|G\|_{r,\rho} \leqslant R\}$ in the space $E_{r,\rho}$. □

The next theorem, which concerns the cases of saddle–node and of saddle, was firstly proved by C. Briot and J. Bouquet [15] and can be called the *analytic Hadamard–Perron theorem*. It will be used in further sections.

THEOREM 3.20. *Consider an analytic planar system*

$$\dot{x} = \lambda_1 x + \cdots, \qquad \dot{y} = \lambda_2 y + \cdots$$

such that $\lambda_2 \neq 0$ *and* $\lambda_1/\lambda_2 \leqslant 0$. *Then there exists an invariant analytic curve tangent to* $x = 0$ *at the origin.*

PROOF. In the case of a saddle–node we can assume that $\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{f(x,y)}{\lambda_2 y + g(x,y)}$ where f, g have second order zero at the origin and $f(0, y) = \mathrm{O}(y^3)$. In the case of a saddle we can assume that $\frac{\mathrm{d}x}{\mathrm{d}y} = \frac{\lambda_1 x + f(x,y)}{\lambda_2 y + g(x,y)}$ with f, g having trird order zero at the origin (because the resonant terms in the Poincaré–Dulac normal form start from cubic monomials).

We put the equation of the invariant curve in the form $x = y^2\phi(y)$. In the case of saddle–node we get the equation $y^2\phi' + 2y\phi = f(y^2\phi, y)/(\lambda_2 y + g(y^2\phi, y))$. After division by y^2 it takes the form $\phi' + 2\phi/y = F(y, \phi)$ (F – analytic) which can be written in the integral form

$$\phi(y) = y^{-2}\int_0^y s^2 F\bigl(s, \phi(s)\bigr)\,\mathrm{d}s.$$

The latter equation is solved using the contraction principle in some ball in the Banach space of analytic functions ϕ defined in a disc $|y| < \epsilon$ with the sup-norm.

In the saddle case we arrive to the equation $\phi' = (\lambda - 2)\phi/y + F(y, \phi)$, $\lambda = \lambda_1/\lambda_2 < 0$, or to the integral equation $\phi = y^{\lambda-2}\int_0^y s^{2-\lambda} F(s, \phi(s))\,\mathrm{d}s$. □

3.3. *Ecalle–Voronin moduli*

After the resolution theorem our next task is to describe the analytic classifications of planar holomorphic foliations near elementary singular points. As we have said this theory is practically completed. However before classifying the two-dimensional vector fields we shall present the classification of analogous one-dimensional objects.

These objects are the germs of analytic diffeomorphisms $f : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$:

$$z \to f(z) = \mu z + a_2 z^2 + \cdots. \tag{3.11}$$

DEFINITION 3.21. Two germs f and $\tilde{f}$ of conformal diffeomorphisms of $(\mathbb{C},0)$ are *analytically equivalent* (respectively, *formally equivalent, topologically equivalent)* if there is a germ $h(z)$ of analytic diffeomorphism (respectively, a formal power series $h(z) \sim h_1 z + h_2 z^2 + \cdots$, a homeomorphism $h(z)$) conjugating f with $\tilde{f}$. It means that

$$\tilde{f} = h^{-1} \circ f \circ h.$$

Analogous definition takes place in the case of multi-dimensional diffeomorphisms of $(\mathbb{C}^n,0)$.

The diffeomorphism h transforms the orbits of $\tilde{f}$, i.e. the sets $\{\ldots, (\tilde{f})^{-1}(z), z, \tilde{f}(z), (\tilde{f})^2(z) = \tilde{f} \circ \tilde{f}(z), \ldots\}$, to the orbits of f. Thus f and $\tilde{f}$ have diffeomorphic sets of fixed points and of periodic points of given period.

The following result is an analogue of the Poincaré–Dulac normal form (Theorem 3.17).

THEOREM 3.22. *If the multiplicator μ in* (3.11) *is not a root of unity then the map f is formally linearizable, i.e. it is formally equivalent to the linear map μz.*

If $|\mu| \neq 1$ then f is analytically linearizable.

PROOF. The analogue to the homological equation for changes of the form $z \to z + g_k z^k$ applied to the maps $\mu z + a_k z^k + \cdots$ is following

$$(\mu - \mu^k) g_k = a_k.$$

This gives the first statement of the theorem.

The case $|\mu| \neq 1$ for a diffeomorphism corresponds to the Poincaré type for a vector field. The proof of analyticity of the linearizing transformation repeats the Brushlinskaya's proof of Theorem 3.19. Instead of the continuous flow $\{g^t\}_{t\in\mathbb{R}}$ one should use the discrete cascade $\{f^n\}_{n\in\mathbb{Z}}$. We omit the details. □

We focus our attention on *germs of analytic diffeomorphisms tangent to identity*

$$f(z) = z + az^{p+1} + \cdots, \quad a \neq 0.$$

Note that, after suitable normalization, we can assume that $a = 1$. Following Il'yashenko in [42] we denote the space of such germs by $\mathcal{A}_p$.

PROPOSITION 3.23. *Any germ from $\mathcal{A}_p$ is formally equivalent to the germ*

$$f_{p,\lambda} = g_w^1,$$

where g_w^t is the phase flow map after time t of the vector field

$$w = w_{p,\lambda} = \frac{z^{p+1}}{1 + \lambda z^p} \frac{\partial}{\partial z}$$

and $\lambda \in \mathbb{C}$ is the invariant of the formal classification.

PROOF. We cancel the suitable terms in f by means of the conjugating maps $z \to z + b_l z^l$. Simple calculations show that the leading contribution from this change is the term

$$(l - p - 1) b_l z^{l+p}.$$

If $l \neq p + 1$ then the corresponding power of z in f can be canceled. There remains only the power z^{2p+1}.

Thus the formal normal form is $z + z^{p+1} + \nu z^{2p+1}$. Expanding $f_{p,\lambda}$ into powers of z we see that the latter form is formally equivalent to $f_{p,\lambda}$ for $\lambda = p + 1 - \nu$. □

The space of germs f which are formally equivalent to $f_{p,\lambda}$ is denoted by $\mathcal{A}_{p,\lambda}$.

PROPOSITION 3.24. (See [18,71].) *Any germ from $\mathcal{A}_p$ is topologically equivalent to $g = f_{p,0} = z(1 - z^p)^{-1/p}$.*

PROOF. This proof relies on drawing the phase portrait of the vector field $\dot{z} = z^{p+1}/p$, whose phase flow (with real time) is $g^t(z) = z(1 - tz^p)^{-1/p}$. It consists of $2p$ elliptic sectors and $2p$ parabolic sectors. Next one shows that if we perturb the map g^1 within the class $\mathcal{A}_p$ then the qualitative behavior of the orbits remain the same. We omit the details (which can be found also in [89]). □

The functional moduli of the analytical classification of germs of conformal mappings are connected with the structure of the space of orbits of the action of such a mapping on a neighborhood of the fixed point. It means that we look at the space U^*/f, where $U^* = U \setminus 0$ is a punctured neighborhood of $z = 0$ and points z and $f(z)$ are treated equivalent. (The space U^*/f is Hausdorff, but the analogous space U/f space is not Hausdorff.)

The structure of the orbit space can be better seen in the t-chart where

$$t = t(z) = -1/pz^p + \lambda \ln z$$

is the (complex) time in the vector field $w_{p,\lambda}$. Because the solution of the equation $\dot{z} = w_{p,\lambda}$ is $t = t(z) + \text{const}$, the map $f_{p,\lambda} = g_w^1$ takes the simple form

$$t \to t + 1.$$

Assume, for example, that $\lambda = 0$ (then $t(z)$ is single-valued) and that $p = 1$. Then the neighborhood $U^* = \{|z| < r\}$ is replaced by $W = \{|t| > R\}$. The set W is divided into the vertical strips of width 1. Identifying two sides of such strip we get a cylinder which can be identified with the doubly punctured sphere $\mathbb{C}^*$: $-i\infty \to 0$, $+i\infty \to \infty$. The space $W/(\mathrm{id} + 1)$ consists of two copies of $\bar{\mathbb{C}}$ with identifications near 0 and ∞, like at Fig. 6. The vertical strips from the t-chart correspond to the crescents in the z-chart.

In the general case (λ and p arbitrary and f close to $f_{p,\lambda}$) the chart $t(z)$ is a well defined diffeomorphism in sectors (of width $> 2\pi/p$)

$$S_j = \left\{ \frac{j-2}{2p}\pi - \delta < \arg z < \frac{j}{2p}\pi + \delta \right\},$$

$j = 1, \dots, 2p$, where $\delta > 0$ is small.

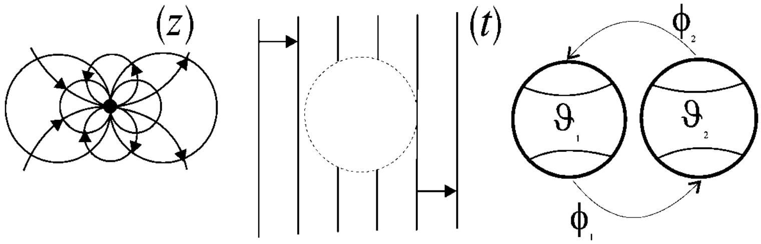


Fig. 6.

The function $t(z)$ sends S_j to $\tilde{S}_j$. In $\tilde{S}_j$ we denote by $t_j = t(z)|_{S_j}$ the local charts, where the branches of $\ln z$ are such that they are real in $S_1 \cap \mathbb{R}$, $t_j = t_{j+1}$ in the intersections $S_j \cap S_{j+1}$ for $j < 2p$ and $t_{2p} = t_1 + 2\pi \mathrm{i}\lambda$.

The maps $f_{p,\lambda}$ and f are replaced by the mappings

$$\tilde{f}_{p,\lambda} : t_j \to t_j + 1,$$

$$\tilde{f} : t_j \to t_j + 1 + R_j(t_j),$$

in $\tilde{S}_j$. Here we can assume that

$$R_j(t) = \mathrm{O}(t^{-M})$$

for some large M. The vertical strips from Fig. 5 and orbit spaces will be only slightly perturbed. The vertical lines are replaced by the iterations of one line, $f^n(l_0)$ where, due to the fact that $R(t)$ is small, there are no intersections between them.

LEMMA 3.25. *Each quotient space $\tilde{S}_j/\tilde{f}$ is conformally equivalent to $\mathbb{C}^*$.*

PROOF. This quotient space, which we denote by T, is obtained from the (non-straight) strip between the lines $l_0 = \{\mathrm{Re}\, t = c\}$ and $\tilde{f}(l_0)$ by identifying its sides. Here $\tilde{f}(c + \mathrm{i}\tau) = c + \mathrm{i}\tau + 1 + R_j(c + \mathrm{i}\tau)$ where $R(t)$ is small.

We treat $\mathbb{C}^*$ as the straight strip $c \leqslant \mathrm{Re}\, t \leqslant c + 1$ with identified sides. It is parametrized by $\theta \in [0, 1]$ and $\tau \in \mathbb{R}$: $t = c + \theta + \mathrm{i}\tau$.

We construct a map from $\mathbb{C}^*$ to T

$$F(c + \theta + \mathrm{i}\tau) = c + \theta + \mathrm{i}\tau + \theta R_j(t).$$

Because $F'_t = 1 + R(t)/2 + \theta R'_j(t)$, $F'_{\bar{t}} = R_j(t)/2$, the map F is quasi-conformal and satisfies the condition of regularity from Proposition 3.74 in Appendix 1. The modulus of quasi-conformality $\mu_F = F'_{\bar{t}}/F'_t$ is small and tends to zero at the ends of the strip $\{c < \mathrm{Re}\, t < c + 1\}$; (they correspond to 0 and ∞ in $\bar{\mathbb{C}}$).

Applying Proposition 3.40 from Appendix 1 we get the assertion of the lemma. □

Therefore, U^*/f is topologically equivalent to $2p$ copies of $\bar{\mathbb{C}}$ with suitable identifications of the neighborhoods of 0 and ∞. More precisely, we have the spheres $\bar{\mathbb{C}} \times \{j\}$, $j = 1, \dots, 2p$, and identifying maps of the form:

$$\phi_j : (\bar{\mathbb{C}}, 0) \times \{j\} \to (\bar{\mathbb{C}}, 0) \times \{j+1\}, \quad j \text{ even},$$
$$\phi_j : (\bar{\mathbb{C}}, \infty) \times \{j\} \to (\bar{\mathbb{C}}, \infty) \times \{j+1\}, \quad j \text{ odd}.$$

The functional invariants of the classification of germs of maps are the invariants of the holomorphic structure on the space U^*/f.

In the spheres one uses conformal coordinates ϑ_j, e.g. $\vartheta_j = \mathrm{e}^{2\pi \mathrm{i} t_j}$ if $f = f_{p,\lambda}$. Generally, $\vartheta_j = \mathrm{e}^{2\pi \mathrm{i}\tau_j}$ for suitable 'times' $\tau_j = t_j + \mathrm{o}(1)$.

ϑ_j are defined uniquely modulo multiplication by a constant. If we have chosen ϑ_1, then ϑ_2 can be taken in such a way that $\vartheta_2 = \phi_1(\vartheta_1) = \vartheta_1 + \mathrm{o}(\vartheta_1)$ as $\vartheta_1 \to \infty$, $\vartheta_3 = \phi_2(\vartheta_2) = \vartheta_2 + o(\vartheta_2)$ is such that $\phi_2'(0) = 1$, etc. However the charts ϑ_{2p} and ϑ_1 near ∞ cannot be correlated in this way. We have $\vartheta_{2p} = \mathrm{e}^{2\pi \mathrm{i} t_{2p}} = \mathrm{e}^{2\pi \mathrm{i} t_1 - 4\pi^2\lambda} = \nu\vartheta_1$. Thus $\phi_{2p}(\vartheta_{2p}) = \nu\vartheta_{2p} + \mathrm{o}(\vartheta_{2p})$, $\nu = \exp(-4\pi^2\lambda)$.

REMARK 3.26. Changing the chart $\vartheta_1 \to C\vartheta_1$ induces the same changes in the spheres $\bar{\mathbb{C}}_j$ and the maps ϕ_j become conjugated by means of C. This non-uniqueness of the charts ϑ_j is equivalent to the non-uniqueness of the 'times', $\tau_j \to \tau_j + \theta$ where $C = \mathrm{e}^{2\pi \mathrm{i}\theta}$, and the non-uniqueness of the coordinate z, $z \to g_w^\theta(z)$.

The above suggests the following definition.

DEFINITION 3.27. Consider the space of collections $\phi = (\phi_1, \dots, \phi_{2p})$ of germs of maps $\phi_j : (\bar{\mathbb{C}}, 0(\infty)) \times \{j\} \to (\bar{\mathbb{C}}, 0(\infty)) \times \{j+1\}$ with identity linear parts for $j < 2p$ and $\phi_{2p}'(\infty) = \nu$. Two such collections ϕ, ϕ' are called *equivalent*, $\phi \sim \phi'$, iff there exists $C \in \mathbb{C}^*$ such that

$$\phi_i \circ C = C\phi_i'.$$

The set of equivalence classes of such ϕ's is *the moduli space* and is denoted by $\mathcal{M}_{p,\lambda}$.

If $f \in \mathcal{A}_{p,\lambda}$, then the construction (with the strips in $\tilde{S}_j$'s and identifications of U^*/f with a collection of glued together Riemann spheres) defines a collection ϕ. Thus we have a mapping from $\mathcal{A}_{p,\lambda}$ to $\mathcal{M}_{p,\lambda}$. It is the first statements of the following *Ecalle–Voronin theorem*.

THEOREM 3.28. (See [28,83].) *There is a map $f \to \mu_f$ from $\mathcal{A}_{p,\lambda}$ to $\mathcal{M}_{p,\lambda}$ with the following properties*:

- *if f and $\tilde{f}$ are holomorphically equivalent then $\mu_f = \mu_{\tilde{f}}$* (*invariance*);
- *if $\mu_f = \mu_{\tilde{f}}$ then f and $\tilde{f}$ are holomorphically equivalent* (*equimodality*);
- *for any class $[\phi] \in \mathcal{M}_{p,\lambda}$ there exists $f \in \mathcal{A}_{p,\lambda}$ such that $[\phi] = \mu_f$* (*realization*).

J. Ecalle proved a variant of this result in 1975. His proof relies on exploiting the Borel transform. S. Voronin's proof appeared in 1981 and it relies on application of analytic methods (almost complex structures and quasi-conformal mappings). Below we present the Voronin's proof (following the Il'yashenko's paper [42]).

However it must be said that the Ecalle–Voronin moduli were already described by G. Birkhoff in 1939 in [12] and partial results appeared in the woks of L. Leau [45] and other mathematicians. Very interesting review of the history of the problem of analytic classification is presented in the lectures of F. Loray [49].

The essential part of the proof of the Ecalle–Voronin theorem is the fact that the dynamics defined by the map f and by the model map are conjugated holomorphically in sectors.

THEOREM 3.29. *In every sector S_j there is an unique analytic diffeomorphism of the form*

$$H_j(z) = z + h_j(z), \quad h_j = \mathrm{o}(z^{p+1}),$$

conjugating f with $f_{p,\lambda}$.

If we skip the assumption $h_j = \mathrm{o}(z^{p+1})$ then we can replace H_j by $g_w^\theta \circ H_j$ for some $\theta \in C$, where g_w^θ is the phase flow generated by the vector field $w = [z^{p+1}/(1 + \lambda z^p)]\partial_z$.

PROOF. It is enough to construct the conjugating diffeomorphisms in the t-charts, i.e. $\tilde{H}_j(t) = t + \tilde{h}_j(t)$, $t \in \tilde{S}_j$. Then $H_i = t^{-1} \circ \tilde{H}_j \circ t$.

The conjugation condition $\tilde{H}_j \circ \tilde{f} = \tilde{f}_{p,\lambda} \circ \tilde{H}_j$, i.e. $t + 1 + R_j + \tilde{h}_j \circ \tilde{f} = t + \tilde{h}_j + 1$, means that

$$\tilde{h}_j = \tilde{h}_j \circ \tilde{f} + R_j.$$

Iterating this equation (with respect to $\tilde{h}_j$) infinitely many times, we find the solution

$$h(t) = \sum_0^\infty R_j \circ \tilde{f}^n(t).$$

The condition $|R_j| < |t|^{-M}$ and the property $\tilde{f}^n(t) \sim \mathrm{const} \cdot n$ ensures that this series is convergent and represents an analytic function of t.

To prove the second statement we compose f with suitable g_w^θ such that $f - f_{p,\lambda} = \mathrm{o}(z^{p+1})$ and apply the first statement. □

Take the system $(H_1, \ldots, H_{2p})$ of germs diffeomorphisms from Theorem 3.29. It can be interpreted as certain Čech cochain. Namely we consider the (non-Abelian) sheaf on S^1 of germs of analytic diffeomorphisms strongly tangent to identity. If U is an open arc in S^1 then the group of sections of this sheaf at U consists of diffeomorphism $H_{S(U)} = \mathrm{id} + h_{S(U)}$, $h_{S(U)}(z) = \mathrm{O}(z^{p+1})$ defines on a sector $S(U)$ with base at U (in the polar coordinates $S(U) = \{z = r\mathrm{e}^{\mathrm{i}\theta}, \theta \in U\}$, see [54]).

The system

$$H = (H_1, \dots, H_{2p})$$

defines the Čech cochain associated with the covering of S^1 induced by the sectors S_i. We consider the coboundary of this cochain, i.e. $\delta H = (\Phi_1, \dots, \Phi_{2p})$

$$\Phi_j = H_{j+1} \circ H_j^{-1}$$

in $S_j \cap S_{j+1}$.

PROPOSITION 3.30.

(a) *The maps Φ_j commute with the standard map $f_{p,\lambda}$.*

(b) *They are exponentially close to identity*

$$\left|\Phi_j(z) - z\right| < \mathrm{e}^{-c/|z|^p};$$

(*equivalently, $H_j - H_{j+1}$ are exponentially small*).

(c) *They are defined uniquely up to conjugation with the map g_w^θ for some $\theta \in \mathbb{C}$.*

(d) *In suitable charts ϑ_j in the spheres $\bar{\mathbb{C}} \times \{j\}$ they define the element μ_f of the moduli space $\mathcal{M}_{p,\lambda}$.*

PROOF. (a) The first statement follows from the identity $H_j \circ f = f_{p,\lambda} \circ H_j$.

(b) The second property follows from the first one. Indeed, in the t-chart the commutativity of the induced map

$$\tilde{\Phi}_j(t) = t + \tilde{\phi}_j(t)$$

with the standard map $t \to t+1$ means that the functions $\tilde{\phi}_j(t)$ are periodic, $\tilde{\phi}_j(t+1) = \tilde{\phi}_j(t)$. Thus $\tilde{\phi}_j$ are expanded into the Fourier series $\sum c_{jl}\mathrm{e}^{2\pi \mathrm{i} l t}$.

The function Φ_j is analytic in $S_j \cap S_{j+1}$, which corresponds to the domain $\operatorname{Im} t > t_0 > 0$ in the $t = t_j$-chart in $\tilde{S}_j$ if j is odd, and to the domain $\operatorname{Im} t < -t_0$ if j is even. Moreover, $\tilde{\Phi}_j(t) = t + \mathrm{o}(t)$ as $t \to \infty$ (because $\tilde{H}_j$ have this property).

This implies that

$$\begin{aligned} \tilde{\Phi}_j(t_j) &= t_j + \sum_{l>0} c_{jl}\mathrm{e}^{2\pi \mathrm{i} l t_j}, && j \text{ odd}, \\ \tilde{\Phi}_j(t_j) &= t_j + \sum_{l<0} c_{jl}\mathrm{e}^{2\pi \mathrm{i} l t_j}, && j \text{ even}, \\ \tilde{\Phi}_{2p}(t_{2p}) &= -2\pi \mathrm{i}\lambda + t_{2p} + \sum_{l<0} c_{jl}\mathrm{e}^{2\pi \mathrm{i} l t_j}. \end{aligned} \tag{3.12}$$

The constant in the latter formula follows from the difference between the local charts $t_{2p} = t_1 + 2\pi \mathrm{i}\lambda$ and t_1.

Now we have $\Phi_j = t_{j+1}^{-1} \circ \tilde{\Phi}_j \circ t_j$ with $t_j = -1/(pz^p) + \lambda \ln z$. If j is odd then we see that $\Phi_j(z) = z + \mathrm{O}(|c_{j1} z^{\alpha} \mathrm{e}^{-2\pi \mathrm{i}/(pz^p)}|)$. Analogous estimates hold in the case of even j.

(c) This point follows from the second statement of Theorem 3.29.

(d) Recall that the moduli space $\mathcal{M}_{p,\lambda}$ is the space of conformal structures on the orbit space U^*/f. Thus we should associate with any system $\tilde{\Phi}$ a system of conformal charts on the spaces S_j/f (or $\tilde{S}_j/\tilde{f}$) and a system of gluing maps between them.

We introduce new charts $\tau_j := \tilde{H}_j(t_j) = t_j + \cdots$ in $\tilde{S}_j$ having the property that they conjugate f with the translation id$+1$. They are also defined by the formulas

$$\tau_j = t_j \circ H_j.$$

One easily checks that $\tau_j \circ f \circ \tau_j^{-1} = \mathrm{id} + 1$. It is also clear that the space $\tilde{S}_j/\tilde{f}$ is conformally equivalent to the τ_j-plane divided by the action of the translation id$+1$.

The identification of $\tilde{S}_j/\tilde{f}$ with $\mathbb{C}^* = \mathbb{C}^* \times \{j\}$ is given by the formula

$$\tau_j \to \vartheta_j = \mathrm{e}^{2\pi \mathrm{i} \tau_j}.$$

The gluing maps $\vartheta_j \to \vartheta_{j+1} = \phi_j(\vartheta_j)$ (see Definition 3.27) arise from the differences between τ_j and τ_{j+1} in $\tilde{S}_j \cap \tilde{S}_{j+1}$. We have $\phi_j = [\exp 2\pi \mathrm{i}(\cdot)] \circ (\tau_{j+1} \circ \tau_j^{-1}) \circ [(1/2\pi \mathrm{i}) \ln(\cdot)]$. Therefore we should find an expression for $\tau_{j+1} \circ \tau_j^{-1}$. We have

$$\tau_{j+1} \circ \tau_j^{-1} = t_{j+1} \circ H_{j+1} \circ H_j^{-1} \circ t_j^{-1} = \tilde{\Phi}_j.$$

If $\tilde{\Phi}_j$ are defined by the formulas (3.12) then $\phi_j(\vartheta_j) = \exp\{2\pi \mathrm{i}[\tau_j + \sum c_{jl} \vartheta_j^l]\}$, or

$$\phi_j = \vartheta_j + \sum d_{j,l} \vartheta_j^l,$$
$$\phi_{2p} = \nu \vartheta_{2p} + \sum d_{jl} \vartheta_{2p}^l.$$

(Above the sums run either over positive or over negative l's.)

Like t_j, the variables τ_j are defined uniquely modulo translation by a constant θ. This means that the charts ϑ_j in $\mathbb{C}^* \times \{j\}$ are defined uniquely modulo multiplication by the constant $C = \mathrm{e}^{2\pi \mathrm{i} \theta}$. □

PROOF OF ECALLE–VORONIN THEOREM 3.28. (a) *Invariance of the class* μ_f. Let f, g be two analytically equivalent germs from the class $\mathcal{A}_{p,\lambda}$, h being the conjugating holomorphism: $h \circ f = g \circ h$. Let H, G be the cochains normalizing f, g respectively. Then the cochain $G \circ h$ normalizes f. Moreover, we find that $g \circ h = g_w^{\theta} \circ H$ (as H and $G \circ h$ conjugate the same f with $f_{p\lambda}$). So

$$G_{j+1} \circ G_j^{-1} = g_w^{\theta} \circ (H_{j+1} \circ H_j^{-1}) \circ g_w^{-\theta}.$$

This means that, if $\tilde{\Phi}$, $\tilde{\Psi}$ are the coboundaries associated with f, g (expressed in the t-charts), then they are equivalent by means of the translation map id$+\theta$. Hence the maps

$\phi_j = [\exp 2\pi \mathrm{i}(\cdot)] \circ \tilde{\Phi}_j \circ [(1/2\pi \mathrm{i}) \ln(\cdot)]$ and $[\exp 2\pi \mathrm{i}(\cdot)] \circ \tilde{\Psi}_j \circ [(1/2\pi \mathrm{i}) \ln(\cdot)]$ are equivalent by means of a rescaling $C \times \mathrm{id}$.

(b) *Equimodality*. Let $\tilde{\Phi}$, $\tilde{\Psi}$ be two collections associated with maps f, g respectively. Assume that they are equivalent by means of translation. Therefore choosing another normalizing cochain we can obtain the coincidence of the cochains $\tilde{\Phi}$ and $\tilde{\Psi}$.

But the identity $G_{j+1} \circ G_j^{-1} = H_{j+1} \circ H_j^{-1}$ is equivalent to

$$H_j^{-1} \circ G_j = H_{j+1}^{-1} \circ G_{j+1}.$$

This means that the maps $H_j^{-1} \circ G_j$ are prolonged to a punctured neighborhood of $z = 0$ as a single valued mapping. By the Riemann removability of singularities theorem we prolong it to an analytic map h in a whole neighborhood of 0.

Moreover, because G_j conjugates g with $f_{p,\lambda}$ and H_j^{-1} conjugates $f_{p,\lambda}$ with f then h conjugates g with f.

(c) *Realization*. Let $\mu \in \mathcal{M}_{p,\lambda}$ be an element of the moduli space. We have to show that $\mu = \mu_f$ for some map f.

We choose a representative $\tilde{\Phi}$ of μ and associated with it the 1-cocycle $\Phi = (\Phi_j)$ of germs of holomorphic maps in the sectors $S_j \cap S_{j+1}$, all but one tangent to the identity.

Consider the disjoint collection of the sectors, i.e. $S_j \times \{j\}$. Denote by z_j the charts induced by (z, j). We glue these sectors using the mappings $\Phi_j : z_{j+1} \sim \Phi_j(z_j)$. We obtain certain topological space $\mathcal{S}$ homeomorphic to a punctured disc. This space has a conformal structure arisen from the coordinates z_j. The space $\mathcal{S}$ (and its closure $\bar{\mathcal{S}}$) admits a homeomorphism f_0 induced from $f_{p,\lambda}$ at each $S_j \times \{j\}$. f_0 is analytic with respect to the conformal structure. Our task is to show that the conformal structure of $\bar{\mathcal{S}}$ coincides with the conformal structure of the punctured disc and that f_0 belongs to $\mathcal{A}_{p,\lambda}$.

Let $\chi_j(z)$ be an infinitely smooth partition of unity associated with the covering S_j. We need that the functions χ_j and their gradients have at worst power type singularities as $z \to 0$.

We construct the map

$$H_0 = \sum \chi_j(z_j) z_j$$

from $\mathcal{S}$ to a punctured disc $D^* \subset \mathbb{C}$.

LEMMA 3.31. *The mapping H_0 is quasi-conformal.*

PROOF. (See Appendix 1.) In $S_j \cap S_{j+1}$ we have $H_0 = \chi_j z_j + \chi_{j+1} z_{j+1} = z_j + (\Phi_j(z_j) - z_j)\chi_{j+1}$ which implies

$$\partial H_0/\partial z_j = 1 + \mathrm{o}(1), \qquad \partial H_0/\partial \bar{z}_j = (\Phi_j - z_j)\partial \chi_{j+1}/\partial \bar{z}_j = \mathrm{o}(1).$$

This means that the almost complex structure, induced on D^* by H_0, is defined by the Beltrami differential $\mu(\mathrm{d}\bar{z}/\mathrm{d}z)$, where $\mu = \mu(H_0^{-1})$ is exponentially small together with its derivatives. □

Let us prolong μ to the whole disc D. By Theorem 3.72 from Appendix 1 the almost complex structure on D is integrable and there exists a diffeomorphic mapping $G : D \to \mathbb{C}$ such that $\mu_G = \mu$. This means that the composition $H = G \circ H_0 : \mathcal{S} \to \mathbb{C}$ is conformal and $\mathcal{S}$ is conformally equivalent to D^*.

The mapping f_0 in $\mathcal{S}$, induced from $f_{p,\lambda}$, is prolonged to a mapping of D. We denote it by f. Of course, it is analytic.

Because the functions μ, G, H are smooth and f is conjugated by means of H with $f_{p,\lambda}$ at the sectors S_j, then their linear parts are also conjugated. So $f'(0) = 1$ and $f \in \mathcal{A}_l$ for some l.

But the index l is a topological invariant (see Proposition 3.24), it is the number of 'elliptic sectors' (divided by 2). So $l = p$.

Finally, the formal classification invariant λ is included into the functional modulus $\tilde{\Phi}_{2p} = -2\pi i\lambda + t_{2p} + \cdots$.

Note that the maps $H_j = z_j \circ H^{-1}$ form the normalizing maps for f (the normalizing cochain). Its coboundary maps are $z_{j+1} \circ z_j{}^{-1} = \Phi_j$. This means that the collection $\tilde{\Phi}$ is the functional modulus for f.

Theorem 3.28 is complete. □

Now we consider conformal maps of the form

$$f(z) = e^{2\pi i m/n} z + \cdots, \tag{3.13}$$

i.e. $f'(0)$ is a primitive root of unity of order n. We call them $m : n$ *resonant maps*.

PROPOSITION 3.32. *Any germ of this form is formally equivalent to*

$$e^{2\pi i m/n} g_w^1, \quad w = \frac{z^{np+1}}{1 + \lambda z^{np}} \frac{\partial}{\partial z}.$$

PROOF. The Poincaré–Dulac theorem for diffeomorphisms allows to reduce f to the form $e^{2\pi i m/n} z(1 + \sum a_l z^{nl})$. Let p be the first index l with non-zero a_l. We can assume that the coefficient $a_p = 1$.

Next, repeating the proof of Proposition 3.21, we successively cancel the terms $a_l z^{nl+1}$ using the changes $z \to z + b_l z^{n(l-p)}$. It is possible to do it in all cases but $l = 2p$. □

One denotes by $\mathcal{A}_{m,n,p,\lambda}$ the space of germs with the formal normal form having the indicated indices.

Let us pass to analytic classification of the germs from $\mathcal{A}_{m,n,p,\lambda}$. Note firstly that f^n is tangent to identity, it belongs to $\mathcal{A}_{q,\beta}$, $q = np$ with some β (determined by λ).

Consider the orbit space U^*/f^n, consisting of $2np$ Riemann spheres $\mathbb{C}^* \times \{j\}$. The map f transforms the orbits of f^n to orbits of f^n and induces the map $\pi_* f : U^*/f^n \to U^*/f^n$. We describe the action of this map on the Riemann spheres.

Note that the sphere $\mathbb{C}^* \times \{j\}$ is the quotient of the sector S_j with the bisectrix $\arg z = \pi(j-1)/(2pn)$ and the transformation f is approximately the rotation by the angle $2\pi m/n$. Therefore $f(S_j) \approx S_{j+2pm}$ and $\pi_* f : \mathbb{C}^* \times \{j\} \to \mathbb{C}^* \times \{j + 2pm\}$.

Moreover, $\pi_* f$ is a holomorphic diffeomorphism and takes the form $\pi_* f(\vartheta_j) = C_j \vartheta_{j+2pm}$. It means also that $\vartheta_{j+2pm} \circ f = C_j^{-1} \vartheta_j$. One can see that $C_j = \mathrm{e}^{2\pi \mathrm{i} r/n}$ is a root of unity.

Indeed, because $\phi_j = \vartheta_{l+1} \circ \vartheta_l^{-1}$ are close to the identity (or to the multiplication by the constant ν for $l = 2pn$) then $\vartheta_{j+2pm+1} \circ \vartheta_{j+2pm}^{-1} = \vartheta_{j+2pm+1} \circ f \circ f^{-1} \circ \vartheta_{j+2pm}^{-1} = C_{j+1}^{-1} \vartheta_{j+1} \circ \vartheta_j^{-1} \circ C_j = C_{j+1}^{-1} \phi_j \circ C_j = \mathrm{id} + \cdots$. So $C_j = C_{j+1} = \cdots = C$. Because $\pi_* f^n = \mathrm{id}$ we get $C^n = 1$.

From the above it follows that only $2p$ first gluing maps are needed to determine the functional modulus $\mu^*_{f^n}$. Indeed, we have $\phi_{j+2pm} = \mathrm{e}^{-2\pi \mathrm{i} r/n} \phi_j \circ \mathrm{e}^{2\pi \mathrm{i} r/n}$. For the last gluing map we get $\phi_{2pn} = \nu \mathrm{e}^{-2\pi \mathrm{i} r/n} \phi_{2p(n-m)} \circ \mathrm{e}^{2\pi \mathrm{i} r/n}$, where $\nu = \phi'_{2pn}(\infty)$ is the constant expressed by the formal modulus λ.

All this leads to the following result.

THEOREM 3.33. (See [42].) *The space $\mathcal{A}_{m,n,p,\lambda}$ of germs of conformal diffeomorphism modulo analytic equivalence is in one-to-one correspondence with the space $\mathcal{M}_{m,n,p,\lambda}$ of systems*

$$\phi = (\phi_1, \dots, \phi_{2p})$$

(*of germs of analytic diffeomorphisms $\phi_j(\bar{\mathbb{C}}, 0) \to (\bar{\mathbb{C}}, 0)$ $(j$ odd$)$ or $\phi_j(\bar{\mathbb{C}}, \infty) \to (\bar{\mathbb{C}}, \infty)$ $(j$ even$)$ with identity linear parts*) *modulo the equivalence $\phi \sim D\phi \circ D^{-1}$.*

3.4. *Martinet–Ramis moduli*

The saddle–node singularity of analytic planar vector field is such that one of its eigenvalues vanishes and the other eigenvalue is non-zero. Firstly we present the formal classification of saddle–nodes and next we describe the moduli of their analytic classification.

PROPOSITION 3.34. *Any complex analytic saddle–node is formally orbitally equivalent to*

$$\dot{x} = x^{p+1}, \qquad \dot{y} = -y(1 + \lambda x^p).$$

PROOF. The Poincaré–Dulac theorem 3.17 and Example 3.18(b) give the following (non-orbital) normal form $\dot{x} = a_{p+1} x^{p+1} + a_{p+2} x^{p+2} + \cdots$, $\dot{y} = \lambda_2 y(1 + b_1 x + b_2 x^2 + \cdots)$, where $a_{p+1} \neq 0$; if all $a_j = 0$ then the singular point would have infinite multiplicity (non-isolated in the complex analytic case). By linear changes of variables and time we can assume that $a_{p+1} = 1$, $\lambda_2 = -1$.

Now we divide the field by $1 + b_1 x + \cdots$ what gives $\dot{y} = y$ and

$$\dot{x} = x^{p+1}(1 + c_1 x + \cdots). \tag{3.14}$$

Therefore the problem is reduced to formal *non-orbital* classification of 1-dimensional saddle–nodes (3.14).

The latter problem is solved in the same way as in the proof of Theorem 3.23 (about formal classification of diffeomorphisms $z+z^{p+1}+\cdots$). Moreover, this special change can be made analytic. Using the changes $x\to x+d_lx^l$, we obtain the (leading) term $(l-p-1)d_lx^{l+p}\partial_x$ in the transformed vector field. If $l\neq p+1$, then the corresponding monomial vector field can be canceled. There remains only the field $w_{p,\nu}=x^{p+1}(1+\nu x^p)\partial_x$ where ν is the formal invariant.

Of course, one can fix also other terms in ∂_x, like $x^{p+1}(1-\lambda x^p+(\lambda x^p)^2-\cdots)=x^{p+1}(1+\lambda x^p)^{-1}$, $\lambda=-\nu$. Multiplication of this vector field by $1+\lambda x^p$ gives the vector field from thesis of the proposition. □

The integer p is called the *codimension* of saddle–node and $\lambda\in\mathbb{C}$ is the *modulus of formal classification*. The space of germs with the formal normal form as in Proposition 3.34 will be denoted by $\mathcal{E}_{p,\lambda}$.

As we shall see, the normal form from Proposition 3.34 is not analytic. However some of its statements have analytic equivalents. Note that it has two formal invariant curves: the *strong manifold* $x=0$ and the formal *center manifold* $y=0$. By Theorem 3.20 the strong manifold is always analytic but usually the center manifold is not analytic.

EXAMPLE 3.35. (See [33].) The equation $\mathrm{d}y/\mathrm{d}x=(y-x)/x^2$ has the following unique formal solution $y=\sum_1^\infty(n-1)!x^n$ passing through 0.

LEMMA 3.36. *There exist local analytic coordinates and an analytic change of time which give the following initial analytic form*

$$\dot{x}=x^{p+1},\qquad \dot{y}=-y(1+\lambda x^p)+g_0(x)+y^2g_2(x,y),$$

where g_0, y^2g_2 *have zero of order* $\geqslant p+2$ *at the origin.*

PROOF. Since the strong manifold is analytic we can assume that the vector field has the form (in some analytic coordinates) $\dot{x}=xf(x,y)$, $\dot{y}=-y+g(x,y)$. This preliminary form is not satisfactory to our purposes. We want to get the factor x^{p+1} in the first component.

This is done using series of the analytic changes $x\to x+x^j\psi_j(y)$, $j=1,\dots,p$. Let $\dot{x}=x[\phi_1(y)+\mathrm{O}(x^2)]$ where $\phi_1(0)=0$. Then we use the change $x=x_1(1+\psi_1(y))$ such that $-y\psi_1'+\phi_1(y)=0$. Similarly, further terms $x^j\phi_j(y)$, $\phi_j(0)=0$, for $j=2,\dots,p$, are removed from $\dot{x}$. Thus we can put $\dot{x}=x^{p+1}$ and $\dot{y}=g_0(x)+yg_1(x)+y^2g_2(x,y)$.

Moreover, using polynomial changes in y we can cancel some initial terms from g_0 and g_2. The function $g_1(x)$ is reduced to $-1-\lambda x^p$ by a change like in the proof of Proposition 3.34; it turns out to be analytic. □

We define the *extended sectors*

$$T_j=S_j'\times\{|y|<\epsilon\}$$

where $S_j'=\{|x|<\epsilon,\ \frac{(2j-1)\pi}{2p}-\alpha<\arg x<\frac{(2j-1)\pi}{2p}+\alpha\}=\mathrm{e}^{\pi\mathrm{i}/2p}S_j$, $\alpha\in(\frac{\pi}{2p},\frac{\pi}{p})$ are sectors rotated with respect to the sectors S_i from the previous subsection. The following important result belongs to M. Hukuhara, T. Kimura and T. Matuda. Its proof is postponed to Appendix 2.

THEOREM 3.37. (See [40].) *In each extended sector T_j there exists a unique analytic diffeomorphism*

$$H_j(x,y) = \bigl(x, y + \tilde{H}_j(x,y)\bigr), \quad \tilde{H}_j = \mathrm{O}(x^{p+1}),$$

conjugating the system from Lemma 3.36 *with its formal normal form from Proposition* 3.34.

We pass to description of the Martinet–Ramis moduli of analytical orbital equivalence of germs from $\mathcal{E}_{p,\lambda}$.

The reader can guess that these moduli are constructed using the differences between the normalizing diffeomorphisms,

$$\Phi_j = H_{j+1} \circ H_j^{-1} \quad \text{in } T_j \cap T_{j+1}.$$

However, the description of the Φ_j's can be simplified.

Notice that the diffeomorphisms Φ_j keep the first coordinate fixed and preserve the formal normal form $\omega_0 = x^{p+1}\,\mathrm{d}y + y(1+\lambda x^p)\,\mathrm{d}x = 0$ (of the corresponding Pfaff equation). The latter fact gives series of restrictions onto possible forms of Φ_j's.

Let us look at the behavior of the leaves of the foliation defined by $\omega_0 = 0$ in the sectors $T_j \cap T_{j+1}$. These leaves are given by

$$y = C\mathrm{e}^{-t(x)}, \quad t(x) = t_j(x) = -1/(px^p) + \lambda \ln x, \quad C \in \mathbb{C},$$

and the first integral takes the form

$$h(x,y) = h_j(x,y) = y\mathrm{e}^{t(x)} = y\mathrm{e}^{t_j(x)}.$$

(Because of $\ln x$ in the definition of $t(x)$ we index the 'times' and first integrals in different sectors. Thus $t_{2p} = t_1 + 2\pi\mathrm{i}\lambda$, $h_{2p} = \mathrm{e}^{-2\pi\mathrm{i}\lambda}h_1$.)

The neighborhood of $x = y = 0$ is divided into sectors of two types:

- the sectors of *jump* where $\operatorname{Re} x^{-p} \to +\infty$ and $\mathrm{e}^{-t(x)} \to \infty$ (in particular, $\{x > 0\}$ lies in a sector of jump $T_{2p} \cap T_1$),
- the sectors of *fall* where $\operatorname{Re} x^{-p} \to -\infty$ and $\mathrm{e}^{-t(x)} \to 0$.

In the sectors of jump only one leaf approaches the singularity, it is the *local center manifold*. Other leaves diverge as $x \to 0$.

In the sectors of fall all the leaves are indistinguishable. All approach the singularity. We see also that the extended sectors $T_{j+1} \cap T_j$ lie wholly either in the sectors of jump or in the sectors of fall.

PROPOSITION 3.38. *The transition diffeomorphisms Φ_j take the forms*

$$\Phi_j(x,y) = \bigl(x, \phi_j(h_j)\mathrm{e}^{-t_{j+1}}\bigr),$$

where

$$\phi_j(h) = a_j + h, \quad a_j \in \mathbb{C},$$

if $T_{j+1} \cap T_j$ *lies in a sector of fall,*

$$\phi_j(h) = h + \alpha_{j,2}h^2 + \cdots,$$

if $T_{j+1} \cap T_j$ *lies in a sector of jump and* $j \neq 2p$, *and*

$$\phi_{2p}(h) = \mathrm{e}^{-2\pi \mathrm{i}\lambda}h + \alpha_{2p,2}h^2 + \cdots.$$

Moreover Φ_j *tend to* id *exponentially fast as* $x \to 0$.

PROOF. The first statement follows from the fact that Φ_j preserve the formal normal form. So if h_{j+1} is the first integral in T_{j+1} then the function $h_{j+1} \circ \Phi_j$ is the first integral in $T_j \cap T_{j+1}$. But also h_j is a first integral. Thus the first integral is a function of the second first integral, $h_{j+1} \circ \Phi_j = \phi_j \circ h_j$.

In the coordinates (x, h) the map Φ_j is equal to $(x, \phi_j(h))$. Passing to the (x, y) coordinates we get the formula for Φ_j.

The proof of the further statements relies on the fact that $\Phi_j - \mathrm{id} = \mathrm{O}(x^{p+1})$.

Let $T_j \cap T_{j+1}$ be a sector of fall. We have $t_j = t_{j+1}$ (because $j \neq 2p$) and $\mathrm{Re}\, t(x) \to \infty$. Let the Taylor expansion of ϕ_j be $\sum_{n=0}^{\infty} \alpha_{j,n}h^n$. Then

$$\Phi_j = \Bigl(x, \alpha_{j,0}\mathrm{e}^{-t(x)} + \alpha_{j,1}y + \sum \alpha_{j,n}y^n \mathrm{e}^{(n-1)t(x)}\Bigr).$$

The condition $\Phi_j \to id$ implies that $\alpha_{j,1} = 1$ and $\alpha_{j,2} = \alpha_{j,3} = \cdots = 0$. This means that ϕ_j is a translation. Note also that in this case $\Phi_j - \mathrm{id} = (0, \mathrm{O}(\mathrm{e}^{-t(x)}))$ is exponentially small.

Let $T_j \cap T_{j+1}$ be a sector of jump and $j \neq 2p$. Then, acting exactly in the same manner as in the previous case, we get $\alpha_{j,0} = 0$ and $\alpha_{j,1} = 1$. It means that ϕ_j is a germ of diffeomorphism of $(\mathbb{C}, 0)$ with $\phi'(0) = 1$. In this case $\Phi_j - \mathrm{id}$ is exponentially small too.

Let $j = 2p$. We are in the sector of jump with $\mathrm{e}^{t_{2p}} = \mathrm{e}^{2\pi \mathrm{i}\lambda}\mathrm{e}^{t_1}$. We have $\Phi_{2p} = (x, \phi_{2p}(h_{2p})\mathrm{e}^{-t_1}) = (x, \mathrm{e}^{2\pi \mathrm{i}\lambda}\phi_{2p}(h_{2p})\mathrm{e}^{-t_{2p}})$. Analogously as in the previous case we get the formula for ϕ_j and the exponential closeness to the identity. □

The fact that the diffeomorphisms Φ_j have different forms can be explained geometrically. In the sector of jump the values of the first integral are bounded, $h \in (\mathbb{C}, 0)$. In the sectors of fall the first integral takes values in the Riemann sphere $\bar{\mathbb{C}}$ and $h = \infty$ along the strong manifold $x = 0$; the corresponding map ϕ_j is an automorphism of this sphere satisfying $\phi_j(\infty) = \infty$.

The collection of ϕ_j's satisfying the properties of Proposition 3.38 do not yet form a complete set of invariants. The reader can notice that the coordinates in the normalized Pfaff equation $\omega_0 = 0$ are not unique. Its non-uniqueness lies in the possibility to multiplication of y by a constant. This gives the non-uniqueness of the choice of the local first integral: $h_j \to Ch_j$, where C is the same in each sector.

DEFINITION 3.39. Two collections $\phi = (\phi_1, \dots, \phi_{2p})$ and $\tilde{\phi} = (\tilde{\phi}_1, \dots, \tilde{\phi}_{2p})$, satisfying the conditions from Proposition 3.38, are called *equivalent* iff they are conjugated by a linear transformation: $\phi_j \circ C = C\tilde{\phi}_j$.

The equivalence class $[\phi]$ of such a collection is called the *Martinet–Ramis modulus*. The space of Martinet–Ramis moduli is denoted by $\mathcal{N}_{p,\lambda}$.

The above construction of the Martinet–Ramis modulus, associated to saddle–node, defines a map from $\mathcal{E}_{p,\lambda} \to \mathcal{N}_{p,\lambda}$

$$V \to \epsilon_V.$$

The following result is known as the *Martinet–Ramis theorem*.

THEOREM 3.40. (See [54].) *The map $V \to \epsilon_V$ has the following properties*:

- *it sends orbitally analytically equivalent germs of vector fields to the same point* (*invariance*);
- *if $\epsilon_V = \epsilon_{V'}$ then V and V' are orbitally analytically equivalent* (*equimodality*);
- *for any $\epsilon \in \mathcal{N}_{p,\lambda}$ there exists $V \in \mathcal{E}_{p,\lambda}$ such that $\epsilon = \epsilon_V$* (*realization*).

PROOF. The first two points are proved in the same way as the analogous points of the Ecalle–Voronin theorem 3.28. Only the realization property needs separate arguments.

Assume that we have an element from $\mathcal{N}_{p,\lambda}$ which is an equivalence class with some representative $\phi = (\phi_j)$. These ϕ_j define the diffeomorphisms Φ_j in the extended sectors $T_j \cap T_{j+1}$ (see Proposition 3.38).

Take disjoint union of the extended sectors $\bigsqcup T_j \times \{j\}$ and glue them using the Φ_j as the gluing maps. We obtain certain (real) 4-dimensional manifold $\mathcal{S}$. In each $T_j \times \{j\}$ we have the ‘coordinates’ $z_j = (x, y_j)$ and $z_{j+1} = \Phi_j(z_j)$ at the glued part.

Note also that $\mathcal{S}$ is equipped with a foliation $\mathcal{F}_0$ such that in each T_j it is the foliation of the vector field in the formal normal form. We have to define certain complex structure on $\mathcal{S}$ such that the foliation $\mathcal{F}_0$ arises from a holomorphic vector field with saddle–node singularity.

Firstly one defines the map $H_0 : \mathcal{S} \to W \subset \mathbb{C}^2 \setminus \{x = 0\}$ as

$$H_0 = \sum \chi_j z_j$$

where $\{\chi_j\}$ is a partition of unity of $\mathcal{S}$ associated with the covering (T_j) with regular properties as $x \to 0$.

H_0 defines an almost complex structure on W, the image of the complex structures on T_j. By definition this almost complex structure is integrable (see Definition 3.67 in Appendix 1). By Theorem of Newlander–Nirenberg 3.70 (in Appendix 1) the torsion of this almost complex structure vanishes identically.

It turns out that the almost complex structure from W can be prolonged smoothly to an almost complex structure on a full neighborhood of 0 in $\mathbb{C}^2$. Indeed, the almost complex structure is defined by $(1,0)$-forms $\mathrm{d}x$ and $\mathrm{d}y$, where $\mathrm{d}x$ is good (with zero derivative). Since $y = \sum \chi_j y_j$. In $T_j \cap T_{j+1}$, we have $y = \chi_j y_j + \chi_{j+1} y_{j+1} = y_j + h_j(x, y_j)\chi_j$

(see Theorem 3.37). Thus $\mathrm{d}y = \mathrm{d}y_j + \mathrm{o}(1)$, where the term $\mathrm{o}(1)$ is exponentially small as $x \to 0$.

The torsion of the latter almost complex structure is also equal to zero. Applying again the Newlander–Nirenberg theorem, we show that this almost complex structure is integrable. It means that it is equal to the transportation of the standard complex structure by means of some differentiable map $G : (\mathbb{C}^2, 0) \to (\mathbb{C}^2, 0)$

The composition $H = G^{-1} \circ H_0 : \mathcal{S} \to \mathbb{C}^2$ is holomorphic on $\mathcal{S}$ and prolongs to the closure of $\mathcal{S}$. It transforms the foliation $\mathcal{F}_0$ to a foliation $\mathcal{F}$ in $(\mathbb{C}^2, 0)$.

The foliation $\mathcal{F}$ is a saddle–node foliation with the Martinet–Ramis invariant equal to $[\phi]$. □

The Martinet–Ramis theorem allows to solve the problem of analyticity of the center manifold and to explain the reason of non-analyticity of the Euler's solution. Note that in each sector of jump the local center manifold is defined uniquely, using topological criteria. It is the only leaf of the holomorphic foliation containing the singular point in its closure. This local center manifold prolongs itself (uniquely) to the adjacent sectors of fall. The gluing map Φ_j in the sectors of jump leave this local center manifold invariant, $\phi_j(0) = 0$.

The problem is whether the local center manifolds from different sectors of jump lie in one leaf of the foliation. It means, whether the gluing maps in the sectors of fall preserve the two prolongations of the local center manifolds (from the adjacent sectors of jump). This condition reads as $\phi_j(0) = 0$, where $\phi_j(h) = h + a_j$.

We have proved the following result.

COROLLARY 3.41. *A saddle–node singularity has analytic center manifold iff all* $a_j = 0$ *in its Martinet–Ramis modulus. In other words this modulus has the form* $(\mathrm{id}, \phi_2, \mathrm{id}, \phi_4, \ldots, \phi_{2p})$.

3.5. *Holonomy maps and normal forms for resonant saddles*

Here we finish the description of analytic normal forms for germs of planar vector fields with elementary singular points.

The singular points from the *Poincaré domain* have analytic Poincaré–Dulac normal form. This means that the normalizing transformation can be chosen analytic (see Theorem 3.19).

Therefore, it remains to investigate the saddles, i.e. singularities with negative ratio of eigenvalues $-\lambda < 0$. The saddles are of two kinds: *resonant* (i.e. with rational $\lambda = m/n$) and *non-resonant*.

The description of moduli of resonant saddles relies on the classification of germs of resonant 1-dimensional diffeomorphisms and was completed firstly by J. Martinet and J.-P. Ramis [55].

The problem of analytic conjugation of non-resonant saddles is analogous to the problem of analytic linearization of germs of non-resonant 1-dimensional diffeomorphisms and was solved by J.-C. Yoccoz [87] (see Subsection 3.7).

In classification of saddle singularities the following notion is useful.

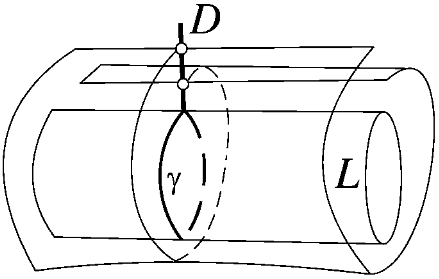


Fig. 7.

DEFINITION 3.42. Let $\mathcal{F}$ be a holomorphic foliation in a 2-dimensional complex manifold, L be a non-singular leaf of $\mathcal{F}$ and $\gamma \subset L$ be a closed loop with beginning at a point x_0. Take a small holomorphic disc D transversal to L at x_0 and parameterized by $z \in (\mathbb{C}, 0)$, $z(x_0) = 0$. We can cover the loop γ by open sets U_i, where the foliation is the flow-box foliation: $\mathcal{F}|_{U_i} = \{dy = 0\}$.

Let $z \in D$. Take a leaf $L(z)$ passing through z. In each U_i we can lift the path $\gamma \cap U_i$ to the leaf $L(z) \cap U_i$. After returning of γ to the point x_0, the lifted curve intersects the disc D again. However the intersection point can differ from z (see Fig. 7). We denote it by $\Delta_\gamma(z)$. We define the *monodromy transformation* (or the *holonomy transformation)* associated with the loop γ as the map

$$z \to \Delta_\gamma(z).$$

The holonomy map Δ_γ is a local holomorphic diffeomorphism (by analytic dependence of solutions on initial conditions). It does not depend on the homotopy class of the loop γ (for fixed D); this is a consequence of the theorem about monodromy. The change of the section D and of the initial point x_0 results in internal conjugations of Δ_γ in the group $\mathrm{Diff}(\mathbb{C}, 0)$ of germs of holomorphic diffeomorphisms of $(\mathbb{C}, 0)$.

The connection between *saddles* and one-dimensional maps lies in the holonomy transformation associated with any of its separatrices. (Recall that, by the analytic Hadamard–Perron theorem, Theorem 3.20, the saddle has both separatrices analytic.) We can assume that

$$\dot{x} = \lambda x(1 + \cdots), \qquad \dot{y} = -y(1 + \cdots).$$

Consider the monodromy transformation $\Delta(x)$ corresponding to a loop in the separatrix $x = 0$. One can easily check that

$$\Delta(x) = \mathrm{e}^{-2\pi \mathrm{i}\lambda} x + \cdots.$$

The following result was firstly proved by J. Mattei and R. Moussu.

PROPOSITION 3.43. (See [56].) *Two germs with saddle singularity with the same ratio* $-\lambda$ *are analytically orbitally equivalent iff the corresponding germs of monodromy maps are analytically equivalent.*

PROOF. If two germs of vector fields are orbitally equivalent then restriction of the conjugating map H to a disc D transversal to $\{x=0\}$ defines a conjugation of monodromies (in D and in $H(D)$).

Let the monodromies be conjugated. Consider two copies of $(\mathbb{C}^2,0)$ with foliations $\mathcal{F}$ and $\mathcal{F}'$ (by the phase curves L, L' of the two vector fields V and V'). In each copy we consider the fibration $\pi:(x,y)\to y$. Their fibers outside $y=0$ are transversal to the leaves of the foliations $\mathcal{F}$ and $\mathcal{F}'$. We denote by $\mathcal{F}_0$ the foliation by the fibers $y=\text{const}$.

Take two discs $D=\{y=\epsilon,\,|x|<\epsilon\}$ and $D'=\{y=\epsilon,\,|x|<2\epsilon\}$. Let $h:D\to D'$ be the map conjugating the monodromies.

We extend the map h to a map $H_0\colon A\to A'$, where $A=\{|x|<\epsilon,\,\epsilon/2<|y|<\epsilon\}$, $A'=\{|x|<2\epsilon,\,\epsilon/2<|y|<\epsilon\}$ are ring-like domains. H_0 preserves the foliation $\mathcal{F}_0$, i.e. $H_0(x,y)=(F(x,y),y)$ and sends leaves L of the foliation $\mathcal{F}$ to leaves L' of the foliation $\mathcal{F}'$. Because of the transversality of the fibers $y=\text{const}$ to the leaves L,L' in the domains A and A' the prolongation H_0 is locally analytic and unique.

H_0 is single-valued in the whole domain A. Indeed, the change of H_0, as the argument turns around the axis $y=0$, is equal to the composition of the monodromy map (defined by $\mathcal{F}$) and of H_0. Because $h=H_0|_D$ conjugates the monodromies the map H_0 is univalent, i.e. $H_0(x,y)$ is meromorphic in y.

Some additional arguments allow to eliminate the possibility that H_0 would have a pole along $y=0$. We refer the reader to [56,54] and [75]. □

Consider firstly the case when $\lambda=-m/n$, $\gcd(m,n)=1$, is a rational number. We call such singular point a *$m:n$ resonant saddle*. Then the monodromy map $\Delta=\mathrm{e}^{2\pi\mathrm{i}m/n}x+\cdots$ is a resonant one-dimensional map considered in Subsection 3.3. Its formal normal form is of one of the following types $\mathrm{e}^{2\pi\mathrm{i}m/n}g_w^1$, $w=[z^{np+1}/(1+\alpha z^{np})]\frac{\partial}{\partial z}$ (see Proposition 3.32). The space of such maps was denoted by $\mathcal{A}_{m,n,p,\alpha}$.

The above form for diffeomorphism corresponds to the following formal orbital normal form for vector fields

$$\dot{x}=x,\quad \dot{y}=y\bigl(-m/n+u^p(1+\lambda u^p)^{-1}\bigr),\quad u=x^my^n, \tag{3.15}$$

where $\lambda=2\pi\mathrm{i}n\alpha$. (The proof relies on passing to the variables x,u in the Poincaré–Dulac normal form and further applying the reduction as in the case of saddle–node.) We denote the class of analytic saddles with the formal orbital normal form (3.15) by $\mathcal{B}_{m,n,p,\lambda}$.

THEOREM 3.44. (See [55].) *The space of classes of analytical orbital equivalence of vector fields from the class* $\mathcal{B}_{m,n,p,\lambda}$ *is the same as the space of classes of analytical equivalence of diffeomorphisms from the class* $\mathcal{A}_{m,n,p,\alpha}$ *described in Theorem* 3.33.

REMARKS ABOUT THE PROOF. In view of Proposition 3.41 the proof of this theorem is reduced to demonstration of the surjectivity of the map $V\to\Delta$ from $\mathcal{B}_{m,n,p,\lambda}$ to $\mathcal{A}_{m,n,p,\alpha}$.

One has to show that any resonant map arises from some resonant saddle as its holonomy. The construction of V is analogous to the construction of the realization parts of the Ecalle–Voronin and Martinet–Ramis classification theorems. We refer the reader to [42]. □

Also with a saddle–node singularity some holonomy transformation can be associated and its relation with the Martinet–Ramis moduli can be studied.

Indeed, the strong invariant manifold of a saddle–node singularity V consists of two leaves of the holomorphic foliation defined by V: the singular point and the leaf L diffeomorphic to a punctured disc. Take a loop $\gamma \subset L$ generating its fundamental group. Using Definition 3.42, we associate with γ the monodromy transformation $\Delta = \Delta_\gamma$.

Recall that $\mathcal{A}_{p,\alpha}$ is the space of germs from $\mathrm{Diff}(\mathbb{C}, 0)$ formally equivalent to the time 1 flow map generated by the vector field $w = [z^{p+1}/(1 + \alpha z^p)]\partial_z$ (see Proposition 3.23).

THEOREM 3.45.

(a) *If* $V \in \mathcal{E}_{p,\lambda}$ *then* $\Delta \in \mathcal{A}_{p,\alpha}$, $\alpha = 2\pi \mathrm{i}\lambda$.

(b) *The Martinet–Ramis modulus of* V *coincides with the Ecalle–Voronin modulus of* Δ. *It means that two vector fields are orbitally analytically equivalent iff their monodromy maps are analytically conjugated. In particular, we have an embedding of the set of equivalence classes of vector fields* $\mathcal{E}_{p,\lambda}/(anal)$ *into the set of equivalence classes of diffeomorphisms* $\mathcal{A}_{p,\alpha}/(anal)$ (*i.e. embedding of the moduli space* $\mathcal{N}_{p,\lambda}$ *into the moduli space* $\mathcal{M}^*_{p,\alpha}$).

PROOF. The first statement can be checked for the vector field in its formal normal form. Then we have $L = \{x = 0 \neq y\}$, $\gamma = \{(0, \mathrm{e}^{\mathrm{i}t})\}$ and the foliation is given by $\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\mathrm{d}x}{\mathrm{d}\ln y} = x^{p+1}/(1 + \lambda x^p) = w_{p,\lambda}$. The monodromy map Δ is given by the time $2\pi\mathrm{i}$ flow map $\Delta = \exp[2\pi \mathrm{i} w_{p,\lambda}] = \exp[w_{p,\alpha}]$.

To prove the second statement let us keep Δ formally equivalent to $g_w^{2\pi\mathrm{i}}$. Then the sectors S_j (from Theorem 3.29) are turned by the angle $\pi/2p$; we denote them by S'_j. The extended sectors T_j are of the form $S'_j \times \{|y| < \epsilon\}$. We fix also the disc $D = \{y = y_0\}$ transversal to the axis $\{x = 0\}$.

The extended sectors contain complete lifts (to the leaves of the foliation) of the loop in the strong manifold appearing in the definition of the holonomy map.

The normalization diffeomorphisms $H_j(x, y)$ in T_j (from Theorem 3.37) define altogether the normalization diffeomorphisms for the monodromy map Δ in $S'_j \times \{y_0\} \subset D$. They define the local parametrization of S_j by means of the first integral $h_j(x) = h \circ H_j|_{S'_j \times \{y_0\}}$. In the charts h_j the map Δ is equal to the formal normal form.

The transition diffeomorphisms $H_{j+1} \circ H_j^{-1}$ are expressed by means of the one-dimensional transition maps $h_{j+1} \circ h_j^{-1}$. The latter define the Martinet–Ramis moduli as well as the Ecalle–Voronin moduli. □

REMARK 3.46. The image of the embedding from the point (b) of the last theorem forms a proper subset of the set $\mathcal{M}^*_{p,\lambda}$. It is because the diffeomorphisms ϕ_j are subject to the restrictions from Proposition 3.36, $\phi_j = h + a_j : (\bar{\mathbb{C}}, \infty) \to (\bar{\mathbb{C}}, \infty)$ (j odd).

It turns out that the correspondence $V \to \Delta$ is not one-to-one even at the level of topological equivalences; i.e. the sets $\mathcal{E}_{p,\lambda}/top$ and $\mathcal{A}_{p,\lambda}/top$ are really different. (At the level of formal equivalences it is one-to-one.)

As the first example we consider the case with the Martinet–Ramis modulus equal to $(\mathrm{id}+a_1, \mathrm{id}, \mathrm{id}+a_2, \mathrm{id}, \dots, \mathrm{id})$ with some $a_j \neq 0$. The monodromy map Δ is topologically equivalent to its formal normal form (see Proposition 3.22) but the vector field is not equivalent to its formal normal form because it does not have analytic center manifold.

Another example concerns the case with $\lambda = 0$ and with all $a_j = 0$, i.e. with analytic center manifold. We consider the monodromy map Δ_c associated with the leaf L_c containing the center manifold. The topology of the foliation determines the topological type of Δ_c, which is determined by the order of tangency of Δ_c to identity (see Proposition 3.24). So the class $[\Delta]_{top}$ is fixed, but there can be infinitely many classes $[\Delta_c]_{top}$.

3.6. *Geometric realization of Martinet–Ramis moduli*

Unlike the Ecalle–Voronin moduli the Martinet–Ramis moduli can be sometimes calculated explicitly. In this subsection we present a result of P. Elizarov about dependence of the Martinet–Ramis moduli on the vector field itself and we give examples (due to the author and L. Teyssier) of explicit calculation of these moduli.

We begin with definitions of the Borel transform and Gevrey series.

DEFINITION 3.47. Let $f(t) = \sum_{k=0}^{\infty} a_k t^k$ be a formal power series. Its *Borel transform* is the series

$$(\mathcal{B}f)(\tau) = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} a_k \tau^k.$$

By analog to the above standard Borel transform we introduce the following modified Borel transforms

$$(\mathcal{B}^l_{SN} f)(\tau) = \sum_{k=1}^{\infty} \frac{k}{\Gamma(1 + (k + \lambda l)/p)} a_k \tau^k$$

and

$$(\mathcal{B}^l_S f)(\tau) = \sum_{k=1}^{\infty} \frac{k}{\Gamma(1 + (km - \lambda l)/(pm^2))} a_k \tau^k.$$

DEFINITION 3.48. Let the series $f(t)$ as above is an asymptotic expansion of a function, denoted also by f, which is holomorphic in a sector S with vertex at $t = 0$. We say that f is of *Gevrey type of order* s if

$$|a_k| \leqslant C(k!)^{s-1} A^k$$

and

$$\left| f(t) - \sum_{j=1j}^{k-1} t^j \right| \cdot |t|^k \leqslant C(k!)^{s-1} A^k,$$

where the constants A, C depend on f and the sector S. The series f of Gevrey type of order

$$s = 1 + 1/p$$

is called *p-summable in the direction θ* if it is an asymptotic expansion of some function f_S holomorphic in a sector S with bisectrix $\arg t = \theta$ and with angle $> \pi/p$.

REMARK 3.49. If a series is of Gevrey type of order $s = 1 + 1/p$ then its Borel transforms $\mathcal{B}^l_{SN}$ or $\mathcal{B}^l_S$ define convergent series.

The transform $\mathcal{B}^l_{SN}$ is used in study of the saddle–node singularities from the class $\mathcal{E}_{p,\lambda}$ and the transform $\mathcal{B}^l_S$ is used in the study of the saddle singularities from the class $\mathcal{B}_{m,n,p,\lambda}$. The sectorial normalizing diffeomorphisms, which are used in the proofs of the Ecalle–Voronin and Martinet–Ramis theorems, have asymptotic expansion in the form of Gevrey series of order $s = 1 + 1/p$; moreover, they turn out p-summable in sectors S_j and T_j respectively (appearing in the proofs). The Gevrey series were used in the works of J. Ecalle [28,29] and J. Martinet and J.-P. Ramis [54,55], see also [53].

Recall that the Martinet–Ramis modulus of a saddle–node singularity from the class $\mathcal{E}_{p,\lambda}$ is a class $[\phi]$ of collections $(\phi_1, \dots, \phi_{2p})$ of germs of holomorphic maps such that $\phi_1(h) = a_1 + h$, $\phi_2(h) = h + a_{2,2}h^2 + a_{2,3}h^3 + \cdots$, $\phi_3(h) = a_3 + h, \dots \phi_{2p-1}(h) = a_{p-1} + h$, $\phi_{2p}(h) = \mathrm{e}^{-2\pi \mathrm{i}\lambda} h + a_{2p}h^2 + \cdots$, and the equivalence relation is defined by means of simultaneous mutual conjugation of the maps ϕ_j by means of a rescaling $h \to Ch$.

Let us pass to the logarithmic chart t, such that $h = \mathrm{e}^{2\pi \mathrm{i} t}$. Then the maps ϕ_j become maps $\Phi_j(t)$ expanded into Fourier series:

$$\Phi_1(t) = t + \sum_{l=-\infty}^{-1} b_{1,l} \mathrm{e}^{2\pi \mathrm{i} l t}, \qquad \Phi_2(t) = t + \sum_{l=1}^{\infty} b_{2,l} \mathrm{e}^{2\pi \mathrm{i} l t},$$

$$\vdots$$

$$\Phi_{2p}(t) = t - 2\pi \mathrm{i}\lambda + \sum_{l=1}^{\infty} b_{2p,l} \mathrm{e}^{2\pi \mathrm{i} l t},$$

where $b_{j,l} = (-a_i)^{-l}/(-l)$, $l = -1, -2, \dots$, for odd j are defined by $b_{j,-1} = -a_1$. In other words, the Martinet–Ramis moduli can be defined by the collection of the coefficients $b_{j,l}$, where $p \in \{1, \dots, 2p\}$ and $l \in \mathbb{N} \cup \{-1\}$.

P. Elizarov considered vector fields of the form

$$V_{p,\lambda} + \varepsilon \cdot \left(y \sum_{l=-1}^{\infty} y^l f_l(x) \right) \partial_y,$$

where $V_{p,\lambda}$ is the formal normal form vector field, f_l are holomorphic germs and ε is a small parameter. By Lemma 3.36 we can assume that $f_0(x) \equiv 0$ and that the other germs f_j have high order zero at $x = 0$. The following result an be regarded as calculation of the tangent to the moduli map.

THEOREM 3.50. (See [30].) *We have*

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} b_{j,l}|_{\varepsilon=0} = \bigl(\operatorname{sign}(l) \mathrm{e}^{-2\pi j\lambda l i/p} l^{\lambda l/p-1}/p\bigr) \cdot (\mathcal{B}^l_{SN} f_l)\bigl(\mathrm{e}^{-(2j+1)\pi \mathrm{i}/p} (l/p)^{1/p}\bigr).$$

EXAMPLE 3.51. As an application of this result we consider the following germ

$$V = V_{p,\lambda} + x^k y^{q+1} \partial_y,$$

where $k > p$ and $q \in \mathbb{N} \cup \{-1\}$ are such that $(k + q\lambda)/p$ is not a negative integer. Then the series normalizing it to the formal orbital normal form is divergent. Moreover, two such germs with different exponents q are not analytically equivalent.

IDEA OF THE PROOF. We do not prove the Elizarov theorem, but we can say few words about the proof. By Theorem 3.45 the Martinet–Ramis modulus for the vector field $V \in \mathcal{E}_{p,\lambda}$ coincides with the Ecalle–Voronin of the monodromy diffeomorphism $\Delta \in \mathcal{A}_{p,\alpha}$. The holonomy map Δ can be calculated modulo $\mathrm{O}(\varepsilon^2)$ by integration of suitable equation in variations. We get $\Delta = \Delta_0 + \varepsilon R$ where Δ_0 is the time 1 phase flow map g^1_w and R is expressed via the series $y \sum y^l f_l(x)$. Recall that the conjugating map from the sectorial normalization theorem (Theorem 3.29) in the t_j-chart was defined by the formula $\tilde{H}_j(t_j) = t_j + \varepsilon \sum_m \tilde{R} \circ \tilde{\Delta}^m$, where $\tilde{\Delta}(t_j) = t_j + 1 + \mathrm{O}(\varepsilon)$. So, it is enough to compute the sums $\sum_m \tilde{R}(t_j + m)$ or rather differences between such sums for pairs of consecutive j's. The result is the above Elizarov's formula. □

Elizarov calculated also the tangent to the moduli map in the case of resonant saddle. Recall that here the modulus associated to a germ of vector field V from the class $\mathcal{B}_{m,n,p,\lambda}$ is the Ecalle–Voronin modulus of the holonomy map Δ associated with a loop in one of the separatrices. This modulus, rewritten in the logarithmic chart $t = \log \vartheta$ is an equivalence class of collection of $2p$ functions

$$\Phi_j(t) = t - \pi \mathrm{i}\lambda/np + \sum_{l=1}^{\infty} b_{j,l} \mathrm{e}^{2\pi \mathrm{i} l t}, \quad j \text{ odd},$$

$$\Phi_j(t) = t - \pi \mathrm{i}\lambda/np + \sum_{l=-\infty}^{-1} b_{j,l} \mathrm{e}^{2\pi \mathrm{i} l t}, \quad j \text{ even}.$$

Elizarov considers following germs

$$V = V_0 - \varepsilon x \frac{u^p}{1+\lambda u^p} P(x,y)\partial_x,$$

where V_0 is in the formal orbital normal form, defined in the previous subsection, $u = x^m y^n$ and $P = \sum c_{k,s} x^k y^s = \sum_l f_l(\tilde{x}) z^l$ and $\tilde{x} = xz^n$, $y = z^m$ and $f_l(\tilde{x}) = \sum_{ms-nk=l} c_{k,s}\tilde{x}^k$. The next theorem is proved in the same way as Theorem 3.50.

THEOREM 3.52. (See [30].) *We have*

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} b_{j,l}|_{\varepsilon=0} = \alpha_{j,l} \cdot (\mathcal{B}_S^l f_l)\big(\mathrm{e}^{-(2j+1)\pi\mathrm{i}/(pm)}(l/pm^2)^{1/pm}\big)$$

where $\alpha_{j,l} \neq 0$ *are some concrete constants.*

We promised to calculate explicitly the Martinet–Ramis moduli in some situations. The following example is a continuation of Example 2.28 from Subsection 2.2.

EXAMPLE 3.53. Consider of the system

$$\dot{x} = x^{p+1}, \qquad \dot{y} = -y(p + ax^p) - x^{p+1}\varphi(x),$$

where $\varphi(x)$ is a germ of analytic function and $a \in \mathbb{C}$. Of course, it is a saddle–node singularity with the first integral

$$H = yx^a \mathrm{e}^{-1/x^p} + H_1(x)$$

where $H_1(x) = \int_0^x s^a \mathrm{e}^{-1/s^p}\varphi(s)\,\mathrm{d}s = \int f$ is the function defined in Example 2.28. Recall that the function $H_1(x)$ is well defined only in some sectors which contain one sector of fall of the function f, where $\operatorname{Re} 1/x^p > 0$ (and is a sector of jump for the foliation), and do not contain whole sectors of jump of f, where $\operatorname{Re} 1/x^p < 0$. These sectors are equal $S_1' \cap S_{2p}'$, where $H = x^a \mathrm{e}^{-1/x^p}(y + \chi_1(x))$, $S_3' \cup S_4'$, where $H = x^a \mathrm{e}^{-1/x^p}(y + \chi_3(x)) + A_0$, etc. The functions $\chi_j(x)$ are analytic in the sectors and the constants A_j are given by integration of $x^a \mathrm{e}^{-1/x^p}\varphi(x)\,\mathrm{d}x$ along so-called *asymptotic cycles* σ_j which begin and end at the origin and surround exactly one sector of jump (like at Fig. 1 in Section 2).

It follows that the Martinet–Ramis modulus is the following

$$\phi = (\mathrm{id} + A_0, \mathrm{id}, \mathrm{id} + A_1, \mathrm{id}, \ldots, \mathrm{id} + A_{p-1}, \mathrm{e}^{-2\pi\mathrm{i}\lambda}\,\mathrm{id}).$$

Consider saddle–node singularities with *analytic center manifold*. By Lemma 3.36 we can write them in the following analytic form

$$\begin{aligned} V &= V_0 + y^2 g_2(x,y)\partial_y = V_0 + RY, \\ V_0 &= x^{p+1}\partial_x - y(1 + \lambda x^p)\partial_y, \end{aligned} \tag{3.16}$$

where $R = yg_2(x, y)$ and $Y = y\partial_y$. So the center manifold is $x = 0$. Note that the vector fields V_0 and Y commute, $[V_0, Y] = 0$. L. Teyssier [80] noticed that one can try to conjugate V with V_0 by means of a local diffeomorphism of the form

$$G = g_Y^{F(x,y)}$$

where $\{g_T^t\}$ is the flow generated by the vector field Y: $G(x, y) = (x, \mathrm{e}^{F(x,y)}y)$. Rather elementary calculations, which use the equivalence of the equation $G^*U = W$ to $U \circ G = \frac{\partial G}{\partial W}$, give the following formula.

LEMMA 3.54. *The condition $G^*V_0 = V$ is equivalent to the equation*

$$\frac{\partial F}{\partial V} = -R.$$

A natural way to solve the equation $V(F) = -R$ is to rewrite it as $\dot{f}(t) = -r(t)$, when restricted to a leaf L of the foliation $\mathcal{F}$ defined by the vector field V. On the leaf L the 1-form $\mathrm{d}t = \tau = \mathrm{d}x/x^{p+1}$ is the 'time' form and we can express the solution in the form

$$F(x, y) = \int_{\gamma(x,y)} (-R)\tau$$

where $\gamma(x, y)$ is a path in the leaf $L(x, y)$ which connects the origin $x = y = 0$ with the point (x, y). There arises the problem of correctness of this definition.

Let us numerate the leaves of the foliation $\mathcal{F}$ by a value h of the first integral $yx^{-\lambda}\mathrm{e}^{1/px^p}$ (of V_0) restricted to a transverse line $x = x_*$, $L = L_h$. Since the vector field V has analytic center manifold, the function $(y|_{L_h})(x)$ on a fixed leaf L_h should tend to zero as $x \to 0$ within a sector of fall. Let $\gamma_j^\infty(h)$ be the lifts to L_h of the loops σ_j in the x-plane which begin and end at the origin in adjacent sectors of fall and surround exactly one sector of jump (like in Fig. 1 in Section 2); they are called the *asymptotic cycles* (see [79,80]). The quantities

$$\chi_j(h) = \int_{\gamma_j^\infty(h)} (-R)\tau$$

are the obstacles to the single validity of the function F.

But in suitable sectors which contain just one sector of fall we can define corresponding functions $F_j(x, y) = \int_{\gamma_j(x,y)} (-R)\tau$, where $\gamma_j(x, y)$ are paths in $L(x, y)$ which begin at the origin in the j-th sector of fall and end at the point (x, y). We have $F_{j+1} - F_j = \chi_j$. These functions define corresponding conjugating sectorial diffeomorphisms $H_j = g_Y^{F_j}$ and the functional cochain $\Phi_j = H_{j+1} \circ H_j^{-1} = (x, y\mathrm{e}^{\chi_j(h)}) = (x, h\mathrm{e}^{\chi_j(h)}\mathrm{e}^{-t(x)})$.

We can generalize the above to the case of resonant saddle, i.e. for $V = V_0 + RY$, where $R = yg(x, y)$ and $Y = y\partial_y$ are as above and

$$V_0 = -\frac{m}{n}x\big(1 + (\lambda - 1)u^p\big)\partial_x + y(1 + \lambda u^p)\partial_u, \quad u = x^n y^m. \tag{3.17}$$

Here also one tries changes of the form $G = g_Y^{F(x,y)}$ and arrives to the functions $\chi_j(h)$ defined in the same way.

Let us summarize this in the following theorem of L. Teyssier. (In his work [80] he uses a version of the Martinet–Ramis moduli when the germs ϕ_j are such that $\phi'_{2j}(0) = \mathrm{e}^{2\pi\mathrm{i}\lambda/p}$, but it is a subject to different choices of first integrals in the extended sectors T_j.)

THEOREM 3.55. (See [80].) *For a saddle–node singularity with analytic center manifold the Martinet–Ramis modulus equals*

$$(\mathrm{id}, h\mathrm{e}^{\chi_1(h)}, \mathrm{id}, h\mathrm{e}^{\chi_2(h)}, \mathrm{id}, \mathrm{e}^{-2\pi\mathrm{i}\lambda} h\mathrm{e}^{\chi_p(h)}).$$

An analogous formula holds for functional moduli for resonant saddle singularity from Theorem 3.44.

We finish this subsection by description of some recent results (due to S. Voronin, A. Grintchy, Yu. Meshcheryakova and L. Teyssier) about classification of saddle–node and resonant saddle singularities with respect to conjugation by means of local diffeomorphisms, but without change of time.

PROPOSITION 3.56. (See [60,38].) *Any saddle–node singularity whose formal orbital normal form is like V_0 in* (3.16) *is formally equivalent to*

$$P(x)V_0,$$

where $P(x)$, $P(0) \neq 0$, is a polynomial of degree at most p. Two such vector fields with polynomials P and P' are equivalent if $P'(x) = P(\mathrm{e}^{2\pi\mathrm{i}k/p}x)$ for some integer k.

Any resonant saddle singularity with formal orbital normal form like V_0 in (3.17) *is formally equivalent to*

$$P(u)V_0,$$

where $P(u)$, $u = x^n y^m$,is a polynomial of degree at most p.

PROOF. Consider the saddle–node case. Of course, we can assume that the vector field is of the form $Q(x,y)V_0$ where Q is a series with non-zero constant term. We try to reduce Q by means of diffeomorphisms which preserve the foliation $\mathcal{F}_0$ defined by V_0 and which are close to identity. They are of the form $G = g_{V_0}^{F(x,y)}$, i.e. the phase flow maps after varying time $F(x,y)$. An analogue of Lemma 3.54 says that $G^*(QV_0) = Q'V_0$, where Q' satisfies the equation

$$V_0(F) = \frac{1}{Q} - \frac{1}{Q'}.$$

Since $V_0(F) = x^{p+1}F_x - (1+\lambda x^p)yF_y$, it is easy to see that we can firstly reduce from $\frac{1}{Q}$ all terms divisible by y and next all terms divisible by x^{p+1}. From this the result easily follows.

In the resonant saddle case analogously we firstly reduce all terms from $\frac{1}{Q}$ different from a power of u and next use the fact that $V(R \circ u) = u^{p+1}\frac{\mathrm{d}R}{\mathrm{d}u}$. □

Let us describe the moduli, which generalize the Martinet–Ramis moduli for saddle–node singularities for fixed the polynomial $P(x)$ from the last proposition. Denote by $\mathcal{E}_{p,\lambda,P}$ the space of saddle–node germs with such formal normal form.

We consider collections $(\phi; \psi) = (\phi_1, \ldots, \phi_{2p}; \psi_1, \ldots, \psi_p)$ where $\phi = (\phi_1, \ldots, \phi_{2p})$ is the collection appearing in Definition 3.39 of the Martinet–Ramis moduli and $\psi_j(h)$ are germ of holomorphic functions in $(\mathbb{C}, 0)$ such that $\psi_j(0) = 0$.

DEFINITION 3.57. Two such collections $(\phi; \psi) = (\phi_1, \ldots, \phi_{2p}; \psi_1, \ldots, \psi_p)$ and $(\phi'; \psi') = (\phi'_1, \ldots, \phi'_{2p}; \psi'_1, \ldots, \psi'_p)$ are called *equivalent* iff ϕ_j and ϕ'_j are simultaneously conjugated by a linear transformation, $\phi_j \circ C = C\phi'_j$, and altogether we have $\psi_j \circ C = \psi'_j$.

The space of equivalence classes $[\phi; \psi]$ of such a collections is the *moduli space* denoted by $\mathcal{N}_{p,\lambda,P}$.

THEOREM 3.58. (See [60,79].) *There exists a map $W \to \varepsilon_W$ from $\mathcal{E}_{p,\lambda,P}$ to $\mathcal{N}_{p,\lambda,P}$ with the following properties*:

- *it sends analytically equivalent germs of vector fields to the same point* (*invariance*);
- *if $\varepsilon_W = \varepsilon_{W'}$ then W and W' are analytically equivalent* (*equimodality*);
- *for any $\varepsilon \in \mathcal{N}_{p,\lambda,P}$ there exists $W \in \mathcal{E}_{p,\lambda,P}$ such that $\varepsilon = \varepsilon_W$* (*realization*).

IDEA OF THE PROOF. We can assume that the vector field is of the form $W = QV$, where V is like in the thesis of Lemma 3.36, i.e. with $\dot{x} = x^{p+1}$. We try to reduce the germ QV to the germ $P(x)V$ using conjugation by means of the map $G = g_V^F$. Recall that this leads to the equation $V(F_j) = \frac{1}{Q} - \frac{1}{P}$. We can solve this equation in the extended sectors T_j. The obstacles to a solution in a whole neighborhood of the origin are given by the integrals

$$\psi_j(h) = \int_{\gamma_j^\infty} \left(\frac{1}{Q} - \frac{1}{P}\right) \frac{\mathrm{d}x}{Qx^{p+1}}$$

along the asymptotic cycles γ_j^∞ in the leaves L_h. The functions ψ_j constitute part of the modulus $(\phi; \psi)$.

By the sectorial normalization theorem (Theorem 3.37) we can reduce the vector field V in extended sectors T_j to V_0 and the obstructions to a univalent reduction form the collection ϕ (the Martinet–Ramis modulus), the other part of (ϕ, ψ).

We omit he other details of the proof. □

Finally we note that there exists a generalization of the latter theorem to the case of resonant saddle singularities. For details we refer the reader to the papers of A. Grintchy and S. Voronin [38] and L. Teyssier [79,80]. Besides the collection of $2p$ germs ϕ_i of local diffeomorphisms, defined in Theorem 3.44 (via Theorem 3.33) we have also collection of p germs of functions ψ_j.

3.7. *Linearization*

Now we pass to the non-resonant saddles. Their one-dimensional equivalents are the germs of non-resonant analytic diffeomorphisms f of $(\mathbb{C}, 0)$ of the form

$$z \to \mathrm{e}^{2\pi \mathrm{i}\alpha} z + \mathrm{O}(z^2),$$

where $\alpha \notin \mathbb{Q}$, $0 < \alpha < 1$. The Poincaré–Dulac theorem for diffeomorphisms says that any such f is formally linearizable (see Theorem 3.22); there exists a formal power series $h(z) \sim z + h_2 z^2 + \cdots$ such that $h^{-1} \circ f \circ h \sim \lambda z$, $\lambda = \mathrm{e}^{2\pi \mathrm{i}\alpha}$. The problem is under which assumptions f is analytically linearizable, i.e. when the series defining h is convergent.

The positive or negative answer to this question depends on the way how quickly the number α can be approximated by rational numbers. If α is slowly approximated by rationals then any such diffeomorphism is analytically linearizable. If α is quickly approximated by rationals then there should exist diffeomorphisms f for which the series h is divergent.

Let $\alpha = 1/a_1 + 1/a_2 + \cdots$ be the continued fraction expansion of the number α and let $1/a_1 + 1/a_2 + \cdots + 1/a_n = \frac{p_n}{q_n}$ be the successive rational approximations of α.

DEFINITION 3.59. We say that the number α satisfies the *Briuno condition* if

$$\sum \frac{\ln q_{n+1}}{q_n} < \infty.$$

The next theorem is the fundamental result of this section. Its first part was firstly obtained by A. Briuno, but the proof we present is due to C. Yoccoz. The second part was obtained by C. Yoccoz.

THEOREM 3.60. (See [16,87].) *If the number α satisfies the Briuno condition then any analytic diffeomorphism f is analytically linearizable and if α does not satisfy the Briuno condition then there exists an analytic diffeomorphism $f = \mathrm{e}^{2\pi \mathrm{i}\alpha} z + \cdots$ which is not analytically linearizable.*

REMARK 3.61. We should mention some previous results in this field. C. Siegel [74] proved that analyticity of the conjugating diffeomorphism holds when there exist two constants C, μ, such that for any integers p, q we have $|\alpha - p/q| > C|q|^{-\mu}$. If $\mu > 2$ then the set of those α, for which there exists a constant C with this property, is a set of full Lebesque measure.

A. Briuno [16] and H. Cremer [25] noticed that the Briuno condition is close to an optimal condition for convergence. They proved that if $\limsup \frac{\ln q_{n+1}}{q_n} = \infty$ then there exists an analytic diffeomorphism f which is not analytically linearizable.

The proofs of these results are based on estimations. The main achievement of J.-C. Yoccoz was to give a geometrical proof of a theorem of Briuno and to show that the Briuno condition is also necessary for convergence.

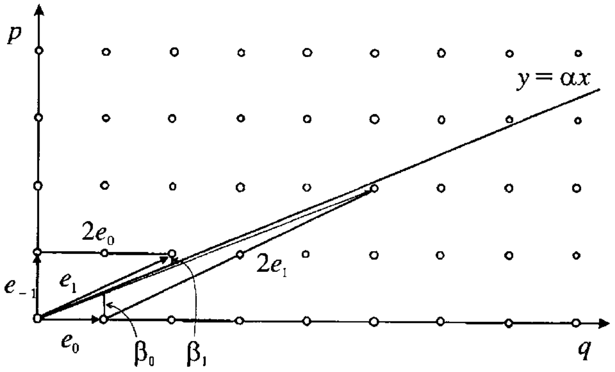


Fig. 8.

Before proving the main theorem of this section we recall some facts about *continued fractions*.

Let $0 < \alpha < 1$ be an irrational. Consider the real plane with coordinates x, y (see Fig. 8). We draw the line $y = \alpha x$. We distinguish the points with integer coordinates (q, p) in the first quadrant. They all, except $(0, 0)$, do not belong to our line. We consider the convex hulls of the sets of integer points of the quadrant, which lie at one side of our line (below and above respectively). (One can imagine the integer points as nails and our line as a thread with fixed one end at infinity; next one stretches the thread in the down direction and in the up direction.) The vertices (q, p) of the broken lines just constructed define the best approximations of the number α by rationals p/q.

We describe the construction of our broken lines in another way. Denote the basic vectors $e_{-1} = (0, 1)$, $e_0 = (1, 0)$. They lie in different sides of the line $y = \alpha x$. We shall construct a sequence of vectors $e_1, e_2, \ldots$ as follows. Let e_{n-1} and e_n be already constructed and lie in different sides of the line. We add the vector e_n to e_{n-1} as many times a_{n+1} as it is possible with the condition that the sum lies in the same side of the line as e_{n-1}.

In this way we obtain the sequence of natural numbers a_n and the sequence of integer vectors

$$e_1 = e_{-1} + a_1 e_0, \ldots, e_{n+1} = e_{n-1} + a_{n+1} e_n.$$

(Note that $a_1 = [1/\alpha]$, the integer part of $1/\alpha$.) The endpoints of the vectors $e_n = (q_n, p_n)$ are the vertices of the above two convex hulls.

We denote by β_n the 'distance' of the point e_n to the line $y = \alpha x$

$$\beta_n = (-1)^n (q_n \alpha - p_n).$$

From Fig. 8 it is seen that $\beta_{-1} = 1$, $\beta_0 = \alpha$ and $\beta_{n+1} = \beta_{n-1} - a_{n+1}\beta_n$, which implies that the ratios p_n/q_n form the best approximations to α, in the sense that $|\alpha - p_n/q_n| < |\alpha - p/q|$ for any p and any $q < q_n$. In fact, we have $\beta_n < |q\alpha - p| < \beta_{n-1}$ for $q_{n-1} < q < q_n$.

We define also the numbers $\alpha_n = \beta_n / \beta_{n-1}$. Thus $\alpha_0 = \alpha$. They lie between 0 and 1. We obtain $\alpha_{n-1}^{-1} = a_n + \alpha_n$. This means that $\alpha_0 = \{\alpha\}, \ldots, \alpha_{n+1} = \{\alpha_n^{-1}\}$ where $\{\cdot\}$ denotes the fractional part. Moreover, iterating the latter identity we recognize the continued fraction

$$\alpha = a_0 + \alpha_0 = a_0 + \frac{1}{a_1 + \alpha_1} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \alpha_2}} = \cdots.$$

Note also that

$$\beta_n = \alpha_0 \cdots \alpha_n.$$

We note also that the (oriented) area of the parallelogram with sides e_n, e_{n-1} is equal to $(-1)^n$. It implies $q_{n+1}\beta_n + q_n\beta_{n+1} = 1$, or

$$\beta_n(q_{n+1} + q_n\alpha_{n+1}) = 1.$$

From this one proves the following (see also [89])

LEMMA 3.62. *The following conditions are equivalent*:

- $\sum q_n^{-1} \ln q_{n+1} < \infty$;
- $\Phi(\alpha) \overset{df}{=} -\sum \beta_n \ln \alpha_{n+1} < \infty$.

3.7.1. *Proof of the analytic linearizability* (We follow [65].) Recall that we have to prove the analyticity of the conjugation series under assumption of the Briuno condition.

We begin with the following criterion of analytic linearizability.

LEMMA. *f is linearizable iff it is stable, i.e. iff for any neighborhood U of* 0 *there is a neighborhood V such that $f^n(V) \subset U$ for all $n \geqslant 0$.*

PROOF. If f is linearizable then it is stable because the linear map is stable.

Assume that f is stable and let V be such that $f^n(V) \subset U = \mathbf{D}$, $n \geqslant 0$. Then the family $f^n|_V$ is normal. Define $h_n = \frac{1}{n}\sum_{i=0}^{n-1} \lambda^{-i} f^i$. It is also normal family. We have $h_n \circ f = \lambda h_n + \frac{1}{n}\lambda(\lambda^{-n} f^n - 1)$.

Now it is enough to choose a convergent subsequence from $\{h_n\}$ and the limit conjugates f with λz. □

After rescaling the variable z we can assume that we are dealing with functions which are analytic diffeomorphisms of the unit disc $\mathbf{D} = \{|z| < 1\}$ into $\mathbb{C}$ with the fixed first term $e^{2\pi i\alpha} z$. We denote the space of such functions by $S(\alpha)$. Of course, we can (and shall) assume that $0 < \alpha < 1$.

The space $S(\alpha)$ has nice properties based on the following *Theorem of Koebé*:

If $f : (\mathbf{D}, 0) \to (\mathbb{C}, 0)$ is an univalent holomorphic function with $|f'(0)| = 1$ then

$$\bigl|f(z)\bigr| < |z|\bigl(1 - |z|\bigr)^{-2}, \qquad \bigl|f'(z)\bigr| < \bigl(1 + |z|\bigr)\bigl(1 - |z|\bigr)^{-3}.$$

The important corollary of the Koebé estimates is the following.

LEMMA. *The spaces $S(\alpha)$ are compact in the topology of almost uniform convergence. In other words, the family $S(\alpha)$ is normal.*

The authorship of next lemma belongs to A. Douady and E. Ghys.

LEMMA. *The set of those α, for which any $f \in S(\alpha)$ is linearizable, is invariant with respect to the action of the group* $\mathrm{SL}(2,\mathbb{Z}) : \alpha \to \frac{a\alpha+b}{c\alpha+d}$.

PROOF. The group $\mathrm{SL}(2,\mathbb{Z})$ is generated by transformations of the form $\alpha \to \alpha + 1$ and $\alpha \to -1/\alpha$. The invariance with respect to the first transformations is obvious.

Let $0 < \alpha < 1$. We shall associate with any $f \in S(\alpha)$ a map $g \in S(\beta)$, $\beta = -1/\alpha$. If l is a ray and $f(l)$ is its image, then we add a curve l' closing-up a triangle-like region Δ. We glue l with $f(l)$ by means of f and obtain a surface, which admits a holomorphic structure of the disc **D** (see Appendix 1 below). This equivalence is of the form $z \to z_1 = z^{1/\alpha} + \cdots$.

Denote by g the map of first return to Δ. It turns out to be holomorphic with respect to the complex structure on $\Delta/\sim$. Expressed in the chart $z_1 = \rho \mathrm{e}^{\mathrm{i}\theta}$ on $\Delta/\sim$ it takes the form

$$\mathrm{e}^{2\pi\mathrm{i}(j\alpha-1)/\alpha+\mathrm{i}\theta}\rho + \cdots = \mathrm{e}^{2\pi\mathrm{i}\beta} z_1 + \cdots,$$

where $j = j(z_1)$ is the moment of the first return. □

It is useful to pass to the chart, where the punctured disc $\mathbf{D}^*$ is replaced by its infinite covering; the upper half-plane $\mathbf{H} = \{\operatorname{Im}\zeta > 0\}$ with the covering map $\zeta \to z = \mathrm{e}^{2\pi\mathrm{i}\zeta}$.

The lift of the map f to **H** has the form

$$F(\zeta) = \zeta + \alpha + \sum a_n \mathrm{e}^{2\pi\mathrm{i}n\zeta},$$

i.e. it commutes with the translation id$+1$. Using the compactness of the space $S(\alpha)$ (or estimates for the coefficients a_n) it is possible to get the following uniform bound for the difference of F from the pure translation.

LEMMA. *There exists a constant C, not depending on α and F, such that the inequality*

$$\left|F(\zeta) - \zeta - \alpha\right| < \alpha/4$$

holds for

$$\operatorname{Im}\zeta > t(\alpha) \overset{df}{=} (2\pi)^{-1} \cdot \ln \alpha^{-1} + C.$$

The construction from the proof of the previous lemma can be implemented for maps on **H**. It is presented at Fig. 9. The line $l = \{\zeta = \mathrm{i}t, t > t(\alpha)\}$ is identified with $F(l)$. The domain $\tilde{\Delta}$ between l, $F(l)$ and the straight segment (joining $\mathrm{i}t(\alpha)$ with $F(\mathrm{i}t(\alpha))$) is replaced

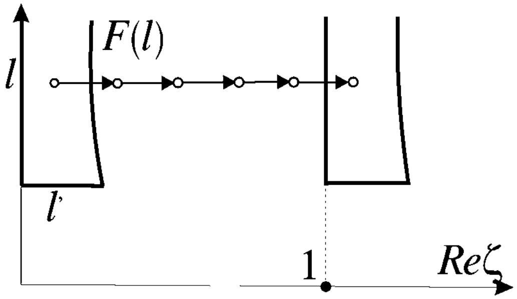


Fig. 9.

by the domain $\{0 \leqslant \operatorname{Re} \zeta_1 \leqslant 1, \operatorname{Im} \zeta_1 \geqslant 0\} \subset \mathbf{H}$: $\zeta_1 \approx \alpha^{-1}(\zeta - \mathrm{i}t(\alpha))$. The map $G \in S(\beta)$, $\beta = -1/\alpha$ is induced by the return map to $\tilde{\Delta}$ with the chart ζ_1.

Because we use the expansions of α into continuous fractions we prefer to get some map from $S(1/\alpha)$. This is done by conjugation of the map G with the map $\zeta = s + \mathrm{i}t \to -s + \mathrm{i}t$ which reverses the orientation. The new map will be denoted also by G.

Therefore we have the transformation from $S(\alpha)$ to $S(1/\alpha)$: $F \to G = G(F)$.

Note that if F is unstable, i.e. iterations of some points escape from the upper half-plane, then G is also unstable and the escape is quicker that for F. The following lemma provides a quantitative description of this.

LEMMA. *If a point* ζ *with* $\operatorname{Im} \zeta > t(\alpha)$ *escapes from* $\mathbf{H}$ *after* m *iterations of* F, *i.e.* $F^m(\zeta) \notin \mathbf{H}$ *and* $F^i(\zeta) \in \mathbf{H}$, $i < m$, *then there exists a point* ζ_1 *with* $\operatorname{Im} \zeta_1 > \alpha^{-1}(\operatorname{Im} \zeta - t(\alpha) - C_1)$ *which escapes from* $\mathbf{H}$ *after* $m_1 < m$ *iterations of* G. *Here* C_1 *is an universal constant such that* G *is well defined for* $\operatorname{Im} \zeta > t(\alpha) + C_1$.

Let us start to iterate the property of the latter lemma. Let the formulas

$$\alpha_0 = \alpha, \qquad \alpha_n^{-1} = a_{n+1} + \alpha_{n+1}$$

define the sequence of natural numbers a_n and $\alpha_n \in (0, 1)$. We put $F_0 = F$, $F_1 = G(F_0) - a_1 \in S(\alpha_1), \ldots, F_{n+1} = G(F_n) - a_n$

If the point ζ_1 (from the latter lemma) satisfies the inequality $\operatorname{Im} \zeta_1 > t(\alpha_1)$, which is equivalent to

$$\operatorname{Im} \zeta > t(\alpha_0) + \alpha_0 t(\alpha_1) + C_1, \tag{3.18}$$

then $F_1^{m_1}(\zeta_1) \notin \mathbf{H}$, $F_2^{m_2}(\zeta_2) \notin \mathbf{H}$ for some ζ_1, ζ_2 and $0 < m_2 < m_1 < m$.

At the n-th repeating of this procedure the inequality (3.18) should be replaced by $\operatorname{Im} \zeta > t(\alpha_0) + \alpha_0 t(\alpha_1) + \alpha_0 \alpha_1 t(\alpha_2) + \cdots + \alpha_0 \cdots \alpha_n t(\alpha_{n+1}) + C_1$. Because $t(\alpha) = -\ln \alpha/(2\pi) + C$, $\beta_n = \alpha_0 \cdots \alpha_n$ (see above), the latter finite series equals (up to constants) to the partial sum defining the series $\Phi(\alpha) = -\sum \beta_n \ln \alpha_{n+1}$ (defined in Proposition 3.62). By the Briuno condition this series is convergent.

We claim that the region $\operatorname{Im} \zeta > \sum \beta_n t(\alpha_{n+1}) + C_1$ is the region of stability of the map F. Assuming contrary, it should contain a point ζ_0 escaping from $\mathbf{H}$ after n_0 iterations

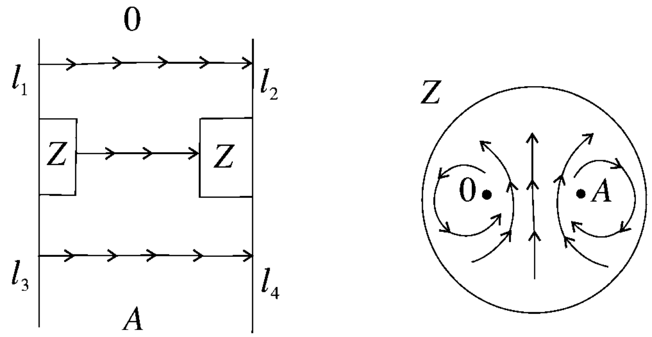


Fig. 10.

of $F = F_0$. This would imply existence of infinite sequence of points ζ_j escaping after n_j iterations of F_j. Moreover one should have infinite sequence of inequalities

$$0 < \cdots < n_2 < n_1 < n_0.$$

Of course, it is not possible.

3.7.2. *Proof of the non-linearizability* The idea of the proof of the second part of Theorem 3.60 is to construct a map from $S(\alpha)$ with a sequence of periodic orbits approaching the fixed point 0. By reversing and modifying the construction of Douady and Ghys one successively inserts fixed points near 0. In the next modification this fixed point becomes periodic and a new fixed point is inserted.

The reversing of the construction $F \to G(F)$, i.e. the transformation $G \to F(G)$, is presented at Fig. 10.

Let $G \in S(\beta)$, $0 < \beta < 1$, and let a be a positive integer. We put $\alpha = (\beta + a)^{-1}$ and let $\tilde{G} = G + a$.

We have the half-line $l_1 = \{\zeta = \mathrm{i}t, t > t_0\}$, its image $l_2 = \tilde{G}(l_1)$, the half-line $l_3 = \{\mathrm{i}t, t < t_0\}$ and the half-line $l_4 = l_3 + [\tilde{G}(\mathrm{i}t_0) - \mathrm{i}t_0]$. These curves bound a domain Λ, close to a vertical strip. Identifying the sides of this non-straight strip, we would obtain a surface $\Lambda/\sim$ with a holomorphic structure of $\mathbb{C}^*$. Because we want to have a disc, we delete form $\Lambda/\sim$ a disc Z around $\mathrm{i}t_0$ of size O(1), like at Fig. 10. Thus $(\Lambda/\sim) \setminus Z$ is the same as a doubly punctured disc $\mathbf{D} \setminus \{0, A\}$: $\mathrm{i}\infty$ is sent to 0 and $-\mathrm{i}\infty$ is sent to a point A. The set Z is sent to the outside of $\mathbf{D}$. (Of course, $(\Lambda/\sim) \setminus Z$ can be next covered by $\mathbf{H}$).

Let us define the dynamics on $(\Lambda/\sim) \setminus Z$, as induced by the translation $\mathrm{id}+1$. Because G commutes with $\mathrm{id}+1$ and l_4 is a translation of l_3 we obtain a well defined map of the quotient $(\Lambda/\sim) \setminus Z$. We denote it by $F = F(G)$. It is easy to see that $F \in S(\alpha)$ and that F has two fixed points in $\mathbf{D}$: 0 and A (see Fig. 10). Therefore we have inserted a new fixed point.

Before exploring this construction we have to perform some estimates. We want to insert the fixed point A as close to 0 as possible. However there are some restrictions implied by the geometry of the closure of $(\Lambda/\sim) \setminus Z$. When we delete from it a disc containing 0 and A then we obtain a domain Λ_1 in form of an annulus.

There is a conformal invariant of annulus, the modulus equal to $m = \ln(R/r)$ (see the Definition 3.75 below).

The modulus of the ring Λ_1 is of order $-\ln|A|$. On the other hand, this modulus can be estimated from Fig. 10. The domain Z is of order $O(1)$ and the domain Λ is of order $O(a) = O(\alpha^{-1})$; (it is the width of Λ). This implies that the modulus of Λ_1 is of order $\ln\alpha$. The rigorous proof uses the estimate from the Theorem 3.76: $m \geqslant L^2/S$, where $S = \iint_{\Lambda-\Lambda_1}\rho$, $L = \inf\int_\gamma \rho$, where $\rho(\zeta) > 0$ is some function and γ's are paths joining the two components of the boundary of Λ_1. Yoccoz [87] uses $\rho(\zeta) = 1/|\zeta - \zeta_j|$ (in the ζ-plane).

This allows to get the estimate $|A| < \text{const}\cdot\alpha$. This estimate, expressed in terms of coordinates at the upper half-plane, says that we have the fixed point ζ_0 of F and satisfying

$$\operatorname{Im}\zeta_0 > t(\alpha) - C_2$$

where $t(\beta) = -(\ln\beta)/(2\pi) + C$ is the same as in the previous proof and C_2 is some universal constant.

Notice also that, if G has some periodic orbit then F also has a periodic orbit. We have the following

LEMMA. *Let* $0 < \beta < 1$, *a be a positive integer,* $\alpha = (\beta + a)^{-1}$ *and* $G \in S(\beta)$. *There exists* $F \in S(\alpha)$ *such that*:

- *F has fixed point ζ_0 with* $\operatorname{Im}\zeta_0 > t(\alpha) - C_2$.
- *If G has a periodic orbit with the rotation number p/q (i.e. $G^q(\zeta) = \zeta + p$) then F has a periodic orbit $\{F^i(\zeta')\}$ with the rotation number $(a + p/q)^{-1}$ and such that*

$$\min \operatorname{Im} F^j(\zeta') > \alpha \cdot \min \operatorname{Im} G^i(\zeta) + t(\alpha) - \text{const}.$$

Let α_n and a_n be the sequences defining the continued fraction expansion. For any natural m we construct series of maps $F_{m,m+1}, F_{m,m}, \ldots, F_{m,0}$ as follows.

$F_{m,m+1} = \zeta + \alpha_{m+1} \in S(\alpha_{m+1})$, $F_{m,m} = F(F_{m,m+1})$ (with $a = a_{m+1}$) and other maps are constructed in the same way: $F_{m,l} = F(F_{m,l+1})$. Thus $F_{m,0} \in S(\alpha)$.

Moreover, each $F_{m,0}$ has periodic orbits with rotation numbers equal to the reduct p_n/q_n of the continuous fraction. The imaginary parts of such orbit are estimated from below by $\sim \sum_{i=1}^n \beta_{i-1}|\ln\alpha_i|$. Because the Briuno condition fails the latter series tends to infinity.

Now we choose a convergent subsequence from the sequence $F_{m,0}$. Its limit is just the F we are looking for. It has infinite series of periodic orbits with arbitrary large imaginary part. These periodic orbits form an obstacle to the linearizability of F. □

We shall apply the theorems of Briuno and of Yoccoz to germs of analytic planar vector fields with singular points of saddle type without resonance

$$\dot{x} = \alpha x(1 + \cdots), \qquad \dot{y} = -y(1 + \cdots), \tag{3.19}$$

where α is irrational.

As in the resonant case one associates with the analytic separatrix $x = 0$ the monodromy map $\Delta : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$, $\Delta(x) = e^{-2\pi i\alpha}x + \cdots$. Theorems 3.60 implies the following.

THEOREM 3.63. *If α satisfies the Briuno condition than any germ* (3.19) *is analytically orbitally linearizable.*

If α does not satisfy the Briuno condition then we cannot simply apply Theorem 3.60. We should have the realization theorem; any germ of non-resonant one-dimensional map is realized as a holonomy map of certain holomorphic foliation with saddle singularity. Such theorem was proved by R. Perez-Marco and C. Yoccoz [66] (see also [31]).

THEOREM 3.64. *Let α be irrational. The correspondence*:

germ of vector field (3.19) $\rightarrow$ *the germ Δ of its monodromy*

is surjective. This implies that if α does not satisfy the Briuno condition, then there exists a germ (3.19) *of vector field which is not analytically linearizable.*

We do not present the proof of this result. It relies on gluing together certain domains with standard foliations and application the theory of almost complex structures to get a holomorphic foliation.

3.8. *Appendix 1. Complex structures and almost complex structures*

Consider the following situation. Let M^n be a n-dimensional complex manifold. Let us treat it as a $2n$-dimensional real manifold ${}^{\mathbb{R}}M^{2n}$. Assume now that the complex atlas on M is replaced by a C^1-smooth atlas on ${}^{\mathbb{R}}M^{2n}$. It turns out that the complex structure from M is not completely lost on ${}^{\mathbb{R}}M^{2n}$. There remains its trace called the almost complex structure.

DEFINITION 3.65. An *almost complex structure* on a real manifold M is a field $J = \{J_x\}$ of endomorphisms of the tangent spaces T_xM satisfying the identity

$$J_x^2 + I = 0.$$

In other words, it is a field of splittings of the complexifications of tangent spaces into eigenspaces of J_x with eigenvalues $\mathrm{i} = \sqrt{-1}$ and $-\mathrm{i}$

$${}^{\mathbb{C}}T_xM = T_xM \otimes \mathbb{C} = T_x^{1,0} \oplus T_x^{0,1}$$

of equal (complex) dimension n.

The standard complex structure on a complex manifold N, with local complex coordinates $z_j = x_j + \mathrm{i}y_j$, induces the almost complex structure by the formulas $J\partial_{x_j} = \partial_{y_j}$, $J\partial_{y_j} = -\partial_{x_j}$.

DEFINITION 3.66. An almost complex structure J is called *integrable* if locally it is an image of complex structure by means of a C^1-mapping. It means that in neighborhood of any point on M there exists a local system of C^1-coordinates x_j, y_j such that the above formulas hold.

If M is a real analytic $2n$-dimensional manifold and ${}^{\mathbb{C}}M$ is its local complexification (a complex $2n$-dimensional manifold containing M as its real part) then the integrability of the almost complex structure means simultaneous integrability of the two distributions $T^{1,0} = \bigcup T_x^{1,0}({}^{\mathbb{C}}M)$ and $T^{0,1}$. The local leaves of the corresponding foliations are $\{x_1 + \mathrm{i}y_1 = c_1, \dots, x_n + \mathrm{i}y_n = c_n\}$ and $\{x_1 - \mathrm{i}y_1 = d_1, \dots, x_n - \mathrm{i}y_n = d_n\}$, $x_j, y_j, c_j, d_j \in \mathbb{C}$.

The splitting of the complexifications of the tangent spaces induces analogous splitting of the complexifications of the cotangent spaces into forms of the type $(1,0)$ and of the type $(0,1)$,

$$ {}^{\mathbb{C}}T^*M = {}^{\mathbb{C}}\mathcal{E}^1 = \mathcal{E}^{1,0} \oplus \mathcal{E}^{0,1}; $$

ω is of the type $(1,0)$ iff it vanishes at $T^{0,1}$. This induces splitting of all differential k-forms (with complex coefficients) into forms of the type (p,q).

As we know from Frobenius theorem (see [43]) the distributions $T^{1,0}$ and $T^{0,1}$ (of vectors of the type $(1,0)$ and $(0,1)$ respectively) are integrable iff they are involutive. This means that the commutator of any two vector fields from $T^{1,0}$ (respectively from $T^{0,1}$) is a vector field from $T^{1,0}$ (respectively from $T^{0,1}$). This condition can be written in the following compact form.

DEFINITION 3.67. The *torsion* of the almost complex structure J is the following bilinear form on the space of vector fields on M

$$ N(X,Y) = 2\big\{[JX, JY] - [X,Y] - J[X, JY] - J[JX, Y]\big\}. $$

Vanishing of the torsion is equivalent to the fact that the derivative of a form ω of the type $(1,0)$ does not have component of the type $(0,2)$. Note that, if $JX = \epsilon X$ and $JY = \epsilon Y$, $\epsilon = \pm\mathrm{i}$, then $N(X,Y) = -4\{[X,Y] + \epsilon \mathrm{i} J[X,Y]\}$.

The following result is known as the *Newlander–Nirenberg theorem*. In the case of real analytic manifold M this result follows from the above remarks and the Frobenius theorem. A. Newlander and L. Nirenberg proved it in the finitely smooth case.

THEOREM 3.68. (See [64].) *An almost complex structure is integrable iff its torsion vanishes.*

Now we pass to the 1-dimensional case. Here the problem of integrability of almost complex structures is treated differently. Assume that we are in a 2-dimensional plane (local chart) which we identify with the complex plane $\mathbb{C}$. Here the splitting of ${}^{\mathbb{C}}T^*\mathbb{C}$ can be defined by means of one 1-form

$$ \omega = \mathrm{d}z + \mu\, \mathrm{d}\bar{z} $$

where $\mu = \mu(z, \bar{z})$ is a function. We have $\mathcal{E}^{1,0} = \mathbb{C}\,\omega$ and the other space $\mathcal{E}^{0,1}$ is its conjugate.

The function μ is not defined uniquely. Notice that, when we replace the complex variable z by $w = \phi(z)$ (ϕ holomorphic), then the analogous 1-form $\tilde{\omega} = \mathrm{d}w + \tilde{\mu}\,\mathrm{d}\bar{w}$ takes the form

$$\phi_z'\,\mathrm{d}z + \tilde{\mu}\bar{\phi_z'}\,\mathrm{d}\bar{z} = \phi_z'\big[\mathrm{d}z + \tilde{\mu}\cdot(\bar{\phi_z'}/\phi_z')\cdot\mathrm{d}\bar{z}\big]. \tag{3.20}$$

This rule of transformation of μ suggests the following definition.

DEFINITION 3.69. The quantity

$$\mu\cdot(\mathrm{d}\bar{z}/\mathrm{d}z)$$

is called the *Beltrami differential* defining the almost complex structure. In particular, $|\mu|$ is a well defined function.

The integrability of the almost complex structure, defined by the Beltrami differential, means existence of a differentiable complex function $\phi = \phi(z, \bar{z})$ which is a homeomorphism and such that the almost complex structure is defined by $\phi^*\,\mathrm{d}u = \mathrm{d}\phi = \phi_z\,\mathrm{d}z + \phi_{\bar{z}}\,\mathrm{d}\bar{z}$, i.e. the following *Beltrami equation*

$$\phi_{\bar{z}} = \mu\phi_z$$

holds. The Beltrami equation is a partial differential equation and is solved in a suitable Sobolev space. The ϕ's satisfying the Beltrami equation are called the *quasi-conformal maps* and $\mu = \mu_\phi$ is called the *modulus* of ϕ.

(Note that the torsion of any almost complex structure in $\mathbb{C}$ vanishes. It is because $\bigwedge^2 \mathcal{E}^1 \otimes \mathbb{C} = \mathcal{E}^{1,0} \wedge \mathcal{E}^{0,1}$ is one dimensional and contains only 2-forms of the type $(1, 1)$. Therefore here we have different kind of problems than in multi-dimensional case. In particular, one condition of integrability would be the preservation of the orientation by the map ϕ; it is equivalent to the property $|\mu| < 1$.)

The next theorem has many authors (L. Ahlfors, B. Bojarski, L. Bers).

THEOREM 3.70. (See [1].) *If μ is continuous and satisfies the estimate*

$$\big|\mu(x)\big| \leqslant k < 1$$

then the Beltrami equation has a solution and the almost complex structure defined by $\mu\,\mathrm{d}\bar{z}/\mathrm{d}z$ is integrable.

REMARK 3.71. If g is a conformal (i.e. analytic) mapping then $\phi \circ g$ is quasi-conformal with the modulus $\mu_{\phi\circ g} = \mu_\phi \circ g \cdot (\bar{g'}/g')$ and $g \circ \phi$ is quasi-conformal with $\mu_{g\circ\phi} = \mu_\phi$ (see (3.20)). This means that the Beltrami equation has many solutions.

If $\|\mu_\phi\|_\infty = k < 1$ then the map ϕ transforms infinitesimally small sphere around any z_0 to an infinitesimally small ellipse with the ratio of lengths of the semi-axes (longer to shorter) $K(z_0)$. We have $\sup K(z_0) = (1+k)/(1-k)$.

This implies that, if $\mu_\phi = \mu_\psi$ then the map $\phi \circ \psi^{-1}$ is conformal; ψ^{-1} maps circles to ellipses which are sent again to circles by ϕ.

Consider now the following situation. Let $f : \mathbb{C}^* = \mathbb{C} \setminus 0 \to S$ be a smooth diffeomorphism onto a Riemann surface S. Define $\mu = f'_{\bar{z}}/f'_z$. Assume that $\mu(z)$ can be continuously prolonged to $\bar{\mathbb{C}} = \mathbb{C} \cup \infty$ with $\mu(0) = \mu(\infty) = 0$.

PROPOSITION 3.72. *In such a case there exists a conformal mapping* $g : \mathbb{C}^* \to S$. *This means that* S *has conformal type of* $\mathbb{C}^*$ (*not of an annulus or of a punctured disc*).

PROOF. We prolong μ to $\bar{\mathbb{C}}$ and integrate the corresponding Beltrami equation in the Riemann sphere. Using an eventual composition with a Möbius automorphism we can assume that the solution $u : \bar{\mathbb{C}} \to \bar{\mathbb{C}}$ is a homeomorphism such that $u(0) = 0$, $u(\infty) = \infty$.

We have $\mu_u = \mu_f$ in $\mathbb{C}^*$ and the map $u \circ f^{-1} : S \to \mathbb{C}^*$ is conformal (see the above remark). □

DEFINITION 3.73. If $Q \subset \mathbb{C}$ is a (curved) quadrangle, considered together with a pair of arcs on its boundary (b-arcs) then it can be transformed conformally onto a rectangle (with sides a, b), such that the b-arcs are sent to the vertical sides. The ratio of the lengths of sides is called the *modulus* of Q (see [1])

$$m(Q) = a/b.$$

If $P \subset \mathbb{C}$ is a domain, which can be conformally transformed to an annulus $\{r < |z| < R\}$, then we define its *modulus* as

$$m(P) = \ln(R/r).$$

$m(Q)$ is an invariant of conformal transformations; (by applying the Schwarz lemma one easily shows that two rectangles are conformally equivalent iff their moduli are the same).

The condition of quasi-conformality (for a map ϕ) says that $m(Q') < Km(Q)$ for any quadrangle Q transformed to a quadrangle Q'. Here $K = (1+k)/(1-k)$ is the constant from Remark 3.73. It means that K measures the distortion of small squares.

Note that the map $\ln z$ transforms the standard annulus, cut along a radius, into the rectangle with the sides $\ln R - \ln r$ and 2π. Thus, up to a constant, the modulus of a ring is the same as the modulus of corresponding quadrangle.

There is another definition of the modulus of quadrangle which does not make use of conformal mapping to rectangle.

THEOREM 3.74. (See [1].) *The modulus* $m(Q)$ *of a rectangle equals* $\sup_\rho L^2(\rho)/A(\rho)$, *where* ρ *are square integrable non-negative functions,* $A(\rho) = \iint_Q \rho^2 \, \mathrm{d}x \, \mathrm{d}y$ *and* $L(\rho) = \inf \int_\gamma \rho |\mathrm{d}z|$ *and* γ *are paths joining the* b*-arcs.*

3.9. *Appendix 2. Proof of the Hukuhara–Kimura–Matuda theorem*

(We follow the book [40] and also [89].) We shall present the proof only in the case $p = 1$. Therefore we assume that $x^2\,\mathrm{d}y/\mathrm{d}x = g_0(x) + y(1+\lambda x) + y^2 g_2(x, y)$ where g_0, $y^2 g_2$ have zero at $x = y = 0$ of order $\geqslant 3$ (compare Lemma 3.36).

Let $S = \{-\pi/2+\alpha < \arg x < 3\pi/2-\alpha, |x| < \epsilon\}$ be a sector in the x-plane (see Fig. 11); the case of opposite sector is considered analogously. We seek a change $(x, y) \to (x, z)$ such that $x^2\,\mathrm{d}z/\mathrm{d}x = z(1 + \lambda z^2)$ in S.

The proof is divided into two parts: existence of analytic central manifold in S_1 and analyticity of $z(x, y)$.

Existence of the center manifold. We seek it in the form of a graph

$$y = u(x)\mathrm{e}^{-1/x}, \quad x \in S,$$

where $u(x)\mathrm{e}^{-1/x} = \mathrm{O}(x)$ in the sector. We obtain the following equation onto u

$$x^2 u' = F(x, u) \tag{3.21}$$

where $F = g_0\mathrm{e}^{1/x} + \lambda x u + u^2\mathrm{e}^{-1/x} g_2(x, u\mathrm{e}^{1/x})$. We seek the solution to (3.21) in the class of functions satisfying the estimate

$$|u| < K|x|\mathrm{e}^{\mathrm{Re}(1/x)} \tag{3.22}$$

for some constant K.

The right-hand side of (3.21) satisfies the estimates: $|F(x, u_1) - F(x, u_2)| < \delta|u_1 - u_2|$ and $|F| < (K_0 + \delta K)|x|\mathrm{e}^{\mathrm{Re}(1/x)}$ for some small δ and fixed K_0. The first (Lipschitz) condition will guarantee the uniqueness of the solution.

As in other similar problems, Eq. (3.21) is replaced by the integral equation

$$u(x) = \int_0^x F\bigl(s, u(s)\bigr)s^{-2}\,\mathrm{d}s,$$

which is a fixed point equation $u = T(u)$ for a non-linear operator T. We use the following *Leray–Schauder–Tikhonov fixed point theorem*:

Any continuous map of a compact and convex subset of a locally convex space into itself contains a fixed point.

In our case the set of u's satisfying (3.22) is convex and compact in the topology of almost uniform convergence. It remains to show that the operator T maps the set defined by (3.22) into itself.

The latter statement is proved by appropriate choosing the path γ of integration in the formula defining T. This path is presented at Fig. 11.

Let $s = r\mathrm{e}^{\mathrm{i}\theta}$. We have the straight segment γ_1 joining 0 with $x_0 = r_0\mathrm{e}^{\mathrm{i}\pi/2+2\eta}$, where $\eta > 0$ is a small constant and r_0 will be defined later. In the second part γ_2 we have we have either $r(\theta) = r_0\sqrt{\sin(\theta - \pi/2 - \eta)}$ or $r_0\sqrt{\sin(\pi/2 + 3\eta - \theta)}$, depending on whether $\arg x > \pi/2 + 2\eta$ or not. If the angle α (defining the sector S, see above) is greater than 4η

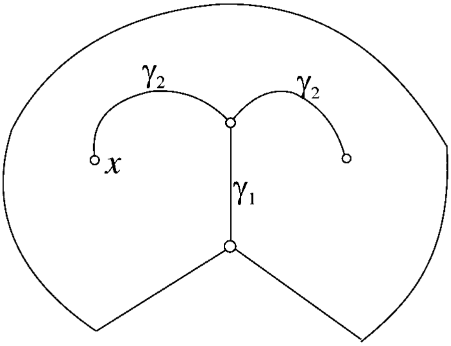


Fig. 11.

then $r_0 < r(\theta) < Mr_0$ (where $M = M(\eta)$). Now we chose r_0 such that Mr_0 is sufficiently small, $< \epsilon$.

One can show that for large K the quantity $|Fs^{-2}|$ is bounded by maximum of the absolute values of some derivatives of $K|s|\mathrm{e}^{\mathrm{Re}(1/s)}$ along $\gamma_{1,2}$ (see [40]). This gives $|T(u)| = |\int Fs^{-2}| \leqslant K|x|\mathrm{e}^{\mathrm{Re}(1/x)}$.

2. Now we can assume that $x^2\,\mathrm{d}y/\mathrm{d}x = y(1+\lambda x + yg_2)$ in S. We have to find the change $y \to z(x, y)$ giving the normal form $x^2 z' = z(1+\lambda x)$.

We represent y as a function of (x, z) : $y = z + \phi(x, z)$. The function ϕ satisfies the following partial differential equation $x^2\phi_x + (1+\lambda x)z\phi_z = G(x, z, \phi)$.

We solve it using the method of characteristics: $\phi(x, z) = \psi(t)$ where $\dot{x} = x^2$, $\dot{z} = z(1+\lambda x)$, $\dot{\psi} = G(x, z, \psi)$. Here t can be replaced by x, i.e. we get $x^2\psi'_x = G(x, z, \psi)$, $z = Cx^\lambda \mathrm{e}^{-1/x}$. This is a family of equations depending on C.

Now we proceed as in the previous point. We put $\psi = w\mathrm{e}^{-1/x}$ where $w = w(x, C)$ satisfies the equation

$$x^2 w' = G(x, z, w), \quad z = Cx^\lambda \mathrm{e}^{-1/x},$$

analogous to (3.21). We replace it by the integral equation

$$w(x, C) = \int_0^x G\bigl(s, Cs^\lambda \mathrm{e}^{-1/s}, w(s, C)\bigr) s^{-2}\,\mathrm{d}s,$$

which we solve in the space of w's satisfying the estimate $|w| < K|z|^2$ for some large constant K.

Using the same estimates as in the previous point (with the same choice of path of integration) one shows that this integral operator satisfies the assumption of the Leray–Schauder–Tikhonov theorem.

We refer the reader to [40] for more details. □

4. Bogdanov–Takens singularity

The subjects of investigation in this section are germs at $0 \in \mathbb{C}^2$ of analytic vector fields of the form

$$\dot{x} = y + \cdots, \qquad \dot{y} = \cdots, \tag{4.1}$$

i.e. with nilpotent linear part. Our aim is to find orbital normal forms for them (see Definition 3.1).

There was done some work on the formal orbital normal forms for the systems with nilpotent linear part. F. Takens [78] in 1974 proved that the system (4.1) can be formally reduced to

$$\dot{x} = y + a(x), \qquad \dot{y} = b(x) \tag{4.2}$$

where

$$a(x) = a_r x^r + a_{r+1} x^{r+1} + \cdots, \qquad b(x) = b_{s-1} x^{s-1} + b_s x^s + \cdots \tag{4.3}$$

are some formal power series (see Theorem 4.2 below). Some authors use the following, equivalent to (4.2), prenormal form (see Remark 4.3 below)

$$\dot{x} = y, \qquad \dot{y} = b(x) + yc(x).$$

Since these forms are not the complete normal forms, we shall call them the *Takens prenormal forms*. In order to obtain the form (4.2) Takens uses only changes of variables x, y but not the time.

DEFINITION 4.1. (See [76].) If $s < 2r$ in (4.3) then we call this case the *generalized cusp case*. The case with $s > 2r$ will be called the *generalized saddle–node case* and the case with $s = 2r$ we call the *generalized saddle case*.

4.1. *Formal orbital normal form*

We begin with the following result of Takens

THEOREM 4.2. (See [78].) *There is a formal change of the variables* (x, y) *which reduces the vector field* (4.1) *to the Takens prenormal form* (4.2).

PROOF. We use a composition of infinite series of transformations of the form $\mathrm{id} + Z$, i.e.

$$(x, y) \to (x_1, y_1) = (x, y) + Z(x, y),$$

$x_1 = x + Z_1(x, y)$, $y_1 = y + Z_2(x, y)$, where $Z_1(x, y)$ and $Z_2(x, y)$ are homogeneous polynomials of degree k. We shall treat also Z as a vector field, $Z = Z_1\partial_x + Z_2\partial_y$.

Let $V = y\partial_x + \sum_{j>1} V^{(j)}(X)$, where $V^{(j)}$ are homogeneous summands of the vector field V. Comparing the terms of degree k in the transformed vector field $(\mathrm{id} + Z)^* V = (\mathrm{id} + Z)_*^{-1} V \circ (\mathrm{id} + Z)$ we obtain the equation

$$[Z, y\partial_x] + V^{(k)} = \left(V^{(k)}\right)^{\text{Takens}} \tag{4.4}$$

where $(V^{(k)})^{\text{Takens}}$ is a term from the Takens prenormal form, $(V^{(k)})^{\text{Takens}} = a_k x^k \partial_x + b_k x^k \partial_y$.

The operator $[\cdot, y\partial_x]$, written in the components, acts on $Z_1\partial_x + Z_2\partial_y$ as follows

$$Z_1\partial_x + Z_2\partial_y \to (Z_2 - y\partial Z_1/\partial x)\partial_x - y\partial Z_2/\partial x \partial_y. \tag{4.5}$$

Equation (4.4) is the *Takens homological equation* and operator (4.5) is the *Takens homological operator*.

Applying formula (4.5) to Eq. (4.4) we see that we can cancel all the terms in the y-component of $V^{(k)}$ divisible by y and, when Z_2 is fixed, we can also cancel the terms divisible by y in the x-component. In this way we obtain the formal Takens prenormal form (4.2). □

REMARK 4.3. The equivalent to (4.1) preliminary normal forms are

$$\dot{x} = y, \qquad \dot{y} = b(x) + yc(x)$$

and

$$\dot{x} = y + \mu x c(x), \quad \dot{y} = b(x) + yc(x), \quad \mu > 0.$$

Indeed, in the first case we make the substitution $y_1 = y + a(x)$. (Note that here we do not need any restrictions).

In the second case we make the substitution $y_1 = y + d(x)$ with $d(x)$ satisfying the equation $\mu x d' = -d + a(x)$. We have $d(x) = \mu^{-1} x^{-1/\mu} \int_0^x s^{1/\mu - 1} a(s)\,\mathrm{d}s$. We see that, if $\mu > 0$, then $d(x)$ is well defined; as a formal power series if $a(x)$ is formal, and as an analytic function if $a(x)$ is analytic. We see also that $d(x) = \mathrm{const} \cdot x^r + \cdots$ if $a(x) = \mathrm{const} \cdot x^r + \cdots$.

Generally, if there exist two analytic separatrices $y = u_1 x^r + \cdots$ and $y = u_2 x^r + \cdots$ then one firstly transforms these separatrices to $y = u_{1,2} x^r$. In such case the vector field takes the form

$$V_H + g(x, y) E_H,$$

where V_H and E_H are given in (4.8) and (4.9) below (see [47] and [76]).

PROPOSITION 4.4. (See [13].) *Let $s < \infty$. There exists a change of the variables (x, y) and of the time which reduces vector field* (4.2) *to the form*

$$\left[y + a(x)\right]\partial_x + x^{s-1}\partial_y. \tag{4.6}$$

Moreover, if vector field (4.2) *is analytic then also the vector field* (4.6) *is analytic.*

PROOF. Assume that $b(x) = x^{s-1}b_1(x), b_1(0) \neq 0$. Our aim is to reduce $b_1(x)$ to 1.

We apply the change $x_1 = x\lambda(x)$, $\mathrm{d}t/\mathrm{d}t_1 = \eta(x)$, which should satisfy the equations $\mathrm{d}x_1/\mathrm{d}t_1 = y + a_1(x_1)$, $\mathrm{d}y/\mathrm{d}t_1 = x_1^{s-1}$. Comparing the terms $x_1^{s-1}\partial_y$ and $y\partial_{x_1}$ we get the conditions $\eta(x) \cdot (x\lambda(x))' = 1$, $\eta(x)b_1(x) = \lambda^{s-1}(x)$, or the Bernoulli equation $x\lambda' + \lambda = b_1(x)\lambda^{1-s}$. It has the solution

$$\lambda(x) = sx^{-1}\left[\int_0^x \tau^{s-1}b_1(\tau)\,\mathrm{d}\tau\right]^{1/s}.$$

This solution is an analytic function when $b_1(x)$ is analytic. □

The form (4.6) will be called the *Bogdanov–Takens prenormal form*.

The formal orbital classification of vector fields (4.2) requires additional partition of the space of such vector fields into some classes. Let

$$n = \min\left(r, \frac{s}{2}\right) \in \frac{1}{2}\mathbb{Z}, \tag{4.7}$$

where r, s are defined in (4.3). Introduce a quasi-homogeneous gradation $\widetilde{\deg}$ by $\widetilde{\deg}\, x = -\widetilde{\deg}\, \partial_x = 1$, $\widetilde{\deg}\, y = -\widetilde{\deg}\, \partial_y = n$. Denote

$$V_H = (y + ax^n)\partial_x + bx^{2n-1}\partial_y, \tag{4.8}$$

the lowest $\widetilde{\deg}$-part of the vector field V in (4.2) with $\widetilde{\deg}\, V_H = n - 1$, and let

$$E_H = x\partial_x + ny\partial_y \tag{4.9}$$

be the *quasi-homogeneous Euler field* with $\widetilde{\deg} E_H = 0$.

Note that one can change the pair (a, b) in (4.8) to $(\nu^2 a, \nu b)$, $\nu \in \mathbb{C}^*$.

Putting formally $z = x^n$ and dividing by x^{n-1}, we get from (4.8) the linear system

$$\dot{z} = naz + ny, \qquad \dot{y} = bz. \tag{4.10}$$

Its eigenvalues are equal $\lambda_{1,2} = \frac{n}{2}(a \mp \sqrt{a^2 + 4b/n})$.

DEFINITION 4.5. We call (4.10) the *principal linear system*.

We make a subtler division (than in Definition 4.1) of the set of Bogdanov–Takens singularities:

- C_{odd}: generalized (odd) cusp with $(a, b) = (0, 1)$ and s odd; here the ratio $\lambda_1/\lambda_2 = -1$.
- $S_{k:-l}$: generalized saddle, or generalized cusp with $s = 2n = 2r$ even, when $\lambda_1/\lambda_2 = -k/l \in \mathbb{Q}_-$.
- S_λ: generalized saddle, when $\lambda_1/\lambda_2 = \lambda \in \mathbb{R}_- \setminus \mathbb{Q}$.
- F: generalized saddle with $\lambda_1/\lambda_2 \in \mathbb{C} \setminus \mathbb{R}$.
- $N_{k:l}$: generalized saddle with $\lambda_1/\lambda_2 = k/l \in \mathbb{Q}_+$.

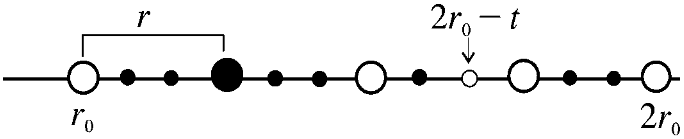


Fig. 12.

- N_λ: generalized saddle with $\lambda_1/\lambda_2 = \lambda \in \mathbb{R}_+$.
- SN: generalized saddle–node, i.e. $(a, b) = (1, 0)$ and $\lambda_1/\lambda_2 = 0$.

Note that in the cases $S_\#$ (respectively F, $N_\#$, SN) the principal linear system (4.10) is a saddle (respectively focus, node, saddle–node).

In the following theorems we present a classification of vector fields from each of the above classes. These fields are of the form $V = V_H + W$, where W is of higher order with respect to the quasi-homogeneous. Each class contains the singularity V_H which is of infinite codimension (as $W \equiv 0$).

Below the notation $j = k(s)$ means $j = k \bmod s$.

THEOREM 4.6. (See [47,76].) *Any vector field from the class* C_{odd} *is formally orbitally equivalent either to* V_H*, or to* $V_H + (x^r + \sum_{j \geqslant r, j \neq 0(s)} a_j x^j)\partial_x$*, where* $r > \frac{s}{2}$*,* $r \neq 0(s)$.

EXAMPLE 4.7. In the usual cusp case, i.e. $s = 3$, equivalent to the above is the following form

$$V_H + \big[A_0(H) + xA_1(H)\big]E_H,$$

where $H = \frac{1}{2}y^2 - \frac{1}{3}x^3$ is the first integral (Hamiltonian) of V_H, used by F. Loray [48] and by M. Canalis-Durand and others [19,20].

THEOREM 4.8. (See [77].) *Any vector field from the class* $S_{k:-l}$, $\gcd(k, l) = 1$, *is formally orbitally equivalent either to* V_H*, or to*

$$V_H + x^t\big(1 + \phi(x)\big)\partial_x,$$

where $t > r$ *and the series* $\phi(x)$ *are of the following forms* (*with* $r_0 = (k + l)r$):

(a) $t \neq 0(r)$ *and* $\phi(x) = \sum_{j \neq 0(r_0),\ j+t \neq 0(r) \text{ or } j+t=r(r_0)} a_j x^j$ (*see Fig.* 12);

(b) $t = p_0 r_0 + r$ *and* $\phi(x) = a_{p_0 r_0} x^{p_0 r_0}$;

(c) $t = p_0 r_0 + r$ *and* $\phi(x) = a_{p_0 r_0} x^{p_0 r_0} + a_{j_0} x^{j_0} + \sum_{j > j_0, j \neq 0(r),\ j \neq j_0 + p_0 r_0} a_j x^j$*, where* $a_{j_0} \neq 0$ *for some* $j_0 > 0$, $j_0 \neq 0(r)$.

THEOREM 4.9. (See [77].) *Any vector field from the classes* S_λ, F, N_λ, $N_{k:l}$ *with* $k, l > 1$, $\gcd(k, l) = 1$, *is formally orbitally equivalent either to* V_H, *or to* $V_H + (x^t + \sum_{j \neq 0(r),\ j > t} a_j x^j)\partial_x$ *where* $t \neq 0(r)$.

THEOREM 4.10. (See [77].) *Any vector field from the class $N_{k:1}$ is formally orbitally equivalent either to V_H, or to $V = V_H + x^{kr}\partial_x$ for $k > 1$, or to $V = V_H + (cx^{kr} + x^t + \sum_{j\neq 0(r),\ j\neq t+kr,\ j>t} a_j x^j)\partial_x$ where $t > r$ and $t \neq 0(r)$, $k \geqslant 1$, and $c = 0$ for $k = 1$.*

THEOREM 4.11. (See [75].) *Any vector field from the class SN is formally orbitally equivalent either to V_H, or to*

$$V_H + x^{s-r-1}\bigl(1+\phi(x)\bigr)E_H,$$

where the series ϕ is of one of the following forms:

(a) $\phi = \sum_{j\neq 0(r)} b_j x^j$ *if* $s \neq 0(r)$;

(b) $\phi = b_{m_0 r} x^{m_0 r}$ *if* $s = (m_0+2)r$;

(c) $\phi = b_{m_0 r} x^{m_0 r} + b_{j_0} x^{j_0} + \sum_{j>j_0,\ j\neq 0(r),\ j\neq j_0+n_0 r} b_j x^j$ *where also* $s = (m_0+2)r$ *and* $b_{j_0} \neq 0$, *for* $j_0 \neq 0(r)$.

REMARK 4.12. The above theorems say nothing about the uniqueness of the presented normal forms. In fact, they are unique relative to an action of a finite group of changes. Each vector field is either of the type $V_H + x^m(1+\phi(x))\partial_x$ or of the type $V_H + x^m(1+\phi(x))E_H$. We consider the changes $x \to \alpha x$, where α is a root of unity of fixed order (e.g. $\alpha^{2r-s} = 1$ for C_{odd}) and we should have $\tilde{\phi}(x) \equiv \phi(\alpha x)$ in order that $V_H + x^m(1+\tilde{\phi}(x))\partial_x$ is equivalent to $V_H + x^m(1+\phi(x))\partial_x$, etc.

We begin the proof of Theorems 4.6, 4.8–4.10 and 4.11. They run along the same scheme. So, we present the general method, but detailed calculations will be provided only for resonant generalized saddle (Theorem 4.8).

The orbital changes rely on application of the conjugations

$$V \to \mathcal{P}_V(Z) := (\mathrm{Ad}_{\exp Z})_* V$$

and multiplications $V \to (1+\chi)V$. Here $Z = Z_1\partial_x + Z_2\partial_y$ is a vector field (formal or analytic), $\exp Z$ is the phase flow diffeomorphism (after the time 1), Ad_h is the adjoint action of a diffeomorphism h on the Lie algebra of vector fields and χ is a function (formal or analytic).

Note that if Z is parallel to V, $Z = \kappa(x,y)V$, then the map $\exp Z$ preserves the phase portrait of V and the field $\mathcal{P}_V(\kappa V)$ is also parallel to V. In order to avoid this ambiguity, one uses the notion of a bi-vector field introduced by R. Bogdanov in [13]. If $V = V_1\partial_x + V_2\partial_y$, then we define the *bi-vector field*

$$Z \wedge V = \Omega \cdot \partial_x \wedge \partial_y \tag{4.11}$$

where $\Omega(x,y) = V_2 Z_1 - V_1 Z_2$. One can say that Ω measures the component of Z transversal to V. If $\Omega = 0$ and V has isolated singularity, then $Z = \kappa V$ for some function κ.

This suggests that one should consider the map

$$Z \to \frac{\mathcal{P}_V(Z) \wedge V}{\partial_x \wedge \partial_y}, \tag{4.12}$$

from the space $(\mathbb{C}[[x,y]])^2$ of formal vector fields to the space $\mathbb{C}[[x,y]]$ of formal functions. Since the map $Z \to \mathcal{P}_V(Z)$ is non-linear, the map (4.12) cannot be factorized to a map from $\mathbb{C}[[x,y]]$ to $\mathbb{C}[[x,y]]$. But the linear part of $\mathcal{P}_V(\cdot)$ equals $-\mathrm{ad}_V(\cdot)$ and we have the following well the defined linear map

$$\mathcal{L}_V\Omega = -\frac{\mathrm{ad}_V Z \wedge V}{\partial_x \wedge \partial_y}.$$

Standard calculations give the following formula

$$\mathcal{L}_V\Omega = \dot{\Omega} - \operatorname{div} V \cdot \Omega \tag{4.13}$$

where $\dot{\Omega} = V(\Omega) = \partial\Omega/\partial V$ and div denotes the divergence.

We shall prove our theorems using the linear operator $\mathcal{L}_V$ (with different V's) and also the quasi-homogeneous gradation $\widetilde{\deg}$; we eliminate recursively terms of growing quasi-homogeneous degree. If $U = U_1\partial_x + U_2\partial_y$ is the part of V ('transversal' to V) devoted to killing, then we have the *bi-vector homological equation*

$$\mathcal{L}_V\Omega + \Theta = 0, \quad \Theta = \frac{U \wedge V}{\partial_x \wedge \partial_y}. \tag{4.14}$$

Having solved the bi-vector homological equation, we get some function Ω. Having the function Ω, we find the vector field from the equation $Z_1V_2 - Z_2V_1 = \Omega$; assuming $V_1 = y + ax^n + \cdots$, $V_2 = bx^{2n-1} + \cdots$ it is seen that the solution exists (provided that the expansion of Ω begins from terms of $\widetilde{\deg} \geqslant n+2$). The application of $(\mathrm{Ad}_{\exp Z})_* V$ destroys U, leaving only terms of the form $\sigma(x,y)V$. The latter will be reduced using the multiplication $(1-\sigma)V$. The terms non-linear in Z are of higher quasi-homogeneous degree and are eliminated in further steps of the recursive process.

The whole algorithm of reduction is divided into three general stages. Firstly we approximate the homological operator $\mathcal{L}_V$ by $\mathcal{L}_{V_H}$: we determine its kernel in $\mathbb{C}[[x,y]]$ and a subspace complementary to its image. Next, we use the operator $\mathcal{L}_{V_H + x^t\partial_x}$ to Ω's from $\ker \mathcal{L}_{V_H}$ to reduce some additional terms; here $x^t\partial_x$ is the first term not reduced in the previous step. In the third step we use the operator $\mathcal{L}_{V_H + x^t\partial_x + \mathrm{const}\cdot x^u\partial_x}$.

Denote by $\mathcal{X}_{m,i} \approx \mathbb{C}^{m+1}$ the space of polynomials Ω of the form

$$\Omega = \omega_0 x^{mn+i} + \omega_1 x^{(m-1)n+i} y + \cdots + \omega_m x^i y^m,$$

i.e. of quasi-homogeneous degree $\widetilde{\deg}\,\Omega = mn + i$.

LEMMA 4.13.

(a) *The equation $\mathcal{L}_V\Omega = 0$ means that the function Ω is the inverse integrating factor for the vector field V, i.e.* $\operatorname{div}(\Omega^{-1}V) = 0$.

(b) *In the case of generalized resonant saddle $S_{k:-l}$ the operator $\mathcal{L}_{V_H}$ acts between the spaces $\mathcal{X}_{m,i}$ and $\mathcal{X}_{m+1,i-1}$ (of dimensions $m+1$ and $m+2$ respectively) when $i > 0$, and between the spaces $\mathcal{X}_{m,0}$ and $\mathcal{X}_{m,r-1}$ (of the same dimension $m+1$) when $i = 0$.*

(c) *In the case of generalized resonant saddle* $S_{k:-l}$ *we have* $\ker \mathcal{L}_{V_H} = (y-u_1x^r)^{pk+1}\times (y-u_2x^r)^{pl+1}\cdot\mathbb{C}$, $r=n$, *if* $mn+i=(pk+pl+2)r$, *i.e. in* $\mathcal{X}_{pk+pl+2,0}$, $p=1,2,\ldots$ (*here* $y=u_{1,2}z$ *are the equations of separatrices in* (4.10)).

PROOF. The first statement follows from the relation $\operatorname{div}(fV)=V(f)+f\cdot\operatorname{div}V$ and the second statement is rather obvious.

If Ω is the inverse integrating multiplier for V_H then the field $X_H=V_H/\Omega$ is Hamiltonian with quasi-homogeneous Hamilton function H, i.e. V_H has quasi-homogeneous first integral. The unique quasi-homogeneous first integrals take the form

$$H=(y-u_1x^r)^{A\lambda_2}(y-u_2x^r)^{-A\lambda_1},\quad A\in\mathbb{C},$$

where $\lambda_{1,2}$ are the eigenvalues of the linear system (4.10). Then

$$X_H=V_H\cdot\left[(y-u_1x^r)^{1-A\lambda_2}(y-u_2x^r)^{1+A\lambda_1}\right]^{-1}.$$

We see that the only possibility that the function $\Omega=(y-u_1x^r)^{1-A\lambda_2}(y-u_2x^r)^{1+A\lambda_1}$ is a polynomial is just as in the statement (c). ☐

COROLLARY 4.14. *In the case* $i>0$ *we have* $\ker\mathcal{L}_{V_H}|\mathcal{X}_{m,i}=0$ *and* $\operatorname{Im}\mathcal{L}_{V_H}\mathcal{X}_{m,i}$ *is of codimension* 1, *with* $x^{(m+1)r+i-1}\cdot\mathbb{C}$ *as a complementary subspace to* $\operatorname{Im}\mathcal{L}_{V_H}\mathcal{X}_{m,i}$.

In the case $i=0$ *either* $\ker\mathcal{L}_{V_H}|\mathcal{X}_{m,0}=0$ *and* $\operatorname{coker}\mathcal{L}_{V_H}|\mathcal{X}_{m,0}=0$, *or* $\ker\mathcal{L}_{V_H}|\mathcal{X}_{m,0}\neq 0$ *and* $\operatorname{Im}\mathcal{L}_{V_H}|\mathcal{X}_{m,i}$ *is of codimension* 1 *with* $x^{mr+r-1}\cdot\mathbb{C}$ *as a complementary subspace.*

PROOF. The only fact we have to prove is the statement about complementary subspaces. On the one hand, it follows from the Bogdanov–Takens prenormal form; i.e. the reduction to $(y+a(x))\partial_x+bx^{2r-1}\partial_y$ (see [76]).

One can see it also after composing $\mathcal{L}_{V_H}$ with the projection π of $\mathcal{X}_{m,0}$ onto the m-dimensional subspace of polynomials divisible by y. In the monomial basis in the case $i>0$ the operator $\pi\circ\mathcal{L}_{V_H}$ becomes upper-triangular with entries $jr+i\neq 0$ on the diagonal. In the case $i=0$ the left-down minor of dimension m is upper-triangular with entries $jr\neq 0$ on the diagonal. ☐

The polynomials from $\ker\mathcal{L}_{V_H}$, which are described in Lemma 4.18, form inverse integrating factors also for some vector fields which are not quasi-homogeneous.

Consider the function

$$F(x,y)=\frac{1}{(y-u_1x^r)^{pk}(y-u_2x^r)^{pl}}+\nu\cdot\ln\left[\frac{yx^{-r}-u_1}{yx^{-r}-u_2}\right]$$

where ν is a constant. It is the first integral of the vector field

$$V_H+\eta\cdot(y-u_1x^r)^{pk}(y-u_2x^r)^{pl}x^{r-1}E_H \tag{4.15}$$

where $\eta = \nu(u_2 - u_1)/p(k+l)$. The inverse integrating factor is

$$-(pk+pl)^{-1}(y-u_1x^r)^{pk+1}(y-u_2x^r)^{pl+1}, \tag{4.16}$$

the same as in Lemma 4.13(c).

Note that the vector field (4.15) consists of two quasi-homogeneous parts (of degrees $r-1$ and $r-1+rp(k+l)$). The first integral also consists of two quasi-homogeneous parts, of degrees $-(k+l)r$ and 0. Note also that the level curves $F(x,y)=f$, $f \neq \infty$ are non-algebraic, only the curve $F=\infty$ is algebraic.

It turns out that this example presents the only situations (in a sense), where a non-quasi-homogeneous vector field has a polynomial inverse integrating factor.

LEMMA 4.15. *Consider a vector field of the form $V = V_H + V_1$ where V_H has the principal linear part a $(k:-l)$-resonant saddle,* $\gcd(k,l)=1$, *and V_1 is a quasi-homogeneous field of degree $d = \widetilde{\deg} V_1 > r-1$and such that*

$$d+1 \neq 0(r).$$

Then V does not admit any inverse integrating factor of the form

$$\Omega_0 + \Omega_1 + \cdots \tag{4.17}$$

where $\Omega_0 \in \ker \mathcal{L}_{V_H}$, Ω_1 is a quasi-homogeneous polynomial of degree $d-r+1$ and "$\cdots$" means higher degree terms.

Moreover, for the vector field (4.15) *the only possible inverse integrating factor of the form* (4.17) (*with $\Omega_0 \in \ker \mathcal{L}_{V_H}$, and Ω_1 polynomial*) *is the function* (4.16).

PROOF. Following Remark 4.3 we can assume that V_1 of the form $g(x,y)E_H$. Then, after applying the quasi-homogeneous blowing-up $(x,u) \to (x,y) = (x, ux^r)$, we get the Bernoulli equation

$$\frac{\mathrm{d}x}{\mathrm{d}u} = \frac{u+a}{-r(u-u_1)(u-u_2)} x + \frac{\tilde{g}(u)}{(u-u_1)(u-u_2)} x^{d-r+2}$$

with $\tilde{g} = -g(1,u)/r$. This equation has the first integral

$$G = \frac{1}{(y-u_1x^r)^{\alpha}(y-u_2x^r)^{\beta}} + c\int^u \frac{\tilde{g}(u)\,\mathrm{d}u}{(u-u_1)^{\alpha+1}(u-u_2)^{\beta+1}},$$

where $\frac{\alpha}{\beta} = \frac{k}{l}$ and $(\alpha+\beta)r = d+1-r$.

The first integral G corresponds to the inverse integrating factor $M = (y-u_1x^r)^{\alpha+1}(y-u_2x^r)^{\beta+1}$ which cannot be polynomial (by our choice of d). Any other first integral for V is a functions of the primitive integral G. But taking it in the form $F = G^{\gamma}$ we find the inverse integrating factor equal $M \cdot G^{1-\gamma}$, which cannot be written in the form (4.17); it would be equal $\Omega_0 + \Omega_1 + \cdots$ where Ω_1 is not a polynomial. It is also clear that other forms of first integral (like $F = G^{\gamma} + \text{const} \cdot G^{\delta}$) do not lead to a right inverse integrating factor.

The second statement is proved in the same way. $\square$

PROOF OF THEOREM 4.8. By Lemma 4.14 the operator $\mathcal{L}_{V_H}$ applied to the space $\mathcal{X}_{m,i}$ (of quasi-homogeneous Ω's) has non-zero kernel only in one case. It occurs when $mr + i = (pk + pl + 2)r$, i.e. $m = pk + pl + 2$, $i = 0$. This kernel is 1-dimensional and is generated by the polynomial $\Omega_0 = (y - u_1 x^r)^{pk+1}(y - u_2 x^r)^{pl+1}$.

The cokernel of $\mathcal{L}_{V_H} : \mathcal{X}_{m,0} \to \mathcal{X}_{m,r-1}$ is equal 0 for $m \neq pk + pl + 2$ and is 1-dimensional otherwise. Also $\operatorname{coker}(\mathcal{L}_{V_H} : \mathcal{X}_{m,i} \to \mathcal{X}_{m+1,i-1})$ is 1-dimensional for $i > 0$.

If $\operatorname{coker} \mathcal{L}_{V_H} \neq 0$ then we can assume that it is generated by a power $\Theta = x^{\beta}$, $\beta = (pk + pl + 2)r + r - 1$ or $\beta = (m + 1)r + i - 1$ $(0 < i < r)$; it corresponds to the field (via $\Theta = W \wedge V_H / \partial_x \wedge \partial_y$):

$$W = x^{p(k+l)r+r} \partial_x \quad \text{or} \quad W = x^{(m-1)r+i} \partial_x$$

respectively. Instead of $x^{p(k+l)r+r-1}\partial_x$ in the first case we can take $W = H^p x^{r-1} E_H$, where $H = (y - u_1 x^r)^k (y - u_2 x^r)^l$. Therefore application of the homological operator, associated with the leading part of V_H, reduces V to

$$V_H + \sum b_{pj} H^p x^{r-1} E_H + \sum_m \sum_{i=1}^{r-1} a_{m,i} x^{mr+i} \partial_x \tag{4.18}$$

where b_p, $a_{m,i}$ are complex coefficients. We have three possibilities:

- $V = V_H$;
- $V = V_H + a_{m_0,i_0} x^{m_0 r + i_0} \partial_x + \cdots$, $i_0 \neq 0$, $a_{m_0,i_0} = 1$;
- $V = V_H + b_{p_0} H^{p_0} x^{r-1} E_H + \cdots$, $b_{p_0} = 1$.

In the second case dominating is the term $a_t x^t \partial_x$ with $t = m_0 r + i_0 \neq 0(r)$. Then we use the operator $\mathcal{L} = \mathcal{L}_{V_H} + \mathcal{L}_{x^t \partial_x}$ to functions

$$\Omega = \Omega_0 + \Omega_1, \qquad \Omega_0 \in \ker \mathcal{L}_{V_H},$$

to reduce further terms. Here, by Lemma 4.13, $\Omega_0 = \text{const} \cdot H^j (y - u_1 x^r)(y - u_2 x^r)$, $j = 1, 2, \ldots$, and by Lemma 4.15 the corresponding operator defines an isomorphism between suitable spaces, of Ω's and of quasi-homogeneous functions Θ's (the kernel of $\mathcal{L}$ is zero). This gives the point (a) of Theorem 4.8.

In the third case dominating is the term $W_0 = H^{p_0} x^{r-1} E_H$, corresponding to $x^t \partial_x$ with

$$t = p_0 r_0 + r, \qquad r_0 = (k + l)r$$

(see the cases (b) and (c) in Theorem 4.8). Thus we apply the operator $\mathcal{L} = \mathcal{L}_{V_H} + \mathcal{L}_{W_0}$ to functions of the form $\Omega_0 + \Omega_1$, where $\Omega_0 = c H^q (y - u_1 x^r)(y - u_2 x^r)$, $c \in \mathbb{C}$, and Ω_1 is quasi-homogeneous with $\widetilde{\deg}\,\Omega_1 - \widetilde{\deg}\,\Omega_0 = p_0 r_0$. Here $q = 1, 2, \ldots$.

Let $q = p_0$. By the example with the vector field (4.15) we see that the corresponding operator $\mathcal{L}$ has non-zero kernel $(\Omega_0 + \Omega_1) \cdot \mathbb{C} = (H^{p_0}(y - u_1 x^r)(y - u_2 x^r) + 0) \cdot \mathbb{C}$ and the operator is not surjective. This means that the term $x^{t+p_0r_0}\partial_x = x^{r+2p_0r_0}\partial_x$ (corresponding to $H^{2p_0}x^{r-1}E_H$ in (4.18)) cannot be killed.

Let $q \neq p_0$. Then Lemma 4.15 asserts that the corresponding kernel is 0. Therefore all the terms $x^{t+qr_0}\partial_x$ with $q \neq p_0$, corresponding to $H^{p_0+q}x^{r-1}E_H$ in (4.18), can be killed.

In the case (b) form the thesis of Theorem 4.8 we have the situation where all other terms $a_{m,i}x^{mr+i}\partial_x$ in (4.18) are equal zero; (there remain only two terms $b_pH^px^{r-1}E_H$ with $p = p_0, 2p_0$).

If there exists a term $W_1 = a_{j_0}x^{j_0+t}\partial_x$, with j_0 not divisible by r, then we use the operator $\mathcal{L} = \mathcal{L}_{V_H} + \mathcal{L}_{W_0} + \mathcal{L}_{W_1}$ applied to

$$\Omega = \text{const} \cdot H^{p_0}(y - u_1x^r)(y - u_2x^r) + \Omega_1 + \Omega_2 = \Omega_0 + \Omega_2,$$

where $\Omega_0 \in \ker \mathcal{L}_{V_H+W_0}$, in order to kill additional term. Here $\ker \mathcal{L} = 0$ and the operator $\mathcal{L}$ is a surjection. This gives the case (c) of Theorem 4.8.

The remaining part of Theorem 4.8 is standard. □

Other theorems of this subsection are proved along the same lines as in the latter proof.

REMARK 4.16. Let us say some words about the classification of the Bogdanov–Takens singularities with respect to conjugation, not orbital equivalence (see Definition 3.1).

Firstly, A. Baider and J. Sanders [7,8] showed that in the generalized cusp case, i.e. $s < 2r$, the unique (non-orbital) normal form is following

$$V_H + \Bigg(\alpha_{2s-1}x^{2s-1} + \sum_{j>s-1,\ j\neq -2,-1,r-2(s)} \alpha_j x^j\Bigg)\partial_y + \Bigg(\sum_{j\geqslant r-1,\ j\neq -1(s)} \beta_j x^j\Bigg)(x\partial_x + y\partial_y).$$

They obtained also a unique normal form in the generalized saddle case, but it is more complicated (and we do not present it here).

Next, H. Kokubu, H. Oka and D. Wang [44] considered the case of non-resonant generalized saddle S_λ for $r = 2$ and $s = 4$. They reduced the system to so-called first order normal form (which is unique)

$$V_H + \Bigg(\sum_{j\geqslant 4} \alpha_j x^j\Bigg)\partial_y.$$

X. Wang, G. Chen and D. Wang [84] considered the generalized cusp singularity with $s = 3$ and $r \geqslant 2$. They obtained the following unique normal forms

$$V_H + \alpha_3x^3\partial_y + \big(xf(x^3) + x^3g(x^3)\big)y\partial_y, \qquad f(0) \neq 0,$$

and

$$V_H + (\alpha_3x^3 + \alpha_5x^5)\partial_y + \big(x^4f(x^3) + x^3g(x^3)\big)y\partial_y,$$

provided $110f(0) \neq 183\alpha_3 g(0)$.

Also A. Algaba, E. Freire and E. Gamaro [2] studied the problem and obtained some list of (non-orbital) normal forms.

Nevertheless the problem of classification of the nilpotent singularities with respect to formal conjugation is still not finished. Probably the list of formal orbital normal forms given above would be helpful to complete this classification.

4.2. *Analytic Takens prenormal form*

The main result of this subsection solves the old problem of analyticity of the prenormal form obtained by F. Takens.

THEOREM 4.17. (See [76,49].) *The Takens prenormal form* $\dot{x} = y + a(x)$, $\dot{y} = b(x)$ *of the singularity* $\dot{x} = y + \cdots$, $\dot{y} = x^{s-1}\cdots$, $s < \infty$, *obtained in the proof of Theorem* 4.2 *is analytic. This implies also that the Bogdanov–Takens prenormal form* (4.6) *is analytic.*

REMARK 4.18. The theorem holds true also when $s = 2$. Then the singular point is a saddle with the $1 : -1$ resonance and Theorem 4.18 gives an analytic prenormal form for such saddles.

The theorem holds true also when $s = \infty$, i.e. in the case of generalized saddle–node of infinite codimension. Here the vector field has formal first integral, $H(x, y) = y$ in the formal Takens coordinates. It is easy to get analyticity of this integral (see [75] and [86]).

Theorem 4.18 has two proofs. One proof from [76] relies on estimates of the successive approximations from the proof of the Takens theorem. It seems that it can be generalized to the cases of higher dimensional nilpotent singularities of vector fields. Another proof from [49] is geometrical and uses a preparation theorem for holomorphic foliations. Below we present the latter proof with details.

PROOF OF THEOREM 4.17. Following [49] we shall prove even more general statement: *if the linear part of a holomorphic germ* $V = V_0$ *of vector field at* $0 \in \mathbb{C}^2$ *is not radial, i.e. not of the form* $\mu(x\partial_x + y\partial_y)$, *then there exist local analytic coordinates* x, y *in which* $V_0 = (y + a(x))\partial_x + b(x)\partial_y$ *for some analytic germs* a, b.

PROOF. By the non-radiality assumption it is easy to find linear coordinates in which $dV_0(0) = \left(\begin{smallmatrix} 0 & 1 \\ \alpha & \beta \end{smallmatrix}\right)$. Moreover, we can assume that $V_0 = f\partial_x + g\partial_y$ with $f(0, y) = y$.

In $\mathbb{C} \times \bar{\mathbb{C}}$ we consider the vertical line $L = 0 \times \bar{\mathbb{C}}$ together with its covering defined by $\Delta_0 = 0 \times \{|y| < r\}$ and $\Delta_\infty = 0 \times \{|y| > r/2\}$. Also we denote by $C = \Delta_0 \cap \Delta_\infty$ the intersection *corona* (see [49]).

If $r > 0$ is small enough then the vector field V is well defined on a neighborhood of the closure of Δ_0 and is transverse to this disc outside $y = 0$. By the flow-box theorem there exists a unique germ of a diffeomorphism $\Phi : (\mathbb{C}^2, C) \to (\mathbb{C}^2, C)$, such that $\Phi(0, y) = (0, y)$ and Φ conjugates V_0 to the horizontal vector field $V_\infty = y\partial_x$.

Let us glue the germs of holomorphic surfaces $(\mathbb{C} \times \bar{\mathbb{C}}, \Delta_0)$ and $(\mathbb{C} \times \bar{\mathbb{C}}, \Delta_\infty)$ along the corona by means of the diffeomorphism Φ. We obtain a germ of smooth complex surface

M along a rational curve L with a meromorphic vector field V. Since the x-component of V_0 at C agrees with $y\partial_x$, it follows that the Jacobian of the gluing map Φ takes the form $\mathrm{d}\Phi(0,y)=\left(\begin{smallmatrix}1&0\\ *&1\end{smallmatrix}\right)$ and the embedded rational curve L has zero self-intersection.

V. Savelev [69] proved that in this situation the line L is a regular fiber of some trivial fibration on M, i.e. there exist global coordinates $(x,y):(M,L)\to(\mathbb{C},0)\times\bar{\mathbb{C}}$ sending L onto $\{x=0\}$.

The vector field V has a simple pole along a curve, which we may assume still given by $\{y=\infty\}$. Since $\partial_y=-(1/y)^2\partial_{1/y}$ has second order zero at $\{y=\infty\}$ we can write

$$V=\big[a_0(x)+a_1(x)y\big]\partial_x+\big[b_0(x)+b_1(x)y+b_2(x)y^2+b_3(x)y^3\big]\partial_y. \tag{4.19}$$

The vector field V has exactly one isolated zero, say at $x=y=0$.

The tangency divisor Σ between the foliation $\mathcal{F}$ induced by V and the fibration $\{x=\text{const}\}$ is a smooth curve $\Sigma=\{a_0(x)+a_1(x)y=0\}$ intersecting transversally the line $L=\{x=0\}$. Indeed, the Jacobian of the change from the first chart to the global coordinates at the singular point fixes the y direction, so that the linear part of the vector field takes the form $\left(\begin{smallmatrix}k&l\\ m&n\end{smallmatrix}\right)$, $l\neq0$. Therefore $a_1(0)\neq0$ in (4.19).

We can further simplify (4.19). Note that the admissible changes of the coordinates in $(\mathbb{C},0)\times\widehat{\mathbb{C}}$ are of the form $(x,y)\to(\phi(x),\frac{\alpha(x)y+\beta(x)}{\gamma(x)y+\delta(x)})$.

By straightening one leaf of the foliation $\mathcal{F}$ (i.e. through $(0,\infty)$) we can assume that $\{y=\infty\}$ is a leaf of $\mathcal{F}$. Thus $b_3(x)\equiv0$ in (4.19).

Next, the change $y\to(\mathrm{e}^{\int b_2/a_1})y$ reduces $b_2(x)$ in (4.19) to zero.

By a change of the x coordinate we can assume that $a_1(x)\equiv1$. The change $y\to y-a_1(x)$ straightens the curve Σ to $\{y=0\}$. Hence we can put $a_0(x)\equiv0$.

We obtain the vector field $V=y\partial_x+[b_0(x)+b_1(x)y]\partial_y$ which is equivalent to the Takens field (4.2) (compare Remark 4.3). ☐

Next to consider is the question of analyticity of the final orbital normal forms given in Theorems 4.6, 4.8–4.10 and 4.11.

But, as we noticed in Remark 4.18, in the case $s=2$ we have a $1:-1$ resonant saddle, where the formal normal form is not analytic (see Subsection 3.5). So there is no hope to expect that for $s\geqslant3$ one could prove analyticity of the final formal orbital normal form.

This question for the standard cusp singularity, i.e. $s=3$ and $r=2$, was firstly considered in [19]. There the authors calculated numerically some initial terms in the Loray's normal form

$$V_H+\big(x+A_0(H)\big)E_H,\quad H=y^2/2-x^3/3, \tag{4.20}$$

i.e. the coefficients a_{3j} in the series $A_0(H)=a_3H+a_6H^2+\cdots$ (compare Remark 4.7). These coefficients turned out growing like the Gevrey series of order 2, i.e.

$$a_{3j}\sim CA^{6j}\Gamma(\alpha+6j) \tag{4.21}$$

(compare Definition 3.48). Approximate values of the constants in (4.21) are $C\approx7678015$, $A\approx0.1803$ and $\alpha\approx-37.27$.

In [20] M. Canalis-Durand and R. Schäfke gave a rigorous proof of the divergence of the form (4.21). They proved the following

THEOREM 4.19. (See [20].) *For an analytic vector field*

$$V = V_H + \Delta(x, y)E_H, \tag{4.22}$$

with $H = \frac{1}{2}y^2 - \frac{1}{3}x^3$ *and analytic function* $\Delta = x + \cdots$, *there exists a unique formal series* $\Phi(x, y) = 1 + \cdots$ *such that the formal transformation* $(x, y) \to (x\Phi^2, y\Phi^3)$ *and division by* Φ *reduces* V *to the form* (4.20) *such that* $|a_{3j}| < CA^{6j}(6j)!$ *for some constants* C, A. *The series* $A_0(z^6)$ *is* 1*-summable if* $\arg z$ *is not congruent to* $\pi/6$ *modulo* $\pi/3$.

Moreover, for typical germs Δ, *i.e. outside a hypersurface in the space of such germs, the series* $A_0(H)$ *is divergent.*

REMARK 4.20. (About the proof of Theorem 4.19.) A general vector field which is tangent to the cusp $\{H(x, y) = 0\}$ is reduced to the form (4.22), after division by a non-zero function. The reduction used in the first statement of Theorem 4.19 comes from the Loray's algorithm [47].

The proof of Theorem 4.19 given in [20] is very short. However we cannot present it here with details. It uses so-called Borel transform $\mathcal{B}$ (see Definition 3.47), or its generalization to two variables, and its inverse (so-called Laplace transform).

The main idea of the proof is to solve the suitable conjugation equation in the 'Borel plane'. It turns out that the Borel transform of the series $\Phi(x, y)$ is a meromorphic function with poles. Namely, these poles are responsible for the divergence of the series Φ.

Finally, we note that in the sequent paper [21] Canalis-Durand and Schäfke proved an analogous result when $\Delta(x, y)$ in (4.22) starts from higher order terms and in (4.20) we have $V_H + (A_0(H) + xA_1(H))E_H$. Then the order of summability of the series A_0, A_1 depends on their first non-zero terms. ☐

REMARK 4.21. Recall that in Subsection 3.5 we defined certain holonomy map associated with a saddle or a saddle–node singularity of planar vector field. This holonomy map (or monodromy map) was associated with a loop in a separatrix of the singularity. Also with the nilpotent singularities one can associate certain 1-dimensional object. It is the so-called hidden holonomy group.

Namely, in the resolution π of the singularity one obtains some special divisor E with 3 singular points p_0, p_1, p_2 of the induced foliation $\mathcal{G} = \pi^*\mathcal{F}$. (In the case of ordinary cusp it is the divisor E_3 from Fig. 4.) The *hidden holonomy group* G is formed by the monodromy maps $\Delta_\gamma : D \to D$, of a small disc $D \approx (\mathbb{C}, 0)$ transversal to a point $q \in E^* = E \setminus \{p_0, p_1, p_2\}$. The loops γ run through the fundamental group $\pi_1(E^*, q)$.

In many situations the (analytic) orbital classification of the vector fields V is reduced to the classification of finitely generated subgroups $G \subset \mathrm{Diff}(\mathbb{C}, 0)$ (obeying some natural relations) with respect to conjugation in the group $\mathrm{Diff}(\mathbb{C}, 0)$.

This is an interesting and advanced theory . The interested reader is referred to [23,31, 48,50,61,76,75] and [89] for details.

References

[1] L.V. Ahlfors, *Lectures on Quasi-Conformal Mappings*, Van Nostrand–Reinhold, Princeton (1966).

[2] A. Algaba, E. Freire and E. Gamaro, *Computing simplest normal forms for the Bogdanov–Takens singularity*, Qualit. Theory Dynam. Systems **3** (2) (2002), 377–435.

[3] V.I. Arnold, *On local problems in analysis*, Vestnik Mosk. Univ. Ser. I. Mat. Mekh. **25** (2) (1970), 52–5 (in Russian).

[4] V.I. Arnold, *Geometrical Methods in the Theory of Differential Equations*, Springer–Verlag, Berlin–Heidelberg–New York (1983) (in Russian: Nauka, Moscow (1978)).

[5] V.I. Arnold and Yu.S. Il'yashenko, *Ordinary Differential Equations*, Ordinary Differential Equations and Smooth Dynamical Systems, Springer-Verlag, Berlin–Heidelberg–New York (1997), 1–148 (in Russian: Fundamental Directions **1**, VINITI, Moscow (1985), 1–146).

[6] D.G. Babbitt and V.S. Varadarajan, *Local moduli for meromorphic differential equations*, Asterisque **169–170** (1989).

[7] A. Baider and J. Sanders, *Unique normal forms: the nilpotent Hamiltonian case*, J. Differential Equations **92** (2) (1991), 282–304.

[8] A. Baider and J. Sanders, *Further reduction of the Bogdanov–Takens normal form*, J. Differential Equations **99** (2) (1992), 205–244.

[9] H. Bateman and A. Erdelyi, *Higher Transcendental Functions*, Vol. 1, McGraw–Hill, New York (1953).

[10] I. Bendixson, *Sur les courbes définies par des équations différentielles*, Acta Math. **25** (1901), 1–88.

[11] G.D. Birkhoff, *The generalized Riemann problem for linear differential equations and the allied problem for linear difference and q-difference equations*, Proc. Amer. Acad. Arts Sci. **49** (1913), 531–568.

[12] G.D. Birkhoff, *Déformations analytiques et fonctions auto-équivalentes*, Annales **9** (1939), 51–122.

[13] R.I. Bogdanov, *Local orbital normal forms of vector fields on a plane*, Trans. Petrovski Sem. **5** (1979), 50–85 (in Russian).

[14] B.L.J. Braaksma, *Multi-summability and Stokes multipliers of linear meromorphic differential equations*, J. Differential Equations **92** (1991), 54–75.

[15] C.A. Briot et J.C. Bouquet, *Propriétés des fonctions définies par des équations différentielles*, J. Éc. Polytech. **36** (1856), 133–198.

[16] A.D. Briuno, *Analytic form of differential equations*, Trans. of Moscow Math. Soc. for the year 1971, Amer. Math. Soc., Providence, RI (1973) (in Russian: **25** (1971), 119–262); *II*, Trans. of Moscow Math. Soc. for the year 1972, Amer. Math. Soc., Providence, RI (1974) (in Russian: **26** (1972), 199–239).

[17] N.N. Brushlinskaya, *The finiteness theorem for families of vector fields in a neighborhood of singular point of the Poincaré type*, Funct. Anal. Applic. **5** (1971), 10–15 (in Russian).

[18] C. Camacho and P. Sad, *Topological classification and bifurcations of flows with resonances in* $\mathbb{C}^2$, Invent. Math. **67** (1982), 447–472.

[19] M. Canalis-Durand, F. Michel and M. Teysseyre, *Algorithms for formal reduction of vector field singularities*, J. Dynam. Contr. Syst. **7** (1) (2001), 101–125.

[20] M. Canalis-Durand and R. Shäfke, *On the normal form of a system of differential equations with nilpotent linear part*, C. R. Acad. Sci. Paris, Ser. I **336** (2003), 129–134.

[21] M. Canalis-Durand and R. Shäfke, *Divergence and summability of normal forms of differential equations with nilpotent linear part*, Ann. Fac. Sci. Touluse Math. (6) **13** (4) (2004), 493–513.

[22] F. Cano, *Reduction of singularities of foliations and applications*, Singularities Symposium – Łojasiewicz 70, Banach Center Publ., Vol. 44, PWN, Warszawa (1998), 51–70.

[23] D. Cerveau and R. Moussu, *Groupes d'automorphismes de (C,0) et équations différentielles* $y\,\mathrm{d}y + \cdots = 0$, Bull. Soc. Mat. France **116** (1988), 459–488.

[24] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill Book C., New York (1955).

[25] H. Cremer, *Über die Häufigkeit der Nichtzentren*, Math. Ann. **115** (1938), 573–580.

[26] H. Dulac, *Solutions d'un système d'équations différentielles dans le voisinage des valeurs singuliéres*, Bull. Soc. Math. France **40** (1912), 1–6.

[27] F. Dumortier, *Singularities of vector fields in the plane*, J. Differential Equations **23** (1977), 53–106.

[28] J. Ecalle, *Théorie itérative. Introduction à la théorie des invariants holomorphes*, J. Math. Pures Appl. **54** (1975), 183–258.

[29] J. Ecalle, *Sur les fonctions résurgentes*, Vols. I, II, III, Publ. Math. d'Orsay, Orsay (1981).

[30] P.M. Elizarov, *Tangents to moduli maps*, Nonlinear Stokes Phenomena, Yu. Il'yashenko, ed., Adv. in Sov. Math., Vol. 14, Amer. Math. Soc., Providence, RI (1993), 107–137.

[31] P.M. Elizarov, Yu.S. Il'yashenko, A.A. Shcherbakov and S.M. Voronin, *Finitely generated groups of germs of one-dimensional conformal mappings and invariants of complex singular points of analytic foliations of the complex plane*, Nonlinear Stokes phenomena, Yu. Il'yashenko, ed., Adv. in Sov. Math., Vol. 14, Amer. Math. Soc., Providence, RI (1993), 57–105.

[32] A. van den Essen, *Reduction of singularities in differential equations* $A\,\mathrm{d}y = B\,\mathrm{d}x$, Équations Différentielles et Systèmes de Pfaff dans le Champ Complexe, R. Gérard et J.-P. Ramis, eds., Lecture Notes in Math., Vol. 712, Springer-Verlag, Berlin–Heidelberg–New York (1979), 44–59.

[33] L. Euler, *De seriebus divergentibus*, Leonhardi Euleri Opera Omnia. I, Vol. 14, Teubner, Leipzig–Berlin (1925), 601–602.

[34] L. Euler, *Specimen transformationinis singularis sewrierum*, Nova Acta Petropolitana **XII** (1801), 58–78.

[35] E. Fabry, *Sur les intégrales des équations différentielles à coefficients rationnels*, Thése, Paris (1885).

[36] P. Godement, *Topologie Algébrique et Théorie des Faisceaux*, Hermann, Paris (1958).

[37] J.M.A. Heading, *An Introduction to Phase-Integral Methods*, J. Wiley & Sons, New York (1962).

[38] A.A. Grintchy and S.M. Voronin, *An analytic classification of saddle resonant singular points of holomorpfic vector fields in the complex plane*, J. Dynam. Contr. Syst. **2** (1996), 21–53.

[39] M. Hukuhara, *Sur les points singuliers des équations différentielles linéaires. II*, J. Fac. Sci. Hokkaido Univ. **5** (1937), 123–166.

[40] M. Hukuhara, T. Kimura and T. Matuda, *Équations différentielles ordinaires du premier ordre dans le champ complexe*, Math. Soc. Japan, Tokyo (1961).

[41] Yu.S. Il'yashenko, *Algebraic unsolvability and almost algebraic solvability of the problem for center–focus*, Funkc. Anal. Prilozh. **6** (3) (1972), 30–37 (in Russian).

[42] Yu.S. Il'yashenko, *Nonlinear Stokes phenomena*, Nonlinear Stokes Phenomena, Yu. Il'yashenko, ed., Adv. in Sov. Math., Vol. 14, Amer. Math. Soc., Providence, RI (1993), 1–53.

[43] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vols. 1, 2, Interscience Publishers, New York (1963, 1969).

[44] H. Kokubu, H. Oka and D. Wang, *Linear grading function and further reduction of normal form*, J. Differential Equations **132** (1996), 293–318, 450–459.

[45] L. Leau, *Étude sur les équations fonctionelles à une ou plusieurs variables*, Ann. Fac. Sci. Toulouse **11** (1897), 1–110.

[46] A.H.M. Levelt, *Jordan decomposition for a class of singular differential operators*, Arkiv Math. **13** (1975), 1–27.

[47] F. Loray, *Réduction formelle des singularités cuspidales de champs de vecteurs analytiques*, J. Differential Equations **158** (1999), 152–173.

[48] F. Loray, *5 leçons sur la structure transverse d'une singularité de feuilletage holomorphe en dimension 2 complexe*, Monogr. Red. TMR Eur. Sing. Ec. Diff. Fol., Vol. 1 (1999), 1–92.

[49] F. Loray, *A preparation theorem for codimension-one foliations*, Ann. Math. (2) **163** (2006), 709–722.

[50] F. Loray and R. Meziani, *Classification des certaines feuilletages associes à un cusp*, Bol. Soc. Bras. Mat. **25** (1994), 93–106.

[51] A.M. Lyapunov, *Stability of Motion*, Academic Press, New York (1966) (in Russian: *General Problem of Stability of Motion*, Gostekhizdat, Moskva (1950)).

[52] B. Malgrange, *Remarques sur les équations différentielles à points singulier irrégulier*, Équations différentielles et systèmes de Pfaff dans le champ complexe, R. Gérard et J.-P. Ramis, eds., Lecture Notes in Math., Vol. 712, Springer-Verlag, Berlin–Heidelberg–New York (1979), 77–86.

[53] B. Malgrange, *Introduction aux travaux de J. Ecalle*, Enseign. Math. **31** (1985), 261–282.

[54] J. Martinet and J.-P. Ramis, *Problèmes de modules pour des équations différentielles nonlinéaires du premier ordre*, Publ. Math. IHES **55** (1982), 63–164.

[55] J. Martinet and J.-P. Ramis, *Classification analytique des équations différentielles nonlinéaires resonantes du premier ordre*, Ann. Sci. École Norm. Sup. **16** (1983), 571–621.

[56] J.F. Mattei and R. Moussu, *Holonomie et intégrales premiéres*, Ann. Sci. École Norm. Sup. **13** (1988), 469–523.

[57] N.B. Medvedeva, *The first focus number of a complex monodromic singular point*, Proc. of Petrovski Seminar **13** (1988), 106–122 (in Russian).

[58] N.B. Medvedeva, *On the analytic solvability of the problem of distinguishing between a center and a focus*, Dokl. Akad. Nauk **394** (2004), 735–738 (in Russian).

[59] N.B. Medvedeva, *On the analytic solvability of the problem of distinguishing between center and focus*, Proc. Steklov Inst. Math. **254** (2006), 7–93 (in Russian: Trudy Mat. Inst. Steklova **254** (2006), 11–100).

[60] Yu.I. Meshcheryakova and S.M. Voronin, *Analytic classification of generic degenerate singular points of germs of holomorphic vector fields on the complex plane*, Izv. VUZ Mat. **1** (2002), 13–16 (in Russian); Translation in: *Russ. Math. (Izv. VUZ)* **46** (2002), 11–14.

[61] R. Meziani, *Classification analytique d'équations différentielles* $y\,\mathrm{d}y + \cdots = 0$ *et espace de modules*, Bol. Soc. Bras. Mat. **21** (1996), 23–53.

[62] R. Moussu, *Symmérie et forme normale des centers et foyers dégénérées*, Ergodic Theory Dynam. Systems **2** (1982), 241–251.

[63] R. Moussu, *Une démonstration géometrique d'un théoreme de Lyapunov–Poincaré*, Bifurcation, Théorie Ergodique et Applications, Dijon 1981, Asterisque **98–99** (1982), 216–223.

[64] A. Newlander and L. Nirenberg, *Complex analytic coordinates in almost complex manifolds*, Ann. Math. **65** (1954), 391–404.

[65] R. Perez-Marco, *Solution complète au problème de Siegel de linéarization d'une application holomorphe au voisinage d'un point fixe. (D'aprés J.-C. Yoccoz)*, Sém. Bourbaki, Fév. 1992, Asterisque **206** (1992), 273–310.

[66] R. Pérez-Marco and J.-C. Yoccoz, *Germes de feuilletages holomorphes à holonomie prescrite*, Asterisque **222** (1994), 345–371.

[67] H. Poincaré, *Mémoire sur les courbes définies par une équation differentielle*, Oeuvres de Henri Poincaré, Vol. 1, Gauthier–Villars, Paris (1951).

[68] A.P. Sadovskii, *On the problem of distinguishing center and focus with nonzero linear part*, Differential Equations **12** (1976), 1237–1246 (in Russian).

[69] V.I. Savelev, *Zero-type imbedding of a sphere into complex surface*, Moscow Univ. Math. Bull. **37** (4) (1982), 34–39 (in Russian: Vestn. Mosk. Univ. Ser. I Mat. Mekh. **37** (4) (1982), 28–32, 85).

[70] A. Seidenberg, *Reduction of singularities of differential equation* $A\,\mathrm{d}y = B\,\mathrm{d}x$, Amer. J. Math. **90** (1968), 248–269.

[71] A.A. Shcherbakov, *Topological classification of germs of conformal mappings with an identical linear part*, Vestnik Mosk. Univ. Ser. I Mat. Mekh. **37** (3) (1982), 52–57 (in Russian).

[72] Y. Sibuya, *Simplification of a system of linear ordinary differential equations about a singular point*, Funkcial. Ekvac. **4** (1962), 29–56.

[73] Y. Sibuya, *Stokes phenomena*, Bull. Amer. Math. Soc. **83** (1977), 1075–1077.

[74] C.L. Siegel, *Iterations of analytic functions*, Ann. Math. **43** (1942), 607–612.

[75] E. Stróżyna, *The analytic and formal normal forms for the nilpotent singularity. The generalized saddle–node case*, Bull. Sci. Math. **126** (2002), 555–579.

[76] E. Stróżyna and H. Żołądek, *The analytic and formal normal forms for the nilpotent singularity*, J. Differential Equations **179** (2002), 479–537 .

[77] E. Stróżyna and H. Żołądek, *Orbital formal normal forms for general Bogdanov–Takens singularity*, J. Differential Equations **193** (2003), 239–259.

[78] F. Takens, *Singularities of vector fields*, Publ. Math. IHES **43** (1974), 47–100.

[79] L. Teyssier, *Analytical classification of addle–node vector fields*, J. Dynam. Control. Syst. **10** (2004), 577–605.

[80] L. Teyssier, *Examples of non-cunjugated holomorphic vector fields and foliations*, J. Differential Equations **205** (2004), 390–407.

[81] H. Turrittin, *Convergent solutions of ordinary differential equations in the neighborhood of an irregular singular point*, Acta Math. **93** (1955), 27–66.

[82] V.S. Varadarajan, *Linear meromorphic differential equations: A modern point of view*, Bull. Amer. Math. Soc. **33** (1996), 1–42.

[83] S.M. Voronin, *Analytic classification of germs of conformal maps* $(\mathbb{C}, 0) \to (\mathbb{C}, 0)$ *with identical linear part*, Funct. Anal. Appl. **15** (1) (1981), 1–17 (in Russian).

[84] X. Wang, G. Chen and D. Wang, *Unique normal forms for the Bogdanov–Takens singularity in a special case*, C. R. Acad. Sci. Paris, Sér. I **332** (2001), 551–555.
[85] W. Wasow, *Asymptotic Expansions for Ordinary Differential Equations*, J. Wiley & Sons, New York (1965).
[86] X. Gong, *Integrable analytic vector field with a nilpotent linear part*, Ann. Inst. Fourier **45** (1995), 1449–1470.
[87] J.-C. Yoccoz, *Théorème de Siegel, nombres de Briuno et polynômes quadratiques*, Asterisque **231** (1995), 3–88.
[88] H. Żołądek, *The extended monodromy group and Liouvillian first integrals*, J. Dynam. Contr. Syst. **4** (1998), 1–28.
[89] H. Żołądek, *The Monodromy Group* , Monografie Matematyczne, Vol. 67, Birkhäuser, Basel (2006).

This page intentionally left blank

Author Index

Roman numbers refer to pages on which the author (or his/her work) is mentioned. Italic numbers refer to reference pages. Numbers between brackets are the reference numbers. No distinction is made between the first author and co-author(s).

Ablowitz, M. 182, *262* [1]
Abramov, A.A. 567, *588* [1]
Agarwal, R.P. 444, *588* [2]
Ahlfors, L.V. 667, 668, *684* [1]
Akhmedov, K.T. 567, *588* [3]
Akhmerov, R.R. 50, 51, *126* [1]
Akhmet, M. 587, *588* [4]
Alber, M. 141, 243, *263* [2]; *263* [3]
Alexander, G.C. 103, *130* [109]
Alexander, J.C. 8, 65, *126* [2]
Algaba, A. 681, *684* [2]
Alonso, A. 138, *263* [4]
Amerio, L. 143, *263* [5]
Andres, J. 166, *263* [6]
Anosov, D. 169, *263* [7]
Antonian, S. 7, *126* [6]
Antonyan, S.A. 7, 26, 52, *126* [3]; *126* [4]; *126* [5]; *126* [7]
Arnold, L. 140, 146, 147, 204, *263* [8]; *263* [9]
Arnold, V.I. 4, 65, *126* [8]; 601, 623, 629, *684* [3]; *684* [4]; *684* [5]
Arzelà, C. 473, *588* [5]
Ascher, U.M. 567, *588* [6]
Atkinson, F.V. 184, *263* [10]
Augustynowicz, A. 586, *588* [7]
Aulaskari, R. 314, *356* [1]

Babbitt, D.G. 608, 609, 611, *684* [6]
Baider, A. 680, *684* [7]; *684* [8]
Baĭnov, D.D. 587, *588* [8]; *588* [9]; *592* [96]
Balanov, Z. 7, 8, 10, 23–25, 27, 28, 36, 52, 53, 66, 75, 84, 98, 100–102, 104, 109, 111, 120, *126* [9]; *126* [10]; *126* [11]; *126* [12]; *126* [13]; *126* [14]; *126* [15]; *126* [16]; *126* [17]; *126* [18]; *126* [19]; *126* [20]; *126* [21]; *129* [98]; *129* [99]
Bank, S. 275, 284, 290, 293–295, 300, 302, 303, 306, 327, 331, 332, 343, 344, 346, 347, 349, 350, *356* [2]; *356* [3]; *356* [4]; *357* [5]; *357* [6]; *357* [7]; *357* [8]; *357* [9]; *357* [10]; *357* [11]; *357* [12]; *357* [13]; *357* [14]; *357* [15]; *357* [16]; *357* [17]; *357* [18]
Barsegian, G. 337, 347, *357* [19]
Bartsch, T. 7, *127* [22]
Barut, A.O. 22, *127* [23]
Bateman, H. 627, *684* [9]
Battelli, F. 368, *438* [1]
Beals, R. 181, 182, *263* [11]
Bebutov, M. 136, *263* [12]
Beesack, P. 310, 311, *357* [20]; *357* [21]
Bendixson, I. 619, *684* [10]
Bersani, A. 166, *263* [6]
Beyn, W.-J. 368, *439* [2]
Bhattacharyya, T. 567, *588* [10]
Bieberbach, L. 276, 326, *357* [22]; *357* [23]
Binding, P.A. 567, *588* [10]
Birkhoff, G.D. 367, *439* [3]; 608, 637, *684* [11]; *684* [12]
Bjerklöv, K. 139, 142, 173, 178–180, *263* [13]; *263* [14]; *263* [15]; *263* [16]; *263* [17]
Blanchard, F. 179, *263* [18]
Bocharov, G. 112, 120, *128* [70]
Bochi, J. 142, 171, 175, 255, *263* [19]
Bochner, S. 139, *263* [20]
Bogdanov, R.I. 672, 675, *684* [13]
Bohr, H. 143, 165, *263* [21]
Borisovich, Yu.G. 7, *127* [24]
Borsuk, K. 7, *127* [25]
Bouquet, J.C. 632, *684* [15]
Bourgain, J. 142, 175, *263* [23]
Bowen, R. 180, *263* [22]
Braaksma, B.L.J. 612, *684* [14]

Bredon, G.E. 11, 20, 22, 23, 28, *127* [26]
Briot, C.A. 632, *684* [15]
Briuno, A.D. 658, *684* [16]
Bröcker, T. 13, 15, 17, 22, 52, *127* [27]
Brodsky, S. 7, 52, *126* [9]
Bronstein, I. 161, *263* [24]
Brushlinskaya, N.N. 631, *684* [17]

Camacho, C. 634, *684* [18]
Cameron, R. 158, *263* [25]
Canalis-Durand, M. 674, 682, 683, *684* [19]; *684* [20]; *684* [21]
Cano, F. 619, *684* [22]
Cao, T. 306, 318, 319, *357* [24]; *357* [25]; *357* [26]
Cerveau, D. 683, *684* [23]
Cesari, L. 585, *588* [11]; *588* [12]
Chan, D. 4, 75, *127* [28]
Chen, G. 680, *687* [84]
Chen, Z. 306, 319, *357* [27]; *357* [28]; *358* [49]
Chevalley, C. 22, 52, *127* [29]
Chiang, Y.-M. 294–296, 347, *357* [29]; *357* [30]; *358* [31]; *358* [32]
Chornyĭ, V.Z. 586, *589* [13]
Chossat, P. 4, 52, 75, *127* [30]; *127* [31]; *127* [32]
Chow, S.-N. 4, 65, 84, 95, 98, 120, *127* [33]; *127* [34]; *127* [35]; *127* [36]
Chulaevsky, V. 142, *263* [26]
Chyzhykov, I. 317, *358* [33]
Cima, J. 309, *358* [34]
Coddington, E. 184, *264* [32]
Coddington, E.A. 599, 602, *684* [24]
Colonius, F. 136, *264* [33]
Conley, C. 165, *263* [27]
Conner, P.E. 7, 23, *127* [37]
Coomes, B.A. 368, 425, 438, *439* [4]
Coppel, W.A. 142, 147, 171, 212, 213, 250, 251, *263* [28]; *263* [29]; 375, 376, 436, *439* [5]
Craig, W. 182, 183, *264* [34]
Crandall, M.G. 419, *439* [6]
Cremer, H. 658, *684* [25]

Damanik, D. 142, *263* [15]; *263* [30]
Dancer, E.N. 4, 7, 8, 52, 66, *127* [38]; *127* [39]; *127* [40]; *127* [41]; *127* [42]
De Concini, C. 183, 212, *264* [31]
de Oliveira, J.C.F. 84, 120, *130* [125]
Demay, Y. 75, *127* [30]
Deng, B. 368, *439* [7]
Devaney, R.L. 407, *439* [8]
Dieci, L. 368, *439* [9]
Dikareva, L.Yu. 585, *589* [14]; *590* [57]
Dinh Cong Nguyen 140, *263* [9]
Donal O'Regan 444, *588* [2]
Drasin, D. 294, *358* [35]
Dubrovin, B.A. 22, 36, 40, *127* [44]; 182, 183, 242, *264* [35]
Dulac, H. 629, *684* [26]
Dumortier, F. 619, *684* [27]
Dunford, N. 207, *264* [36]
Duren, P. 218, 262, *264* [37]; 313, *358* [36]
Dylawerski, G. 6–8, 53, 66, *127* [45]; *127* [46]

Ecalle, J. 636, 652, *684* [28]; *685* [29]
Eilenberg, S. 52, *127* [47]
Eliasson, L.H. 142, 169, 183, *264* [38]; *264* [39]
Elizarov, P.M. 653, 654, 665, 683, *685* [30]; *685* [31]
Ellis, R. 158, 160, 168, *264* [40]; *264* [41]; *264* [42]
Engelking, R. 11, *127* [48]
Erbe, L.H. 53, 66, 85, *128* [49]; *128* [50]
Erdelyi, A. 627, *684* [9]
Eremenko, A. 289, 326, 327, 343, 344, 354, 355, *358* [37]; *358* [38]; *358* [39]; *358* [40]; *358* [41]
Euler, L. 606, 643, *685* [33]; *685* [34]
Evhuta, N.A. 585, *589* [15]; *589* [16]

Fabbri, R. 142, 154, 255–258, 261, 262, *264* [43]; *264* [44]; *264* [45]; *264* [46]; *264* [47]; *264* [48]; *264* [49]; *264* [50]
Fabry, E. 603, *685* [35]
Farkas, M. 445, 446, 450, 452, 471–473, 475, 477, 480, *589* [17]
Farzamirad, M. 52, 53, 75, 84, 120, *126* [10]; *126* [11]
Fathi, A. 169, *264* [51]
Favard, J. 161, 165, *264* [52]
Fay, J.D. 226, 230, 231, 233, 236, 237, *264* [53]
Fečkan, M. 509, 567, *589* [18]; *589* [19]; *589* [20]; *589* [21]; *589* [22]; *589* [23]
Fedorov, Y. 141, 243, *263* [2]; *263* [3]
Fiedler, B. 75, *128* [51]; *128* [52]; *128* [53]
Fink, A. 143, 160, *264* [54]
Floyd, E.E. 7, 23, *127* [37]
Fokas, A. 335, *358* [42]
Fomenko, A.T. 22, 23, 36, 40, *127* [44]; *128* [54]
Fomenko, T.N. 7, *127* [24]; *128* [55]
Fowler, K. 322, *358* [43]
Freedman, H.I. 120, *128* [56]
Frei, M. 284, 296, *358* [44]
Freire, E. 681, *684* [2]
Fuks, D.B. 23, *128* [54]
Fuller, F.B. 4, 8, 52, *128* [57]
Fulton, W. 22, *128* [58]
Furstenberg, H. 145, 167, 168, *264* [55]

Gackstatter, F. 330, 353, *358* [45]; *358* [46]
Gaines, R.E. 567, 585, *589* [24]
Gamaro, E. 681, *684* [2]
Gao, L. 355, *358* [47]; *358* [48]
Gao, S. 306, *357* [27]; *358* [49]
Gao, S.-A. 296, *357* [30]
Gardner, C. 181, *264* [56]
Gęba, K. 6–8, 27, 52, 53, 66, *127* [46]; *128* [49]; *128* [59]; *128* [60]; *128* [61]; *128* [62]
Geronimo, J.S. 183, *264* [57]
Gesztesy, F. 141, *264* [58]
Giachetti, R. 212, 252, 253, *264* [59]
Gilbert, D.J. 140, 191, 192, *264* [60]
Glasner, E. 179, *263* [18]
Glasner, S. 139, 160, 169, *264* [42]; *264* [61]
Godement, P. 609, *685* [36]
Gol'dberg, A. 328, 344, *358* [50]
Goldstein, M. 174, *265* [62]
Golubitsky, M. 4, 28, 52, 65, 75, 84, 120, *128* [63]; *128* [64]; *128* [65]; *128* [66]; *128* [67]; *128* [74]
Goma, I.A. 567, *589* [25]; *589* [26]; *589* [27]
Gong, X. 681, *687* [86]
Gottschalk, W. 158, *265* [63]
Grande, F. 166, *263* [6]
Greene, J. 181, *264* [56]
Greene, R. 216, 217, 220, *265* [64]
Griffiths, P. 231, *265* [65]
Grintchy, A.A. 656, 657, *685* [38]
Gromak, V. 337, 340, 342, *358* [51]
Gromyak, M. 586, *590* [50]
Guckenheimer, J. 4, 65, 84, 120, *128* [68]
Gundersen, G. 275, 287–291, 297, 304, 305, 317, 331, 332, *357* [9]; *358* [33]; *358* [52]; *358* [53]; *358* [54]; *358* [55]; *358* [56]; *358* [57]
Guo, S. 84, *128* [69]

Hadass, R. 317, *359* [58]
Hadeler, K.P. 112, 120, *128* [70]
Halburd, R. 347, *358* [31]
Hale, J.K. 4, 65, 78, *127* [33]; *128* [71]; 367, 386, 413, *439* [10]; *439* [11]; 509, 533, 585, *588* [12]; *589* [28]
Harris, J. 22, *128* [58]; 231, *265* [65]
Hartman, P. 171, 210, *265* [66]; *265* [67]
Hayman, W. 324, 347, *358* [32]; *359* [59]; *359* [60]
He, Y. 271, 324, 341, 342, *359* [61]; *359* [62]; *359* [63]; *361* [119]
Heading, J.M.A. 611, 612, *685* [37]
Heck, A. 578, *589* [29]
Hedlund, G. 158, *265* [63]
Heittokangas, J. 310, 314–321, *358* [33]; *359* [64]; *359* [65]; *359* [66]; *359* [67]; *359* [68]; *359* [69]; *359* [70]; *359* [71]
Hellerstein, S. 300, *359* [72]
Henry, D. 432, *439* [12]
Herman, M. 151, 169, 171, *264* [51]; *265* [68]
Herold, H. 271, 273–275, 283, 310, 311, 316, *359* [73]; *359* [74]; *359* [75]
Hille, E. 271, 273, 285, 287, 307, 342, *359* [76]; *359* [77]; *359* [78]; *359* [79]
Hinkkanen, A. 336–339, *359* [80]; *359* [81]; *359* [82]; *359* [83]; *359* [84]
Holden, H. 141, *264* [58]
Holmes, P.J. 4, 65, 84, 120, *128* [68]; *130* [129]
Hörmander, L. 272, 273, *359* [85]
Hosabekov, O. 567, *589* [30]
Hotzel, R. 327, 334, 354, 355, *359* [86]; *359* [87]
Hukuhara, M. 336, *360* [88]; 604, 644, 669, 670, *685* [39]; *685* [40]
Husemoller, D. 23, *128* [72]
Hutchinson, G.E. 112, 120, *128* [73]

Ihrig, E. 28, 52, *128* [74]
Il'yashenko, Yu.S. 303, *360* [89]; 601, 626, 627, 633, 637, 642, 650, 665, 683, *684* [5]; *685* [31]; *685* [41]; *685* [42]
Imkeller, P. 135, *265* [69]
Iooss, G. 65, 75, *127* [30]; *128* [75]
Ishizaki, K. 323, 325, 330, 332–334, *360* [90]; *360* [91]; *360* [92]; *360* [93]; *360* [94]; *360* [95]
Its, A. 335, *358* [42]; *360* [96]
Iwasaki, K. 284, *360* [97]
Ize, J. 6–8, 49, 52, 53, 66, 75, *128* [76]; *129* [77]; *129* [78]; *129* [79]; *129* [80]; *129* [81]

Jäger, T. 176, 178, *263* [16]; *265* [70]; *265* [71]
Jambois, T. 233, *265* [72]
Jank, G. 271, 285, 287, 288, 290, 334, *359* [87]; *360* [98]
Jankowski, T. 587, *589* [31]; *589* [32]; *589* [33]; *589* [34]
Jarnik, V. 271, *360* [99]
Jaworowski, J. 7, *129* [82]; *129* [83]
Jitomirskaya, S. 142, 175, *263* [23]
Jodel, J. 6–8, 53, 66, *127* [46]
Johnson, R.A. 136, 139, 140, 142, 147–149, 151, 152, 154, 158, 163, 165–170, 172–174, 178–180, 182, 183, 195, 197, 198, 212, 215, 219, 220, 237, 250–253, 255–258, 261, 262, *263* [15]; *263* [17]; *264* [31]; *264* [41]; *264* [44]; *264* [45]; *264* [46]; *264* [47]; *264* [48]; *264* [49]; *264* [50]; *264* [57]; *264* [59]; *265* [73]; *265* [74]; *265* [75]; *265* [76]; *265* [77]; *265* [78]; *265* [79]; *265* [80]; *265* [81]; *265* [82]; *265* [83]; *265* [84]; *265* [85]; *265* [86]; *265* [87];

265 [88]; *265* [89]; *265* [90]; *266* [91]; *266* [92]; *266* [93]; *266* [94]; *266* [95]; *266* [96]; *266* [97]; 376, *439* [13]
Jorba, À. 142, 175–177, *266* [98]; *266* [99]

Kamenskiĭ, M.I. 50, 51, *126* [1]
Kapaev, A. 335, *358* [42]
Katajamäki, K. 352–355, *360* [100]
Kato, M. 355, *363* [181]
Katok, A. 169, *263* [7]
Kaufman, R. 327, 343, 349, *357* [10]; *357* [11]; *357* [12]
Kaup, D. 182, *262* [1]
Kawakubo, K. 20, 22, 23, *129* [84]
Keller, H.B. 567, *589* [35]
Khoma, G. 586, *590* [50]
Kiguradze, I.T. 444, *589* [36]
Kim, W. 317, *360* [101]
Kimura, H. 284, *360* [97]
Kimura, T. 644, 669, 670, *685* [40]
Kinnunen, L. 297, *360* [102]
Kirchgraber, U. 368, *439* [14]
Kirillov, A.A. 13, 17, *129* [85]
Kitaev, A. 337, *360* [103]
Kliemann, W. 136, *264* [33]
Klimyk, A.U. 22, *131* [142]
Knill, O. 142, *266* [100]
Kobayashi, S. 666, *685* [43]
Koçak, H. 368, 425, 438, *439* [4]
Kokubu, H. 680, *685* [44]
Kolyada, S. 179, *263* [18]
Komiya, K. 7, *129* [86]
Korhonen, R. 310, 315–321, *359* [68]; *359* [69]; *359* [70]; *359* [71]; *360* [104]; *360* [105]
Kosniowski, C. 7, *129* [87]
Kotani, S. 184, 216, *266* [101]
Krantz, S. 216, 217, 220, *265* [64]
Krasnosel'skiĭ, M.A. 52, 83, *129* [88]; *129* [89]; 562, 576, *589* [37]
Krawcewicz, W. 7, 8, 10, 16–18, 23, 27, 29, 36, 41, 47, 49–53, 66, 75, 83–86, 95, 98, 100–102, 104, 109, 111, 120, *126* [10]; *126* [11]; *126* [12]; *126* [13]; *126* [14]; *126* [15]; *126* [16]; *126* [17]; *126* [18]; *126* [19]; *128* [49]; *128* [50]; *128* [60]; *129* [90]; *129* [91]; *129* [92]; *129* [93]; *129* [94]; *129* [95]; *129* [96]
Krein, M.G. 447, 523, 585, *589* [38]
Krikorian, R. 142, *266* [102]
Kruskal, M. 181, *264* [56]
Kuratowski, K. 60, *129* [97]
Kurbanbaev, O.O. 587, *589* [39]; *589* [40]
Kurpel′, N.S. 567, 572, *590* [41]
Kushkuley, A. 7, 8, 24, 25, 27, 28, 36, 52, *126* [20]; *126* [21]; *129* [98]; *129* [99]
Kwapisz, M. 585, 586, *588* [7]; *590* [42]; *590* [43]; *590* [44]
Kwon, K.H. 296, *360* [106]

Laine, I. 271, 275, 284, 285, 287, 293–295, 297, 298, 301–304, 322–324, 330–332, 336–340, 342, 344–347, 350, 353, *357* [9]; *357* [13]; *357* [14]; *357* [15]; *357* [16]; *357* [17]; *357* [19]; *358* [45]; *358* [46]; *358* [51]; *359* [62]; *359* [80]; *359* [81]; *359* [82]; *359* [83]; *359* [84]; *360* [107]; *360* [108]; *360* [109]; *360* [110]
Lamb, J.S.W. 4, 75, 84, *128* [69]; *129* [100]; *129* [101]; *129* [102]; *131* [147]
Langford, W.F. 4, 75, *128* [63]
Langley, J. 294, 300, 301, *357* [18]; *358* [35]; *360* [111]; *360* [112]; *360* [113]; *360* [114]; *360* [115]
Lappan, P. 314, *356* [1]
Laptinskii, V.N. 585, *591* [86]
Lashof, R. 7, *129* [103]
Lauterbach, R. 4, 52, 75, *127* [31]; *127* [32]
Lavie, M. 317, *360* [116]; *360* [117]
Law, C. 337, *360* [103]
Lê, D. 337, 347, *357* [19]
Le Lyong Taĭ 539, 542, *592* [92]
Leau, L. 637, *685* [45]
Lederer, C. 135, *265* [69]
Lenz, D. 142, *263* [30]
Levelt, A.H.M. 604, *685* [46]
Levins, R. 112, *129* [104]
Levinson, N. 184, *264* [32]; 599, 602, *684* [24]
Levitan, B. 139, 161, 165, 166, *268* [158]
Lewis, L.G., Jr. 7, *130* [105]
Li, T. 179, *266* [103]
Li, Y. 337, 341, 342, *360* [118]; *361* [119]; *363* [195]
Liao, L. 328, 330, 334, *361* [120]; *361* [121]; *361* [122]; *361* [123]
Lin, X.B. 367, *439* [11]
Lions, J.L. 91, 92, *130* [106]
Listopadova, V.V. 567, *592* [93]; *592* [94]
London, D. 309, 310, *361* [124]
Loray, F. 637, 672, 674, 681, 683, *685* [47]; *685* [48]; *685* [49]; *685* [50]
Luchka, A.Yu. 567, *590* [45]; *590* [46]; *592* [93]; *592* [94]
Lyapunov, A.M. 624, 625, *685* [51]

Maass, A. 179, *263* [18]
Madirimov, M. 7, *130* [107]; *130* [108]
Magenes, E. 91, 92, *130* [106]
Magnus, W. 253, *266* [104]

Magnusson, P.C. 103, *130* [109]
Malgrange, B. 608, 652, *685* [52]; *685* [53]
Mallet-Paret, J. 4, 65, 84, 95, 98, 120, *127* [34]; *127* [35]; *127* [36]; *130* [110]; *130* [111]
Malmquist, J. 324, 326, *361* [125]; *361* [126]
Mañe, R. 156, *266* [105]
Mann, L.N. 11, 12, *130* [112]
Marchenko, V. 183, *266* [106]
Marcus, L. 171, 172, *266* [107]
Markus, L. 349, *361* [127]
Marsden, J.E. 4, 65, *130* [113]
Martinet, J. 637, 646, 647, 649, 652, *685* [54]; *685* [55]
Martynyuk, D.I. 586, *590* [47]; *590* [51]
Martynyuk, O.M. 586, *591* [73]
Marusjak, A.G. 567, 572, *590* [41]
Marzantowicz, W. 6–8, 53, 66, *127* [46]; *128* [61]; *130* [114]
Massabò, I. 6–8, 53, *128* [62]; *129* [78]; *129* [79]
Mattei, J.F. 619, 649, *685* [56]
Mattheij, R.M.M. 567, *588* [6]
Matuda, T. 644, 669, 670, *685* [40]
Matveev, V. 182, 183, 242, *264* [35]
Mawhin, J. 471, 473, 480, 484, 509, 585, *590* [48]; *590* [49]; *591* [81]
Mawhin, J.L. 567, 585, *589* [24]
May, J.P. 7, *130* [105]
McClure, J.E. 7, *130* [105]
McCracken, M. 4, 65, *130* [113]
McCrory, C. 7, 52, *130* [115]
McLeod, J. 337, *360* [103]
Medvedeva, N.B. 628, 629, *686* [57]; *686* [58]; *686* [59]
Melbourne, I. 4, 52, 75, *127* [28]; *127* [32]; *129* [100]; *129* [101]; *129* [102]; *130* [116]; *131* [147]
Melnikov, V.K. 368, *439* [15]
Memory, M.C. 95, 98, *130* [117]; *130* [118]; *130* [119]
Merenkov, S. 289, *358* [41]
Meshcheryakova, Yu.I. 656, 657, *686* [60]
Mészáros, J. 448, 555, 556, 585, *591* [69]
Meziani, R. 683, *685* [50]; *686* [61]
Michel, F. 674, 682, *684* [19]
Miles, J. 300, 324, *359* [60]; *359* [72]
Miller, M. 165, *263* [27]
Millionščikov, V. 139, 141, 151, 170, 171, *266* [108]; *266* [109]
Mitropol′skiĭ, Yu.A. 586, *590* [50]; *590* [51]
Miura, R. 181, *264* [56]
Mohon’ko, A. 352–354, *361* [128]; *361* [129]
Moore, R. 171, 172, *266* [107]
Morita, K. 11, 24, *130* [120]
Moser, J. 141, 151, 152, 174, 197, 198, 212, 215, 252, 257, *265* [89]; *266* [110]; *266* [111]
Moussu, R. 619, 625, 649, 683, *684* [23]; *685* [56]; *686* [62]; *686* [63]
Mues, E. 336, *361* [130]
Mumford, D. 226, 242, *266* [112]
Murata, Y. 337, *361* [131]

Nehari, Z. 308, 309, *361* [132]; *361* [133]; *361* [134]; *361* [135]
Nemytskii, V. 145, 252, *266* [113]
Nerurkar, M. 154, 169, *265* [90]; *266* [91]; *266* [114]; *266* [115]
Nevanlinna, F. 284, *361* [136]
Newell, A. 182, *262* [1]
Newlander, A. 666, *686* [64]
Nguyen Dinh Cong 141, 142, *266* [116]; *266* [117]
Nirenberg, L. 7, *130* [121]; 666, *686* [64]
Nomizu, K. 666, *685* [43]
Noumi, M. 335, *361* [137]
Novikov, D. 303, 304, *361* [138]
Novikov, S. 141, 171, 182, 183, 242, *264* [35]; *266* [118]
Novikov, S.P. 22, 36, 40, *127* [44]
Novo, S. 138–140, 154, 169, 173, *266* [92]; *266* [93]; *266* [119]; *267* [120]; *267* [121]; *267* [122]; *267* [123]; *267* [124]; *267* [125]
Novokshenov, V. 335, *358* [42]; *360* [96]
Ntouyas, S.K. 444, *590* [52]
Núñez, C. 138, 154, 175–177, *263* [4]; *264* [46]; *264* [47]; *264* [48]; *264* [49]; *266* [99]; *266* [119]; *267* [120]
Nurzhanov, O.D. 587, *592* [106]
Nussbaum, R.D. 84, 120, *130* [122]; *130* [123]; *130* [124]

Obaya, R. 138–140, 154, 169, 173, 175–177, *263* [4]; *266* [92]; *266* [93]; *266* [99]; *266* [119]; *267* [120]; *267* [121]; *267* [122]; *267* [123]; *267* [124]; *267* [125]
Oka, H. 680, *685* [44]
Okamoto, K. 335, 336, *361* [139]; *361* [140]
Oleinikova, S.A. 585, *590* [57]
Ortega, R. 139, 165, 179, *267* [126]; *267* [127]; *267* [128]
Oseledets, V. 140, 147, 204, *263* [9]; *267* [129]
Ostrovsky, V. 183, *266* [106]

Palmer, K.J. 136, 147, 148, 250, 251, *266* [94]; *267* [130]; 367, 368, 372, 376, 397, 403, 405–407, 419, 425, 438, *438* [1]; *439* [4]; *439* [16]; *439* [17]; *439* [18]; *439* [19]
Pampel, T. 368, *439* [20]

Pavani, R. 142, 261, *264* [50]
Pearson, D.B. 140, 191, 192, *264* [60]
Perestyuk, N.A. 586, 587, *590* [53]; *590* [54]; *591* [87]
Perez-Marco, R. 660; *686* [65]
Pérez-Marco, R. 665; *686* [66]
Perov, A.I. 585, *590* [55]; *590* [56]; *590* [57]
Peschke, G. 8, 52, *130* [126]
Petrenko, V. 271, *361* [141]
Petrovski, I.G. 473, *590* [58]
Pfaltzgraff, J. 309, *358* [34]
Phelps, R.R. 150, 157, 204, *267* [131]
Poeschl, J. 182, *267* [132]
Poincaré, H. 367, *439* [21]; 606, 624, 625, 629, 631, *686* [67]
Pokornyi, V. 308, *361* [142]
Pommerenke, C. 310, *361* [143]
Portnov, M.M. 585, *590* [57]
Pöschel, J. 257, *266* [111]
Potapov, A.S. 50, 51, *126* [1]
Prieto, C. 7, *130* [114]
Puig, J. 142, 169, *267* [133]; *267* [134]; *267* [135]

Rabier, P. 7, *130* [127]
Rabinowitz, P.H. 83, *130* [128]; 419, *439* [6]
Rachůnková, I. 444, *590* [59]
Rączka, R. 22, *127* [23]
Ramis, J.-P. 637, 646, 647, 649, 652, *685* [54]; *685* [55]
Rand, R.H. 84, 120, *130* [129]; *131* [140]
Rättyä, J. 310, 315–322, *359* [68]; *359* [69]; *359* [70]; *359* [71]; *360* [104]; *360* [105]; *361* [144]
Rebaza, J. 368, *439* [9]
Redheffer, R. 336, *361* [130]
Rieth, J.v. 354, *361* [145]
Rodkina, A.E. 50, 51, *126* [1]
Rontó, A. 510, 523, 529–532, 543, 551, 554, 555, 568, 585–588, *590* [53]; *590* [60]; *590* [61]; *590* [62]; *590* [63]; *591* [64]; *591* [65]; *591* [66]; *591* [67]
Rontó, M. 444, 445, 448, 455, 474, 488, 509, 510, 523–525, 532, 538–543, 548, 551, 554–556, 560–562, 568, 571, 575, 576, 585–588, *590* [60]; *590* [63]; *591* [64]; *591* [65]; *591* [66]; *591* [67]; *591* [68]; *591* [69]; *591* [70]; *591* [71]; *591* [72]; *591* [73]; *591* [74]; *591* [75]; *591* [76]; *591* [77]; *591* [78]; *591* [79]; *591* [80]; *592* [88]; *592* [89]; *592* [90]; *592* [91]
Rossi, J. 289, 293, 300, *359* [72]; *361* [146]; *361* [147]
Rouche, N. 471, 473, 480, 484, 585, *591* [81]
Ruan, H. 52, 53, 84, 98, 120, *126* [11]; *126* [13]; *126* [14]; *126* [15]; *126* [16]
Rubel, L. 343, *361* [148]; *362* [149]; *362* [150]
Rubinsztein, R.L. 7, *130* [130]
Rudin, W. 216, *267* [136]
Rudyak, Y. 7, *130* [131]
Russell, R.D. 567, *588* [6]
Rutitskii, Ya.B. 562, 576, *589* [37]
Rutman, M.A. 447, 523, 585, *589* [38]
Rybicki, S. 8, 52, *131* [132]; *131* [133]; *131* [134]

Sacker, R.J. 147, 156, 158, 160, 161, 195, 251, 254, *267* [137]; *267* [138]; *267* [139]
Sad, P. 634, *684* [18]
Sadovskii, A.P. 625, *686* [68]
Sadovskiĭ, B.N. 50, 51, *126* [1]
Sakamoto, K. 368, *439* [7]
Samoilenko, A.M. 444, 445, 455, 474, 488, 509, 523–525, 529, 538–543, 548, 551, 554, 555, 560–562, 567, 568, 575, 576, 584–588, *590* [47]; *590* [51]; *591* [64]; *591* [70]; *591* [74]; *591* [75]; *591* [76]; *591* [77]; *591* [78]; *591* [79]; *591* [80]; *591* [82]; *591* [83]; *591* [84]; *591* [85]; *591* [86]; *591* [87]; *592* [88]; *592* [89]; *592* [90]; *592* [91]; *592* [92]; *592* [93]; *592* [94]
Sanders, J. 680, *684* [7]; *684* [8]
Sans, A. 138, *267* [124]; *267* [125]
Sarafova, G.H. 587, *588* [8]; *588* [9]; *592* [95]; *592* [96]
Sattinger, D. 181, 182, *263* [11]
Savelev, V.I. 682, *686* [69]
Schaeffer, D.G. 4, 52, 65, 75, 84, 120, *128* [64]; *128* [67]
Schlag, W. 174, *265* [62]
Schubart, H. 337, *362* [151]
Schwartz, J. 207, *264* [36]
Schwartzman, S. 201, *267* [140]
Schwarz, B. 308, 310, 311, *357* [21]; *362* [152]
Seddighi, K. 567, *588* [10]
Segal, G. 183, *267* [141]
Segal, G.B. 7, *131* [135]
Segur, H. 182, *262* [1]
Seidenberg, A. 619, *686* [70]
Selberg, H. 351, 352, *362* [153]
Selgrade, J. 147, 156, 195, *267* [142]
Sell, G.R. 136, 147, 148, 156, 158, 160, 161, 195, 251, 254, *266* [94]; *267* [137]; *267* [138]; *267* [139]; *267* [143]; 376, *439* [13]
Shäfke, R. 674, 683, *684* [20]; *684* [21]
Shapino, L. 160, *264* [42]
Shcherbakov, A.A. 634, 665, 683, *685* [31]; *686* [71]
Shchobak, N.M. 571, 587, *591* [65]; *591* [71]
Shen, L.-C. 293, *362* [154]

Shen, W. 144, *267* [144]; *267* [145]
Shimomura, S. 284, 302, 303, 331, 336–342, *358* [51]; *360* [97]; *362* [155]; *362* [156]; *362* [157]; *362* [158]; *362* [159]; *362* [160]; *362* [161]; *362* [162]; *362* [163]; *362* [164]; *362* [165]; *362* [166]
Shlapak, Ju.D. 586, *592* [97]; *592* [98]
Shon, K. 319, *357* [28]
Shovkoplyas, V.N. 587, *590* [54]
Shtern, A.I. 22, *131* [152]
Sibuya, Y. 604, 608, *686* [72]; *686* [73]
Siegel, C.L. 226, 228, 229, 237, *267* [146]; 658, *686* [74]
Sil'nikov, L.P. 367, *439* [22]; *439* [23]
Simó, C. 142, *266* [98]; *267* [135]
Simon, B. 151, 183, *267* [147]; *267* [148]
Sinai, Y. 142, *263* [26]
Slyusarchuk, V.E. 432, *439* [24]
Smale, S. 367, *439* [25]
Smith, J.M. 120, *131* [136]
Sobkovich, R.I. 572, 586, *592* [99]; *592* [100]; *592* [101]
Sobolev, S.L. 56, *131* [137]
Sons, L. 322, 347, *358* [43]; *362* [167]
Sorvali, T. 331, *360* [108]
Spanily, T. 85, 86, 95, 98, *129* [90]
Staněk, S. 444, *590* [59]
Stark, J. 178, *265* [71]
Steinbart, E. 289–292, 304, 305, *358* [54]; *358* [55]; *358* [56]; *358* [57]; *362* [168]
Steinberger, M. 7, *130* [105]
Steinlein, H. 7, 8, 10, 23, 52, 53, 66, 75, 84, 98, 100–102, 104, 109, 111, 120, *126* [17]; *126* [18]; *126* [19]; *131* [138]
Steinmetz, N. 324, 325, 328, 330, 336–339, *362* [169]; *362* [170]; *362* [171]; *362* [172]; *362* [173]; *362* [174]; *362* [175]
Steinmetz, S. 345, *362* [176]
Stepanov, V. 145, 252, *266* [113]
Stetsenko, V.Ya. 562, 576, *589* [37]
Stewart, I.N. 4, 52, 75, 84, 120, *128* [65]; *128* [66]; *128* [67]
Stoffer, D. 368, *439* [14]
Stong, R.E. 7, *131* [139]
Storti, D.W. 84, 120, *131* [140]
Strelitz, S. 344, *362* [177]
Strizhak, T.G. 586, *592* [102]
Stróżyna, E. 649, 671, 672, 674, 675, 677, 681, 683, *686* [75]; *686* [76]; *686* [77]
Su, W. 328, *361* [120]
Sussmann, H. 169, *266* [115]
Svarichevskaya, N.A. 567, *588* [3]
Szmigielski, J. 181, 182, *263* [11]
Szmolyan, P. 368, *439* [26]

Takano, K. 336, *361* [140]
Takens, F. 671, *686* [78]
Tananika, A.A. 585, *590* [56]
Tarallo, M. 139, 165, 179, *267* [126]; *267* [127]; *267* [128]
Tatjèr, C. 175–177, *266* [99]
Teysseyre, M. 674, 682, *684* [19]
Teyssier, L. 655–657, *686* [79]; *686* [80]
Titchmarsh, T.C. 185, *267* [149]
Toda, N. 345, 355, *363* [178]; *363* [179]; *363* [180]; *363* [181]
Tohge, K. 294, 356, *360* [109]; *363* [182]
Toland, J.F. 4, 52, 66, *127* [40]; *127* [41]; *127* [42]
Tolstov, G.P. 510, *592* [103]
tom Dieck, T. 7, 13, 15, 17, 22, 23, 52, *127* [27]; *127* [43]
Tripathi, V.K. 103, *130* [109]
Trofimchuk, E.P. 585, *592* [104]
Trofimchuk, S.I. 445, 523, 551, 554, 555, 585–588, *591* [64]; *591* [72]; *591* [74]; *591* [75]; *591* [76]; *591* [77]; *591* [78]; *591* [79]; *591* [80]
Trubowitz, E. 182, *267* [132]
Turrittin, H. 604, *686* [81]
Tvrdý, M. 444, *590* [59]

Ullrich, E. 351, 352, *363* [183]
Ulrich, H. 7, 53, *131* [141]
Ul'yanova, V.I. 567, *588* [1]

Vainikko, G.M. 562, 576, *589* [37]
Valiron, G. 351, 352, *363* [184]
van den Essen, A. 619, *685* [32]
Vanderbauwhede, A. 509, *592* [105]
Varadarajan, V.S. 603–605, 608, 609, 611, *684* [6]; *686* [82]
Veech, W. 139, 144, *267* [150]; *268* [151]
Viana, M. 142, 171, 175, 255, *263* [19]
Vignoli, A. 6–8, 49, 52, 53, 66, 75, *128* [62]; *129* [78]; *129* [79]; *129* [80]; *129* [81]
Vilenkin, N.Ja. 22, *131* [142]
Vinberg, È.B. 22, *131* [143]
Vinograd, R. 139, 151, 170, 171, 179, *268* [152]
Vivi, P. 7, 52, 53, 75, 84, 120, *129* [91]; *129* [92]
Volkmann, L. 271, 285, 287, 288, 290, *360* [98]
von Rieth, J. 324, 330, *363* [185]
Voronin, S.M. 636, 656, 657, 665, 683, *685* [31]; *685* [38]; *686* [60]; *686* [83]

Walter, W. 509, *590* [49]
Wang, D. 680, *685* [44]; *687* [84]

Wang, S. 289–292, 304–306, *358* [49]; *358* [55]; *358* [56]; *358* [57]; *361* [147]; *363* [186]
Wang, X. 680, *687* [84]
Wang, Y. 325, *360* [94]
Warner, F.W. 22, *131* [144]
Wasow, W. 612, 615, *687* [85]
Weikard, R. 303, *363* [187]
Weinberger, H. 184, *268* [153]
Weiss, B. 139, 169, *264* [61]
Wilson, G. 183, *267* [141]
Winkler, S. 253, *266* [104]
Wittich, H. 271, 284, 296, 330, 337, *362* [151]; *363* [188]; *363* [189]; *363* [190]
Wu, J. 7, 8, 16–18, 29, 36, 41, 47, 49–53, 66, 75, 77, 83–86, 94, 95, 98, 104, 105, 111, 120, 122, *128* [49]; *128* [50]; *129* [90]; *129* [91]; *129* [92]; *129* [93]; *129* [94]; *129* [95]; *131* [145]; *131* [146]
Wu, J.H. 7, 27, 52, 53, *128* [60]
Wulff, C. 4, 75, *129* [101]; *129* [102]; *131* [147]

Xia, H. 7, 27, 52, 66, 83, 84, 120, *129* [95]; *129* [96]; *131* [148]
Xiao, X. 271, *359* [63]

Yagubov, M.A. 567, *588* [3]
Yakovenko, S. 303, 304, *360* [89]; *361* [138]
Yamashita, S. 332, *363* [191]
Yanagihara, N. 323, *360* [95]
Yang, C-C. 328, *361* [121]
Yang, C.-C. 322, 328, *361* [120]; *363* [192]
Yang, L.Z. 299, *363* [193]
Yang, R. 297, 298, *360* [110]
Ye, Z. 322, 328, 330, 334, *361* [122]; *361* [123]; *363* [192]
Yi, H.-X. 318, 319, *357* [25]; *357* [26]
Yi, Y. 144, 178, *267* [144]; *267* [145]; *268* [154]
Yoccoz, J.-C. 647, 658, 664, 665, *686* [66]; *687* [87]
Yorke, J.A. 4, 8, 65, 84, 95, 98, 120, *126* [2]; *127* [36]; *130* [111]; 179, *266* [103]
Yoshida, M. 284, *360* [97]
Yosida, K. 353, *363* [194]
Yuan, W. 337, *363* [195]
Yukhno, L.F. 567, *588* [1]

Zabreĭko, P.P. 52, 83, *126* [21]; *129* [89]; *131* [149]; 562, 576, 585, *589* [15]; *589* [16]; *589* [37]
Zafer, A. 587, *588* [4]
Zakhari͡ichenko, Yu.O. 567, *590* [46]
Zames, G. 182, *268* [155]
Zampogni, L. 140, 141, 182, 183, 219, 220, 237, 241, 250, *266* [95]; *266* [96]; *266* [97]; *268* [156]; *268* [157]
Zavalykut, G.D. 587, *592* [106]; *592* [107]
Zeeman, E.C. 7, 52, *131* [150]
Zhelobenko, D.P. 13, 22, *131* [151]; *131* [152]
Zhikov, V. 139, 161, 165, 166, *268* [158]
Zhu, K. 314, *363* [196]
Zimogljad, V. 344, *363* [197]
Żołądek, H. 596, 604, 609, 619, 634, 660, 669, 671, 672, 674, 675, 677, 681, 683, *686* [76]; *686* [77]; *687* [88]; *687* [89]

Subject Index

$1/3.41613062\ldots$ – convergence radius of series (5.64), 529
– spectral radius of operator (7.12), 555
– value of $R_{K,1}/r(K)$ for constant K, 523
β-neighbourhood, 443
$\delta_\Omega(f)$, 540, 557
$\mathrm{GL}_n(\mathbb{R})$, 443
μ-Lipschitzian map, 50
$\mathbb{1}_n$, 443
$\partial\Omega$, 443
$\Sigma^{\mathrm{ed}}(L)$, 194
Θ-divisor, 229

$A(\Gamma)$-module $A_1^t(G)$, 32
Abel identity, 293
Abel map, 228
absolutely irreducible, 13
adjoint, 373, 416
Airy differential equation, 289
algebraic solvability, 624, 627
algebroid solution, 350, 352–356
almost automorphic
– extension, 144, 162, 168
– flow, 144
– function, 144
almost complex structure, 665
almost periodic flow, 143
alternating group, 98
analytic Hadamard–Perron theorem, 632
analytical solvability, 629
arbitrary growth theorem, 344
arranging equivariant spectral data, 74
Arzelà–Ascoli theorem, 473
asymptotic cycle, 609, 654
asymptotic phase, 367, 387, 394–396, 401
auxiliary function, 88
auxiliary G-invariant function, 7

Baker–Akhiezer function, 241
Banach G-representations, 17
Banach vector bundle, 19
basic degree, 45
basic map, 37, 44, 45
Bebutov construction, 136–138, 143, 145
Bebutov flow, 136
Beltrami differential, 667
Beltrami equation, 667
Bessel equation, 611
bi-orientable, 32
bifurcation, 87
Birkhoff Ergodic Theorem, 145
Birkhoff recurrence, 143
BL-condition, 294
BL-conjecture, 294
BL-function, 294
Blaschke-oscillatory, 308, 310
Bloch, 315
Bogdanov–Takens prenormal form, 673
Bogdanov–Takens singularity, 618, 671
Bohr almost periodicity, 143
Borel transform, 651
Borsuk theorem, 473
boundary conditions
– linear
– – periodic, 444
– – three-point, 444
– non-linear
– – separated, 444, 551
– – two-point, 444
boundary value problem
– regular, 443
– singular, 444
bounded mean motion, 178
Briot–Bouquet differential equation, 271, 283, 342, 343
Briuno condition, 658
Brouwer degree, 471
Burnside ring, 28, 29
$B(y,\beta)$, 443

$\mathbb{C}$-complementing, 44
– map, 40, 45
– pair, 40, 45
$C_T^n(\tau, E)$, 510
$C([0,T]\times D, \mathbb{R}^n)$, 443
Camassa–Holm equation, 181

Cameron’s theorem, 158, 161, 167
Carathéodory conditions, 529
center, 3, 94, 624, 625
– isolated, 3
– manifold, 643, 647
center–focus problem, 623
chaotic behaviour, 367, 402, 407
character, 143
– of representation, 15
characteristic
– equation, 54, 77, 93, 114
– operator, 54, 93
– root, 54, 93
– trajectory, 623
Choquet theory, 157, 204
closed
– gap, 253
– operator, 21
Clunie, 337, 339, 345, 352
– lemma, 322
cocycle, 137, 145
coincidence problems, 84
community matrix, 113
complementing function, 60
complete reducibility theorem, 13
completely continuous map, 50
complex function spaces, 307, 312, 314, 320
– Bergman space, 313, 314
– – weighted, 313, 314
– Bloch space, 307, 314
– BMOA, 315
– Dirichlet space, 314
– – weighted, 314
– Hardy space, 307, 310, 313, 332
– – weighted, 332
– Korenblum space, 313, 318
– Nevanlinna class, 313
– normal functions, 315, 322
– Q_p-space, 307, 314
complex isotypical decomposition, 17
complexification, 14
condensing
– field, 50
– map, 50
cone construction, 257
conjugacy class, 11
conjugation, 14
continuation, 278, 356
– along the curve, 276
– of local solutions, 275–277
continued, 336
continuous family of equivariant Fredholm operators of index zero, 85
crossing numbers, 63, 71, 77, 95, 96, 101, 109
cusp, 618, 674, 682

D_β, 443
defect, 337, 338
defining equation, 602
Denjoy cocycle, 180
density of exponential dichotomy, 251
desingularization, 619
dicritical
– davison, 618
– node, 617
diffeomorphism tangent to identity, 633
differential
– field, 343, 348–350
– independence, 349
– of the first kind, 226
– of the second kind, 231
– of the third kind, 229, 231
differentially
– algebraic, 349
– elementary functions, 349
– transcendental, 349
diffusion equation, 113
dihedral group, 98, 122
direct growth problem, 320
disconjugate, 308–310
distal
– extension, 161
– pair, 161
distance, 310
divisor, 227
dominating, 67
– orbit types, 9, 67
Duffing equation, 477
dynamical spectrum, 147, 148, 151

Ecalle–Voronin moduli, 632, 636, 650
Ecalle–Voronin theorem, 636, 639
eigenvalue, eigenfunction, 184
elementary singular point, 617
equilibrium point, 104
equivalence of
– diffeomorphisms, 633, 636, 641, 642, 649
– of linear systems, 596, 602, 608
– of vector fields, 616, 656, 657
equivariant Dugundji theorem, 26
ergodic
– measure, 144
– theory, 144
exceptional divisor, 617
exponent of convergence, 288, 289, 293
exponential dichotomy, 141, 147, 194, 200, 250

Favard property, 139, 161
Favard theory, 165
field, 49
finite oscillation property, 290, 306
finiteness degree, 318
– of growth, 297
first focus quantity, 628
fixed singularities, 326, 335
Floquet exponent, 201
Floquet matrix, 153
Floquet theory, 153, 155, 303
flow, 142, 143
– homomorphism, 143
– isomorphism, 143
folding homomorphism Θ_l, 74
Fredholm operator, 20
Frei theorem, 318
frequency module, 144
function
– (τ, E)-proper, 510
– τ-even, 510
– τ-odd, 510
– approximate determining, 470
– determining, 470
– even, 510
– odd, 510
– T-periodic, 443
fundamental domain, 24

G-action, 10
G-equivariant field, 48
G-equivariant homotopy of compact fields, 49
G-equivariantly homotopic, 49
G-fundamental, 51
G-homotopy, 25
G-invariant norm, 17
G-manifold, 19
G-representation, 12
G-representation conjugate, 14
G-space, 10
G-vector bundle, 19
Gauss hypergeometric equation, 596, 602
generalized Θ-function, 235
generalized cusp, 671
generalized Jacobian variety, 234
generalized Riemann vector, 235
generalized saddle, 671
generalized saddle–node, 671
Gevrey series, 651
Gilbert–Pearson theory, 140, 191, 205
global bifurcation problems, 81
global continuation of bifurcating branches, 109
global Hopf bifurcation, 111
– theorem, 82
global oscillation property, 290, 306
good resolution, 617, 619
Green’s function, 195

H-fixed-point subset, 11
(H)-normality, 27
Haar integral, 15
Hartman argument, 210
Hausdorff metric, 21
Herglotz function, 207
Hermite–Weber differential equation, 339
Hölder theorem, 349
holomorphic differential, 226, 231
holomorphic foliations, 616
holonomy transformation, 647, 648
homoclinic orbit, 403
homotopy factorization, 40
Hopf bifurcation, 54
Hukuhara–Kimura–Matuda theorem, 669
hull, 137, 143
Hutchinson model, 112
hyperbolic, 310, 367, 386, 405
hyperbolic distance, 311
hyperbolic periodic orbits, 381

icosahedral group, 98, 124
index of L, 21
indicial equation, 301
induction over orbit types, 23
intristic dimension, 16
invariant
– measure, 144
– set, 143
inverse growth problem, 320
irreducible representation, 12
irregular singular point, 596, 603
isolated, 94
– center, 77
isometric Hilbert G-representation, 17
isotropy, 11
isotypical component, 18
isotypical decomposition, 15, 18, 55
iterated order, 297, 306, 317, 318
iterated type, 319

Jacobian variety, 227

Korteweg–de Vries (K-dV) equation, 181
Kotani theory, 212, 215
Krein–Rutman theorem, 523
Kronecker flow, 144
Kronecker winding, 142, 144
Krylov–Bogoliubov construction, 145, 152

l-folding, 38
l-th isotypical crossing number, 63
$L_1([0,T],\mathbb{R}^n)$, 530
Lamé equation, 302, 303
lemma of the logarithmic derivative, 286, 351
Leray–Schauder twisted degree, 48
$\lfloor\cdot\rfloor$, 519
Li–Yorke chaos, 139, 179
lifting homeomorphism, 25
limit periodic, 144
limit point case, 188
linearization of diffeomorphism, 633, 658
Liouville-type frequencies, 250
Lipschitz condition, 518, 537, 539
local bifurcation
– $\Gamma \times S^1$-invariant, 116
– invariant, 7, 68, 78, 80, 88
– result, 97
local existence, 271, 273, 276
local index, 35
local solution, 336, 356
Lyapunov exponent, 146, 151, 175, 201, 202, 253

Malgrange–Sibuya theorem, 608
Malmquist, 325, 328, 332, 345
Malmquist theorem, 271,, 324, 326, 327, 330, 353, 354
map, 26, 27
Maple© input data, 102
Marcus and Moore disconjugacy, 171
Martinet–Ramis moduli, 642, 646, 650, 651, 654, 656
Martinet–Ramis theorem, 646
Melnikov function, 417
Millionscikov–Vinograd examples, 139, 175
Millionscikov–Vinograd type, 170
minimal flow, 143
minimal support, 191
modulus
– of quadrangle, 668
– of unnulus, 668
Mohon’ko, 322, 337, 339, 352
– lemma, 323
monodromic singular point, 623
monodromy
– operator, 598
– transformation, 648
movable singularities, 326
multiplicativity property, 47, 48

$\mathbb{N}$, 443
necessary condition
– for Hopf bifurcation, 58
– for the occurrence of Hopf bifurcation, 57
negative spectrum, 95, 101, 108, 118
Nevanlinna theory, 284–287, 307, 312, 319, 324, 328, 351, 352
– characteristic function, 285, 286, 323, 324, 345, 351, 353
– counting function, 285, 351, 352
– deficiency, 286, 288
– first main theorem, 285, 286, 351
– logarithmic derivative lemma, 292, 297, 352
– non-integrated counting function, 288, 351
– proximity function, 285, 351
– ramification index, 287, 337–340
– second main theorem, 286, 287, 310, 340, 341
Newlander–Nirenberg theorem, 666
Newton–Puiseux diagram, 290
node, 617
non-oscillatory, 308–310
non-trivial solutions, 87
nonautonomous dynamical systems, 135
nonlinear Schrödinger (NLS) equation, 181
normal, 27
normal form
– for diffeomorphisms, 634, 641, 649
– for linear system, 603
– for vector fields, 680
normal homotopy, 27
n species ecosystem, 112
numbers $n(L,H)$, 28
numerical shadowing, 368, 425

Ω-admissible, 26, 49
– homotopy, 26
$\Omega_{P,\sigma}$, 520
$\Omega(\beta)$, 557
orbit, 11, 143
– space, 11
– type, 11
orbital equivalence, 616, 642, 646, 649, 674
orbital linearization, 665
orbital normal form, 649, 656
order, 286, 288, 290–292, 294–296
Ortega and Tarallo examples, 165
oscillatory, 307, 308
Oseledets
– spectrum, 147, 151
– theory, 146

p-summability, 652
Painlevé differential equation, 271, 274, 275, 334
– Airy solutions, 337
– fifth Painlevé equation, 336, 337
– first Painlevé equation, 336
– first Painlevé hierarchy, 341

– fourth Painlevé equation, 336–338
– higher order Painlevé equations, 340
– rational solutions, 337, 339
– second Painlevé equation, 336, 338
– second Painlevé hierarchy, 341
– sixth Painlevé equation, 336, 337
– third Painlevé equation, 336–338
Painlevé equation, 335
Painlevé property, 334, 335, 340
parametrisation
– method of, 567
parametrized equivariant coincidence problem, 86
period matrix, 226
periodic, 294, 295, 302, 303, 331
– orbit, 367, 386
periodic-to-periodic homoclinic orbit, 367, 402, 412, 415
perturbed system, 367, 413, 416
Picard–Lindelöf theorem, 475
Poincaré map, 367, 382, 385, 386, 403, 405
Poincaré–Dulac normal form, 629
Poincaré–Dulac theorem, 601, 629
Poisson recurrence, 193, 194
principal solution, 171
projective flow, 139, 146, 154, 155, 197
proximal extension, 161, 168
property (τ, E), 510
proximal flow, 161
proximal pair, 161
pseudo orbit, 425
– homoclinic, 426
pseudo periodic orbit, 425

quasi-conformal map, 667
quasi-periodic flow, 142, 144
quasi-section, 258

$\mathbb{R}$, 443
$\mathbb{R}_+$, 443
$R^{\text{ed}}(L)$, 194
$r(A)$, 443
random dynamics, 135
ray of division, 606
$\text{RE}^{\pm}(L_\omega)$, 208, 209
reaction equation, 113
Recurrence Formula, 40, 43, 44
regular fundamental domain, 25
regular normal
– homotopy, 27
– map, 27
regular singular point, 596–598, 600
rescaling, 336
resolvent, 191
resonance, 601, 629
resonant diffeomorphism, 641, 649
resonant one-dimensional map, 649
resonant saddle, 630, 647, 656
resonant singular point, 617
result due to Frei, 296
Riccati differential equation, 271, 274, 275, 324, 330–333, 338, 339
Riccati equation, 196
Riemann inversion problem, 228
Riemann theta function, 228
robustness, 412
rotation number, 151, 152, 197, 202, 251
roughness theorem, 373

Sacker–Sell spectrum, 147
saddle, 617, 664
saddle–node, 617, 642, 656
Schwarzian derivative, 332
Schwarzian differential equation, 271, 332–334
Schwarzmann homomorphism, 141, 201
sequence
– equicontinuous, 474
– uniformly bounded, 473
shadow, 425
shooting method, 567
Siegel lemma, 350
singular point of Poincaré type, 629
singularity of the Fuchs type, 598
small function, 324
small meromorphic coefficients, 354
smooth G-vector bundle, 20
smoothness of the projections, 376
SNA (strange nonchaotic attractor), 175, 176
solutions of zero order, 344
space, 314
spectral matrix, 190, 191
spectral measure, 189
spectrum, 191
Splitting lemma, 40, 41
stable and unstable manifolds, 367, 386, 395, 396
stable fibre, 402
stochastic dynamical systems, 135
Stokes cocycle, 608
Stokes matrix, 607
Stokes operator, 607
Stokes sheaf, 607
strictly ergodic flow, 145
strong manifold, 643
strongly elliptic case, 167
Sturm–Liouville operator, 140, 184, 191
subordinate solution, 191
subrepresentation, 12
subset, 51

successive approximations, 453, 457, 470, 486, 488, 492, 498, 523, 540, 548, 557, 570
sufficient condition
– for Hopf bifurcation, 79, 81
– for symmetric Hopf bifurcation, 67
– for the occurrence of Hopf bifurcation, 58
Suspension Procedure, 40
symmetric configuration of transmission lines, 105
symmetric Hopf bifurcation, 66
– problem, 4
symmetric system of the Hutchinson model, 114

Takens prenormal form, 671, 681
telegrapher's equation, 103
theorem
– Arzelà–Ascoli, *see* Arzelà–Ascoli theorem
– Borsuk, *see* Borsuk theorem
– Krein–Rutman, *see* Krein–Rutman theorem
– Picard–Lindelöf, *see* Picard–Lindelöf theorem
theorem due to L. Fuchs, 326
topological support, 145, 193
topological transformation group, 10
transmission lines, 102
transversal, 367, 405
trichotomy, 367, 368, 405
trivial solution, 87
tubular map, 35
twisted (by the homomorphism $\varphi : K \to S^1$) l-folded subgroup, 30
twisted conjugacy class, 31
twisted equivariant degree, 33
twisted G-equivariant degree, 6
twisted subgroups, 30

$\mathcal{U}_j$-multiplicity, 77, 96
unbounded mean motion, 177, 178
unitary Hilbert G-representation, 17
univalence of solutions, 307, 308
univalent, 332
unstable fibre, 402

$\mathcal{V}_i$-multiplicity, 96
$\mathcal{V}_{j,l}$-isotypical crossing number, 72
Valiron–Mohon'ko, 322, 352
– theorem, 323
Van der Pol equation, 481
vector of Riemann constants, 228

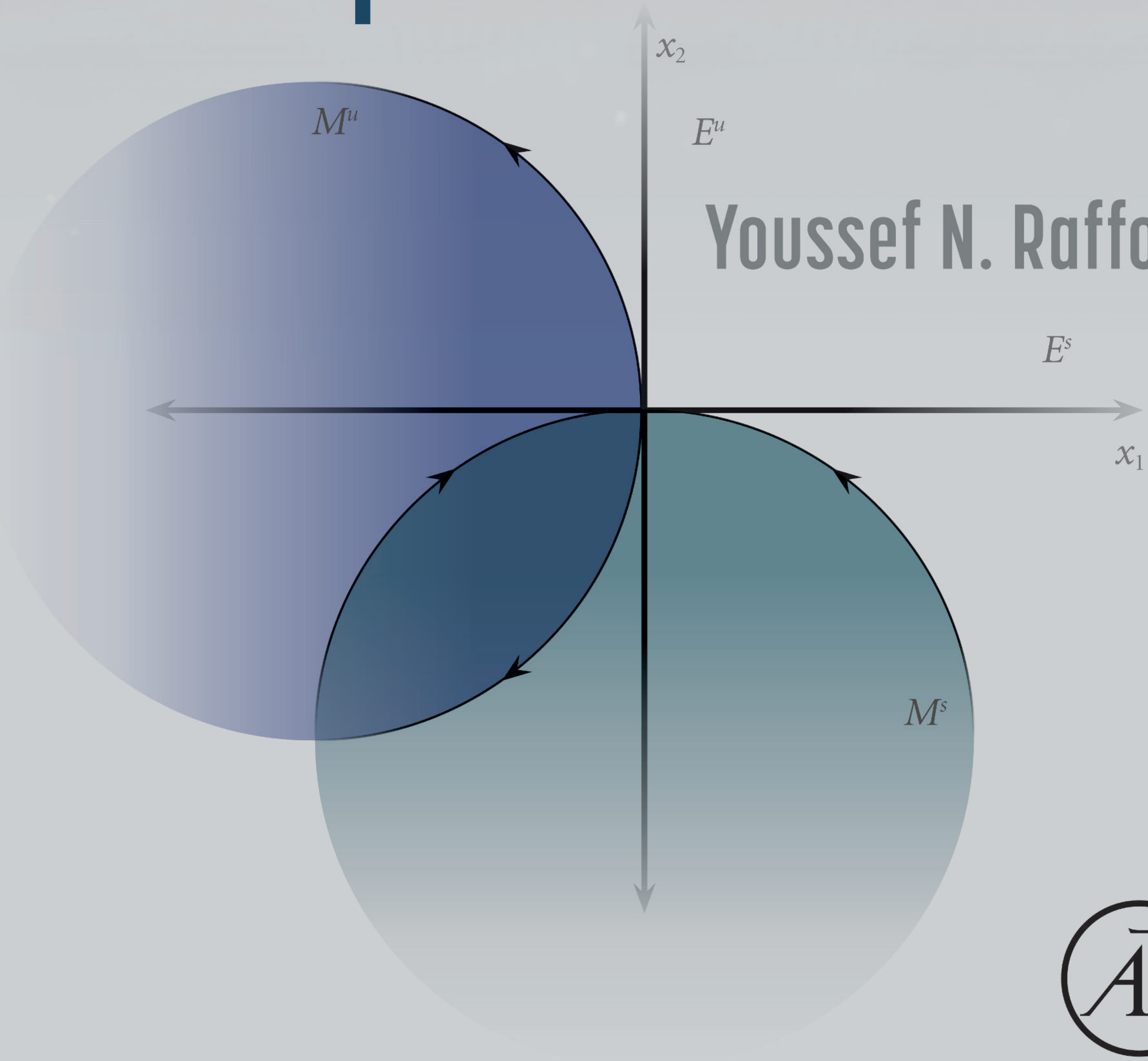
W-singular point, 57
weakly elliptic case, 167, 169
weakly hyperbolic case, 167, 169
Weierstraß $\wp$-function, 302
Weyl group, 12
Weyl m-function, 188, 190, 194
Wiman–Valiron, 290
– theory, 288
Wronskian, 293, 294

$X \sqsupset_\Omega Y$, 531

$\mathbb{Z}$, 443
zeros, 310

Advanced Differential Equations

Youssef N. Raffoul


x_2
E^u
M^u
E^s
x_1
M^s

Contents

Preface xi

CHAPTER 1 Preliminaries and Banach spaces **1**
- **1.1** Preliminaries 1
- **1.2** Escape velocity 6
- **1.3** Applications to epidemics 8
- **1.4** Metrics and Banach spaces 10
- **1.5** Variation of parameters 18
  - 1.5.1 RC circuit 20
- **1.6** Special differential equations 21
- **1.7** Exercises 23

CHAPTER 2 Existence and uniqueness **27**
- **2.1** Existence and uniqueness of solutions 27
- **2.2** Existence on Banach spaces 41
- **2.3** Existence theorem for linear equations 44
- **2.4** Continuation of solutions 45
- **2.5** Dependence on initial conditions 48
- **2.6** Exercises 53

CHAPTER 3 Systems of ordinary differential equations **57**
- **3.1** Existence and uniqueness 57
- **3.2** $x' = A(t)x$ 64
  - 3.2.1 Fundamental matrix 71
  - 3.2.2 $x' = Ax$ 77
  - 3.2.3 Exponential matrix e^{At} 82
- **3.3** $x' = A(t)x + g(t)$ 92
- **3.4** Discussion 95
- **3.5** Exercises 98

CHAPTER 4 Stability of linear systems **103**
- **4.1** Definitions and examples 103
- **4.2** $x' = A(t)x$ 108
- **4.3** Floquet theory 111
  - 4.3.1 Mathieu's equation 122
  - 4.3.2 Applications to Mathieu's equation 125
- **4.4** Exercises 126

CHAPTER 5 Qualitative analysis of linear systems **129**
- **5.1** Preliminary theorems 129
- **5.2** Near-constant systems 133
- **5.3** Perturbed linear systems 137

Copyright Elsevier 2022

5.4 Autonomous systems in the plane 139
5.5 Hamiltonian and gradient systems 145
5.6 Exercises 155

CHAPTER 6 Nonlinear systems **159**
6.1 Bifurcations in scalar systems 159
6.2 Stability of systems by linearization 168
6.3 An SIR epidemic model 177
6.4 Limit cycle 181
6.5 Lotka–Volterra competition model 185
6.6 Bifurcation in planar systems 192
6.7 Manifolds and Hartman–Grobman theorem 198
  6.7.1 The stable manifold theorem 201
  6.7.2 Global manifolds 210
  6.7.3 Center manifold 211
  6.7.4 Center manifold and reduced systems 213
  6.7.5 Hartman–Grobman theorem 218
6.8 Exercises 224

CHAPTER 7 Lyapunov functions **231**
7.1 Lyapunov method 231
  7.1.1 Stability of autonomous systems 238
  7.1.2 Time-varying systems; non-autonomous 249
7.2 Global asymptotic stability 256
7.3 Instability 259
7.4 ω-limit set 260
7.5 Connection between eigenvalues and Lyapunov functions 267
7.6 Exponential stability 273
7.7 Exercises 277

CHAPTER 8 Delay differential equations **287**
8.1 Introduction 287
8.2 Method of steps 289
8.3 Existence and uniqueness 292
8.4 Stability using Lyapunov functions 294
8.5 Stability using fixed point theory 299
  8.5.1 Neutral differential equations 300
  8.5.2 Neutral Volterra integro-differential equations 306
8.6 Exponential stability 309
  8.6.1 Instability 315
8.7 Existence of positive periodic solutions 319
8.8 Exercises 326

CHAPTER 9 New variation of parameters **329**
9.1 Applications to ordinary differential equations 329
  9.1.1 Periodic solutions 334

Copyright Elsevier 2022

9.2 Applications to delay differential equations 335
9.2.1 The main inversion 338
9.2.2 Variable time delay 340
9.3 Exercises 341

Bibliography 343
Index 347

Student Resources
For the Partial Solutions Manual, visit the companion site:
https://www.elsevier.com/books-and-journals/book-companion/9780323992800

Copyright Elsevier 2022

Preface

Differential equations are widely used by mathematicians, physicists, engineers, biologists, chemists, and scientists who work in relevant fields. They encounter the use of differential equations in the study of Newton's law of cooling, Maxwell's equations, Newton's laws of motion, fluid dynamics equations, equations in plasma dynamics, equations in stellar dynamics, Hook's law, Schrödinger's equation, acoustic wave equation, equations in chemical kinetics, equations in thermodynamics, Einstein's equations for general relativity, population models, epidemics, and so on. For many years, the author has been encouraged by the graduate students at the University of Dayton to write a concise and reader-friendly book on the subject of advanced differential equations. So this book grew out of lecture notes that the author has been constantly revising and using for a graduate course in differential equations. The book should serve as a two-semester graduate textbook in exploring the theory and applications of ordinary differential equations and differential equations with delays. It is intended for students who have basic knowledge of ordinary differential equations and real analysis. While writing this book, the author tried to balance rigor and presenting the most difficult material in an elementary format by adopting easier and friendlier notations that make the book accessible to a wide audience. It was the author's main intention to provide many examples to illustrate the theory conveyed in the theorems. The author made every effort to include contemporary topics such as the use of *fixed point theory* in several places to prove the existence and uniqueness, various notions of stability, and the existence of positive periodic solutions on Banach spaces. What makes the book appealing and distinguished from other books is the addition of Chapters 8 and 9 on delay differential equations with advanced topics. The author is convinced that any student who completes the whole book, especially Chapters 8 and 9, should be ready to carry on with meaningful research in delay differential systems.

Much of the pedagogical and mathematical development of this book is influenced by the author's style of presentation. The literature on differential equations is vast and well established, and some of the ideas found their way into this book.

Since stability is the central part of this book, namely by the Lyapunov method, we must mention some history. Lyapunov functions are named after Alexander Lyapunov, a Russian mathematician, who in 1892 published his book *The General Problem of Stability of Motion*. Lyapunov was the first to consider the modifications necessary in nonlinear systems to the linear theory of stability based on linearizing near a point of equilibrium. His work, initially published in Russian and then translated to French, received little attention for many years. Interest in Lyapunov stability started suddenly during the Cold War period when his method was found to be applicable to the stability of aerospace guidance systems, which typically contain

Copyright Elsevier 2022

strong nonlinearities not treatable by other methods. More recently, the concept of the Lyapunov exponent related to Lyapunov's first method of discussing stability has received wide interest in connection with chaos theory.

Chapter 1 deals with various introductory topics, including variation of parameters formula, metric spaces, and Banach spaces.

In Chapter 2 we introduce Gronwall's inequality that we make use of to prove the uniqueness of solutions. We introduce theorems on the existence and uniqueness of solutions, their dependence on initial data, and their continuation on maximal interval.

In Chapter 3 we introduce systems of differential equations. We briefly discuss how the existence and uniqueness theorems of Chapter 2 are extended to suit systems. Then we develop the notion of the fundamental matrix as a solution and utilize it to write solutions of non-homogeneous systems so that they can be analyzed.

Stability theory is the central part of this book. Chapters 4–8 are totally devoted to stability. In Chapter 4 we are mainly concerned with the stability of linear systems via the variation of parameters. The chapter also includes a nice section of Floquet theory with its application to Mathieu's equation.

Chapters 5 and 6 are deeply devoted to the study of the stability of linear systems, near-linear systems, perturbed systems, autonomous systems in the plane, and stability by linearization. Chapter 5 is ended with the study of Hamiltonians and gradient systems. We begin Chapter 6 by looking at stability diagrams in scalar equations and move into the study of bifurcations as it naturally arises while looking at stability. Bifurcation occurs when the dynamics abruptly change as certain parameters move across certain values. We end Chapter 6 by considering stable and unstable manifolds, which then delves into the Hartman–Grobman theorem. The theorem says that the behavior of a dynamical system in the domain near a hyperbolic equilibrium point is qualitatively the same as the behavior of its linearization near this equilibrium point, where hyperbolicity means that no eigenvalue of the linearization has a zero real part.

Chapter 7 delves deeply into the stability of general systems using Lyapunov functions. We prove general theorems regarding the stability of autonomous and non-autonomous systems by assuming the existence of such Lyapunov function. We touch on the notion of ω-limit set and its correlation to Lyapunov functions. The chapter is concluded with a detailed discussion on exponential stability.

Chapter 8 is solely devoted to the study of delay differential equations. It contains recent development in the research of delay differential equations. We begin the chapter by pointing out how basic results from ordinary differential equations are easily extended to delay differential equations. We introduce the *method of steps* and show how to piece together a solution. The transition of moving from ordinary differential equations to delay differential equations was made simple through the extension of Lyapunov functions to Lyapunov functionals. Then we move on to

Copyright Elsevier 2022

a whole new concept, *fixed point theory*. The use of fixed point theory alleviates some of the difficulties that arise from the use of Lyapunov functionals when studying stability. Later on, we apply fixed point theory to the study of stability and the existence of positive periodic solutions of neutral differential equations and neutral Volterra integro-differential equations, respectively. We end the chapter with the use of Lyapunov functionals to obtain necessary conditions for the exponential stability of Volterra integro-differential equations with finite delay.

Chapter 9 deals with current research concerning the use of a *new variation of parameters formula*. The objective is to introduce a new method for inverting first-order ordinary differential equations with time-delay terms to obtain a new variation of parameters formula that we use to study the stability, boundedness, and periodicity of general equations in ordinary and delay differential equations.

A combination of Chapters 1–3, 6, and 7 can be used to deliver a course on nonlinear systems for engineers.

The author has not attempted to give the historical origin of the theory, except in very rare cases. This resulted in the situation that not every reference listed in References is mentioned in the text or the body of the book.

Exercises play an essential learning tool of the course and accompany each chapter. They range from routine calculations to solving more difficult problems to open-ended ones. Students must read the relevant material before attempting to do the exercises.

I am indebted to Dr. Mohamed Aburakhis, who fully developed all the codes for all the figures in the book. I like to thank the hundreds of graduate students at the University of Dayton whom the author taught for the last 20 years and who helped the polishing and refining of the lecture notes, most of which have become this book.

A heartfelt appreciation to Jeff Hemmelgarn from the University of Dayton for carefully reading the whole book and pointing out many typos.

Youssef N. Raffoul
University of Dayton
Dayton, Ohio
June 2021

Copyright Elsevier 2022

CHAPTER

Preliminaries and Banach spaces

1

We briefly discuss basic topics of ordinary differential equations and provide examples that illustrate the need for a comprehensive and systematic theory of differential equations. In addition, we introduce metrics and Banach spaces, which we will use throughout the book.

1.1 Preliminaries

Let $\mathbb{R}$ denote the set of real numbers, let I be an interval in $\mathbb{R}$, and consider a function $x : I \to \mathbb{R}$. We say the function x has a derivative at a point $t^* \in I$ if the limit

$$\lim_{t\in I,\, t\to t^*} \frac{x(t) - x(t^*)}{t - t^*}$$

exists as a finite number. In this case, we adopt the notation

$$x'(t^*) = \lim_{t\in I,\, t\to t^*} \frac{x(t) - x(t^*)}{t - t^*},$$

where $x'(t^*)$ is the instantaneous rate of change of the function x at t^*. If t^* is one of the endpoints of the interval I, then the above definition of derivative becomes that of a one-sided derivative. If $x'(t^*)$ exists at every point $t^* \in I$, then we say that x is differentiable on I and write $x'(t)$. Throughout the book, we might use $\frac{dx}{dt}$ to indicate $x'(t)$. Similarly, if $x'(t)$ has a derivative function, then we call it the second derivative of the function $x(t)$ and denote it by $x''(t)$. For higher-order derivatives, we use the notations

$$x'''(t), x^{(4)}(t), \dots, x^{(n)}(t), \ \text{ or } \ \frac{d^n x}{dt^n}.$$

The *order of a differential equation* is defined by the highest derivative present in the equation. An *nth-order ordinary differential equation* is a functional relation of the form

$$F\big(t, x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \frac{d^3x}{dt^3}, \dots, \frac{d^nx}{dt^n}\big) = 0,\ t \in \mathbb{R} \tag{1.1.1}$$

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00007-3

Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

between the independent variable t and the dependent variable x, and its derivatives

$$\frac{dx}{dt}, \frac{d^2x}{dt^2}, \frac{d^3x}{dt^3}, ..., \frac{d^nx}{dt^n}.$$

Loosely speaking, by a solution of (1.1.1) on an interval I, we mean a function $x(t) = \varphi(t)$ such that

$$F\big(t, \varphi(t), \varphi'(t), \dots, \varphi^{(n)}(t)\big)$$

is defined for all $t \in I$ and

$$F\big(t, \varphi(t), \varphi'(t), \dots, \varphi^{(n)}(t)\big) = 0$$

for all $t \in I$. If we require, for some initial time $t_0 \in \mathbb{R}$, a solution $x(t)$ to satisfy the initial conditions

$$x(t_0) = a_0, \ \frac{dx}{dt}(t_0) = a_1, \ \frac{d^2x}{dt^2}(t_0) = a_2, \ \dots, \ \frac{d^{n-1}x}{dt^{n-1}}(t_0) = a_{n-1}, \tag{1.1.2}$$

for constants a_i, $i = 0, 1, 2, ..., n-1$, then (1.1.1) along with (1.1.2) is called an initial value problem (IVP).

Following the notation of (1.1.1), a first-order differential equation takes the form

$$F\big(t, x, \frac{dx}{dt}\big) = 0.$$

Hence, if we assume that we can solve for $\dfrac{dx}{dt}$, then we have

$$x'(t) = f(t, x)$$

for some function f that satisfies certain continuity conditions.

Let $x : I \to \mathbb{R}$ be a function. A differentiable function $z : I \to \mathbb{R}$ is called an *antiderivative* of the function x on the interval I if

$$z'(t) = x(t) \ \text{ for all } \ t \in I.$$

The set of all antiderivatives of x is denoted by

$$\int x(t)dt$$

and called the indefinite integral of the function x. When we calculate the indefinite integral $\int t^2 dt$, we in fact solve the first-order differential equation $x'(t) = t^2$. The family of its solutions is given by $t^3/3 + c$, where c is any constant. Thus we may write the solution as $x(t) = t^3/3 + c$, which is a one-parameter family of solutions, the same as the family of all the antiderivatives of t^3. Now, if we impose an initial condition on the differential equation, say $x(t_0) = x_0$, for some initial time

Copyright Elsevier 2022

t_0 and real number x_0, then the constant c is uniquely determined by the relation $x_0 = t_0^3/3 + c$. In this case the differential equation has the unique solution given by $x(t) = t^3/3 + x_0 - t_0^3/3$. However, without imposing the condition $x(t_0) = x_0$, the differential equation would have infinitely many solutions given by $x(t) = t^3/3 + c$.

Differential equations play an important role in modeling the behavior of physical systems such as falling bodies, vibration of a mass on a spring, and swinging pendulum. To illustrate the need for the theoretical study of differential equations and in particular nonlinear ones, we examine a few examples.

Consider the first-order differential equation

$$x'(t) = h(t)g(x), \ x(t_0) = x_0, \ t \geq t_0,$$

where $h, g : \mathbb{R} \to \mathbb{R}$ are continuous.

1. If $g(x_0) = 0$, then $x(t) = x_0$ is a solution.
2. In a region where $g(x) \neq 0$, we can divide by $g(x)$ so that

$$\frac{x'(t)}{g(x(t))} = h(t).$$

Separating the variables and then integrating both sides from t_0 to t give

$$\int_{t_0}^{t} \frac{x'(s)ds}{g(x(s))} = \int_{t_0}^{t} h(s)ds.$$

Using the transformation $u = x(s)$ with $x(t_0) = x_0$, we arrive at

$$\int_{x_0}^{x(t)} \frac{du}{g(u)} = \int_{t_0}^{t} h(s)ds.$$

If for some function G, we have $\frac{dG}{dx} = \frac{1}{g}$, then the above expression implies that

$$G(x(t)) - G(x(t_0)) = \int_{t_0}^{t} h(s)ds$$

or

$$x(t) = G^{-1}\Big(G(x_0) + \int_{t_0}^{t} h(s)ds\Big),$$

provided that the inverse of G exists. Note that the right side of the above expression depends on the initial time t_0 and the initial value x_0. Therefore to emphasize the dependence of solutions on the initial data, we may write a solution $x(t)$ in the form $x(t) = x(t, t_0, x_0)$.

Copyright Elsevier 2022

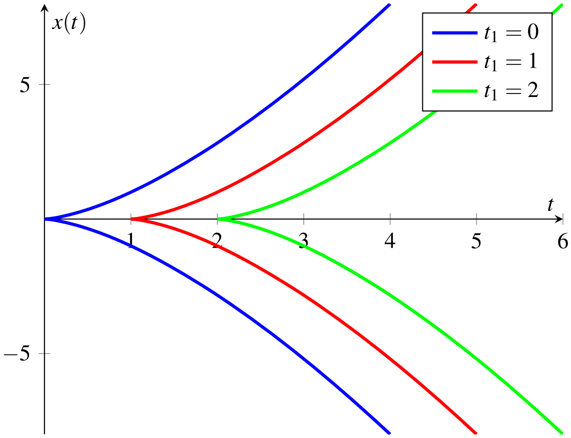


FIGURE 1.1

This example shows infinitely many solutions.

In the next example, we illustrate the existence of more than one solution. Consider the differential equation

$$x'(t) = \frac{3}{2}x^{1/3}(t),\ x(0) = 0,\ t \in \mathbb{R}.$$

It is clear that $x(t) = 0$ is a solution. Hence we may consider a solution $x_1(t) = 0$ and let

$$x_2(t) = \begin{cases} 0 & \text{for } t \leq 0, \\ t^{3/2} & \text{for } t > 0, \end{cases}$$

which is also a continuous and differentiable solution. Likewise, for $t_1 > 0$, we have

$$x_3(t) = \begin{cases} 0 & \text{for } t \leq t_1, \\ (t - t_1)^{3/2} & \text{for } t > t_1. \end{cases}$$

Continuing in this way, we see that the differential equation has infinitely many solutions. Similarly, if x is a solution, then $-x$ is also a solution (see Fig. 1.1).

In the next example, we show that solutions may escape (become unbounded) in finite time. The differential equation

$$x'(t) = x^3(t),\ x(t_0) = x_0 > 0,\ t \geq t_0$$

has the solution

$$x(t) = \sqrt{\frac{1}{\frac{1}{x_0^2} + 2(t_0 - t)}}.$$

Copyright Elsevier 2022

We can easily see that the solution is only valid for $t < t_0 + \dfrac{1}{2x_0^2}$ and becomes unbounded (escapes) as t approaches $t_0 + \frac{1}{2x_0^2}$ from the left.

The most important application in engineering problems is Newton's law

$$m\frac{d^2x(t)}{dt^2} = F[t, x(t), \frac{dx(t)}{dt}]$$

for the position $x(t)$ of a particle with mass m acted on by a force F, which may be a function to time t, the position $x(t)$, and the velocity $\frac{dx(t)}{dt}$. For example, if the force is only due to gravity, then we have the second-order differential equation

$$m\frac{d^2x(t)}{dt^2} = -mg,$$

which has the solution

$$x(t) = -\frac{1}{2}gt^2 + c_1t + c_2,$$

where c_1 and c_2 are constants that can be uniquely determined by specifying the position and velocity of the particle at some instant of time.

In the next example, we consider the problem of *leaky bucket*.

Example 1.1. We have a bucket with no water flowing into it and having a hole in the bottom. If Q is the volumetric flow rate, then $Q_{\text{in}} - Q_{\text{out}} = Q_{\text{stored}}$. Since $Q_{\text{in}} = 0$, we obtain $Q_{\text{stored}} = -Q_{\text{out}}$. Let $h(t)$ be the height of the water in the bucket at time t, and assume that the initial height at time $t = 0$ is h_0. Let $v(t)$ be the velocity of the leaked water (flow velocity). The volumetric flow rate Q_{stored} can be calculated by multiplying the velocity by the area of the bucket, that is, $Q_{\text{stored}} = A_{\text{bucket}}\frac{dh(t)}{dt}$. Similarly, Q_{out} is the flow velocity multiplied by the area of the hole A_{hole}. It follows that $Q_{\text{out}} = A_{\text{hole}}v(t)$. For fluids of height $h(t)$, the velocity of water coming out at the bottom is $v(t) = \sqrt{2gh(t)}$. By rearranging the terms, we arrive at the first-order differential equation in h given by

$$\frac{dh(t)}{dt} = -k\sqrt{h}, \quad h(0) = h_0, \tag{1.1.3}$$

where $k = \dfrac{A_{\text{hole}}}{A_{\text{bucket}}}\sqrt{2g} > 0$. By separating the variables in (1.1.3) and then integrating, we arrive at the solution

$$h(t) = \left(\sqrt{h_0} - \frac{kt}{2}\right)^2. \tag{1.1.4}$$

Note that the solution $h(t)$ given by (1.1.4) decreases from the initial height h_0. Moreover, at time $t^* = \frac{2\sqrt{h_0}}{k}$, the water is completely drained (by gravity), $h(t^*) = 0$, and

Copyright Elsevier 2022

the bucket will remain empty or the height will remain zero after $t^* = \frac{2\sqrt{h_0}}{k}$. Therefore we may write the solution as

$$h(t,0,h_0) = \begin{cases} \left(\sqrt{h_0} - \frac{kt}{2}\right)^2, & 0 \le t \le \frac{2\sqrt{h_0}}{k}, \\ 0, & t > \frac{2\sqrt{h_0}}{k}. \end{cases}$$

Another interesting application is the projectile problem that we analyze in the next section.

1.2 Escape velocity

Let M and R be the mass and radius of the Earth, respectively. We are interested in finding the smallest initial velocity for a mass m to exit the Earth's gravitational field, the so-called *escape velocity*. We assume that no external forces are acting on the system other than the gravitational force. Newton's universal gravitational law states that the force between two massive bodies is proportional to the product of the masses and inversely proportional to the square of the distance between them, where the mass of each body can be considered as concentrated at its center. For a mass m with position x above the surface of the Earth, the force F on the mass is given by

$$F = -G\frac{Mm}{(R+x)^2},$$

where G is the proportionality constant in the universal gravitational law. The minus sign means the force on the mass m points in the direction of decreasing x. By Newton's second law, force $F = m\frac{d^2x}{dt^2}$ (mass times acceleration). When $x = 0$, that is, at the Earth's surface, the gravitational force equals $-mg$, where g is the gravitational constant. Therefore

$$\frac{GM}{R^2} = g, \quad g \approx 9.8\,\mathrm{m/s}^2.$$

We have the second-order differential equation

$$\begin{aligned} \frac{d^2x}{dt^2} &= -G\frac{M}{(R+x)^2} \\ &= -\frac{g}{(1+\frac{x}{R})^2}, \end{aligned} \tag{1.2.1}$$

where the radius of the Earth is known to be $R \approx 6350\,\mathrm{km}$. We transform Eq. (1.2.1) into a first-order differential equation in terms of the velocity v by noting that $\frac{d^2x}{dt^2} = \frac{dv}{dt}$. If we write $v(t) = v(x(t))$, that is, considering the velocity of the mass m as a

Copyright Elsevier 2022

function of its distance above the Earth. Then using the chain rule, we have

$$\begin{aligned}\frac{dv}{dt} &= \frac{dv}{dx}\frac{dx}{dt} \\ &= v\frac{dv}{dx},\end{aligned}$$

since $v = \frac{dx}{dt}$. As a consequence, (1.2.1) becomes the first-order differential equation

$$\frac{dv}{dx} = -\frac{g}{(1+\frac{x}{R})^2}\frac{1}{v}.$$

Suppose the mass is shot vertically from the Earth's surface with initial velocity $v(x=0) = v_0$. Separating the variables and integrating both sides, we obtain

$$\int v dv = -g\int \frac{dx}{(1+\frac{x}{R})^2},$$

which gives

$$\frac{v^2}{2} = \frac{gR^2}{R+x} + c$$

for some constant c. Using the initial velocity condition, we find

$$c = -gR + \frac{v_0^2}{2}.$$

Substituting c back into the solution and simplifying, we arrive at the solution

$$v^2 = v_0^2 - \frac{gRx}{R+x}.$$

The escape velocity is defined as the minimum initial velocity v_0, such that the mass can escape to infinity. Therefore $v_0 = v_{\text{escape}}$ when $v \to 0$ as $x \to \infty$. Taking the limit, we have

$$0 = v_0^2 - \lim_{x\to\infty}\frac{gRx}{R+x} = v_0^2 - 2gR,$$

or

$$v_{\text{escape}}^2 = 2gR.$$

With $R \approx 6350\,\text{km}$ and $g = 127{,}008\,\text{km/h}^2$, we get $v_{\text{escape}} = \sqrt{2gR} \approx 40{,}000\,\text{km/h}$. In contrast, the muzzle velocity of a modern high-performance rifle is 4300 km/h, which is not enough for a bullet shot into the sky to escape from Earth's gravity.

Now we formally attempt to qualitatively analyze differential equations.

Copyright Elsevier 2022

Definition 1.2.1. Let D be an open subset of $\mathbb{R}^2$, and let $f : D \to \mathbb{R}$ be a continuous function. Let $(t_0, x_0) \in D$. We say that $x(t) = x(t, t_0, x_0)$ is a solution of

$$x' = f(t, x), \quad x(t_0) = x_0, \tag{1.2.2}$$

on an interval I if $t_0 \in I$, $x : I \to \mathbb{R}$ is differentiable, $(t, x(t)) \in D$ for $t \in I$, $x'(t) = f(t, x(t))$ for $t \in I$, and $x(t_0) = x_0$.

Definition 1.2.2. A solution $x(t)$ of (1.2.2) is said to be bounded on the interval $I = [0, \infty)$ if for any $t_0 \in [0, \infty)$ and $r > 0$, there exists a *positive number* $\alpha(t_0, r)$ depending on t_0 and r such that $|x(t, t_0, x_0)| \leq \alpha(t_0, r)$ for all $t \geq t_0$ and x_0 such that $|x_0| < r$. It is uniformly bounded if α is independent of the initial time t_0.

Definition 1.2.3. Let $x(t)$ and $y(t)$ be solutions of (1.2.2) with respect to initial conditions x_0 and y_0, respectively. The solution $x(t)$ is then said to be stable if for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon, t_0) > 0$ such that

$$|x(t) - y(t)| < \varepsilon \text{ whenever } |x_0 - y_0| < \delta.$$

Consider the linear differential equation

$$x'(t) = 1, \; x(t_0) = x_0. \tag{1.2.3}$$

It is easy to check that $x(t) = x_0 + (t - t_0)$ is the solution of (1.2.3). If $y(t)$ is another solution with $y(t_0) = y_0$, then we have $y(t) = y_0 + (t - t_0)$. For any $\varepsilon > 0$, let $\delta = \varepsilon$. Then

$$|x(t) - y(t)| = |x_0 + (t - t_0) - y_0 - (t - t_0)| = |x_0 - y_0| < \varepsilon$$

whenever $|x_0 - y_0| < \delta$. Hence the solution $x(t)$ is stable but clearly unbounded. This simple example shows that the notion of a solution being unbounded does not automatically imply that the same solution is unstable with respect to another solution starting at a different initial point. In Chapters 6 and 7 we will discuss boundedness and stability in more detail. The previous examples illustrate the need for a coherent theory for addressing the following issues:

1. Existence and uniqueness.
2. Boundedness and stability.
3. The dependence of solutions on the initial data.

1.3 Applications to epidemics

The *law of mass action* is a useful concept that describes the behavior of a system that consists of many interacting parts, such as molecules, that react with each other, or viruses that are passed along from a population of infected individuals to non-immune ones. The law of mass action was first derived for chemical systems but

Copyright Elsevier 2022

subsequently found wide use in epidemiology and ecology. To describe the law of mass action, we assume that m substances $s_1, s_2, \dots, s_m$ together form a product with concentration p. Then the law of mass action states that $\frac{dp}{dt}$ is proportional to the product of the m concentrations s_i, $i = 1, \dots, m$, that is,

$$\frac{dp}{dt} = k s_1 s_2 \dots s_m.$$

Suppose we have a homogeneous population of fixed size divided into two groups. Those who have the disease are called infective, and those who do not have the disease are called susceptible. Let $S = S(t)$ be the susceptible portion of the population, and let $I = I(t)$ be the infective portion. Then by assumption, we may normalize the population and have $S + I = 1$. We further assume that the dynamics of this epidemic satisfies the law of mass action. Hence, for positive constant λ, we have the nonlinear differential equation

$$I'(t) = \lambda S I. \tag{1.3.1}$$

Let $I(0) = I_0$, $0 < I(0) < 1$, be a given initial condition. By substituting $S = 1 - I$ into (1.3.1) it follows that

$$I'(t) = \lambda I(1 - I), \quad I(0) = I_0. \tag{1.3.2}$$

If we can solve (1.3.2) for $I(t)$, then $S(t)$ can be found from the relation $I + S = 1$. We separate the variables in (1.3.2) and obtain

$$\frac{dI}{I(1-I)} = \lambda dt.$$

Using partial fraction on the left side of the equation and then integrating both sides yield

$$\text{Ln}|I| - \text{Ln}|1 - I| = \lambda t + c,$$

or for positive constant c_1, we have

$$I(t) = \frac{c_1 e^{\lambda t}}{1 + c_1 e^{\lambda t}}.$$

Applying $I(0) = I_0$ gives the solution

$$I(t) = \frac{I_0 e^{\lambda t}}{1 - I_0 + I_0 e^{\lambda t}}. \tag{1.3.3}$$

Now for $0 < I(0) < 1$, the solution given by (1.3.3) is increasing with time as expected. Moreover, using L'Hospital's rule, we have

$$\lim_{t \to \infty} I(t) = \lim_{t \to \infty} \frac{I_0 e^{\lambda t}}{1 - I_0 + I_0 e^{\lambda t}} = 1.$$

Copyright Elsevier 2022

Hence the infection will grow, and everyone in the population will eventually get infected.

1.4 Metrics and Banach spaces

This section is devoted to introductory materials related to Cauchy sequences, metric spaces, contraction, compactness, contraction mapping principle, and Banach spaces. Materials in this section will be of use in several places of the book, especially in Chapters 2, 8, and 9. Throughout the book, by $C(I, \mathbb{R}^n)$, we denote the space of all continuous functions $f : I \to \mathbb{R}^n$ on an interval I, possibly infinite.

Definition 1.4.1. A pair (E, ρ) is a metric space if E is a set and $\rho : E \times E \to [0, \infty)$ such that for all y, z, and u in E, we have

(a) $\rho(y, z) \geq 0$, $\rho(y, y) = 0$, and $\rho(y, z) = 0$ implies $y = z$;
(b) $\rho(y, z) = \rho(z, y)$; and
(c) $\rho(y, z) \leq \rho(y, u) + \rho(u, z)$.

The next definition is concerned with Cauchy sequences.

Definition 1.4.2. (Cauchy sequence) A sequence $\{x_n\} \subseteq E$ is a Cauchy sequence if for each $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $n, m > N \implies \rho(x_n, x_m) < \varepsilon$.

Complete metric spaces play a major role when showing that a fixed point belongs to the metric space of interest.

Definition 1.4.3. (Completeness of metric space) A metric space (E, ρ) is said to be complete if every Cauchy sequence in E converges to a point in E.

Definition 1.4.4. A set L in a metric space (E, ρ) is compact if each sequence in L has a subsequence with a limit in L.

Definition 1.4.5. Let $\{f_n\}$ be a sequence of real functions with $f_n : [a, b] \to \mathbb{R}$.

1. $\{f_n\}$ is uniformly bounded on $[a, b]$ if there exists $M > 0$ such that $|f_n(t)| \leq M$ for all $n \in \mathbb{N}$ and $t \in [a, b]$.
2. $\{f_n\}$ is equicontinuous at t_0 if for each $\varepsilon > 0$, there exists $\delta > 0$ such that for all $n \in \mathbb{N}$, if $t \in [a, b]$ and $|t_0 - t| < \delta$, then $|f_n(t_0) - f_n(t)| < \varepsilon$. Also, $\{f_n\}$ is equicontinuous if $\{f_n\}$ is equicontinuous at each $t_0 \in [a, b]$.
3. $\{f_n\}$ is uniformly equicontinuous if for each $\varepsilon > 0$, there exists $\delta > 0$ such that for all $n \in \mathbb{N}$, if $t_1, t_2 \in [a, b]$ and $|t_1 - t_2| < \delta$, then $|f_n(t_1) - f_n(t_2)| < \varepsilon$.

It is easy to see that $\{f_n\} = \{x^n\}$ is not an equicontinuous sequence of functions on $[0, 1]$ but each f_n is uniformly continuous.

Proposition 1.1. *[Cauchy criterion for uniform convergence] If $\{F_n\}$ is a sequence of bounded functions that is Cauchy in the uniform norm, then $\{F_n\}$ converges uniformly.*

Copyright Elsevier 2022

Definition 1.4.6. A real-valued function f defined on $E \subseteq \mathbb{R}$ is said to be Lipschitz continuous with Lipschitz constant K if $|f(x) - f(y)| \leq K|x - y|$ for all $x, y \in E$.

It is easy to see that the function $f(x) = x^2$ is not Lipschitz on $\mathbb{R}$. This is due to the fact that for any x and y in $\mathbb{R}$, we have that $f(x) - f(y) = |x^2 - y^2| = |x + y||x - y|$, and so there is no constant K such that $|x^2 - y^2| \leq K|x - y|$. Definition 1.4.6 implies that f is globally Lipschitz since the constant K is uniform for all x and y in $\mathbb{R}$.

Remark 1.1. It is an easy exercise that a Lipschitz continuous function is uniformly continuous. Also, if each f_n in a sequence of functions $\{f_n\}$ has the same Lipschitz constant, then the sequence is uniformly equicontinuous.

Lemma 1.1. *If $\{f_n\}$ is an equicontinuous sequence of functions on a closed bounded interval, then $\{f_n\}$ is uniformly equicontinuous.*

Proof. Suppose $\{f_n\}$ is equicontinuous on $[a, b]$. Let $\varepsilon > 0$. For each $x \in K$, let $\delta_x > 0$ be such that $|y - x| < \delta_x \implies |f_n(x) - f_n(y)| < \varepsilon/2$ for all $n \in \mathbb{N}$. The collection $\{B(x, \delta_x/2) : x \in [a, b]\}$ is an open cover of $[a, b]$, so it has a finite subcover $\{B(x_i, \delta_{x_i}/2) : i = 1, \ldots, k\}$. Let $\delta = \min\{\delta_{x_i}/2 : i = 1, \ldots, k\}$. Then, if $x, y \in [a, b]$ with $|x - y| < \delta$, then there is some i with $x \in B(x_i, \delta_{x_i}/2)$. Since $|x - y| < \delta \leq \delta_{x_i}/2$, we have $|x_i - y| \leq |x_i - x| + |x - y| < \delta_{x_i}/2 + \delta_{x_i}/2 = \delta_{x_i}$. Hence $|x_i - y| < \delta_{x_i}$ and $|x_i - x| < \delta_{x_i}$. So, for any $n \in \mathbb{N}$, we have $|f_n(x) - f_n(y)| \leq |f_n(x) - f_n(x_i)| + |f_n(x_i) - f_n(y)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. So $\{f_n\}$ is uniformly equicontinuous. ☐

The next theorem gives us the main method of proving compactness in the spaces we are interested in.

Theorem 1.4.1. *[Ascoli–Arzelà] If $\{f_n(t)\}$ is a uniformly bounded and equicontinuous sequence of real-valued functions on an interval $[a, b]$, then there is a subsequence that converges uniformly on $[a, b]$ to a continuous function.*

Proof. Since $\{f_n(t)\}$ is equicontinuous on $[a, b]$, by Lemma (1.1) $\{f_n(t)\}$ is uniformly equicontinuous. Let $t_1, t_2, \ldots$ be a listing of the rational numbers in $[a, b]$ (note that the set of rational numbers is countable, so this enumeration is possible). The sequence $\{f_n(t_1)\}_{n=1}^{\infty}$ is a bounded sequence of real numbers (since $\{f_n\}$ is uniformly bounded), so it has a subsequence $\{f_{n_k}(t_1)\}$ converging to a number, which we denote $\phi(t_1)$. It will be more convenient to represent this subsequence without subsubscripts, so we write f_k^1 for f_{n_k} and switch the index from k to n. So the subsequence is written as $\{f_n^1(t_1)\}_{n=1}^{\infty}$. Now, the sequence $\{f_n^1(t_2)\}$ is bounded, so it has a convergent subsequence, say $\{f_n^2(t_2)\}$, with limit $\phi(t_2)$. We continue in this way obtaining a sequence of sequences $\{f_n^m(t)\}_{n=1}^{\infty}$ (one sequence for each m), each of which is a subsequence of the previous one. Furthermore, we have $f_n^m(t_m) \to \phi(t_m)$ as $n \to \infty$ for each $m \in \mathbb{N}$. Now, consider the "diagonal" functions defined $F_k(t) = f_k^k(t)$. Since $f_n^m(t_m) \to \phi(t_m)$, it follows that $F_r(t_m) \to \phi(t_m)$ as $r \to \infty$ for each $m \in \mathbb{N}$ (in other words, the sequence $\{F_r(t)\}$ converges pointwise at

Copyright Elsevier 2022

each t_m). We now show that $\{F_k(t)\}$ converges uniformly on $[a,b]$ by showing that it is Cauchy in the uniform norm. Let $\varepsilon > 0$, and let $\delta > 0$ be as in the definition of uniformly equicontinuous for $\{f_n(t)\}$ applied with $\varepsilon/3$. Divide $[a,b]$ into p intervals, where $p > \frac{b-a}{\delta}$. Let ξ_j be a rational number in the jth interval for $j = 1, \dots, p$. Recall that $\{F_r(t)\}$ converges at each of the points ξ_j, since they are rational numbers. So, for each j, there is $M_j \in \mathbb{N}$ such that $|F_r(\xi_j) - F_s(\xi_j)| < \varepsilon/3$ whenever $r, s > M_j$. Let $M = \max\{M_j : j = 1, \dots, p\}$. If $t \in [a,b]$, then it is in one of the p intervals, say the jth. So $|t - \xi_j| < \delta$, and thus $|f_r^r(t) - f_r^r(\xi_j)| = |F_r(t) - F_r(\xi_j)| < \varepsilon/3$ for every r. Also, if $r, s > M$, then $|F_r(\xi_j) - F_s(\xi_j)| < \varepsilon/3$ (since M is the maximum of the M_i). So for $r, s > M$, we have

$$\begin{aligned}
|F_r(t) - F_s(t)| &= |F_r(t) - F_r(\xi_j) + F_r(\xi_j) - F_s(\xi_j) + F_s(\xi_j) - F_s(t)| \\
&\le |F_r(t) - F_r(\xi_j)| + |F_r(\xi_j) - F_s(\xi_j)| + |F_s(\xi_j) - F_s(t)| \\
&\le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

By the Cauchy criterion for convergence, the sequence $\{F_r(t)\}$ converges uniformly on $[a,b]$. Since each $F_r(t)$ is continuous, the limit function $\phi(t)$ is also continuous. ☐

Remark 1.2. The Ascoli–Arzelà theorem can be generalized to a sequence of functions from $[a,b]$ to $\mathbb{R}^n$. Apply the Ascoli–Arzelà theorem to the first coordinate function to get a uniformly convergent subsequence. Then apply the theorem again, this time to the corresponding subsequence of functions restricted to the second coordinate, getting a subsubsequence, and so on.

The next criterion, known as the Weierstrass M-test plays an important role in showing the existence of solutions.

Lemma 1.2. *(Weierstrass M-test) Let $\{f_n\}$ be a sequence of functions defined on a set E. Suppose that for all $n = 1, \dots,$ there is a constant M_n such that $|f_n(t)| \le M_n$ for all $t \in E$. If*

$$\sum_{n=1}^{\infty} M_n < \infty, \ \ \textit{then} \ \ \sum_{n=1}^{\infty} f_n(t)$$

converges absolutely and uniformly on the E.

We remark that the Weierstrass M-test can be easily generalized if the domain of the sequence of functions is a subset of Banach space endowed with an appropriate norm.

Here is an example of the Weierstrass M-test.

Example 1.2. For $n = 1, 2, \dots,$ define the sequence of functions $\{f_n\}$ on $\mathbb{R}$ by $f_n(t) = \dfrac{1}{t^2 + n^2}$. Then $|f_n(t)| = |\dfrac{1}{t^2+n^2}| \le \dfrac{1}{n^2} := M_n$ for all $t \in \mathbb{R}$ and $n \ge 1$. Since

Copyright Elsevier 2022

the series $\sum_{n=1}^{\infty}\frac{1}{n^2}$ converges, by the Weierstrass M-test the series $\sum_{n=1}^{\infty}\frac{1}{t^2+n^2}$ converges uniformly on $\mathbb{R}$. Moreover, as each term of the series is continuous and the convergence is uniform, the sum function is also continuous. (As the uniform limit of continuous functions is continuous.)

Here is another example with a simple twist to it.

Example 1.3. We prove that the series

$$\sum_{n=1}^{\infty}\frac{n^2+x^4}{n^4+x^2}$$

converges to a continuous function $f:\mathbb{R}\to\mathbb{R}$.

Let c be a positive constant. Then for all $x\in[-c,c]$, we have that

$$\left|\frac{n^2+x^4}{n^4+x^2}\right|\le\frac{n^2+x^4}{n^4}\le\frac{1}{n^2}+\frac{c^4}{n^4}:=M_n.$$

On the other hand, the series

$$\sum_{n=1}^{\infty}M_n=\sum_{n=1}^{\infty}\frac{1}{n^2}+c^4\sum_{n=1}^{\infty}\frac{1}{n^4}$$

converges, so Weierstrass M-test implies that the series converges uniformly to a function f on the bounded interval $[-c,c]$. Each term in the series is continuous and since the uniform limit of continuous functions is continuous, the limit function f is continuous on $[-c,c]$ for every $c>0$. Now since every $x\in\mathbb{R}$ lies in such an interval for sufficiently large c, it follows that f is continuous on $\mathbb{R}$. Note that the series does not converge uniformly on $\mathbb{R}$, so we cannot use the argument that the sum is continuous on $\mathbb{R}$ because the series converges uniformly on $\mathbb{R}$.

Banach spaces form an important class of metric spaces. We now define Banach spaces in several steps.

Definition 1.4.7. A triple $(V,+,\cdot)$ is said to be a linear (or vector) space over a field F if V is a set and the following are true.

1. Properties of $+$
   - **a.** $+$ is a function from $V\times V$ to V. Outputs are denoted $x+y$.
   - **b.** for all $x,y\in V$, $x+y=y+x$ ($+$ is commutative).
   - **c.** for all $x,y,w\in V$, $x+(y+w)=(x+y)+w$ ($+$ is associative).
   - **d.** there is a unique element of V, which we denote 0, such that for all $x\in V$, $0+x=x+0=x$ (additive identity).
   - **e.** for each $x\in V$, there is a unique element of V, which we denote $-x$, such that $x+(-x)=-x+x=0$ (additive inverse).

Copyright Elsevier 2022

2. Scalar multiplication
  a. $\cdot$ is a function from $F \times V$ to V. Outputs are denoted $\alpha \cdot x$ or αx.
  b. for all $\alpha, \beta \in F$ and $x \in V$, $\alpha(\beta x) = (\alpha\beta)x$.
  c. for all $x \in V$, $1 \cdot x = x$.
  d. for all $\alpha, \beta \in F$ and $x \in V$, $(\alpha + \beta)x = \alpha x + \beta x$.
  e. for all $\alpha \in F$ and $x, y \in V$, $\alpha(x + y) = \alpha x + \alpha y$.

Commonly, the real numbers and complex numbers are fields in the above definition. For our purposes, we only consider the field of real numbers $F = \mathbb{R}$.

Definition 1.4.8. (Normed spaces) A vector space $(V, +, \cdot)$ is a *normed space* if for each $x \in V$, there is a nonnegative real number $\|x\|$, called the *norm* of x, such that for all $x, y \in V$ and $\alpha \in \mathbb{R}$,

1. $\|x\| = 0$ if and only if $x = 0$,
2. $\|\alpha x\| = |\alpha|\|x\|$,
3. $\|x + y\| \leq \|x\| + \|y\|$.

Remark 1.3. A norm on a vector space always defines a metric $\rho(x, y) = \|x - y\|$ on the vector space. Given a metric ρ defined on a vector space, it is tempting to define $\|v\| = \rho(v, 0)$. But this is not always a norm.

Definition 1.4.9. A Banach space is a complete normed vector space, that is, a vector space $(X, +, \cdot)$ with norm $\|\cdot\|$ for which the metric $\rho(x, y) = \|x - y\|$ is complete.

Example 1.4. The space $(\mathbb{R}^n, +, \cdot)$ over the field $\mathbb{R}$ is a vector space (with the usual vector addition $+$ and scalar multiplication $\cdot$), and there are many suitable norms for it. For example, if $x = (x_1, x_2, \ldots, x_n)$, then

1. $\|x\| = \max_{1 \leq i \leq n} |x_i|$,
2. $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$,
3. $\|x\| = \sum_{i=1}^{n} |x_i|$,
4. $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$, $p \geq 1$

are all suitable norms. Norm 2 is the Euclidean norm: the norm of a vector is its Euclidean distance to the zero vector, and the metric defined from this norm is the usual Euclidean metric. Norm 3 generates the "taxi-cab" metric on $\mathbb{R}^2$, and Norm 4 is the l^p norm.

Throughout the book, it should cause no confusion to use $|\cdot|$ instead of $||\cdot||$ to denote a particular norm.

Remark 1.4. Consider the vector space $(\mathbb{R}^n, +, \cdot)$ as a metric space with its metric defined by $\rho(x, y) = \|x - y\|$, where $\|\cdot\|$ is any of the norms in Example 1.4. The

Copyright Elsevier 2022

completeness of this metric space comes directly from the completeness of $\mathbb{R}$, and hence $(\mathbb{R}^n, \|\cdot\|)$ is a Banach space.

Remark 1.5. In the Euclidean space $\mathbb{R}^n$, compactness is equivalent to closedness and boundedness (Heine–Borel theorem). In fact, the metrics generated from any of the norms in Example 1.4 are equivalent in the sense that they generate the same topologies. Moreover, compactness is equivalent to closedness and boundedness in each of those metrics.

Example 1.5. Let $C([a,b],\mathbb{R}^n)$ denote the space of all continuous functions $f:[a,b]\to\mathbb{R}^n$.

1. $C([a,b],\mathbb{R}^n)$ is a vector space over $\mathbb{R}$.
2. If $\|f\| = \max\limits_{a\le t\le b} |f(t)|$, where $|\cdot|$ is a norm on $\mathbb{R}^n$, then $(C([a,b],\mathbb{R}^n), \|\cdot\|)$ is a Banach space.
3. Let M and K be two positive constants and define

$$L = \{f \in C([a,b],\mathbb{R}^n) : \|f\| \le M; |f(u) - f(v)| \le K|u-v|\}.$$

Then L is compact.

Proof (of part 3). Let $\{f_n\}$ be any sequence in L. The functions are uniformly bounded by M and have the same Lipschitz constant, K. So the sequence is uniformly equicontinuous. By the Ascoli–Arzelà theorem there is a subsequence $\{f_{n_k}\}$ that converges uniformly to a continuous function $f:[a,b]\to\mathbb{R}^n$. We now show that $f\in L$. Well, $|f_n(t)|\le M$ for each $t\in[a,b]$, so $|f(t)|\le M$ for each $t\in[a,b]$ and hence $\|f\|\le M$. Now fix $u,v\in[a,b]$ and $\varepsilon>0$. Since $\{f_{n_k}\}$ converges uniformly to f, there is $N\in\mathbb{N}$ such that $|f_{n_k}(t)-f(t)|<\varepsilon/2$ for all $t\in[a,b]$ and $k\ge N$. So fixing any $k\ge N$, we have

$$\begin{aligned}|f(u)-f(v)| &= |f(u)-f_{n_k}(u)+f_{n_k}(u)-f_{n_k}(v)+f_{n_k}(v)-f(v)|\\ &\le |f(u)-f_{n_k}(u)|+|f_{n_k}(u)-f_{n_k}(v)|+|f_{n_k}(v)-f(v)|\\ &< \varepsilon/2+K|u-v|+\varepsilon/2 = K|u-v|+\varepsilon.\end{aligned}$$

Since $\varepsilon>0$ was arbitrary, $|f(u)-f(v)|\le K|u-v|$. Hence $f\in L$. We have demonstrated that $\{f_n\}$ has a subsequence converging to an element of L. Hence L is compact. □

Example 1.6. Consider $\mathbb{R}$ as a vector space over $\mathbb{R}$ and define the metric $d(x,y) = \dfrac{|x-y|}{1+|x-y|}$. For each $x\in\mathbb{R}$, we define $\|x\| = d(x,0)$. Explain why $\|\cdot\|$ is not a norm on $\mathbb{R}$.

Example 1.7. Let $\phi:[a,b]\to\mathbb{R}^n$ be continuous, and let S be the set of continuous functions $f:[a,c]\to\mathbb{R}^n$ with $c>b$ and $f(t)=\phi(t)$ for $a\le t\le b$. Define $\rho(f,g) = \|f-g\| = \sup\limits_{a\le t\le c<} |f(t)-g(t)|$ for $f,g\in S$. Then (S,ρ) is a complete metric space but not a Banach space since $f+g$ is not in S.

Copyright Elsevier 2022

Example 1.8. Let (S, ρ) be the space of continuous bounded functions $f : (-\infty, 0] \to \mathbb{R}$ with $\rho(f, g) = \|f - g\| = \sup_{-\infty < t \leq 0} |f(t) - g(t)|$. Then:

1. (S, ρ) is a Banach space.
2. The set $L = \{f \in S : \|f\| \leq 1, |f(u)f(v)| \leq |u - v|\}$ is not compact in (S, ρ).

Proof (of 2). Consider the sequence of functions defined

$$f_n(t) = \begin{cases} 0 & \text{if } t \leq -n, \\ \frac{t}{n} + 1 & \text{if } -n < t \leq 0. \end{cases}$$

Then the sequence converges pointwise to $f = 1$, but $\rho(f_n, f) = 1$ for all $n \in \mathbb{N}$. So there is no subsequence of $\{f_n\}$ converging in the norm $\|\cdot\|$ (i.e., converging uniformly) to f. ☐

Example 1.9. Let (S, ρ) be the space of continuous functions $f : (-\infty, 0] \to \mathbb{R}^n$ with

$$\rho(f, g) = \sum_{n=1}^{\infty} 2^{-n} \rho_n(f, g)/\{1 + \rho_n(f, g)\},$$

where

$$\rho_n(f, g) = \max_{-n \leq s \leq 0} |f(s) - g(s)|,$$

and $|\cdot|$ is the Euclidean norm on $\mathbb{R}^n$. Then:

1. (S, ρ) is a complete metric space. The distance between all functions is bounded by 1.
2. $(S, +, \cdot)$ is a vector space over $\mathbb{R}$.
3. (S, ρ) is not a Banach space because ρ does not define a norm, since $\rho(x, 0) = \|x\|$ does not satisfy $\|\alpha x\| = |\alpha|\|x\|$.
4. Let M and K be given positive constants. Then the set

$$L = \{f \in S : \|f\| \leq M \text{ on } (-\infty, 0], |f(u) - f(v)| \leq K|u - v|\}$$

is compact in (S, ρ).

Proof (of 4). Let $\{f_n\}$ be a sequence in L. It is clear that if $f_n \to f$ uniformly on compact subsets of $(-\infty, 0]$, then we have $\rho(f_n, f) \to 0$ as $n \to \infty$. Let us begin by considering $\{f_n\}$ on $[-1, 0]$. Then the sequence is uniformly bounded and equicontinuous, and so there is a subsequence, say $\{f_n^1\}$, converging uniformly to some continuous f on $[-1, 0]$. Moreover, the argument of Example 1.5 shows that $|f(t)| \leq M$ and $|f(u) - f(v)| \leq K|u - v|$. Next, we consider $\{f_n^1\}$ on $[-2, 0]$. Then the sequence is uniformly bounded and equicontinuous, and so there is a subsequence, say $\{f_n^2\}$, converging uniformly, say, to some continuous f on $[-2, 0]$.

Copyright Elsevier 2022

Continuing this way, we arrive at $F_n = f_n^n$, which has a subsequence of $\{f_n\}$ converging uniformly on compact subsets of $(-\infty, 0]$ to a function $f \in L$. This proves that L is compact. □

We leave the proof of next result to the reader.

Theorem 1.4.2. *Let $g : (-\infty, 0] \to [1, \infty)$ be a continuous strictly decreasing function with $g(0) = 1$ and $g(r) \to \infty$ as $r \to -\infty$. Let $(S, |\cdot|_g)$ be the space of continuous functions $f : (-\infty, 0] \to \mathbb{R}^n$ for which*

$$|f|_g := \sup_{-\infty < t \le 0} \frac{|f(t)|}{|g(t)|}$$

exists. Then:

1. *$(S, |\cdot|_g)$ is a Banach space.*
2. *Let M and K be given positive constants. Then the set*

$$L = \{f \in S : \|f\| \le M \text{ on } (-\infty, 0], |f(u) - f(v)| \le K|u - v|\}$$

is compact in (S, ρ).

Definition 1.4.10. Let (E, ρ) be a metric space, and let $D : E \to E$. The operator or mapping D is a contraction if there exists $\alpha \in (0, 1)$ such that

$$\rho\Big(D(x), D(y)\Big) \le \alpha\rho(x, y).$$

The next theorem is known by the name of Caccioppoli theorem, or the Banach contraction mapping principle [9]. A proof can be found in many places, such as Burton [15] or Smart [58].

Theorem 1.4.3. *(Contraction mapping principle) Let (E, ρ) be a complete metric space, and let $D : E \to E$ be a contraction operator. Then there exists a unique $\phi \in E$ such that $D(\phi) = \phi$. Moreover, if $\psi \in E$ and if $\{\psi_n\}$ is defined inductively by $\psi_1 = D(\psi)$ and $\psi_{n+1} = D(\psi_n)$, then $\psi_n \to \phi$, the unique fixed point.*

Proof. Let $y_0 \in E$ and define the sequence $\{y_n\}$ in E by $y_1 = Dy_0$, $y_2 = Dy_1 = D(Dy_0) = D^2 y_0, \ldots, y_n = Dy_{n-1} = D^n y_0$. Next, we show that $\{y_n\}$ is a Cauchy sequence. Indeed, if $m > n$, then

$$\begin{aligned}
\rho(y_n, y_m) &= \rho(D^n y_0, D^m y_0) \\
&\le \alpha\rho(D^{n-1} y_0, D^{m-1} y_0) \\
&\vdots \\
&\le \alpha^n \rho(y_0, y_{m-1})
\end{aligned}$$

Copyright Elsevier 2022

$$
\begin{aligned}
&\le \alpha^n \big\{\rho(y_0, y_1) + \rho(y_1, y_2) + \ldots + \rho(y_{m-n-1}, y_{m-n})\big\} \\
&\le \alpha^n \big\{\rho(y_0, y_1) + \alpha\rho(y_0, y_1) + \ldots + \alpha^{m-n-1}\rho(y_0, y_1)\big\} \\
&\le \alpha^n \rho(y_0, y_1)\big\{1 + \alpha + \ldots + \alpha^{m-n-1}\big\} \\
&\le \alpha^n \rho(y_0, y_1)\frac{1}{1-\alpha}.
\end{aligned}
$$

Thus, since $\alpha \in (0, 1)$, we have that

$$\rho(y_n, y_m) \to 0 \text{ as } n \to \infty.$$

This shows that the sequence $\{y_n\}$ is Cauchy. Since (E, ρ) is a complete metric space, $\{y_n\}$ has a limit, say y in E. Since the mapping D is continuous, we have that

$$D(y) = D(\lim_{n\to\infty} y_n) = \lim_{n\to\infty} D(y_n) = \lim_{n\to\infty} y_{n+1} = y,$$

and y is a fixed point. It remains to show that y is unique. Let $x, y \in E$ be such that $D(x) = x$ and $D(y) = y$. Then

$$0 \le \rho(x, y) = \rho\big(D(x), D(y)\big) \le \alpha\rho(x, y),$$

which implies that

$$0 \le (1-\alpha)\rho(x, y) \le 0.$$

Since $1 - \alpha \neq 0$, we must have $\rho(x, y) = 0$, and hence $x = y$. This completes the proof. □

Another form of the contraction mapping principle:

Theorem 1.4.4. *(Contraction mapping principle, Banach fixed point theorem) Let (E, ρ) be a complete metric space, and let $P : E \to E$ be such that P^m is a contraction for some fixed positive integer m. Then there is a unique $x \in E$ such that $P(x) = x$.*

1.5 Variation of parameters

In this section we develop the variation of parameters formula, which we use, in one form or another, throughout this book. Consider the nonhomogeneous differential equation

$$x'(t) = a(t)x(t) + g(t, x(t)),\ x(t_0) = x_0,\ t \ge t_0 \ge 0, \tag{1.5.1}$$

where $g \in C([0, \infty) \times \mathbb{R}, \mathbb{R})$ and $a \in C([0, \infty), \mathbb{R})$.

Copyright Elsevier 2022

Multiplying both sides of (1.5.1) by the integrating factor $e^{-\int_{t_0}^t a(u)du}$ and observing that

$$\frac{d}{dt}\Big(x(t)e^{-\int_{t_0}^t a(u)du}\Big) = x'(t)e^{-\int_{t_0}^t a(u)du} - a(t)x(t)e^{-\int_{t_0}^t a(u)du},$$

we arrive at

$$\frac{d}{dt}\Big(x(t)e^{-\int_{t_0}^t a(u)du}\Big) = g(t, x(t))e^{-\int_{t_0}^t a(u)du}.$$

Integration of the above expression from t_0 to t and using $x(t_0) = x_0$ yield

$$x(t)e^{-\int_{t_0}^t a(u)du} = x_0 + \int_{t_0}^t g(s, x(s))e^{-\int_{t_0}^s a(u)du} ds,$$

from which we get

$$x(t) = x_0 e^{\int_{t_0}^t a(u)du} + \int_{t_0}^t g(s, x(s))e^{\int_s^t a(u)du} ds,\ t \ge t_0 \ge 0. \tag{1.5.2}$$

It can be easily shown that if $x(t)$ satisfies (1.5.2), then it satisfies (1.5.1). Expression (1.5.2) is known as the variation of parameters formula. Note that (1.5.2) is a functional equation in x since the integrand is a function of x. If we replace the function g with a function $h \in C([0, \infty), \mathbb{R})$, then (1.5.2) takes the special form

$$x(t) = x_0 e^{\int_{t_0}^t a(u)du} + \int_{t_0}^t h(s)e^{\int_s^t a(u)du} ds,\ t \ge t_0 \ge 0. \tag{1.5.3}$$

Another special form of (1.5.2) is that if the function $a(t)$ is constant for all $t \ge t_0$ and g is replaced with h as before, then from (1.5.3), we have that

$$x(t) = x_0 e^{a(t-t_0)} + \int_{t_0}^t e^{a(t-s)} h(s) ds,\ t \ge t_0 \ge 0. \tag{1.5.4}$$

It is easy to compute, using (1.5.4), that the differential equation

$$x'(t) = 2x(t) + t,\ x(0) = 3,$$

has the solution

$$x(t) = \frac{13}{4}e^{2t} - \frac{t}{2} - \frac{1}{4}.$$

Remark 1.6. The variation of parameters formula given by (1.5.2) is valid for

$$g \in C(\mathbb{R} \times \mathbb{R}, \mathbb{R}) \quad \text{and} \quad a \in C(\mathbb{R}, \mathbb{R}).$$

Next, we consider the following application regarding the variation of parameters.

Copyright Elsevier 2022

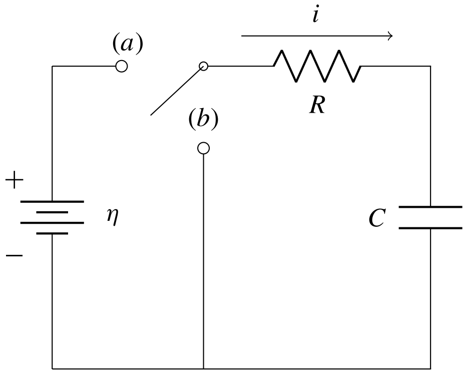


FIGURE 1.2

RC circuit.

1.5.1 RC circuit

Fig. 1.2 shows a resistor R and capacitor C connected in a series. The battery connected to this circuit by a switch provides an electromotive force, or emf force η. Initially, there is no charge on the capacitor. When the switch is flipped to (a), the battery connects, and the capacitor charges. Similarly, when the switch is flipped to (b), the battery disconnects, and the capacitor discharges with energy dissipated in the resistor. Our aim is to determine the voltage drop across the capacitor during charging and discharging.

We start by observing that the voltage drops across a capacitor V_C and resistor V_R are given by

$$V_C = \frac{q}{C}, \quad V_R = iR,$$

respectively, where C is the capacitance and R is the resistance. The current i and the charge q are related by the relation

$$i = \frac{dq}{dt}.$$

From the first equation, we have

$$\frac{dV_C}{dt} = \frac{dq}{dt}.\frac{1}{C} = \frac{i}{C} = \frac{1}{C}\frac{V_R}{R},$$

which implies that

$$V_R = RC\frac{dV_C}{dt}.$$

Copyright Elsevier 2022

Kirchhoff’s voltage law states that the *emf* η in any closed loop is equal to the sum of the voltage drops in that loop. When the switch is thrown to (a), this gives

$$V_R + V_C = \eta.$$

Substituting V_R into $V_R + V_C = \eta$ gives the linear differential equation

$$\frac{dV_C}{dt} + V_C\frac{1}{RC} = \eta\frac{1}{RC}, \quad V_C(0) = 0.$$

An application of the variation of parameters formula gives the solution

$$V_C(t) = e^{-t/RC}\int_0^t \frac{\eta}{RC}e^{s/RC}\,ds,$$

or

$$V_C(t) = \eta\left(1 - e^{-t/RC}\right).$$

Thus the voltage starts at zero and rises slowly but exponentially to η with characteristic time scale given by RC. In other words,

$$V_C(t) \to \eta \text{ as } t \to \infty.$$

Now suppose the switch is thrown to (b). Then by Kirchhoff’s voltage law

$$V_R + V_C = 0,$$

which results in the differential equation

$$\frac{dV_C}{dt} + V_C\frac{1}{RC} = 0, \quad V_C(0) = 0.$$

Then the solution during the discharge phase is given by

$$V_C(t) = \eta e^{-t/RC},$$

and the voltage decays exponentially to zero with the characteristic time scale given by RC.

1.6 Special differential equations

Next, we consider second-order nonlinear differential equations of the form

$$x'' = f(t, x, x'), \tag{1.6.1}$$

where $f(t, x, y)$ is a function of three variables defined in the region $D = \{(t, x, y) : a_1 < t < a_2, b_1 < x < b_2, c_1 < y < c_2\}$ with f, $\frac{\partial f}{x}$, and $\frac{\partial f}{\partial y}$ continuous in D. Little is

Copyright Elsevier 2022

known about how to solve (1.6.1) except in certain particular cases. Two interesting and useful cases occur when (1.6.1) has one of the following forms:

$$F(t, x', x'') = 0 \tag{1.6.2}$$

or

$$F(x, x', x'') = 0. \tag{1.6.3}$$

For (1.6.2), the transformation $x' = u$ transforms the original second-order differential equation into a first-order differential equation $F(t, u, u') = 0$, which can be solved. To see this, we consider the nonlinear second-order differential equation

$$tx'' + 2x' + t = 1, \; t > 0.$$

For $x' = u$, we have the first-order differential equation $tu' + 2u + t = 1$. Using the variation of parameters formula given by (1.5.2), we arrive at

$$u = \frac{c_1}{t^2} + \frac{1}{2} - \frac{t}{3}$$

for some constant c_1. Since $x(t) = \int u(t)dt$, we arrive at the general solution

$$x(t) = -\frac{c_1}{t} + \frac{t}{2} - \frac{t^2}{6} + c_2$$

for some constant c_2.

Equations of the form (1.6.3) can be reduced to first-order equations by the substitution $x' = u$. Eq. (1.6.3) is now replaced by the equivalent system of two differential equations

$$x' = u, \quad F(x, u, u') = 0.$$

By using x as the independent variable we have

$$x'' = \frac{du}{dt} = \frac{du}{dx}\frac{dx}{dt} = \frac{du}{dx}u.$$

Hence (1.6.3) reduces to the first-order differential equation

$$F(x, u, u\frac{du}{dx}) = 0.$$

As an application, we consider the differential equation

$$x'' + (x')^2 x + x'x = 0, \;\; x(0) = 0, \; x'(0) = -1.$$

Then, under the mentioned substitution, we have the new first-order differential equation

$$u\frac{du}{dx} + u^2 x + ux = 0$$

Copyright Elsevier 2022

or

$$\frac{du}{dx} + xu = -x.$$

By the variation of parameters, we arrive at the solution

$$u = -1 + ce^{\frac{x^2}{2}}.$$

Using $u = -1$ when $x = 0$, we have $c = 0$. Now using the transformation $x' = u$, it follows that

$$\frac{dx}{dt} = -1.$$

Using $x(0) = 0$, we arrive at the solution

$$x(t) = -t.$$

1.7 Exercises

Exercise 1.1. Show that the differential equation

$$x'(t) = \frac{3}{2}x^{1/3}(t),\ x(0) = 0,\ t \in \mathbb{R}$$

has infinitely many solutions.

Exercise 1.2. Show that the differential equation

$$x'(t) = x^{\alpha}(t),\ x(0) = 0,\ t \in \mathbb{R}$$

has infinitely many solutions for $\alpha \in (0, 1)$.

Exercise 1.3. Find the escape time of the solutions for the differential equation

$$x'(t) = x^2(t),\ x(t_0) = x_0 \neq 0,\ t \in \mathbb{R}.$$

Exercise 1.4. Show that the solution $x(t, t_0, x_0)$ of

$$x'(t) = t,\ x(t_0) = x_0 \neq 0,\ t \geq 0$$

is stable but unbounded.

Exercise 1.5. Let $x = (x_1, x_2, \ldots, x_n)$, $x_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$.

Show that $\|x\| = \sum_{i=1}^{n} x_i^2$ does not define a norm.

Copyright Elsevier 2022

Exercise 1.6. Consider the following power series $L(x)$, which is also known as Euler's dilogarithm function:

$$L(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^2}.$$

(a) Compute the domain of convergence for $L(x)$. Be sure to give a full analysis of the endpoints.
(b) Show that $L(x)$ uniformly converges on its entire domain of convergence. (Use Weierstrass M-test.)
(c) Explain why the $L(x)$ is continuous on its domain of convergence.

Exercise 1.7. Consider the sequence $\{f_n\}$ of functions $f_n : \mathbb{R} \to \mathbb{R}$ defined by

$$f_n(x) = \frac{nx}{\sqrt{1+n^2x^2}}.$$

Find the pointwise limit of this sequence as $n \to \infty$. Does the sequence converge uniformly on $\mathbb{R}$?

Exercise 1.8. Show that the space $C(I, \mathbb{R})$, where I is any compact subset of $\mathbb{R}$, endowed with the supremum norm

$$||f||_\infty = \sup_{t \in I} |f(t)|$$

is a Banach space.

Exercise 1.9. Show that the space Ω_T of all continuous T-periodic functions given by

$$\Omega_T = \{f \in C(\mathbb{R}, \mathbb{R}) : f(t+T) = f(T) \text{ for all } t \in \mathbb{R}\}$$

with the maximum norm

$$||f|| = \max_{t \in [0,T]} |f(t)|$$

is a Banach space.

Exercise 1.10. Show that the space $C([0, 1], \mathbb{R})$ endowed with the L^1 norm

$$||f||_1 = \int_0^1 |f(t)| dt$$

is a metric space but is not complete.

Exercise 1.11. For $1 \le p < \infty$, let

$$l_p = \{a = (a_1, a_2, \ldots) : a_k \in \mathbb{R}, \sum_k |a_k|^p < \infty\}$$

Copyright Elsevier 2022

and

$$||a||_p = \Big(\sum_k |a_k|^p\Big)^{1/p}.$$

Show that l_p endowed with $||\cdot||_p$ is a Banach space.

Exercise 1.12. Prove Remark 1.1.

Exercise 1.13. Find a continuous solution satisfying

$$x'(t) + x(t) = f(t),\ x(0) = 0,$$

where $f(t) = \begin{cases} 1, & 0 \le t \le 1, \\ 0, & t > 1. \end{cases}$

Is the solution differentiable at $t = 1$?

Exercise 1.14. Find a continuous solution satisfying

$$x'(t) + 2tx(t) = f(t),\ x(0) = 2,$$

where $f(t) = \begin{cases} t, & 0 \le t < 1, \\ 0, & t \ge 1. \end{cases}$

Is the solution differentiable at $t = 1$?

Exercise 1.15. For the SI epidemic model of Section 1.3, use Eq. (1.3.1) and the relation $I + S = 1$ to obtain a first-order differential equation in terms of S and S'. Assume a positive condition $S(0) = 1 - I(0) := S_0$ and find the solution $S(t)$ and $\lim\limits_{t\to\infty} S(t)$.

Exercise 1.16. Let $d > 0$, and let $f : [0, d] \to \mathbb{R}$ be continuous. Let k be any constant and suppose $x(t)$ solves the differential inequality

$$x'(t) \le kx(t) + f(t),\ t \in [0, d].$$

Show that

$$x(t) \le x(0)e^{kt} + \int_0^t e^{k(t-s)} f(s)ds,\ t \in [0, d].$$

Hint: There is a continuous and nonnegative function $g(t)$ such that

$$x'(t) + g(t) = kx(t) + f(t),\ t \in [0, d].$$

Exercise 1.17. Suppose a function $f : [0, d] \times \mathbb{R} \to \mathbb{R}$ is continuous and that $f(t, x)$ is bounded for all $(t, x) \in [0, d] \times \mathbb{R}$, where d is a positive constant. Suppose $x = \varphi(t)$ solves the differential equation

$$x'(t) = xf(t, x),\ x(0) = 1,\ t \in [0, d].$$

Show that there is a constant K such that $\varphi(t) \le e^{Kt}$ for all $t \in [0, d]$.

Copyright Elsevier 2022

Exercise 1.18. In Problems 1–5 solve the differential equation. If no initial conditions are given, find the general solution.

1. $t^2x'' + t = 1$, $t > 0$.
2. $t^2x'' + 2tx' = 1$, $t > 0$.
3. $xx'' - 2(x')^2 + 4x^2 = 0$, $x(1) = 1$, $x'(1) = 2$.
4. $\sin(x')x'' = \sin(x)$, $x(1) = 2$, $x'(1) = 1$.
5. $tx'' - x' = t^2e^t$, $t > 0$.

Copyright Elsevier 2022

CHAPTER

Existence and uniqueness

2

In Chapter 1 we laid out the basics we need to use in this and the following chapters. This chapter is devoted to the existence and uniqueness and the continuation of solutions. We limit our discussions and proofs to scalar differential equations, and in Chapter 3 we indicate how the theory can be naturally extended to vector equations.

2.1 Existence and uniqueness of solutions

This section is devoted to the existence and uniqueness of solutions of the initial value problem (IVP)

$$x' = f(t,x), \quad x(t_0) = x_0, \tag{2.1.1}$$

where we assume that $f : D \to \mathbb{R}$ is continuous and D is a subset of $\mathbb{R} \times \mathbb{R}$. In the case the differential equation (2.1.1) is linear, a solution can be found. However, in general, this approach is not feasible when the differential equation is not linear, and hence another approach must be indirectly adopted that establishes the existence of a solution of (2.1.1). For the development of the existence theory, we need a broader definition of Lipschitz condition.

Definition 2.1.1. The function $f : D \to \mathbb{R}$ is said to satisfy the *global* Lipschitz condition in x if there exists a Lipschitz constant $k > 0$ such that

$$|f(t,x) - f(t,y)| \le k|x-y| \text{ for } (t,x),\ (t,y) \in D. \tag{2.1.2}$$

Definition 2.1.2. The function $f : D \to \mathbb{R}$ is said to satisfy a *local* Lipschitz condition in x if for any $(t_1, x_1) \in D$, there exists a domain $D_1 \subset D$ such that $f(t,x)$ satisfies a Lipschitz condition in x on D_1, that is, there exists a positive constant k_1 such that

$$|f(t,x) - f(t,y)| \le K_1|x-y|, \text{ for } (t,x),\ (t,y) \in D_1. \tag{2.1.3}$$

Definition 2.1.1 can be easily extended to functions $f : D \to \mathbb{R}^n$, where $D \subset \mathbb{R} \times \mathbb{R}^n$ under a proper norm. Let $R = \{(t,x) : |t| \le a,\ |x| \le b\}$ be any rectangle in D. If f and $\frac{\partial f}{\partial x}$ are continuous on R, which is the case in this chapter, then f and $\frac{\partial f}{\partial x}$ are bounded on R. Therefore there exist positive constants $\mathscr{M}$ and $\mathscr{K}$ such that

$$|f(t,x)| \le \mathscr{M} \text{ and } |\frac{\partial f}{\partial x}| \le \mathscr{K} \tag{2.1.4}$$

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00008-5
Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

for all points (t, x) in R. Now for any two points (t, x_1), (t, x_2) in R, by the mean value theorem there exists a constant $\eta \in (x_1, x_2)$ such that

$$f(t, x_1) - f(t, x_2) = \frac{\partial f}{\partial x}(t, \eta)(x_1 - x_2),$$

from which it follows that

$$\begin{aligned} |f(t, x_1) - f(t, x_2 &\le |\frac{\partial f}{\partial x}(t, \eta)||x_1 - x_2| \\ &\le \mathscr{K}|x_1 - x_2|. \end{aligned} \tag{2.1.5}$$

We have shown that if f and $\frac{\partial f}{\partial y}$ are continuous on R, then f satisfies a global Lipschitz condition on R.

Example 2.1. Consider $f(t, x) = x^{2/3}$ in the rectangle $R = \{(t, x) : |t| \le 1,\ |x| \le 1\}$. We claim f does not satisfy the Lipschitz condition on R. Consider the pair of points (t, x_1) and $(t, 0)$ in R where $x_1 > 0$. Then

$$\frac{|f(t, x_1) - f(t, 0)}{x_1 - 0}| = x_1^{-1/3} \to \infty \text{ as } x_1 \to 0^+,$$

and hence there exists no $\mathscr{K}$ such that (2.1.5) holds.

Remark 2.1. Consider (2.1.1) on the rectangle $D \subset \mathbb{R} \times \mathbb{R}$ defined by

$$D = \{(t, x) : |t - t_0| \le a,\ |x - x_0| \le b\},$$

where a and b are positive constants. Let

$$\tau = t - t_0,\ u = x - x_0, \text{ and } g(\tau, u) = f(\tau + t_0, u + x_0).$$

Then the function g is defined on the rectangle

$$D^* = \{(\tau, u) : -a \le \tau \le a,\ -b \le u \le b\},$$

and the IVP (2.1.1) is equivalent to the IVP problem

$$u' = g(\tau, u),\ u(0) = 0.$$

For emphasis, we restate the following definition.

Definition 2.1.3. We say that x is a solution of (2.1.1) on an interval I if $x : I \to \mathbb{R}$ is differentiable, $(t, x(t)) \in D$ for $t \in I$, $x'(t) = f(t, x(t))$ for $t \in I$, and $x(t_0) = x_0$ for $(t_0, x_0) \in D$.

In preparation for the next theorem, we observe that the IVP (2.1.1) is equivalent to

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds. \tag{2.1.6}$$

Copyright Elsevier 2022

Relation (2.1.6) is an *integral equation* since it contains an integral of the unknown function x. This integral is not a formula for the solution, but rather it provides another relation satisfied by a solution of (2.1.1). We have the following definition.

Definition 2.1.4. We say $x : I \to \mathbb{R}$ is a solution of the integral equation (2.1.6) on an interval I if $t_0 \in I$, x is continuous on I, $(t, x(t)) \in D$ for $t \in I$, and (2.1.6) is satisfied for $t \in I$.

The next theorem is fundamental for the proof of the existence theorems.

Theorem 2.1.1. *Let D be an open subset of $\mathbb{R}^2$, and let $(t_0, x_0) \in D$. Then x is a solution of (2.1.1) on an interval I if and only if x satisfies the integral equation given by (2.1.6) on I.*

Proof. Let $x(t)$ be a solution of (2.1.1) on an interval I. Then $t_0 \in I$, x is differentiable on I, and hence x is continuous on I. Moreover, $(t, x(t)) \in D$ for $t \in I$, $x(t_0) = x_0$, and $x'(t) = f(t, x(t))$ for $t \in I$. Now an integration of $x'(t) = f(t, x(t))$ from t_0 to t gives (2.1.6) for $t \in I$. For the converse, if x satisfies (2.1.6) for $t \in I$, then $t_0 \in I$, and x is continuous on I. Moreover, $(t, x(t)) \in D$ for $t \in I$, and (2.1.6) is satisfied for $t \in I$. Thus $x(t)$ is differentiable on I. By differentiating (2.1.6) with respect to t, we arrive at $x'(t) = f(t, x(t))$ for all $t \in I$ and $x(t_0) = x_0 + \int_{t_0}^{t_0} f(s, x(s))ds = x_0$. This completes the proof. $\square$

A graduate textbook in differential equations would be incomplete without the statement and proof of Picard's local existence and uniqueness using successive approximations. In the next section we prove similar theorems on Banach spaces using fixed point theory.

We note that Picard's local existence and uniqueness theorem and Cauchy–Peano existence theorem are widely used, and variant proofs of the theorems can be found in [17], [29], [31], [54], and [70].

The heart of proving our next results is the construction of a sequence of functions that converge to a limit function that satisfies the IVP (2.1.1), although the members of the sequence individually do not.

Theorem 2.1.2 addresses the three basic issues:
(1) Members of the constructed sequence $\{x_n\}$ exist for all time.
(2) The sequence converges, and the limiting function satisfies the integral equation (2.1.6) and hence the IVP (2.1.1).
(3) The limiting function, which is a solution of (2.1.1), is unique.

Theorem 2.1.2. *(Picard's local existence and uniqueness) Let $D \subset \mathbb{R} \times \mathbb{R}$ be defined as*

$$D = \{(t, x) : |t - t_0| \le a,\ |x - x_0| \le b\},$$

where a and b are positive constants. Assume that $f \in C(D, \mathbb{R})$ and f satisfies the Lipschitz condition (2.1.2). Let

$$M = \max_{(t,x)\in D} |f(t, x)| \tag{2.1.7}$$

Copyright Elsevier 2022

and

$$h = \min\{a, \frac{b}{M}\}. \tag{2.1.8}$$

Then the IVP (2.1.1) has a unique solution, denoted by $x(t, t_0, x_0)$, on the interval $|t - t_0| \leq h$ and passing through (t_0, x_0). Furthermore,

$$|x(t) - x_0| \leq b \textit{ for } |t - t_0| \leq h.$$

Proof. Before we begin the proof, we recommend to look at Fig. 2.1.

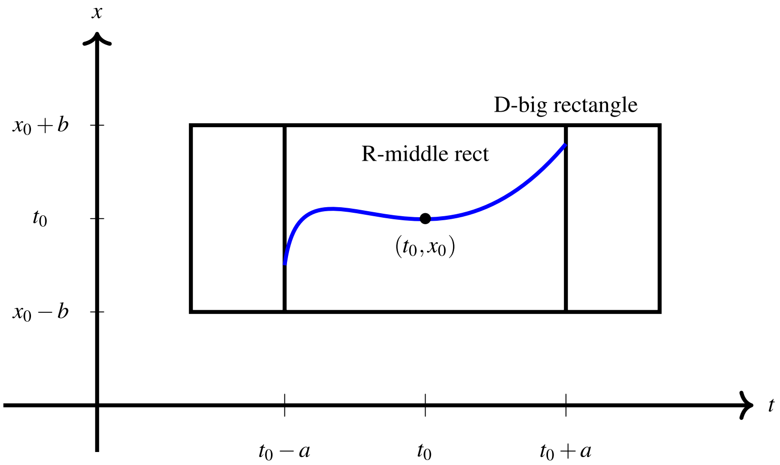


FIGURE 2.1

Interval of existence.

First, we note that since f is continuous on D, condition (2.1.7) is automatically implied. The idea of the proof is to construct a sequence of functions $\{x_n\}$ that converges uniformly to a unique x on the interval $|t - t_0| \leq h$ and is a solution of (2.1.1). We begin by successively defining a sequence of functions $\{x_n\}$ for $|t - t_0| \leq h$ by setting

$$\begin{aligned} x_0(t) &= x_0, \\ x_1(t) &= x_0 + \int_{t_0}^{t} f(s, x_0(s))ds, \\ &\vdots \\ x_n(t) &= x_0 + \int_{t_0}^{t} f(s, x_{n-1}(s))ds. \end{aligned} \tag{2.1.9}$$

According to (2.1.7) and (2.1.8), we have that

$$|x_1(t) - x_0| \leq hM \leq b \text{ for } |t - t_0| \leq h.$$

Copyright Elsevier 2022

Therefore $\int_{t_0}^{t} f(s, x_1(s))ds$ can be defined for $|t - t_0| \leq h$, and hence

$$|x_2(t) - x_0| \leq hM \leq b.$$

Similarly, we may show that $x_3(t), \ldots, x_n(t)$ are well defined on $|t - t_0| \leq h$. Thus

$$|x_n(t) - x_0| = |\int_{t_0}^{t} f(s, x_{n-1}(s))ds| \leq M|t - t_0|. \tag{2.1.10}$$

Next, we show that $\{x_n\}$ converges uniformly to a function, say x, on $|t - t_0| \leq h$. Note that

$$x_n(t) = x_0(t) + [x_1(t) - x_0(t)] + [x_2(t) - x_1(t)] + \cdots + [x_n(t) - x_{n-1}(t)],$$

or

$$x_n(t) = x_0(t) + \sum_{j=0}^{n-1} [x_{j+1}(t) - x_j(t)].$$

Taking the limit if it exists, we have

$$\lim_{n\to\infty} x_n(t) = x_0(t) + \lim_{n\to\infty} \sum_{j=0}^{n-1} [x_{j+1}(t) - x_j(t)]. \tag{2.1.11}$$

Our next task is to compute $|x_{j+1}(t) - x_j(t)|$. Using (2.1.9) and then (2.1.2) followed by (2.1.10) give

$$\begin{aligned}
|x_2(t) - x_1(t)| &\leq \int_{t_0}^{t} \big|f(s, x_1(s)) - f(s, x_0(s))\big| ds \\
&\leq K \int_{t_0}^{t} \big|x_1(s) - x_0(s)\big| ds \\
&\leq KM \int_{t_0}^{t} |s - t_0| ds = KM \int_{t_0}^{t} (s - t_0) ds \\
&= KM \frac{(t - t_0)^2}{2}, \ |t - t_0| \leq a.
\end{aligned}$$

In a similar fashion, we obtain

$$\begin{aligned}
|x_3(t) - x_2(t)| &\leq \int_{t_0}^{t} \big|f(s, x_2(s)) - f(s, x_1(s))\big| ds \\
&\leq K \int_{t_0}^{t} \big|x_2(s) - x_1(s)\big| ds
\end{aligned}$$

Copyright Elsevier 2022

$$\begin{aligned}&\leq K\int_{t_0}^{t} KM\frac{(s-t_0)^2}{2}ds\\&= MK^2\frac{(t-t_0)^3}{3!},\ |t-t_0|\leq a.\end{aligned}$$

Continuing this way, we arrive at

$$|x_j(t)-x_{j-1}(t)|\leq MK^{j-1}\frac{(t-t_0)^j}{j!}. \tag{2.1.12}$$

To complete the induction argument, we assume that (2.1.12) holds for j and show that it holds for $j+1$. Using (2.1.12), we arrive at

$$\begin{aligned}|x_{j+1}(t)-x_j(t)| &\leq \int_{t_0}^{t}\big|f(s,x_j(s))-f(s,x_{j-1}(s))\big|ds\\&\leq K\int_{t_0}^{t}\big|x_j(s)-x_{j-1}(s)\big|ds\\&\leq K\int_{t_0}^{t}MK^{j-1}\frac{(s-t_0)^j}{j!}ds\\&= MK^j\frac{(t-t_0)^{j+1}}{j!(j+1)} = MK^j\frac{(t-t_0)^{j+1}}{(j+1)!},\ |t-t_0|\leq a.\end{aligned}$$

Next, we substitute (2.1.12) into (2.1.11) and get

$$\begin{aligned}\lim_{n\to\infty}x_n(t) &= x_0+\frac{M}{K}\lim_{n\to\infty}\sum_{j=0}^{n-1}\frac{|K(t-t_0)|^{j+1}}{(j+1)!}\\&\leq x_0+\frac{M}{K}\lim_{n\to\infty}\sum_{j=0}^{n-1}\frac{(Kh)^{j+1}}{(j+1)!}\\&\leq x_0+\frac{M}{K}\Big(e^{Kh}-1\Big)<\infty.\end{aligned}$$

Hence, by the Weierstrass M-test, $\{x_n(t)\}$ converges uniformly, say to a function $x(t)$. Next, we show that $x(t)$ is a solution of (2.1.1) on D. We have

$$\begin{aligned}|x(t)-x_0| &= |x(t)-x_n(t)+x_n(t)-x_0|\\&\leq |x(t)-x_n(t)|+|x_n(t)-x_0|\\&\leq |x(t)-x_n(t)|+M|t-t_0|\end{aligned}$$

for every fixed $t\in(t_0-h,t_0+h)$. As a result,

$$|x(t)-x_0| = \lim_{x\to\infty}|x(t)-x_n(t)|+M|t-t_0|$$

Copyright Elsevier 2022

$$
\begin{aligned}
&= 0 + M|t - t_0| \\
&\leq Mh \leq b
\end{aligned}
$$

for $t \in (t_0 - h, t_0 + h)$. So $x(t) \in D$. If the sequence $\{x_n(t)\}$ converges uniformly and $\{x_n(t)\}$ is continuous on the interval $|t - t_0| \leq h$, then

$$\lim_{n\to\infty} \int_{t_0}^{t} f(s, x_n(s))ds = \int_{t_0}^{t} \lim_{n\to\infty} f(s, x_n(s))ds$$

by the uniform continuity of f. As a consequence,

$$
\begin{aligned}
x(t) &= \lim_{n\to\infty} x_{n+1}(t) \\
&= x_0 + \lim_{n\to\infty} \int_{t_0}^{t} f(s, x_n(s))ds \\
&= x_0 + \int_{t_0}^{t} \lim_{n\to\infty} f(s, x_n(s))ds \\
&= x_0 + \int_{t_0}^{t} f(s, \lim_{n\to\infty} x_n(s))ds \\
&= x_0 + \int_{t_0}^{t} f(s, x(s))ds,
\end{aligned}
$$

that is,

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds, \ \ |t - t_0| \leq h. \tag{2.1.13}$$

The integrand $f(s, x(s))$ in (2.1.13) is a continuous function, and hence $x(t)$ is differentiable with respect to t, and its derivative is equal to $f(t, x(t))$. So the proof of the existence is complete. It remains to show that the solution $x(t)$ on D is unique. Let $y(t)$ be another solution of (2.1.1) with $y(t_0) = x_0$. Then

$$y(t) = x_0 + \int_{t_0}^{t} f(s, y(s))ds.$$

Let

$$N = \sup_{|t - t_0| \leq h} |x(t) - y(t)|.$$

Then by the Lipschitz condition, we have

$$|x(t) - y(t)| \leq K \int_{t_0}^{t} \big|x(s) - y(s)\big|ds \tag{2.1.14}$$

or

$$|x(t) - y(t)| \leq KN|t - t_0|.$$

Copyright Elsevier 2022

The substitution of this estimate into (2.1.14) yields

$$|x(t) - y(t)| \leq KN|t - t_0|^2/(2!) \text{ for } |t - t_0| \leq h.$$

Repeating this substitution, we obtain

$$|x(t) - y(t)| \leq KN\frac{|t - t_0|^m}{m!}, m = 1, 2, \ldots,$$

for $|t - t_0| \leq h$. Since the right side of this inequality tends to zero as $m \to \infty$, we have that

$$N = \sup_{|t-t_0|\leq h} |x(t) - y(t)| = 0.$$

This completes the proof. ☐

We will give another proof of the uniqueness once Gronwall's inequality is introduced. To illustrate the above procedure, we provide the following examples.

Example 2.2. We consider the IVP

$$x'(t) = 2t(1 + x), \; x(0) = 0.$$

Set $x_0(t) = 0$. Then for any $n \geq 1$, we have the recurrent formula

$$x_n(t) = \int_0^t 2s(1 + x_{n-1}(s))ds.$$

For $n = 1$, we have

$$x_1(t) = \int_0^t 2sds = t^2,$$

and for $n = 2$,

$$x_2(t) = \int_0^t 2s(1 + s^2)ds = t^2 + \frac{t^4}{2}.$$

We leave it to the reader to verify that

$$x_n(t) = t^2 + \frac{t^4}{2} + \frac{t^6}{3!} + \ldots, \frac{t^{2n}}{n!} = \sum_{i=0}^{n} \frac{(t^2)^i}{i!} - 1.$$

Definition 2.1.5. A sequence $\{x_n\}$ of functions in $C([a, b], \mathbb{R})$ converges uniformly to $x \in C([a, b], \mathbb{R})$ if $\lim_{n\to\infty} ||x_n - x|| = 0$.

It was established in Example 2.2 that

$$x_n(t) = \sum_{i=0}^{n} \frac{(t^2)^i}{i!} - 1$$

Copyright Elsevier 2022

and the true solution of the IVP is given by $x(t) = e^{t^2} - 1$ on the interval [0, 1]. Thus

$$\lim_{n\to\infty} ||x_n - x|| = \lim_{n\to\infty} \Big[\max_{t\in[0,1]} \Big| \sum_{i=0}^{n} \frac{(t^2)^i}{i!} - 1 - (e^{t^2} - 1) \Big| \Big] = \lim_{n\to\infty} \sum_{i=n+1}^{\infty} \frac{1}{i!} = 0.$$

Therefore

$$\lim_{n\to\infty} x_n(t) = \lim_{n\to\infty} \sum_{i=0}^{n} \frac{(t^2)^i}{i!} - 1 = e^{t^2} - 1.$$

Example 2.3. Next, we use Theorem 2.1.2 to find the interval of existence of the unique solution of

$$x'(t) = 1 + x^2,\ x(0) = 1.$$

Let

$$D = \{(t, x) : |t| \le a,\ |x - 1| \le b\},$$

where a and b are constants. It is clear that

$$M = \max_{(t,x)\in D} |f(t, x)| = 1 + (1 + b)^2$$

and

$$h = \min\{a, \frac{b}{M}\} = \min\{a, \frac{b}{1 + (1 + b)^2}\}.$$

So we must find the maximum of the function $g(b) = \dfrac{b}{1 + (1 + b)^2}$. Using calculus, we see that the function g has its maximum at $b = \sqrt{2}$ and is given by $g(\sqrt{2}) = \frac{\sqrt{2}}{1+(1+\sqrt{2})^2}$. Thus the interval of existence and uniqueness is

$$|t| \le \frac{\sqrt{2}}{1 + (1 + \sqrt{2})^2}, \quad \text{or } t \in \Big[-\frac{\sqrt{2}}{1 + (1 + \sqrt{2})^2}, \frac{\sqrt{2}}{1 + (1 + \sqrt{2})^2}\Big].$$

Now suppose in Example 2.3 we chose $D = \{(t, x) : |t| \le 10,\ |x - 1| \le b\}$, that is, $a = 10$. Then the interval of existence found above is much smaller or included in $(-10, 10)$. Recall that the interval of existence is included in D.

The next corollary is an immediate consequence of Theorem 2.1.2 and the discussion leading to (2.1.5).

Corollary 2.1. *Suppose $D \subset \mathbb{R} \times \mathbb{R}$, $f \in C(D, \mathbb{R})$, and $\frac{\partial f}{\partial x}$ is continuous on D. Then for any $(t_0, x_0) \in D$, the IVP (2.1.1) has a unique solution on an interval containing t_0 in its domain.*

As an example, consider

$$x'(t) = t x^{1/2}.$$

Copyright Elsevier 2022

Then

$$f(t,x) = tx^{1/2} \text{ and } \frac{\partial f}{\partial x} = \frac{t}{2x^{1/2}}$$

are continuous in the upper half-plane defined by $x > 0$. We conclude from Corollary 2.1 that for any point $(t_0, x_0),\ x_0 > 0$, there is some interval around t_0 on which the given differential equation has a unique solution.

Unlike Theorem 2.1.2, the next Cauchy–Peano theorem only guarantees the existence of a solution and not uniqueness. This is due to the fact that f is not required to be Lipschitz.

Based on Remark 2.1, we may restate Theorem 2.1.2.

Theorem 2.1.3. *If f and $\frac{\partial f}{\partial x}$ are continuous on the rectangle*

$$D^* = \{(t,x) : -a \le t \le a,\ -b \le x \le b\},$$

where a and b are positive constants, then there is a number h in the interval $(0, a]$ such that the IVP

$$x' = f(t,x),\ \ x(0) = 0,$$

has a unique solution, denoted by $x(t,0,0)$, on the interval $-h \le t \le h$. Furthermore,

$$|x(t)| \le b \textit{ for } |t| \le h.$$

Example 2.4. As an application, we consider the IVP

$$x' = t(x+1)^2,\ \ x(0) = 0,$$

on the rectangle

$$D^* = \{(t,x) : -1 \le t \le 1,\ -2 \le x \le 2\}.$$

It readily follows that the IVP has the solution $x(t) = \dfrac{t^2}{2-t^2}$. We want this solution to be in D^*. This means that $|t| \le 1$ and $\dfrac{t^2}{2-t^2} \le 2$. This is true if and only if $-\dfrac{2}{\sqrt{3}} \le t \le \dfrac{2}{\sqrt{3}}$, from which it follows that $h = \dfrac{2}{\sqrt{3}}$. If we choose a larger b, then we can arrive at a larger h. However, h must be chosen so that $h < \sqrt{2}$. This is due to the fact that

$$\lim_{t \to (\sqrt{2})^-} \frac{t^2}{2-t^2} = \infty.$$

Theorem 2.1.4. *(Cauchy–Peano existence theorem) Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be continuous in a neighborhood of a point $(t_0, x_0) \in \mathbb{R}^2$. Then there exists $\alpha > 0$ such that the*

Copyright Elsevier 2022

IVP (2.1.1) has a solution ϕ on the interval $I := [t_0 - \alpha, t_0 + \alpha]$, that is, there exists a function (not necessarily unique) $\phi = \phi(t)$ defined on I such that

$$\phi'(t) = f(t, \phi(t)),\ t \in I,$$

and $\phi(t_0) = x_0$.

Proof. Since f is continuous in a neighborhood of the point $(t_0, x_0) \in \mathbb{R}^2$, there exists a positive constant a such that the rectangle

$$D = \{(t, x) : |t - t_0| \le a,\ |x - x_0| \le a\}$$

is centered at (t_0, x_0). Moreover, $M = \max_{(t,x)\in D} |f(t, x)|$ exists. Next, we set $\alpha := \dfrac{a}{M}$. By Theorem 2.1.1, we know that $\phi = \phi(t)$ satisfies the IVP (2.1.1) if and only if

$$\phi(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))ds. \tag{2.1.15}$$

Since f is uniformly continuous on the compact set D, given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon)$ such that

$$|t - \tilde{t}| < \delta \text{ and } |x - \tilde{x}| < \delta \text{ imply } |f(t, x) - f(\tilde{t}, \tilde{x})| < \varepsilon$$

for all $(t, x), (\tilde{t}, \tilde{x}) \in D$. Therefore let $\varepsilon = \varepsilon_n = \dfrac{1}{n}$ and $\delta = \delta_n$, $n = 1, 2, \ldots$. Choose points $t_j^{(n)}$ with $t_0^{(n)} = t_0$ and $t_{k(n)} = t_0 + \alpha$ for $0 \le j \le k(n)$ with $|t_{j+1}^{(n)} - t_j^{(n)}| \le \dfrac{\delta_n}{M}$. We define the approximation ϕ_n on $[x_0, x_0 + \alpha]$ by $\phi_n(t_0) = x_0$ and $\phi_n'(t) = f(t_0, x_0)$, $t_0 \le t \le t_1^{(n)}$. Then it is clear that $x_1^{(n)} = \phi_n(t_1^{(n)})$. Continuing in this manner, we get that

$$\phi_n'(t) = f(t_j^{(n)}, x_j^{(n)}) \text{ for } t_j^{(n)} \le t \le t_{j+1}^{(n)}. \tag{2.1.16}$$

Note that ϕ_n is piecewise continuous and may have a jump discontinuity in its derivative at the partition points, and hence we define

$$\triangle_n(t) = \begin{cases} \phi_n'(t) - f(t, \phi_n(t)), & t_j^{(n)} < t < t_{j+1}^{(n)}, \\ 0, & t = x_j^{(n)}. \end{cases}$$

As a consequence, we may define

$$\phi_n(t) = x_0 + \int_{t_0}^{t} \phi_n'(s)ds = x_0 + \int_{t_0}^{t} [f(s, \phi_n(s)) + \triangle_n(s)]ds.$$

Next, we try to have an upper bound on $\triangle_n(t)$. Note that from (2.1.16) and the definition of $\triangle_n(t)$ we have that for $t_j^{(n)} < t < t_{j+1}^{(n)}$,

Copyright Elsevier 2022

$$\begin{aligned}|\triangle_n(t)| &= \left|f(t_j^{(n)}, x_j^{(n)}) - f(t, \phi_n(t))\right|, \; t_j^{(n)} < t < t_{j+1}^{(n)} \\ &\leq |x_j^{(n)} - \phi_n(t)| \leq |x_j^{(n)} - x_{j+1}^{(n)}| \\ &\leq M|t_j^{(n)} - t_{j+1}^{(n)}| \\ &\leq M\frac{\delta_n}{M} = \delta_n = \varepsilon_n = \frac{1}{n},\end{aligned}$$

where we have used that

$$|s - t_j^{(n)}| \leq |x_{j+1}^{(n)} - x_j^{(n)}| \leq \delta_n.$$

By the Ascoli–Arzelà theorem, there exists a uniformly convergent subsequence of ϕ_n that converges to

$$\phi(t) = \lim_{k\to\infty} \phi_{n_k}(t) \; \text{ for } \; t \in [t_0 - \alpha, t_0 + \alpha],$$

that is,

$$\phi_{n_k}(t) = x_0 + \int_{t_0}^{t} [f(s, \phi_{n_k}(s)) + \triangle_n(s)]ds,$$

which converges to

$$\phi(t) = x_0 + \int_{t_0}^{t} [f(s, \phi(s))ds.$$

This completes the proof. □

Example 2.5. Consider the IVP

$$x'(t) = 2tx^{2/3}, \; x(0) = 0.$$

Then $f(t,x) = 2tx^{2/3}$ is continuous everywhere, whereas $\dfrac{\partial f(t,x)}{\partial x} = \dfrac{4t}{3x^{1/3}}$ is not continuous at the points where $x = 0$. Here $x_0 = 0$, and no matter how small we chose $b > 0$, the interval $[-b, b]$ will always contain 0. Therefore we cannot apply Theorem 2.1.2. Instead, we will use Theorem 2.1.4. Let $a = 1$. Then on the rectangle

$$D = \{(t,x) : |t| \leq 1, \; |x| \leq 1\},$$

we have

$$M = \max_{(t,x)\in D} |f(t,x)| = 2 \; \text{ and } \; \alpha = \frac{1}{2} = 0.5.$$

Therefore by Peano's existence theorem, we have at least one solution $x : [-0.5, 0.5] \to [-1, 1]$. As a matter of fact, the IVP has the two solutions $x(t) = 0$ and $x(t) = \frac{t^6}{27}$, both going through the initial point $(0,0)$.

Copyright Elsevier 2022

Next, we state and prove Gronwall's inequality, which plays an important role throughout this book in proving the boundedness, stability, and uniqueness.

Theorem 2.1.5. *(Gronwall's inequality) Let C be a nonnegative constant, and let u, v be nonnegative continuous functions on $[a, b]$ such that*

$$v(t) \leq C + \int_a^t v(s)u(s)ds, \ a \leq t \leq b. \tag{2.1.17}$$

Then

$$v(t) \leq Ce^{\int_a^t u(s)ds}, \ a \leq t \leq b. \tag{2.1.18}$$

In particular, if $C = 0$, then $v = 0$.

Proof. Let $C > 0$, and let $h(t) = C + \int_a^t v(s)u(s)ds$. Then

$$h'(t) = v(t)u(t) \leq h(t)u(t).$$

So we have the differential inequality

$$h'(t) - h(t)u(t) \leq 0.$$

Multiplying both sides of the above expression by the integrating factor $e^{-\int_a^t u(s)ds}$, we get

$$\left(h(t)e^{-\int_a^t u(s)ds}\right)' \leq 0.$$

Integrating both sides from a to t gives

$$h(t)e^{-\int_a^t u(s)ds} - h(a) \leq 0, \ \text{ or } \ h(t) \leq e^{\int_a^t u(s)ds}.$$

Finally,

$$v(t) \leq h(t) \leq Ce^{\int_a^t u(s)ds}.$$

If $C = 0$, then from (2.1.17) it follows that

$$v(t) \leq \int_a^t v(s)u(s)ds \leq \frac{1}{m} + \int_a^t v(s)u(s)ds, \ a \leq t \leq b$$

for any $m \geq 1$. Then from what we have just proved we arrive at

$$v(t) \leq \frac{1}{m}e^{\int_a^t u(s)ds}, \ a \leq t \leq b.$$

Thus for any fixed $t \in [a, b]$, we can let $m \to \infty$ to conclude that $v(t) \leq 0$, and it follows that $v(t) = 0$ for all $t \in [a, b]$. This completes the proof. ☐

Copyright Elsevier 2022

Now we are in a position to revisit the proof of the uniqueness of the solution in Theorem 2.1.2 by using Gronwall's inequality.

Just for a reminder, we assume that the hypotheses of Theorem 2.1.2 hold. Suppose (2.1.1) has two solutions $x(t)$ and $y(t)$ and set $w(t) = x(t) - y(t)$. Then $w'(t) = x'(t) - y'(t)$. Therefore

$$\int_{t_0}^{t} w'(s)ds = w(t) - w(t_0) = \int_{t_0}^{t} \Big(f(s, x(s)) - f(s, y(s))\Big)ds \text{ with}$$

$$w(t_0) = x(t_0) - y(t_0) = 0.$$

So we arrive at

$$\begin{aligned}|w(t)| &\le \int_{t_0}^{t} |f(s, x(s)) - f(s, y(s))|ds \\ &\le K \int_{t_0}^{t} |x(s) - y(s)|ds \\ &= K \int_{t_0}^{t} |w(s)|ds.\end{aligned}$$

Applying Gronwall's inequality with $v(t) = |w(t)|$, $C = 0$, and $u(s) = 1$, we get $|w(t)| = 0$ for all t such that $|t - t_0| \le h$. We conclude that $w(t) = 0$ and hence $x(t) = y(t)$.

The next theorem shows the significance of Theorem 2.1.5.

Theorem 2.1.6. *Suppose* $\phi(t)$ *is continuous on* $[0, \infty)$ *and differentiable on* $(0, \infty)$. *If*

$$\lim_{t\to\infty} \phi(t) = 0$$

and

$$\int_{0}^{\infty} |\phi'|dt < \infty,$$

then all solutions of

$$x'' + (1 + \phi(t))x = 0 \tag{2.1.19}$$

are bounded.

Proof. Multiply both sides of (2.1.19) by x' and integrate the resulting equation from 0 to t to arrive at

$$\frac{(x'(t))^2}{2} + \frac{(x(t))^2}{2} + \int_{0}^{t} \phi(s)x(s)x'(s)ds = c_1,$$

Copyright Elsevier 2022

where c_1 is a constant. Integrating by parts the third term from the left and rearranging give

$$
\begin{aligned}
x^2(t)\big(1+\phi(t)\big) &= 2c_2 - (x'(t))^2 + \int_0^t \phi'(s)x^2(s)ds \\
&\le 2c_2 + \int_0^t \phi'(s)x^2(s)ds,
\end{aligned}
$$

where c_2 is a constant. Since $\lim_{t\to\infty}\phi(t) = 0$, we may choose t_0 large enough so that $1+\phi(t) \ge \frac{1}{2}$ for all $t \ge t_0$. This translates into

$$
x^2(t)(\frac{1}{2}) \le x^2(t)\big(1+\phi(t)\big) \le 2c_2 + \int_0^t \phi'(s)x^2(s)ds
$$

or

$$
x^2(t) \le 4c_2 + 2\int_0^t |\phi'(s)|x^2(s)ds,\ t \ge t_0.
$$

Thus by Theorem 2.1.5 we have that

$$
\begin{aligned}
x^2(t) &\le 4c_2 e^{2\int_0^t |\phi'(s)|ds} \\
&\le 4c_2 e^{2\int_0^\infty |\phi'(s)|ds} = M < \infty \text{ for some positive constant } M
\end{aligned}
$$

and for all $t \ge 0$. Since $x(t)$ is continuous on $0 \le t \le t_0$ and t_0 is arbitrary, it follows that $x(t)$ is bounded. ☐

2.2 Existence on Banach spaces

In this brief section we discuss the existence of solutions of the IVP (2.1.1) on Banach spaces using the contraction mapping principle. For reminder, we restate the IVP

$$
x' = f(t,x),\ \ x(t_0) = x_0, \tag{2.2.1}
$$

where we assume that $f : D \to \mathbb{R}$ is continuous and D is an open subset of $\mathbb{R}\times\mathbb{R}$. The next theorem is another version of Picard's theorem where f satisfies global Lipschitz condition.

Theorem 2.2.1. *(Picard theorem; global version) Let $f \in C(\mathbb{R}\times\mathbb{R},\mathbb{R})$, and let f satisfy the global Lipschitz condition (2.1.2), that is, there exists a constant $K > 0$ such that*

$$
|f(t,x) - f(t,y)| \le K|x-y| \text{ for all } (t,x),(t,y) \in D.
$$

Let $h = \dfrac{1}{2K}$ and $t_0 \in (-h,h)$. Then there exists a unique $x : (-h,h) \to \mathbb{R}$ satisfying the IVP (2.2.1).

Copyright Elsevier 2022

Proof. Define the space

$$\mathbb{B} = C([-h, h], \mathbb{R}).$$

Then $\mathbb{B}$ endowed with the norm

$$||x|| = \max_{t \in [-h,h]} |x(t)|$$

is a Banach space. By Theorem 2.1.1, we have that

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds, \quad t \geq t_0. \tag{2.2.2}$$

The problem comes down to finding a fixed point of the operator $\mathscr{P} : \mathbb{B} \to \mathbb{B}$ defined by

$$\mathscr{P}(x)(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds.$$

By a fixed point we mean that there is an element $x \in \mathbb{B}$ such that $\mathscr{P}(x) = x$, which is equivalent to finding a solution of (2.2.2) and therefore of (2.2.1). Note that $\mathscr{P}$ assigns to any function $x \in \mathbb{B}$ another function, which we denote by $\mathscr{P}x$. Thus for $\mathscr{P}x$ to be defined, we need to evaluate its value at some $t \in [-h, h]$. It is clear that $\mathscr{P}$ is continuous in x. To show that $\mathscr{P}(x)(t)$ is a contraction, we let $x, y \in \mathbb{B}$. Then for $t \in [-h, h]$, we have that

$$\begin{aligned}
\left|\mathscr{P}(x)(t) - \mathscr{P}(y)(t)\right| &\leq \int_{t_0}^{t} |f(s, x(s)) - f(s, y(s))|ds \\
&\leq K \int_{t_0}^{t} |x(s) - y(s)|ds \\
&\leq hK\|x - y\| \\
&\leq \frac{1}{2}\|x - y\|.
\end{aligned}$$

Thus

$$\begin{aligned}
\|\mathscr{P}(x) - \mathscr{P}(y)\| &\leq \sup_{t \in [-h,h]} \left|\mathscr{P}(x)(t) - \mathscr{P}(y)(t)\right| \\
&\leq K \int_{t_0}^{t} |x(s) - y(s)|ds \\
&\leq \frac{1}{2}\|x - y\|.
\end{aligned}$$

Thus $\mathscr{P}$ defines a contraction mapping on the Banach space $\mathbb{B}$, and hence it has a unique $x \in \mathbb{B}$ that is a solution of (2.2.1) by Theorem 2.1.1. This completes the proof. □

Copyright Elsevier 2022

Theorem 2.2.2. *(Picard Theorem; local version) Let* $\mathscr{D} \subset \mathbb{R} \times \mathbb{R}$ *be an open set containing* $(0, x_0)$*, and let* $f \in C(\mathscr{D}, \mathbb{R})$ *be such that* $|f| \le M$ *for a positive constant* M*. Suppose* f *satisfies the local Lipschitz condition in* x *on* $\mathscr{D}$ *with Lipschitz constant* $K > 0$*. Choose*

$$h \in (0, \frac{1}{2K}) \ \text{ so that } [-h, h] \times \{|x - x_0| \le hM\} \subset \mathscr{D}$$

(note that such a choice of h is possible since the set $\mathscr{D}$ *is open). Then there exists a unique* $x : (-h, h) \to \mathbb{R}$ *satisfying the IVP* (2.2.1).

Proof. Let $I = [-h, h]$, where h is as chosen in the hypothesis. Let $\mathbb{B} = \{x \in \mathbb{R} : |y - x_0| \le hM\}$. Define $\mathscr{P} : C(I, \mathbb{B}) \to C(I, \mathbb{B})$ by

$$\mathscr{P}(x)(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds.$$

The problem reduces to showing that $\mathscr{P}(x) = x$. First, we show that $\mathscr{P}$ is well-defined. It is clear from the choice of h that if $x \in C(I, \mathbb{B})$, then $f(t, x)$ is well-defined for all $t \in [-h, h]$, and hence $x_0 + \int_{t_0}^{t} f(s, x(s))ds$ is well-defined. Moreover, since $|f| \le M$, we have that $x_0 + \int_{t_0}^{t} f(s, x(s))ds$ is continuous. Thus $\mathscr{P}(x)(t)$ is a continuous map. Finally,

$$|\mathscr{P}(x)(t) - x_0| \le \int_{t_0}^{t} |f(s, x(s))|ds \le hM.$$

Thus $\mathscr{P}(x)(t) \in \mathbb{B}$ for all $t \in [-h, h]$, and hence $\mathscr{P}(x) \in C(I, \mathbb{B})$ assigns to any function $x \in \mathbb{B}$ another function, which we denote by $\mathscr{P}x$. Thus for $\mathscr{P}x$ to be defined, we need to evaluate its value at some $t \in [-h, h]$. It is clear that $\mathscr{P}$ is continuous in x. To show that $\mathscr{P}(x)(t)$ is a contraction, we let $x, y \in C(I, \mathbb{B})$. Then for each $t \in [-h, h]$, we have that

$$\begin{aligned}
\|\mathscr{P}(x) - \mathscr{P}(y)\| &\le \sup_{t\in[-h,h]} \left|\mathscr{P}(x)(t) - \mathscr{P}(y)(t)\right| \\
&\le K \int_{t_0}^{t} |x(s) - y(s)|ds \\
&\le Kt \max_{|s|\le t} |x(s) - y(s)| \\
&\le Kh\|x - y\| = \frac{1}{2}\|x - y\|.
\end{aligned}$$

Thus $\mathscr{P}$ defines a contraction mapping on the Banach space $C(I, \mathbb{B})$, and hence it has a unique $x : (-h, h) \to \mathbb{R}$ that is a solution of (2.2.1) by Theorem 2.1.1. This completes the proof. □

Consider the differential equation $x'(t) = 1 + x^2$. Theorem 2.2.1 cannot be applied here since $f(t, x) = 1 + x^2$ does not satisfy the global Lipschitz condition on $\mathbb{R}$,

Copyright Elsevier 2022

whereas Theorem 2.2.2 is applicable. Moreover, Theorem 2.2.2 only requires finding $\mathscr{D} \in \mathbb{R} \times \mathbb{R}$ containing $(0, x_0)$. In addition, in both Theorems 2.2.1 and 2.2.2 the initial time t_0 can be chosen so that $t_0 \neq 0$.

2.3 Existence theorem for linear equations

In this section we eloquently apply Theorem 1.4.4 to show that the linear nonhomogeneous differential equation

$$x'(t) = a(t)x(t) + f(t), \tag{2.3.1}$$

where $a, f \in C(\mathbb{R}, \mathbb{R})$, has a unique solution $\phi \in C(\mathbb{R}, \mathbb{R})$ with $\phi(t_0) = x_0$. We have the following theorem.

Theorem 2.3.1. *Let* $(t_0, x_0) \in \mathbb{R}^2$, *and let* $a, f \in C(\mathbb{R}, \mathbb{R})$. *Then* (2.3.1) *has a unique solution* $\phi : (-\infty, \infty) \to \mathbb{R}$ *with* $\phi(t_0) = x_0$.

Proof. Let $s, M > 0$. Find M such that for a given s, we have $|a(t)| \leq M$ for $t \in [t_0 - s, t_0 + s]$. Let $\mathbb{B} = C\big([t_0 - s, t_0 + s], \mathbb{R}\big)$. Then the space $\mathbb{B}$ endowed with the supremum norm $\| \cdot \|$ is a Banach space. Theorem 2.1.1 implies that

$$x(t) = x_0 + \int_{t_0}^{t} a(u)x(u)du + \int_{t_0}^{t} f(u)du.$$

For $x \in \mathbb{B}$, define $\mathscr{P} : \mathbb{B} \to \mathbb{B}$ by

$$\mathscr{P}(x)(t) = x_0 + \int_{t_0}^{t} a(u)x(u)du + \int_{t_0}^{t} f(u)ds.$$

Clearly, $\mathscr{P}(x)(t)$ is a continuous map. Let $x, y \in \mathbb{B}$. Then

$$\begin{aligned}\|\mathscr{P}(x) - \mathscr{P}(y)\| &\leq \sup_{t \in [t_0 - s,\, t_0 + s]} \big|\mathscr{P}(x)(t) - \mathscr{P}(y)(t)\big| \\ &\leq M \int_{t_0}^{t} |x(u) - y(u)| du \\ &\leq Ms\|x - y\|\end{aligned}$$

or

$$\|\mathscr{P}(x) - \mathscr{P}(y)\| \leq Ms\|x - y\|.$$

Now if $Ms = \alpha < 1$, then $\mathscr{P}$ is a contraction, and by Theorem 1.4.3, there is a unique fixed point $\phi \in \mathbb{B}$, that is, $\mathscr{P}(\phi) = \phi$, which is a solution of (2.3.1) on $[t_0 - s, t_0 + s]$. However, we claimed that the solution exists and is unique on $(-\infty, \infty)$. To show

Copyright Elsevier 2022

this, we use Theorem 1.4.4 by letting s be an arbitrary and find an integer $m > 0$ such that $\mathscr{P}^m$ is a contraction:

$$\begin{aligned}|\mathscr{P}^2(x)(t) - \mathscr{P}^2(y)(t)| &= |\mathscr{P}(\mathscr{P}(x))(t) - \mathscr{P}(\mathscr{P}(y))(t)| \\ &\leq \sup_{t\in[t_0-s,\, t_0+s]} \Big| \int_{t_0}^{t} |a(u)||\mathscr{P}(x)(u) - \mathscr{P}(y)(u)|du\Big| \\ &\leq M^2||x-y||\Big|\int_{t_0}^{t} |u-t_0|du| \\ &\leq M^2||x-y||\,|t-t_0|^2/2!.\end{aligned}$$

Continuing this process, we claim that

$$|\mathscr{P}^m(x)(t) - \mathscr{P}^m(y)(t)| \leq M^m|t-t_0|^m||x-y||/m!$$

for positive integer m. The claim is already proved for $m = 1, 2$. We assume that it holds for a positive integer k and then show that the result holds for $k+1$. Indeed,

$$\begin{aligned}|\mathscr{P}^{k+1}(x)(t) - \mathscr{P}^{k+1}(y)(t)| &= |\mathscr{P}(\mathscr{P}^k(x))(t) - \mathscr{P}(\mathscr{P}^k(y))(t)| \\ &\leq \sup_{t\in[t_0-s,\, t_0+s]} \Big| \int_{t_0}^{t} |a(u)||\mathscr{P}^k(x)(u) - \mathscr{P}^k(y)(u)|du\Big| \\ &\leq M^{k+1}||x-y||\Big|\int_{t_0}^{t} |u-t_0|^k du/k!| \\ &\leq M^{k+1}||x-y||\,|t-t_0|^{k+1}/(k+1)!.\end{aligned}$$

Thus

$$|\mathscr{P}^m(x)(t) - \mathscr{P}^m(y)(t)| \leq M^m s^m||x-y||/m!,$$

and for large m, we have

$$\frac{(ms)^m}{m!} < 1,$$

and the proof is complete by Theorem 1.4.4. ☐

2.4 Continuation of solutions

As we mentioned before, the interval of existence and uniqueness guaranteed by Theorem 2.1.1 is, in most cases, shorter than the actual interval of existence and uniqueness. This is misleading, since usually it is not the *maximal interval* of

Copyright Elsevier 2022

existence. To see this, we consider

$$x'(t) = tx,\ x(0) = 1.$$

Let

$$D = \{(t,x) : |t| \le a,\ |x-1| \le b\},$$

where a and b are constants. It is clear that

$$h = \min\{a, \frac{b}{M}\} = \min\{a, \frac{b}{a(1+b)}\}.$$

Thus for any positive constants a and b, we will always have $h < 1$, since either $h \le a$ or $h \le \frac{b}{a(1+b)} < 1/a$. On the other hand, the true solution of the IVP is $x(t) = e^{t^2/2}$, which exists for all $t \in \mathbb{R}$.

Consider the IVP

$$x' = f(t,x), \tag{2.4.1}$$

$$x(0) = x_0. \tag{2.4.2}$$

In the following theorems, we only consider $t \ge 0$ ($t \le 0$ is left as an exercise; see Exercise 2.17).

Theorem 2.4.1. *Suppose* $D = \mathbb{R} \times \mathbb{R}$, $f \in C(D, \mathbb{R})$, *and* $\frac{\partial f}{\partial x}$ *is continuous in* D. *If a solution* ξ *of* (2.4.1)–(2.4.2) *exists on* $[0, \alpha]$ *for some positive* α, *then it can be continued to a solution on some longer interval* $[0, \alpha + r]$ *for some positive constant* r.

Proof. Without loss of generality, we take $x_0 = 0$. Let

$$R = [\alpha - a, \alpha + a] \times [\xi(\alpha) - a, \xi(\alpha) + a],$$

where a is small enough so that $\alpha - a > 0$. Since f is continuous on the rectangle R containing the point $(\alpha, \xi(\alpha))$, we have, by applying Theorem 2.1.2 to the IVP

$$x' = f(t,x) \tag{2.4.3}$$

$$x(\alpha) = \xi(\alpha), \tag{2.4.4}$$

that a solution ϕ of (2.4.3)–(2.4.4) exists on some interval $[\alpha, \alpha + r]$. We claim that the function $\tilde{\xi}$ given by

$$\tilde{\xi}(t) = \begin{cases} \xi(t) & \text{if } 0 \le t \le \alpha, \\ \phi(t) & \text{if } \alpha < t \le \alpha + r, \end{cases}$$

solves (2.4.3)–(2.4.4) on $[0, \alpha + r]$. To do so, we must show that the left and right derivatives exist at α and are equal. Since $\phi(\alpha) = \xi(\alpha)$, the right derivative is given

Copyright Elsevier 2022

by

$$\begin{aligned}\lim_{s\to 0^+}\frac{\tilde{\xi}(\alpha+s)-\tilde{\xi}(\alpha)}{s} &= \lim_{s\to 0^+}\frac{\phi(\alpha+s)-\xi(\alpha)}{s}\\ &= \lim_{s\to 0^+}\frac{\phi(\alpha+s)-\phi(\alpha)}{s}\\ &= \phi'(\alpha)=f(\alpha,\phi(\alpha))\\ &= f(\alpha,\xi(\alpha))=\xi'(\alpha).\end{aligned}$$

Similarly, the left derivative is $\xi'(\alpha)$ due to the fact that ξ is a solution on $[0,\alpha]$. This completes the proof. □

Theorem 2.4.2. *The maximal interval of existence is open.*

Proof. Let J be the maximal interval of existence. If α of Theorem 2.4.1 is in J, then the same theorem shows that the interval $(\alpha-r,\alpha+r)$ is also in J. Hence J is open. This completes the proof. □

We have the following theorem.

Theorem 2.4.3. *Suppose* $D=\mathbb{R}\times\mathbb{R}$, $f\in C(D,\mathbb{R})$, *and* $\frac{\partial f}{\partial x}$ *is continuous in* D. *If* x *is a solution of* (2.4.1)–(2.4.2) *on some interval* J *containing* $t_0=0$, *then* x *is the unique solution on* J.

Proof. Suppose $J=(\alpha,w)$. Without loss of generality, let $x_0=0$. Then Theorem 2.2.1 implies that for some positive constant h, x is the unique solution on the closed interval $[-h,h]$. We will prove the uniqueness on the half-open interval $[0,w)$. Suppose there are two distinct solutions z_1 and z_2 of (2.4.1)–(2.4.2) on $[0,w)$. Then we have $z_1=z_2$ on $[0,h]$, and there exists $t^*\in(h,w)$ such that $z_1(t^*)\neq z_2(t^*)$. Let

$$\xi=\inf\{t\in(0,w): z_1(t^*)\neq z_2(t^*).$$

It follows that $h<\xi\le t^*$. We claim that $\xi<t^*$. Since $z_1(t^*)\neq z_2(t^*)$ and since z_1 and z_2 are continuous, $z_1\neq z_2$ on some interval to the left of t^*. This proves the claim. Due to the continuity of z_1 and z_2 and the equality $z_1=z_2$ on $[0,t^*]$, we must have $z_1(t^*)=z_2(t^*)$. In addition, we know from Theorem 2.4.1 that the IVP

$$x'=f(t,x),\ \ x(\xi)=z_1(\xi)$$

has a unique solution on some interval, say $[\xi-\beta,\xi+\beta]$, for some positive constant β. It follows that $z_1=z_2$ on $[\xi,\xi+\beta]$, which contradicts the definition of ξ. This shows that the solution x is unique on $[0,w)$. Similarly, we may prove the uniqueness on $(\alpha,0]$. This completes the proof. □

Theorem 2.4.4. *Assume the hypotheses of Theorems* 2.4.1 *and* 2.4.3. *If the maximal interval of existence* $J=[0,w)$, *where* $w<\infty$, *then the solution* $\phi(t)$ *of* $x'=f(t,x)$, $x(0)=0$, *is unbounded in* $[0,w)$.

Copyright Elsevier 2022

Proof. We prove by contradiction by assuming that the solution is bounded on $[0, w)$. Since $f(t, \phi)$ is bounded on $[0, w)$, there is a positive constant M such that $f(t, \phi)| \leq M$. Then by Theorem 2.1.1, we have that

$$\phi(t) = \int_{t_0}^{t} f(s, \phi(s))ds.$$

Let $\{t_j\} \uparrow w$. Then for any j and $k > j$, we have that

$$|\phi(t_k) - \phi(t_j)| = \left| \int_{t_j}^{t_k} f(s, \phi(s))ds \right| \leq M|t_j - t_k|.$$

Now the sequence $\{t_j\}$ is convergent and hence is a Cauchy sequence. Therefore $\{\phi(t_j\}$ must be a Cauchy sequence and so also convergent. Since the sequence $\{t_j\}$ was arbitrary and converging to w, we have that $\lim\limits_{t\to w^-} \phi(t)$ exists. Thus we may extend ϕ to w by letting $\phi(w)$ be the limit. Now consider the IVP

$$x' = f(t, x),$$
$$x(w) = \phi(w).$$

This IVP has a unique solution on $[w - r, w + r]$ for some positive constant r, which is an extension of the solution ϕ a little further. Thus $[0, w]$ is not the maximal positive interval of existence, which is a contradiction. This completes the proof. ☐

Consider the IVP

$$x'(t) = x^2, \;\; x(0) = 1.$$

The function $f(t, x) = x^2 \in C^1(\mathbb{R})$, and according to Theorem 2.1.2, the IVP has a unique solution on an open interval containing 0. Solving the IVP, we obtain the solution $x(t) = \frac{1}{t-1}$. The maximal interval of existence must include $t = 0$. Thus the solution is defined on its maximal interval of existence $(-\infty, 1)$. Furthermore,

$$\lim_{t\to 1^-} x(t) = -\infty.$$

According to Theorem 2.4.4, the IVP of Example 2.4 has the maximal interval of existence $(-\sqrt{2}, \sqrt{2})$. It is the case since

$$\lim_{t\to(\sqrt{2})^-} \frac{t^2}{2 - t^2} = \infty \quad \text{and} \quad \lim_{t\to(\sqrt{2})^+} \frac{t^2}{2 - t^2} = -\infty.$$

2.5 Dependence on initial conditions

One of the main characteristics a mathematical model should enjoy is that a solution should depend continuously on initial conditions. In other words, a small change in

Copyright Elsevier 2022

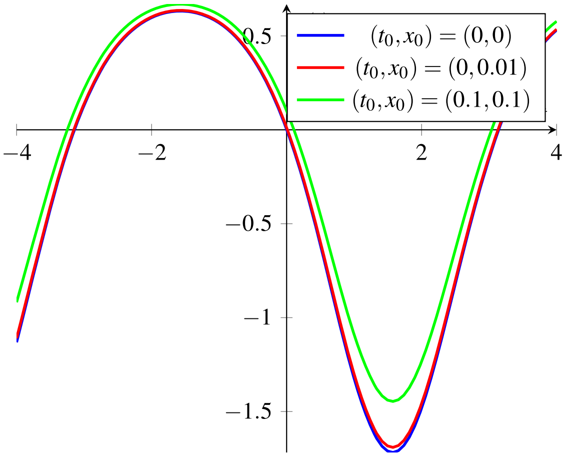


FIGURE 2.2

The three graphs stay close with three different initial conditions $(0, 0)$, $(0, 0.1)$, and $(0.1, 0.1)$.

the initial conditions should result in a small deviation in the solutions. We have seen that a solution $x(t)$ of the IVP

$$x' = f(t, x), \; x(t_0) = x_0 \tag{2.5.1}$$

depends not only on the time t, but also on the initial point (t_0, x_0). Thus, to emphasize such dependence, we use the notation $x(t, t_0, x_0)$ to denote $x(t)$. Consider

$$x'(t) = (x - 1)\cos(t), x(t_0) = x_0,$$

which has the solution

$$x(t) = \frac{e^{\sin(t)}}{e^{\sin(t_0)}}(x_0 - 1) + 1.$$

In Fig. 2.2 we plot the graphs of the solution for three different initial conditions, which shows the sensitivity of the solution to the initial data.

Example 2.6. The scalar (IVP)

$$x'(t) = x^2, \; x(0) = x_0, \;\; t \geq 0$$

has the solution

$$x(t) = \frac{x_0}{1 - x_0 t}$$

with maximal interval $[0, \frac{1}{x_0})$. As the initial value x_0 increases, the maximal interval of existence shrinks as depicted in Fig. 2.3.

Copyright Elsevier 2022

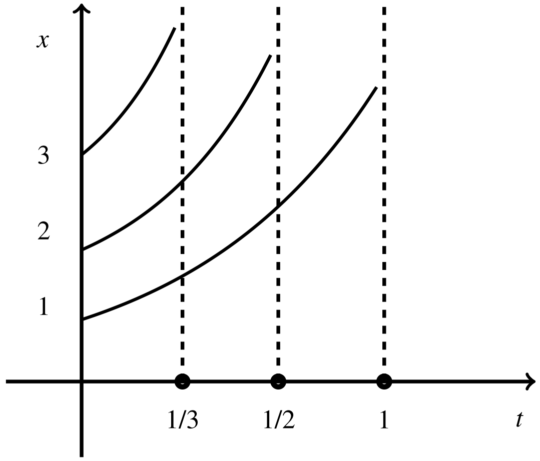


FIGURE 2.3

Dependence on x_0.

Example 2.7. In Fig. 2.4 we display the graph of

$$x' = x\sin(x) + t$$

and show its sensitivity to the initial data.

We have the following theorem.

Theorem 2.5.1. *Let $D \subset \mathbb{R} \times \mathbb{R}$, and suppose f and $\frac{\partial f}{\partial x}$ are continuous and bounded in D. Let x and y be two solutions of (2.5.1) passing through the initial conditions (t_0, x_0) and (t_0^*, x_0^*), respectively. Assume that the two solutions exist on a common interval (α, β), where α and β are finite. Then to each $\varepsilon > 0$, there corresponds $\delta > 0$ such that*

$$|t_0 - t_0^*| < \delta,\ |x_0 - x_0^*| < \delta \ \textit{imply}\ |x(t) - y(t)| < \varepsilon,\ t, t^* \in (\alpha, \beta).$$

Proof. Since f and $\frac{\partial f}{\partial x}$ are continuous and bounded on D, there are positive constants M and K such that

$$|f(t,x)| \le M,\ \ |\frac{\partial f}{\partial x}| \le K.$$

By Theorem 2.1.1, we have that for $t, t^* \in (\alpha, \beta)$,

$$x(t, t_0, x_0) = x_0 + \int_{t_0}^{t} f(s, x(s))ds,\ \ y(t^*, t_0^*, x_0^*) = x_0^* + \int_{t_0^*}^{t^*} f(s, y(s))ds.$$

Copyright Elsevier 2022

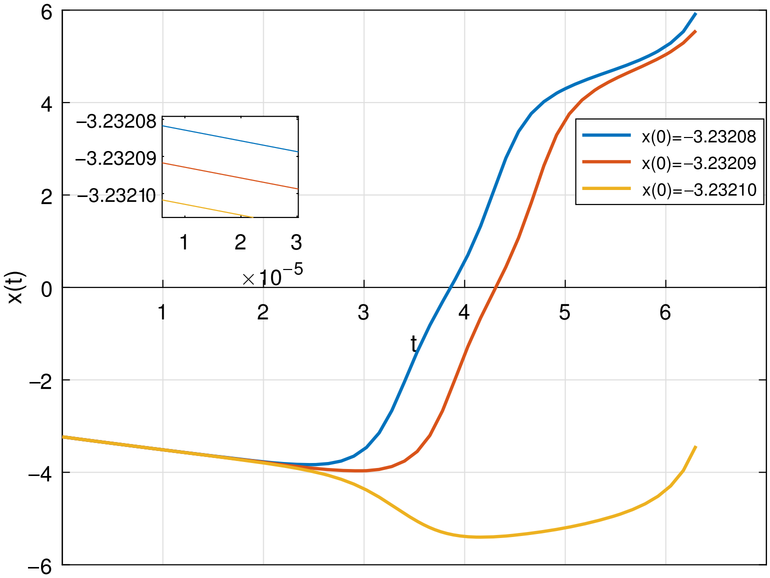


FIGURE 2.4

Sensitivity of solutions of $x' = x\sin(x) + t$ to the initial data.

By noticing that

$$\int_{t_0}^{t} f(s, x(s))ds = \int_{t_0^*}^{t} f(s, x(s))ds + \int_{t_0}^{t_0^*} f(s, x(s))ds,$$

$$\int_{t_0^*}^{t^*} f(s, y(s))ds = \int_{t_0^*}^{t} f(s, y(s))ds + \int_{t}^{t^*} f(s, y(s))ds,$$

we arrive at

$$\begin{aligned}
|x(t, t_0, x_0) - y(t^*, t_0^*, x_0^*)| &\leq |x_0 - x_0^*| + \left| \int_{t_0^*}^{t} |f(s, x(s)) - f(s, y(s))| ds \right| \\
&+ \left| \int_{t_0}^{t_0^*} |f(s, x(s))| ds \right| + \left| \int_{t}^{t^*} |f(s, y(s))| ds \right| \\
&\leq |x_0 - x_0^*| + K \left| \int_{t_0^*}^{t} |x(s)) - y(s)| ds \right| \\
&+ M|t_0^* - t_0| + M|t^* - t| \\
&\leq \delta + K \left| \int_{t_0^*}^{t} |x(s)) - y(s)| ds \right| + 2M\delta.
\end{aligned}$$

Copyright Elsevier 2022

Letting $v(t) = |x(t) - y(t)|$, $C = \delta + 2M\delta$, and $u(s) = K$, we obtain from Theorem 2.1.5 that

$$|x(t, t_0, x_0) - y(t^*, t_0^*, x_0^*)| \le \delta(1 + 2M)e^{K|t-t_0^*|} \le \delta(1 + 2M)e^{K(\alpha-\beta)}$$

or

$$|x(t, t_0, x_0) - y(t^*, t_0^*, x_0^*)| \le \varepsilon$$

for $\delta < \frac{\varepsilon e^{K(\alpha-\beta)}}{1+2M}$. This completes the proof. □

Theorem 2.5.1 shows that the solution $x(t, t_0, x_0)$ of (2.5.1) passing through the initial data (t_0, x_0) is a continuous function in (t, t_0, x_0). We end the chapter with the following theorem.

Theorem 2.5.2. *Let f, g be defined in a region $D \subset \mathbb{R} \times \mathbb{R}$ and satisfy the hypotheses of Theorem 2.5.1. Set*

$$M = \sup_{(t,u)\in D} |f(t, u) - g(t, u)|.$$

Let x and y be the solutions of the IVPs

$$x' = f(t, x),\ x(t_0) = x_0 \quad \text{and} \quad y' = g(t, y),\ y(t_0) = y_0,$$

respectively. Then the solutions x and y satisfy the estimate

$$|x(t) - y(t)| \le |x_0 - y_0|e^{K|t-t_0|} + \frac{M}{K}\Big(e^{K|t-t_0|} - 1\Big), \tag{2.5.2}$$

where K is given in the proof of Theorem 2.5.1.

Proof. Using a similar setup as in the proof of Theorem 2.5.1, we get

$$\begin{aligned}
|x(t) - y(t)| &\le |x_0 - y_0| + \int_{t_0}^{t} \Big|f(s, x(s)) - g(s, y(s))\Big|ds \\
&\le |x_0 - y_0| + \int_{t_0}^{t} \Big|f(s, x(s)) - f(s, y(s))\Big|ds \\
&\quad + \int_{t_0}^{t} \Big|f(s, x(s)) - g(s, y(s))\Big|ds \\
&\le |x_0 - y_0| + \int_{t_0}^{t} \Big[K|x(s) - y(s)| + M\Big]ds.
\end{aligned}$$

Let $U(t) = |x(t) - y(t)|$ and use an improvisation of Exercise 2.9 to obtain the estimate given by (2.5.2). This completes the proof. □

Copyright Elsevier 2022

Suppose M in Theorem 2.5.2 is small, that is, $M \leq \varepsilon > 0$. If $|t - t_0| \leq a$ for positive constant a, then inequality (2.5.2) yields

$$|x(t) - y(t)| \leq |x_0 - y_0|e^{Ka} + \frac{\varepsilon}{K}\left(e^{Ka} - 1\right),$$

and thus if $|x_0 - y_0| \leq \delta$ for some $\delta > 0$, then the two solutions remain close to each other. In other words, Theorem 2.5.2 says that if two differential equations have their right-hand sides "close together," then their solutions cannot differ very much.

2.6 Exercises

Exercise 2.1. Show that $f(t, x) = \dfrac{x}{x^2 + 1}$ is uniformly Lipschitz on $\mathbb{R}$ and identify the Lipschitz constant K.

Exercise 2.2. Compute the Lipschitz constant K for each of the functions in the regions indicated.

1. $f(t, x) = t^3 + x^4$, $\{(t, x) : |t| \leq 2, |x| \leq 3\}$.
2. $f(t, x) = t^2 \cos(x) + t \sin(x)$, $\{(t, x) : |t| \leq 2, |x| < \infty\}$.
3. $f(t, x) = t|x|$, $\{(t, x) : |t| \leq 1, |x| < \infty\}$.
4. $f(t, x) = t^2 \sin(x) \cos(x)$, $\{(t, x) : |t| \leq a, |x| < \infty\}$ for a positive constant a.
   Hint: when you perform $|f(t, x) - f(t, y)|$, add and subtract the term $\sin(x)\cos(y)$.

Exercise 2.3. Apply the Picard iterations to the IVPs

1. $x'(t) = tx$, $x(0) = 1$,
2. $x'(t) = 2t(1 + x)$, $x(0) = 2$,

and show that the obtained $\{x_n(t)\}$ of each iterate converges to the true solution of each IVP (the true solution is the solution found by solving the IVP).

Exercise 2.4. Use Theorem 2.1.2 to find the interval of existence of the unique solution of the following differential equations:

1. $x'(t) = x^2$, $x(0) = 2$.
2. $x'(t) = 5 + x^2$, $x(1) = 2$.
3. $x'(t) = x^3$, $x(0) = 2$.
4. $x'(t) = t + x^2$, $x(0) = 0$.
5. $x'(t) = 1 + x^2$, $x(0) = 0$.

Exercise 2.5. Explain why the IVP $x'(t) = tx - \sin(x)$, $x(0) = 2$, has a unique solution in the neighborhood of the point $(0, 2)$.

Copyright Elsevier 2022

Exercise 2.6. Let f be scalar and continuous in both arguments. Show that if x is a solution of the second-order differential equations $x''(t) = f(t, x(t))$, $x(t_0) = x_0$, $x'(t_0) = x_1$, then it satisfies

$$x(t) = x_0 + x_1(t - t_0) + \int_{t_0}^{t} (t - s) f(s, x(s)) ds. \tag{2.6.1}$$

Hint: Integrate twice and use the general formula

$$\int_0^t \Big[\int_0^s (g(\tau, \phi)\tau) d\tau\Big] ds = \int_0^t \Big[\int_\tau^t ds\Big] (g(\tau, \phi)\tau) d\tau.$$

Exercise 2.7. Explain why neither Theorem 2.1.2 nor Theorem 2.1.4 can be applied to the IVP

$$x'(t) = \frac{x^2}{t}, \ \ x(0) = 1.$$

Use separation of variables and find the family of solutions

$$x(t) = \frac{-1}{\text{Ln}(t) + c},$$

for some constant c. Is it possible to find c?

Exercise 2.8. Check for the existence and uniqueness of solutions of the following IVPs and sketch the corresponding rectangle D:

1. $(t - 1)x'(t) = x^2 + t$, $x(0) = 1$.
2. $(t - 1)x'(t) = x^2 + t$, $x(1) = 0$.
3. $x'(t) = t^{1/3}x + x^2$, $x(0) = 3$.
4. $(t^2 - 1)x'(t) = tx$, $x(0) = 1$.
5. $tx^2x'(t) = x^3 - t^3$, $x(1) = 1$.

Exercise 2.9. (*Generalized Gronwall's inequality*) Show that if f, u, v are nonnegative continuous functions on $[a, b]$ such that

$$u(t) \le f(t) + \int_a^t v(s)u(s)ds, \ a \le t \le b,$$

then

$$u(t) \le f(t) + \int_a^t v(s) f(s) e^{\int_s^t v(r)dr} ds, \ a \le t \le b.$$

Exercise 2.10. Find all nonnegative continuous functions $f : [0, 1] \to \mathbb{R}$ such that

$$f(t) \le \int_0^t f(s)ds, \ 0 \le t \le 1.$$

Copyright Elsevier 2022

Exercise 2.11. Let K_1, K_2 be positive constants, and let f be a nonnegative continuous function on $[a, b]$ such that

$$f(t) \le K_1 + K_2 \int_a^t f(s)ds, \; a \le t \le b.$$

Show that f satisfies

$$f(t) \le K_1 e^{K_2(b-a)}.$$

Exercise 2.12. (*Another form of Gronwall's inequality*) Let K_1, K_2 be positive constants, and let C_0 be a nonnegative constant. Suppose f is a nonnegative continuous function on $[0, h]$ such that

$$f(t) \le C_0 + \int_0^t [K_1 f(s) + K_2]ds, \; 0 \le t \le h.$$

Show that f satisfies

$$f(t) \le C_0 e^{K_1 t} + \frac{K_2}{K_1}\Big(e^{K_1 t} - 1\Big).$$

Exercise 2.13. Let x be a solution of (2.1.1) and suppose f satisfies the global Lipschitz condition with constant K for all $t \in \mathbb{R}$. Suppose that for all $t \in \mathbb{R}$, $\int_0^t |f(s,0)|ds \le M$ for positive constant M. Then $x(t)$ satisfies the inequality

$$|x(t)| \le (|x_0| + M)e^{K(t-t_0)}.$$

Exercise 2.14. Consider the system

$$x' = -x^3.$$

$$y' = -cy + x^{2\lambda} y \cos^2(y).$$

Use Gronwall's inequality to show that
(a) $y(t)$ is bounded for all $t \ge 0$ whenever $c = 0$ and $\lambda > 1$,
(b) $y(t) \to 0$ as $t \to \infty$ whenever $c > 0$ and $\lambda = 1$.

Exercise 2.15. Consider the scalar differential equation

$$x' = (a + \frac{1}{1+t}b)x, \; t \ge 0, \tag{2.6.2}$$

where a is constant. Show that if $y(t) = e^{-at}x$ and x is a solution of (2.6.2), then

$$|y(t)| \le |x(0)|e^{c\ln(1+t)} \quad \text{for some constant } c > 0.$$

In addition, show that if $a < 0$, then

$$|x(t)| \to 0 \;\text{ as }\; t \to \infty.$$

Copyright Elsevier 2022

Exercise 2.16. Let $g : [0, \infty) \to [0, \infty)$ be given by $g(x) = \frac{1}{2}e^{-x}$. Use the contraction mapping principle to prove that the graphs of g and the line $y = x$ intersect at a unique point (a, a).

Exercise 2.17. Prove Theorem 2.4.1 for $t \le 0$.

Exercise 2.18. Consider the IVP

$$x'(t) = 2tx^2,\ x(t_0) = x_0.$$

(a) Use Theorem 2.1.2 to find the largest interval of existence and uniqueness of solutions.
(b) Solve the IVP.
(c) Find the maximal interval of existence when $t_0 = 0$ and $x_0 > 0$ and find the limits at the end points.

Exercise 2.19. Consider the IVP

$$x'(t) = 2(t + 1)(1 + x^2),\ x(0) = 0.$$

(a) Solve the IVP.
(b) Find the maximal interval of existence.

Exercise 2.20. Finish the proof of Theorem 2.5.2 to obtain (2.5.2).

Copyright Elsevier 2022

CHAPTER

Systems of ordinary differential equations

3

We briefly discuss how the existence and uniqueness theorems of Chapter 2 are extended to systems. Then we develop the notion of the fundamental matrix as a solution and utilize it to write solutions of non-homogeneous systems. The author made a conscious decision not to dwell too much on the topic of *Jordan canonical forms* since it is a well-developed subject in linear algebra.

3.1 Existence and uniqueness

For motivational purpose, we consider the second-order differential equation

$$x''(t) + a^2 x(t) = 0, \tag{3.1.1}$$

where a is some nonzero constant. Eq. (3.1.1) is commonly used as a model describing a mechanical process that periodically oscillates. For example, it is used to describe oscillations of a pendulum, provided that the angle change is small or of a spring with a small deflection. We use the transformation $x_1(t) = x(t)$, which is the displacement of the oscillating object. (For the pendulum, it is the angle made with the vertical.) Introducing the second transformation, the velocity or angular velocity $x_2(t) = x'(t)$, and differentiating, we get $x_1'(t) = x'(t) = x_2(t)$ and $x_2'(t) = x''(t) = -a^2 x(t) = -a^2 x_1(t)$. We therefore arrive at two first-order differential equations

$$x_1' = x_2,$$

$$x_2' = -a^2 x_1.$$

Thus we consider the general system of ordinary differential equations

$$\begin{aligned} x_1' &= f_1(t, x_1, \dots, x_n), \\ x_2' &= f_2(t, x_1, \dots, x_n), \\ &\vdots \\ x_n' &= f_n(t, x_1, \dots, x_n). \end{aligned}$$

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00009-7
Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

Using the vector notations

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and

$$f(t, \mathbf{x}) = \begin{pmatrix} f_1(t, \mathbf{x}) \\ f_2(t, \mathbf{x}) \\ \vdots \\ f_n(t, \mathbf{x}) \end{pmatrix},$$

the above system can be written in the vector form

$$\mathbf{x}' = f(t, \mathbf{x}). \tag{3.1.2}$$

Different norms on $\mathbb{R}^n$ were discussed in Chapter 1. For $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, we consider the norm

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

However, in this section, it is more convenient to use the norm that we denote by $|\mathbf{x}|$ given by

$$|\mathbf{x}| = \sum_{i=1}^{n} |x_i|.$$

It is assumed that $|x_i|$ denotes the absolute value. The two norms are topologically equivalent in the sense that the open sets in $\mathbb{R}^n$ are the same under either norm. A set S in $\mathbb{R}^n$ is said to be open if for each $\mathbf{x}^* \in S$, there is a $\delta(\varepsilon) > 0$ such that $\mathbf{x} \in S$ whenever $|\mathbf{x} - \mathbf{x}^*| < \delta(\varepsilon)$. It can be easily shown that $\|\mathbf{x}\| \le |\mathbf{x}| \le \sqrt{n}\|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$. This shows the two norms are equivalent. As a direct result of this, if ϕ^i is a sequence of vector-valued functions such that for some function ϕ, $||\phi^i - \phi|| \to 0$, $i \to \infty$, then $|\phi^i - \phi| \to 0$, $i \to \infty$, and the converse holds too.

Next, we extend some definitions and theorems from Chapter 1 to systems.

Definition 3.1.1. The function $f : D \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is said to satisfy the *global* Lipschitz condition in $\mathbf{x}$ if there exists a positive constant K (called the Lipschitz constant) such that

$$|f(t, \mathbf{x}) - f(t, \mathbf{y})| \le K|\mathbf{x} - \mathbf{y}| \text{ for } (t, \mathbf{x}),\ (t, \mathbf{y}) \in D. \tag{3.1.3}$$

If for every subset $Q \subset D$, there is a positive constant K such that (3.1.3) holds, then the function f satisfies a *local* Lipschitz condition in $\mathbf{x}$. Now we discuss how to

Copyright Elsevier 2022

determine the Lipschitz constant K. Let $R \subset D$ be such that

$$R = \{(t, \mathbf{x}) : |t - t_0| \le a,\ |\mathbf{x} - \mathbf{x_0}| \le b\}$$

and assume that f and all its partial derivatives $\frac{\partial f_i}{\partial x_j}$ are continuous in R. Then the Lipschitz constant K can be defined as

$$K = \sum_{i=1}^{N} \sum_{j=1}^{N} \max_{(t,\mathbf{x}) \in R} \left| \frac{\partial f_i}{\partial x_j} \right|. \tag{3.1.4}$$

Throughout this section, we assume $f : D \to \mathbb{R}^n$ is continuous, where $D \subset \mathbb{R} \times \mathbb{R}^n$ is open.

Definition 3.1.2. Let $(t_0, \mathbf{x_0}) \in D$. We say that $x \in \mathbb{R}^n$ is a solution of the IVP

$$\mathbf{x}' = f(t, \mathbf{x}),\ \mathbf{x}(t_0) = \mathbf{x_0}, \tag{3.1.5}$$

where $\mathbf{x_0} = (x_{01}, x_{02}, \dots, x_{0n})^T$, on an interval I if $t_0 \in I$, $\mathbf{x} : I \to \mathbb{R}^n$ is differentiable, $(t, \mathbf{x}(\mathbf{t})) \in D$ for $t \in I$, $\mathbf{x}' = f(t, \mathbf{x})$ for $t \in I$, and $\mathbf{x}(t_0) = \mathbf{x_0}$.

In preparation for the next theorem we observe that the IVP (3.1.5) given in Definition 3.1.2 is equivalent to the integral relation

$$\mathbf{x}(t) = \mathbf{x_0} + \int_{t_0}^{t} f(s, \mathbf{x}(s))ds. \tag{3.1.6}$$

We have the following definition.

Definition 3.1.3. We say $\mathbf{x} : I \to \mathbb{R}^n$ is a solution of the integral equation given by (3.1.6) on an interval I if $t_0 \in I$, $\mathbf{x}$ is continuous on I, $(t, \mathbf{x}(\mathbf{t})) \in D$ for $t \in I$, and (3.1.6) is satisfied for $t \in I$.

The proof of the next theorem is identical to that of Theorem 2.1.1.

Theorem 3.1.1. *Let D be an open subset of $\mathbb{R}^{n+1}$, and let $(t_0, x_0) \in D$. Then $\mathbf{x}$ is a solution of (3.1.5) on an interval I if and only if $\mathbf{x}$ satisfies the integral equation given by (3.1.6) on I.*

Theorem 3.1.2. *(Picard's local existence and uniqueness) Let $D \subset \mathbb{R} \times \mathbb{R}^n$ be defined by*

$$D = \{(t, \mathbf{x}) : |t - t_0| \le a,\ |\mathbf{x} - \mathbf{x_0}| \le b\},$$

where a and b are positive constants. Assume that $f \in C(D, \mathbb{R}^n)$ satisfies the Lipschitz condition (3.1.3) with Lipschitz constant K given by (3.1.4). Let

$$M = \max_{(t,\mathbf{x}) \in D} |f(t, \mathbf{x})| \tag{3.1.7}$$

Copyright Elsevier 2022

and

$$h = \min\{a, \frac{b}{M}\}. \tag{3.1.8}$$

Then the IVP

$$\mathbf{x}' = f(t, \mathbf{x}),\ \mathbf{x}(t_0) = \mathbf{x_0}, \tag{3.1.9}$$

where $\mathbf{x_0} = (x_{01}, x_{02}, \dots, x_{0n})^T$*, has a unique solution, denoted by* $\mathbf{x}(t, t_0, \mathbf{x_0})$*, on the interval* $|t - t_0| \le h$ *and passing through* $(t_0, \mathbf{x_0})$ *with* $|\mathbf{x} - \mathbf{x_0}| \le b$*.*

Proof. We only give a sketch of the proof since it is identical to that of Theorem 2.1.2. Using Theorem 2.1.1 we have

$$\mathbf{x}(t) = \mathbf{x_0} + \int_{t_0}^{t} f(s, \mathbf{x}(s))ds.$$

In other words, if $\mathbf{x}(t)$ is a solution to the above integral, then it is to be understood that

$$\begin{aligned}
\mathbf{x}_1(t) &= x_{01} + \int_{t_0}^{t} f_1(s, x_1(s), \dots, x_n(s))ds,\\
\mathbf{x}_2(t) &= x_{02} + \int_{t_0}^{t} f_2(s, x_1(s), \dots, x_n(s))ds,\\
&\vdots\\
\mathbf{x}_n(t) &= x_{0n} + \int_{t_0}^{t} f_n(s, x_1(s), \dots, x_n(s))ds.
\end{aligned}$$

As for Picard's iterations, we define

$$\begin{aligned}
\mathbf{x}_{k+1,1}(t) &= x_{01} + \int_{t_0}^{t} f_1(s, x_{k,1}(s), \dots, x_{k,n}(s))ds,\\
\mathbf{x}_{k+1,2}(t) &= x_{02} + \int_{t_0}^{t} f_2(s, x_{k,1}(s), \dots, x_{k,n}(s))ds,\\
&\vdots\\
\mathbf{x}_{k+1,n}(t) &= x_{0n} + \int_{t_0}^{t} f_n(s, x_{k,1}(s), \dots, x_{k,n}(s))ds,
\end{aligned}$$

or

$$\mathbf{x}_{k+1}(t) = \mathbf{x_0} + \int_{t_0}^{t} f(s, \mathbf{x}_k(s))ds,\ k = 0, 1, \dots, n, \tag{3.1.10}$$

where

$$\mathbf{x_0} = (x_{01}, x_{02}, \dots, x_{0n})^T,$$

Copyright Elsevier 2022

and

$$\mathbf{x}_k(t) = (\mathbf{x}_{k,1}(t), \mathbf{x}_{k,2}(t), \ldots, \mathbf{x}_{k,n}(t))^T.$$

Now use (3.1.10) and Theorem 2.1.2 to finish the proof. □

To illustrate Picard's iteration, we present the following example.

Example 3.1. Consider the linear 2×2 system

$$x_1'(t) = 2x_1 - 3x_2,$$
$$x_2'(t) = x_1 - 2x_2,$$

with the initial conditions given by the vector

$$\mathbf{x}(0) = \mathbf{x_0} = \begin{pmatrix} x_{01} \\ x_{02} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Setting $k = 0$, the first iteration is

$$\mathbf{x}_{1,1}(t) = 1 + \int_0^t (2-3)ds = 1 - t,$$
$$\mathbf{x}_{1,2}(t) = -1 + \int_0^t (1-2)ds = -1 - t.$$

Next, setting $k = 1$ gives

$$\mathbf{x}_{2,1}(t) = 1 + \int_0^t [2(1-s) - 3(-1-s)]ds = 1 + 5t + t^2/2,$$
$$\mathbf{x}_{2,2}(t) = -1 + \int_0^t [(1-s) - 2(-1-s)]ds = -1 + 3t + t^2/2.$$

Using the method of the next section, we can show that the true solution is

$$\mathbf{x} = \begin{pmatrix} 3e^t - 2e^{-t} \\ e^t - 2e^{-t} \end{pmatrix}.$$

Remark 3.1. It is essential to know that Theorems 2.1.4, 2.2.1, 2.2.2, 2.4.1, 2.4.2, 2.4.3, and 2.5.1 can be easily extended to the IVP (3.1.9). The instructor may assign them as exercises.

Next, we consider the nth-order differential equation

$$x^{(n)}(t) = g\big(t, x, x', \ldots, x^{(n-1)}\big)$$

and use the transformation

$$x_i(t) = x^{(i-1)}(t), \quad i = 1, 2, \ldots, n$$

Copyright Elsevier 2022

to arrive at the following system:

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= x_3, \\ &\vdots \\ x_n' &= g\big(t, x_1, x_2, \dots, x_n\big). \end{aligned}$$

The right side of this system defines a vector function $f(t, \mathbf{x})$. To fully describe the solution, we impose the initial conditions

$$x^{(i-1)}(t_0) = x_{0i} \ \text{ for } \ i = 1, 2, \dots, n,$$

where x_{0i} for $i = 0, 1, \dots, n-1$ are given constants. Thus our existence and uniqueness results mean that we can impose initial conditions on the function $\mathbf{x}$ and its first $n-1$ derivatives.

Example 3.2. Consider the system

$$\begin{aligned} x_1' &= 1, \\ x_2' &= \frac{x_3}{x_1^2}, \\ x_3' &= -\frac{x_2}{x_1^2} \end{aligned}$$

with initial vector

$$\begin{pmatrix} x_1(\frac{1}{\pi}) \\ x_2(\frac{1}{\pi}) \\ x_3(\frac{1}{\pi}) \end{pmatrix} = \begin{pmatrix} \frac{1}{\pi} \\ 0 \\ -1 \end{pmatrix}.$$

Then the function

$$f(t, x) = \begin{pmatrix} 1 \\ \frac{x_3}{x_1^2} \\ -\frac{x_2}{x_1^2} \end{pmatrix}$$

is continuous with continuous partial derivatives on the set $D = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 > 0\}$, and hence D is the region of existence of solutions.

To find the maximal interval, we try to solve the system. From $x_1' = 1$ we get $x_1(t) = t$, where we have applied the given initial condition. Substituting $x_1 = t$ into the second and third components of the system, and solving for t^2, we arrive at

$$\frac{x_3}{x_2'} = -\frac{x_2}{x_3'},$$

Copyright Elsevier 2022

from which we have

$$\frac{d}{dt}(x_2^2 + x_3^2) = 0, \quad \text{or} \quad x_2^2 + x_3^2 = C.$$

Applying the initial conditions, we get $C = 1$, that is,

$$x_2^2 + x_3^2 = 1.$$

It is clear that $x_2(t) = \sin(\frac{1}{t})$ and $x_3(t) = \cos(\frac{1}{t})$ satisfy $x_2^2 + x_3^2 = 1$ and the initial conditions. Thus the solution is given by

$$\mathbf{x}(t) = \begin{pmatrix} t \\ \sin(\frac{1}{t}) \\ \cos(\frac{1}{t}) \end{pmatrix},$$

and hence the maximal interval of existence is $(0, \infty)$. Note that $\lim_{t \to 0^+} \mathbf{x}(t)$ does not exist.

Theorem 3.1.3. *(Global existence) Let β be a positive constant, and let $D = \mathbb{R}^n$. Let $t_0 \geq 0$, and let $J = [t_0, \beta]$ be an interval in $\mathbb{R}$. Suppose $f : J \times D \to \mathbb{R}^n$ is continuous and satisfies the global Lipschitz condition on $J \times D$ with Lipschitz constant $k > 0$. Assume the existence of two positive continuous functions $M(t)$ and $N(t)$ such that*

$$|f(t, \mathbf{x})| \leq M(t) + N(t)|\mathbf{x}| \quad \textit{on} \quad J \times D. \tag{3.1.11}$$

Then the nonlinear initial value problem

$$\mathbf{x}' = f(t, \mathbf{x}), \ \mathbf{x}(t_0) = \mathbf{x_0} \tag{3.1.12}$$

has a unique solution on the entire interval J.

Proof. First, we show a solution $\mathbf{x} : J \to \mathbb{R}^n$ exists on $[t_0, \beta]$. Due to the continuity of the functions $M(t)$ and $N(t)$, there are two positive constants M_1 and N_1 such that $M(s) \leq M_1$ and $N(s) \leq N_1$ on $[t_0, \beta]$. Thus, for $t_0 \leq t \leq \beta$, we have that

$$\begin{aligned} |\mathbf{x}(t)| &= \Big|\mathbf{x_0} + \int_{t_0}^{t} f(s, \mathbf{x}(s))ds\Big| \\ &\leq |\mathbf{x_0}| + \int_{t_0}^{t} \big(M_1 + N_1|\mathbf{x}(s)|\big)ds \\ &\leq |\mathbf{x_0}| + M_1(\beta_1 - t_0) + \int_{t_0}^{t} N_1|\mathbf{x}(s)|ds. \end{aligned}$$

Applying Theorem 2.1.5, we arrive at the inequality

$$|\mathbf{x}(t)| \leq \big[|\mathbf{x_0}| + M_1(\beta_1 - t_0)\big]e^{N_1(\beta_1 - t_0)}.$$

Copyright Elsevier 2022

This shows $|\mathbf{x}(t)|$ is bounded on $[t_0, \beta]$. Now we show uniqueness. Suppose (3.1.12) has two solutions $\mathbf{x}(t)$ and $\mathbf{y}(t)$ and set $\mathbf{w}(t) = \mathbf{x}(t) - \mathbf{y}(t)$. Then $\mathbf{w}'(t) = \mathbf{x}'(t) - \mathbf{y}'(t)$. Therefore

$$\int_{t_0}^{t} \mathbf{w}'(s)ds = \mathbf{w}(t) - \mathbf{w}(t_0) = \int_{t_0}^{t} \Big(f(s, \mathbf{x}(s)) - f(s, \mathbf{y}(s)) \Big) ds \ \text{ with}$$

$$\mathbf{w}(t_0) = \mathbf{x}(t_0) - \mathbf{y}(t_0) = 0.$$

So we arrive at

$$\begin{aligned} |\mathbf{w}(t)| &\le \int_{t_0}^{t} |f(s, \mathbf{x}(s)) - f(s, \mathbf{y}(s))| ds \\ &\le k \int_{t_0}^{t} |\mathbf{x}(s) - \mathbf{y}(s)| ds \\ &= k \int_{t_0}^{t} |\mathbf{w}(s)| ds. \end{aligned}$$

It follows from Gronwall's inequality that $|\mathbf{w}(t)| = 0$ for all $t \in [t_0, \beta]$ and the solution is unique on $[t_0, \beta]$. This completes the proof. □

3.2 $x' = A(t)x$

Consider the linear system of ordinary differential equations of the form

$$\begin{aligned} x_1' &= a_{11}(t)x_1(t) + a_{12}(t)x_2(t) + \ldots, a_{1n}(t)x_n(t) + g_1(t), \\ x_2' &= a_{21}(t)x_1(t) + a_{22}(t)x_2(t) + \ldots, a_{2n}(t)x_n(t) + g_2(t), \\ &\ldots \\ x_n' &= a_{n1}(t)x_1(t) + a_{n2}(t)x_2(t) + \ldots, a_{nn}(t)x_n(t) + g_n(t), \end{aligned}$$

where the functions a_{ij}, g_i, $1 \le i \le n$, $1 \le j \le n$, are continuous real-valued functions on an interval I. Using vector and matrix notations, the above system is equivalent to the vector equation

$$x'(t) = A(t)x(t) + g(t), \tag{3.2.1}$$

where

$$x := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad x' := \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix},$$

Copyright Elsevier 2022

and

$$A(t) := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad g(t) := \begin{pmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{pmatrix}$$

for $t \in I$. Note that the matrix functions A and g are continuous on an interval I if and only if all their entries are continuous on I.

Definition 3.2.1. We say the $n \times 1$ vector y is a solution of (3.2.1) on I if y is continuously differentiable on I and

$$y'(t) = A(t)y(t) + g(t)$$

for all $t \in I$.

We denote the space of n-tuple continuously differentiable functions $x : I \to \mathbb{R}^n$ by $C^1(I, \mathbb{R}^n)$. Next, we define matrix norm.

Definition 3.2.2. A norm of a square matrix A is a non-negative real number denoted by $\|A\|$. There are several different ways of defining a matrix norm, but they all share the following properties:

1. $\|A\| \geq 0$ for any square matrix.
2. $\|A\| = 0$ if and only if $A = 0$ (zero matrix).
3. $\|kA\| = |k|\|A\|$ for any constant k.
4. $\|A + B\| \leq \|A\| + \|B\|$ for square matrices A and B.
5. $\|AB\| \leq \|A\|\|B\|$.

Example 3.3. Let A be a matrix with entries a_{ij}, $1 \leq i \leq n$, $1 \leq j \leq n$, that is, $A = (a_{ij})$. We define the following possible norms on the matrix A.

1. The 1-norm:

$$||A||_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^{n} |a_{ij}| \right)$$

   (the maximum absolute column sum).
2. The infinity norm:

$$||A||_\infty = \max_{1 \leq i \leq n} \left(\sum_{j=1}^{n} |a_{ij}| \right)$$

   (the maximum of the sums of the absolute values along each row).
3. The Euclidean norm:

$$||A||_E = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij})^2}$$

   (the square root of the sum of squares of all entries).

Copyright Elsevier 2022

Note that if I is the identity matrix, then $||I||_1 = ||I||_\infty = 1$.

Next, we state and prove our first existence and uniqueness result regarding (3.2.1).

Theorem 3.2.1. *Suppose $A(t)$ and $g(t)$ are continuous on some interval $a \le t \le b$. Then* (3.2.1) *has a unique solution $\phi(t)$ on the interval $a \le t \le b$ with $\phi(t_0) = \eta$ and $a \le t_0 \le b$.*

Proof. Let $\|\cdot\|$ be a suitable matrix norm. For the existence, it suffices to show that solutions remain bounded for all $a \le t \le b$. We know from Theorem 3.1.1 that ϕ given by

$$\phi(t) = \eta + \int_{t_0}^{t} [A(s)\phi(s) + g(s)]ds$$

is a solution of (3.2.1) with $\phi(t_0) = \eta$. Let

$$K_1 = |\eta|, \ K_2 = \max_{t\in[a,b]} |g(t)|(b-a), \ \text{ and } \ K_3 = \max_{t\in[a,b]} ||A(t)||.$$

Then

$$\begin{aligned} |\phi(t)| &\le |\eta| + \left| \int_{t_0}^{t} [A(s)\phi(s) + g(s)]ds \right| \\ &\le K_1 + \int_{t_0}^{t} ||A(s)|||\phi(s)|ds + \int_{t_0}^{t} |g(s)|ds \\ &\le K_1 + \max_{t\in[a,b]} ||A(t)|| \int_{t_0}^{t} |\phi(s)|ds + \max_{t\in[a,b]} |g(t)| \int_{t_0}^{t} ds \\ &\le K_1 + K_3 \int_{t_0}^{t} |\phi(s)|ds + \max_{t\in[a,b]} |g(t)|(b-a) \\ &= \Big(K_1 + K_2\Big) + K_3 \int_{t_0}^{t} |\phi(s)|ds. \end{aligned} \tag{3.2.2}$$

Applying Theorem 2.1.5 (Gronwall's inequality) we arrive at the inequality

$$\begin{aligned} |\phi(t)| &\le \Big(K_1 + K_2\Big)e^{K_3(t-t_0)} \\ &\le \Big(K_1 + K_2\Big)e^{K_3(b-a)}. \end{aligned} \tag{3.2.3}$$

From (3.2.3) we see that solutions are bounded for all $t \in [a, b]$ and hence exist. This completes the proof of the existence. For the uniqueness, we assume the existence of two solutions ϕ_1 and ϕ_2 of (3.2.1) passing through (t_0, η). Then imitating the proof

Copyright Elsevier 2022

of the existence, we arrive at

$$\begin{aligned}|\phi_1(t)-\phi_2(t)| &\le \left|\int_{t_0}^{t} A(s)[\phi_1(s)-\phi_2(s)]ds\right| \\ &\le \int_{t_0}^{t} ||A(s)|||\phi_1(s)-\phi_2(s)|ds \\ &\le \max_{t\in[a,b]} ||A(t)|| \int_{t_0}^{t} |\phi_1(s)-\phi_2(s)|ds \\ &\le K_3 \int_{t_0}^{t} |\phi_1(s)-\phi_2(s)|ds.\end{aligned}$$

By Theorem 2.1.5 (Gronwall's inequality) it follows that

$$|\phi_1(t)-\phi_2(t)| = 0 \text{ for all } t\in[a,b],$$

and hence $\phi_1(t)=\phi_2(t)$ for all $t\in[a,b]$. The proof is complete. □

As a consequence of Theorem 3.2.1, we have the following corollary.

Corollary 3.1. *If $A(t)$ and $g(t)$ are continuous on $\mathbb{R}$, then the unique solution $\phi(t)$ of (3.2.1) is defined for all $t\in\mathbb{R}$.*

In the next discussion, we try to estimate the error between two solutions. Let ϕ_1 and ϕ_2 be two solutions of (3.2.1) with $\phi_1(t_0)=\eta_1$ and $\phi_2(t_0)=\eta_2$, respectively. Then

$$\begin{aligned}|\phi_1(t)-\phi_2(t)| &\le |\phi_1(t_0)-\phi_2(t_0)| + \left|\int_{t_0}^{t} A(s)[\phi_1(s)-\phi_2(s)]ds\right| \\ &\le |\eta_1-\eta_2| + \int_{t_0}^{t} ||A(s)|||\phi_1(s)-\phi_2(s)|ds.\end{aligned}$$

Using Gronwall's inequality, we arrive at

$$|\phi_1(t)-\phi_2(t)| \le |\eta_1-\eta_2| e^{\int_{t_0}^{t}||A(s)||ds}, \tag{3.2.4}$$

which is an estimate of the error between the two solutions at two different initial conditions.

Example 3.4. Consider the differential system

$$\begin{aligned}x_1' &= -x_2\sin(t)+4,\\ x_2' &= -x_1+2tx_2-x_3+e^t,\\ x_3' &= 3x_1\cos(t)+x_2+\frac{1}{t}x_3-5t^2.\end{aligned}$$

Copyright Elsevier 2022

Consider two solutions $\phi_1(t)$ and $\phi_2(t)$ on the interval $(1, 3)$ with initial vectors

$$\phi_1(2) = \begin{pmatrix} 7 \\ 3 \\ -2 \end{pmatrix}$$

and

$$\phi_2(2) = \begin{pmatrix} 6.7 \\ 3.2 \\ -1.9 \end{pmatrix}.$$

Next, we compute the norm of the matrix

$$A(t) = \begin{pmatrix} 0 & -\sin(t) & 0 \\ -1 & 2t & -1 \\ 3\cos(t) & 1 & 1/t \end{pmatrix}.$$

Using the 1-norm, we get

$$||A(t)|| = \max\{1 + 3|\cos(t)|, |\sin(t)| + 2t + 1, 1 + \frac{1}{t}\}$$

or

$$||A(t)|| \le \max\{4, 2 + 2t, 1 + \frac{1}{t}\} = 2 + 2t \le 8,\ 1 < t < 3.$$

Using the estimate formula given by (3.2.4), we get

$$\begin{aligned} |\phi_1(t) - \phi_2(t)| &\le (0.3 + 0.2 + 0.1)e^{\int_2^t ||A(s)||ds} \\ &\le 0.6e^{\int_2^t 8ds} \\ &\le 0.6e^{8(t-2)} < 0.6e^{8(3-2)} \approx 1790. \end{aligned}$$

Example 3.5. (*Pendulum with friction*) Fig. 3.1 shows a simple pendulum with massless string of length L and a bob of mass m acting under the external influence of gravity. Let g be the acceleration of gravity, m be the mass of the bob, L be the length of the rod, and b be the coefficient of friction (damping), which all are positive constants. The period is the time it takes for the bob to swing from its farthest right position to its left farthest position and back. Let $s(t)$ be the distance along the arc from the lowest point of the bob at time t, with displacement to the right considered positive. Let $\theta(t)$ be the corresponding angle with respect to the vertical. The gravitational force is directed downward and has the magnitude mg. The force acting in the tangential direction is $-mg\sin(\theta)$. Then the motion is modeled by the second-order differential equation

$$\frac{d^2s}{dt^2} = -g\sin(\theta).$$

Copyright Elsevier 2022

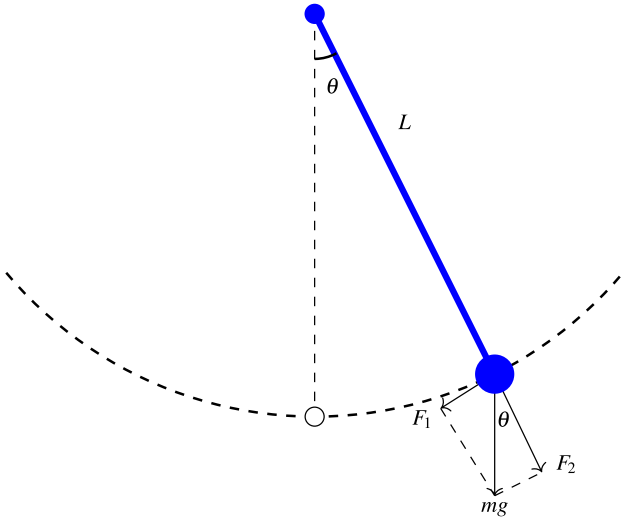


FIGURE 3.1

Pendulum with friction.

Now $s = L\theta$ as L is the radius of the circle, and so

$$\frac{d^2\theta}{dt^2} = -\frac{g}{L}\sin(\theta).$$

We add damping to the model and assume that the damping force is proportional to the velocity. As θ is our independent variable, the damping is represented as the force proportional to the angular velocity, say, $-b\dfrac{d\theta}{dt}$. The negative sign indicates that the damping force is in the opposite direction of the motion. Then the model takes the form

$$mL\theta'' + mg\sin(\theta) + bL\theta' = 0. \tag{3.2.5}$$

Assume the initial data

$$\theta(t_0) = \theta_0, \quad \theta'(t_0) = \omega_0. \tag{3.2.6}$$

To apply Theorem 2.4.4, we put (3.2.5) into a system. Thus we let $x_1(t) = \theta(t)$ and $x_2(t) = \theta'(t)$ and obtain the nonlinear two-dimensional system

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -\frac{g}{L}\sin(x_1) - \frac{b}{m}x_2, \end{aligned} \tag{3.2.7}$$

with corresponding initial conditions

$$x_1(t_0) = \theta_0, \quad x_2(t_0) = \omega_0. \tag{3.2.8}$$

Copyright Elsevier 2022

Let

$$f(t,x)=\begin{pmatrix} x_2 \\ -\frac{g}{L}\sin(x_1)-\frac{b}{m}x_2 \end{pmatrix}.$$

Clearly, f satisfies the global Lipschitz condition on $\mathbb{R}^2$. Thus by Theorems 2.4.3 and 2.4.4, the solution exists on some interval $[t_0,\alpha)$, $\alpha>0$. We will show that $\alpha=\infty$ by appealing to Theorem 2.4.4. Define the *energy function* $E:[t_0,\alpha)\times\mathbb{R}^2\to[0,\infty)$ by

$$E(t):=E(t,x_1,x_2)=\frac{1}{2}x_2^2(t)+\frac{g}{L}\Big(1-\cos(x_1(t))\Big). \tag{3.2.9}$$

Then along the solutions of (3.2.7) we have that

$$\begin{aligned}\frac{d}{dt}E(t)&=x_2\,x_2'+\frac{g}{L}\sin(x_1)\,x_1'\\&=x_2\Big(-\frac{g}{L}\sin(x_1)-\frac{b}{m}x_2\Big)+\frac{g}{L}\sin(x_1)x_2\\&=-\frac{b}{m}x_2^2\le 0.\end{aligned}$$

Let $t_0\le t<\alpha$ and integrate the above inequality from t_0 to t to get

$$E(t)-E(t_0)\le 0$$

or

$$E(t)\le\frac{1}{2}\omega_0^2+\frac{g}{L}\Big(1-\cos(\theta_0)\Big).$$

We have from (3.2.9) that $\frac{1}{2}x_2^2(t)\le E(t)$, so that

$$|x_2(t)|\le\sqrt{2\Big[\frac{1}{2}\omega_0^2+\frac{g}{L}\Big(1-\cos(\theta_0)\Big)\Big]}:=M.$$

Using this in the first equation of (3.2.7), after integrating, we find

$$|x_1(t)|\le|\theta_0|+M(t-t_0)\ \text{ for }\ t_0\le t<\alpha.$$

Now suppose for contradiction that $\alpha<\infty$. Then on $[t_0,\alpha)$,

$$|x_1(t)|\le|\theta_0|+M(\alpha-t_0)\ \text{ and }\ |x_2(t)|\le M.$$

Thus

$$|x|:=\sqrt{x_1^2+x_2^2}\le\sqrt{M+|\theta_0|+M(\alpha-t_0)},$$

which is bounded on $[t_0,\alpha)$. This is a contradiction to Theorem 2.4.4, and hence $\alpha=\infty$. So the maximal interval of existence is $[t_0,\infty)$. It is worth noting that the *energy function* will be referred to as a *Lyapunov function*, which we will study in more detail in Chapter 7.

Copyright Elsevier 2022

In Example 3.5 the energy function had to be chosen just right in order to yield the right results. Generally speaking, for any second-order differential equation of the form

$$x'' + f(x)x' + g(x) = 0 \tag{3.2.10}$$

and under more conditions on g, the energy function may take the form

$$E(t) := E(t, x_1, x_2) = \frac{1}{2}x_2^2(t) + \int_0^{x_1} g(s)ds \tag{3.2.11}$$

by transforming (3.2.10) into a system using the transformations

$$x_i(t) = x^{(i-1)}(t), \quad i = 1, 2.$$

On the other hand, if we do not transform (3.2.10) into a system, then we may use the energy function

$$E(t) := E(t, x) = \frac{1}{2}(x')^2 + \int_0^{x(t)} g(s)ds, \tag{3.2.12}$$

where $\frac{1}{2}(x')^2$ corresponds to the *kinetic energy*, and $\int_0^{x(t)} g(s)ds$ represents the *potential energy*.

3.2.1 Fundamental matrix

Solving the non-homogeneous equation

$$x'(t) = A(t)x(t) + g(t) \tag{3.2.13}$$

requires solving the homogeneous system

$$x'(t) = A(t)x(t), \tag{3.2.14}$$

where $A(t)$ is an $n \times n$ matrix of coefficients $a_{ij}(t)$, which are assumed to be continuous on an interval I. Recall that a solution $x(t)$ of (3.2.14) is an n-tuple of C^1 functions $x_i : I \to \mathbb{R}$. We adopt the notation

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

The solution x maybe considered as a C^1 vector-valued function $x : I \to \mathbb{R}^n$. The space of such functions is denoted by $C^1(I, \mathbb{R}^n)$. If $\mathscr{S}$ is the *solution space* of (3.2.14), then $\mathscr{S} \subset C^1(I, \mathbb{R}^n)$.

Copyright Elsevier 2022

Definition 3.2.3. We say that a set $\{f_1, f_2, \ldots, f_n\}$ of $n \times 1$ vector functions defined on an interval I is *linearly dependent* on I if there exist constants $c_1, c_2, \ldots, c_n$, not all zero, such that

$$c_1 f_1 + c_2 f_2 + \ldots + c_n f_n = 0$$

for every $t \in I$. If the set of functions is not *linearly dependent* on the interval I, it is said to be linearly independent.

We have the following important theorem regarding the solution space $\mathscr{S}$.

Theorem 3.2.2. *The solution space $\mathscr{S}$ is a linear vector space.*

Proof. Let $x \in C^1(I, \mathbb{R}^n)$ be a solution of (3.2.14). It is clear that $x = 0$ is a solution of (3.2.14). If $x(t_*) = 0$ for some $t_* \in I$, then $x(t) = 0$ for all $t \in I$. This is true due to the uniqueness of the solution. Thus the zero solution is in $\mathscr{S}$. Next, we show that $\mathscr{S}$ is closed under addition and scalar multiplication. Let $\phi_1(t)$ and $\phi_2(t)$ be two arbitrary functions in $\mathscr{S}$, and let c_1 c_2 be constants in $\mathbb{R}$. Set $y(t) = c_1\phi_1(t) + c_2\phi_2(t)$. Then

$$\begin{aligned} y'(t) &= (c_1\phi_1(t) + c_2\phi_2(t))' \\ &= c_1\phi_1'(t) + c_2\phi_2'(t) \\ &= c_1 A(t)\phi_1(t) + c_2 A(t)\phi_2(t) \\ &= A(t)\Big(c_1\phi_1(t) + c_2\phi_2(t)\Big) \\ &= A(t)y(t). \end{aligned}$$

Next, we show that the set of solutions of (3.2.14) is linearly independent. Let

$$\{\phi_1(t), \phi_2(t), \ldots, \phi_n(t)\}$$

be a set of solutions of (3.2.14). We show that if $\{\phi_1(t), \phi_2(t), \ldots, \phi_n(t)\}$ is linearly independent at some $t_* \in I$, then the set is linearly independent for all $t \in I$, that is,

$$c_1\phi_1(t) + c_2\phi_2(t) + \ldots + c_n\phi_n(t) = 0$$

only in the case $c_1 = c_2 = \ldots = c_n = 0$. Suppose the contrary, that is, there are c_i, $i = 1, 2, \ldots, n$, not all zero such that

$$c_1\phi_1(t_1) + c_2\phi_2(t_1) + \ldots + c_n\phi_n(t_1) = 0$$

for some $t_1 \in I$, $t_1 \neq t_*$. Now we use the same constants c_i to define a function for all $t \in I$,

$$x(t) = c_1\phi_1(t) + c_2\phi_2(t) + \ldots + c_n\phi_n(t). \tag{3.2.15}$$

Then by the above relation we must have $x(t_1) = 0$, since it is a solution of (3.2.14). Thus by uniqueness $x(t) = 0$ for all $t \in I$. This is a contradiction to the fact that

$$c_1\phi_1(t_*) + c_2\phi_2(t_*) + \ldots + c_n\phi_n(t_*) = 0$$

Copyright Elsevier 2022

only in the case $c_1 = c_2 = \ldots = c_n = 0$. This completes the proof of the linear independence. Left to show that $\mathscr{S}$ is an n-dimensional subspace of the vector space $C^1(I, \mathbb{R}^n)$. Let $e_1, e_2, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$. Then by Corollary 3.1 there exists a set $\{\phi_1(t), \phi_2(t), \ldots, \phi_n(t)\}$ of n solutions that form a set of solutions of (3.2.14) satisfying

$$\phi_1(t_0) = e_1, \phi_2(t_0) = e_2, \ldots, \phi_n(t_0) = e_n. \tag{3.2.16}$$

This set of solutions is linearly independent for all $t \in I$ and hence forms a linearly independent set in the solution space $\mathscr{S}$. Let $\phi(t)$ be an arbitrary solution of (3.2.14) and define

$$\phi_0 = \phi(t_0). \tag{3.2.17}$$

Now since $\phi_0 \in \mathbb{R}^n$, there are constants $c_i, i = 1, 2, \ldots, n$, not all zero, such that

$$\phi_0 = c_1 e_1 + c_2 e_2 + \ldots + c_n e_n. \tag{3.2.18}$$

Using the same constants of (3.2.18), we may define a new function $\psi : I \to \mathbb{R}^n$,

$$\psi(t) = c_1\phi_1(t) + c_2\phi_2(t) + \ldots + c_n\phi_n(t). \tag{3.2.19}$$

Since each $\phi_i, i = 1, 2, \ldots, n$, is a solution of (3.2.14), so is $\psi(t)$. By (3.2.16) and (3.2.18), we have that $\psi(t_0) = \phi_0$. This and (3.2.17) imply that ψ and ϕ have the same initial condition, and hence by the uniqueness of the solution we have that

$$\psi(t) = \phi(t) \text{ for all } t \in \mathbb{R}.$$

From (3.2.19) it follows that the arbitrary solution may be written as

$$\psi(t) = c_1\phi_1(t) + c_2\phi_2(t) + \cdots + c_n\phi_n(t),$$

which is a linear combination of the solutions $\phi_i, i = 1, 2, \ldots, n$. Therefore the set of solutions

$$\{\phi_1(t), \phi_2(t), \ldots, \phi_n(t)\}$$

is a basis for the solution space $\mathscr{S}$. This implies that the dimension of $\mathscr{S}$ is n. This completes the proof. ☐

Definition 3.2.4. (*Fundamental set of solutions*) A set of n solutions of the linear differential system (3.2.14), all defined on the same open interval I, is called a fundamental set of solutions on I if the solutions are linearly independent functions on I.

We have the following corollary.

Corollary 3.2. *Let $A(t)$ be an $n \times n$ matrix of continuous coefficients $a_{ij}(t)$ on an interval I. If $\{\phi_1(t), \phi_2(t), \ldots, \phi_n(t)\}$ form a fundamental set of solutions on I, then*

Copyright Elsevier 2022

the general solution of (3.2.14) *is given by*

$$x(t) = c_1\phi_1(t) + c_2\phi_2(t) + \cdots + c_n\phi_n(t)$$

with constants $c_i, i = 1, 2, \dots, n$*.*

The following definition is needed for the next theorem.

Definition 3.2.5. Let f and g be two continuous functions on $\mathbb{R}$. Then we write

$$f(x) = O\big(g(x)\big)$$

if there exists a positive real constant M and a real number x^* such

$$|f(x)| \le Mg(x) \quad \text{for all} \quad x \ge x^*.$$

Theorem 3.2.3. *If* $\{\phi_1(t), \phi_2(t), \dots, \phi_n(t)\}$ *form a fundamental set of solutions of* (3.2.14) *on* I *and* X *is the matrix function with columns* $\{\phi_1(t), \phi_2(t), \dots, \phi_n(t)\}$*, then*

$$\det X(t) = \det X(t_0) e^{\int_{t_0}^{t} \operatorname{tr}(A(s))ds} \quad \text{for } t \in I. \tag{3.2.20}$$

Proof. Let the Wronskian $w(t) := \det X(t)$. Let $t_0 \in I$ and expand $X(t)$ around t_0:

$$\begin{aligned} X(t) &= X(t_0) + (t-t_0)X'(t_0) + O\big((t-t_0)^2\big) \\ &= X(t_0) + (t-t_0)A(t_0)X(t_0) + O\big((t-t_0)^2\big) \\ &= \big[I + (t-t_0)A(t_0)\big]X(t_0) + O\big((t-t_0)^2\big). \end{aligned}$$

As a consequence, we have

$$\det X(t) = \det\Big(\big[I + (t-t_0)A(t_0)\big]\Big)\det X(t_0)$$

or

$$W(t) = \det\big[I + (t-t_0)A(t_0)\big]W(t_0).$$

By the Terence Tao matrix identity, which says

$$\det(C + hB) = \det(C) + h\operatorname{tr}\big(\operatorname{adj}(C)B\big) + O(h^2),$$

we have that

$$\det(I + (t-t_0)A(t_0))W(t_0) = W(t_0)\big[1 + (t-t_0)\operatorname{tr}\big(A(t_0)\big)\big]$$

or

$$W(t) = W(t_0)\big[1 + (t-t_0)\operatorname{tr}\big(A(t_0)\big)\big]. \tag{3.2.21}$$

Copyright Elsevier 2022

Next, we expand $W(t)$ in Taylor series around t_0:

$$W(t) = W(t_0) + (t - t_0)W'(t_0) + O\big((t - t_0)^2\big).$$

Substituting the value of $W(t)$ given by (3.2.21) into the above equality gives

$$W'(t_0) = W(t_0)\mathrm{tr}\big(A(t_0)\big).$$

Since t_0 is arbitrary, for any time $t \in \mathbb{R}$, we have the equation

$$W'(t) = W(t)\mathrm{tr}\big(A(t)\big),$$

which has the solution

$$W(t) = W(t_0)e^{\int_{t_0}^{t} \mathrm{tr}(A(s)ds} \quad \text{for } \ t \in I.$$

This completes the proof. □

Definition 3.2.6. (*Fundamental matrix*) A matrix solution Φ is called a *fundamental matrix solution* (or, shortly, fundamental matrix) of (3.2.14) on an interval I if its columns form a fundamental set of solutions. If, in addition, $\Phi(t_0) = I$, then a fundamental matrix solution is called the *principal fundamental matrix solution* (or, shortly, principal matrix).

Theorem 3.2.4. *The matrix Φ is a fundamental matrix of $x' = Ax$ at t_0 if and only if*

(a) *Φ is a solution of $x' = Ax$*
and
(b) $\det\Phi(t_0) \neq 0$.

Proof. We know that $\Phi' = A\Phi$. Let

$$\Phi = \big[\phi_1(t), \phi_2(t), \dots, \phi_n(t)\big],$$

where the columns of Φ consist of $\phi_1(t), \phi_2(t), \dots, \phi_n(t)$. Then

$$\Phi' = \big[\phi_1'(t), \phi_2'(t), \dots, \phi_n'(t)\big],$$

and

$$A\Phi = \big[A\phi_1(t), A\phi_2(t), \dots, A\phi_n(t)\big].$$

Since two matrices are equal if and only if their corresponding columns are the same,

$$\Phi'(t) = A\Phi(t)$$

Copyright Elsevier 2022

if and only if

$$x_j' = Ax_j$$

or if and only if each column of Φ is a solution.

Condition (b) is equivalent to the linear independence of

$$\phi_1(t_0), \phi_2(t_0), \ldots, \phi_n(t_0).$$

This completes the proof. □

We claim that the matrix

$$\Phi(t) = \begin{pmatrix} e^{-t} & te^{-t} \\ -e^{-t} & (1-t)e^{-t} \end{pmatrix}$$

is the fundamental matrix for the system

$$x'(t) = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

To see this, we have

$$\Phi'(t) = \begin{pmatrix} -e^{-t} & e^{-t} - te^{-t} \\ e^{-t} & -e^{-t} - (1-t)e^{-t} \end{pmatrix} = \begin{pmatrix} -e^{-t} & (1-t)e^{-t} \\ e^{-t} & -2e^{-t} + te^{-t} \end{pmatrix}.$$

On the other hand,

$$\begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} \Phi = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} e^{-t} & te^{-t} \\ -e^{-t} & (1-t)e^{-t} \end{pmatrix} = \begin{pmatrix} -e^{-t} & (1-t)e^{-t} \\ e^{-t} & te^{-t} - 2e^{-t} \end{pmatrix}.$$

This proves the claim. For the next two theorems, we suppress the argument t.

Theorem 3.2.5. *If Φ is a fundamental matrix of (3.2.14), then ΦC, where C is an arbitrary $n \times n$ nonsingular constant matrix, is a fundamental matrix of (3.2.14). Conversely, if Ψ is another fundamental matrix to (3.2.14), then there exists a constant nonsingular $n \times n$ matrix C such that $\Psi = \Phi C$ for all $t \in I$.*

Proof. Let Φ be a fundamental matrix of (3.2.14) on I. Then

$$(\Phi C)' = \Phi' C = (A(t)\Phi)C = A(t)(\Phi C).$$

So ΦC solves (3.2.14). Next,

$$\det(\Phi C) = \det\Phi \det C \neq 0,$$

and by Theorem 3.2.4, ΦC is a fundamental matrix of (3.2.14). Conversely, let Φ and Ψ be two fundamental matrix solutions of (3.2.14). Since $\Phi\Phi^{-1} = I$, we have that

Copyright Elsevier 2022

$\Phi'\Phi^{-1} + \Phi(\Phi^{-1})' = 0$ or $(\Phi^{-1})' = -\Phi^{-1}\Phi'\Phi^{-1}$. Now we differentiate $\Phi^{-1}\Psi$ and obtain

$$\begin{aligned}(\Phi^{-1}\Psi)' &= (\Phi^{-1})'\Psi + \Phi^{-1}\Psi' \\ &= -\Phi^{-1}\Phi'\Phi^{-1}\Psi + \Phi^{-1}\Psi' \\ &= \big(-\Phi^{-1}A(t)\Phi\Phi^{-1} + \Phi^{-1}A(t)\big)\Psi \\ &= \big(-\Phi^{-1}A(t) + \Phi^{-1}A(t)\big)\Psi \\ &= 0.\end{aligned}$$

Therefore integrating the above expression gives $\Phi^{-1}\Psi = C$ for some $n \times n$ constant matrix C, or $\Psi = \Phi C$. Furthermore, as Ψ and Φ are fundamental matrix solutions, we have $\det\Phi \neq 0$ and $\det\Psi \neq 0$, and hence $\det C \neq 0$. This completes the proof. □

We remark that it is not true in general that $C\Phi$ is a fundamental matrix of (3.2.14).

Theorem 3.2.6. *If Φ is a fundamental matrix of* (3.2.14) *on an interval I, then Φc solves* (3.2.14) *on I for every $n \times 1$ constant vector c.*

Proof. Let Φ be a fundamental matrix of (3.2.14) on I. Then

$$(\Phi c)' = \Phi' c = (A(t)\Phi)c = A(t)(\Phi c).$$

So Φc solves (3.2.14). □

Theorem 3.2.7. *If $\Phi(t)$ is a fundamental matrix of* (3.2.14) *on an interval I, then $\Phi(t)c$ with $c = \Phi^{-1}(t_0)x_0$ is a solution of* (3.2.14) *with $x(t_0) = x_0$.*

Proof. First, $\Phi^{-1}(t_0)$ exists by the definition of a fundamental matrix. Hence by Theorem 3.2.6 $x(t) = \Phi(t)\Phi^{-1}(t_0)x_0$ solves (3.2.14). Moreover, $x(t_0) = \Phi(t_0) \times \Phi^{-1}(t_0)x_0 = Ix_0 = x_0$. This completes the proof. □

3.2.2 $x' = Ax$

Now we consider homogeneous systems with constant matrices. In particular, we look at

$$x'(t) = Ax(t), \tag{3.2.22}$$

where A is an $n \times n$ constant matrix. The fundamental solution of (3.2.22) can be constructed from a knowledge of the eigenvalues and eigenvectors of A. We begin with the following definition.

Definition 3.2.7. Let A be an $n \times n$ constant matrix, shortly, a “matrix.” A number λ is said to be an *eigenvalue* of A if there exists a nonzero vector v such that

$$Av = \lambda v. \tag{3.2.23}$$

Copyright Elsevier 2022

The solution vector v is said to be an *eigenvector* corresponding to the eigenvalue λ. We may refer to λ and v as an *eigenpair*.

Now consider a solution x of (3.2.22) of the form

$$x(t) = ve^{\lambda t}$$

for $v \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. Then we have

$$\lambda ve^{\lambda t} = Ave^{\lambda t},$$

or

$$(A - \lambda I)v = 0 \text{ if and only if } \det(A - \lambda I) = 0.$$

Thus we have the following straightforward result.

Theorem 3.2.8. *If λ_0, v_0 is an eigenpair of A, then*

$$x(t) = e^{\lambda_0 t} v_0 = v_0 e^{\lambda_0 t}$$

is a solution of $x' = Ax$.

Proof. Let $x(t) = e^{\lambda_0 t} v_0$. Then

$$x'(t) = \lambda_0 e^{\lambda_0 t} v_0 = \lambda_0 v_0 e^{\lambda_0 t} = Av_0 e^{\lambda_0 t} = Ae^{\lambda_0 t} v_0 = Ax,$$

as desired. This completes the proof. □

Theorem 3.2.9. *(Independent eigenvectors) Let $v_1, v_2, \ldots, v_p$ be the corresponding eigenvectors to the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ of a matrix A. Then $v_1, v_2, \ldots, v_p$ are linearly independent.*

Proof. Suppose $v_1, v_2, \ldots, v_j$ are linearly independent for positive integer j. If $j < p$, then v_{j+1} can be written as a linear combination of the vectors $v_1, v_2, \ldots, v_j$, that is, there are constants $c_1, c_2, \ldots, c_j$ such that

$$v_{j+1} = c_1 v_1 + c_2 v_2 + \cdots + c_j v_j.$$

Multiply from the left by the matrix A and apply the fact that $Av_i = \lambda_i v_i$ for $i = 1, 2, \ldots j$ to arrive at

$$\begin{aligned} Av_{j+1} &= \lambda_{j+1} v_{j+1} \\ &= \lambda_{j+1}\left(c_1 v_1 + c_2 v_2 + \ldots + c_j v_j\right) \\ &= c_1 \lambda_{j+1} v_1 + c_2 \lambda_{j+1} v_2 + \ldots + c_j \lambda_{j+1} v_j. \end{aligned}$$

Copyright Elsevier 2022

On the other hand,

$$\begin{aligned} Av_{j+1} &= A\big(c_1v_1 + c_2v_2 + \ldots + c_jv_j\big) \\ &= c_1Av_1 + c_2Av_2 + \ldots + c_jAv_j \\ &= c_1\lambda_1v_1 + c_2\lambda_2v_2 + \ldots + c_j\lambda_jv_j. \end{aligned}$$

Subtracting the two equations gives

$$c_1(\lambda_{j+1} - \lambda_1)v_1 + c_2(\lambda_{j+1} - \lambda_2)v_2 + \ldots + c_j(\lambda_{j+1} - \lambda_j)v_j = 0.$$

Since $v_1, v_2, \ldots, v_j$ are linearly independent, we must have that

$$c_1(\lambda_{j+1} - \lambda_1) = 0,\ c_2(\lambda_{j+1} - \lambda_2) = 0,\ \ldots,\ c_j(\lambda_{j+1} - \lambda_j) = 0.$$

Since

$$\lambda_{j+1} - \lambda_j \neq 0, \quad \text{for all} \quad j = 1, 2, \ldots, n,$$

this could only be possible if

$$c_1 = c_2, \ldots, c_j = 0.$$

This implies that

$$v_{j+1} = 0 \ \text{(zero vector)},$$

which is a contradiction, since v_{j+1} is the eigenvector corresponding to λ_{j+1}. This completes the proof. □

As a result of Definition 3.2.4, Corollary 3.2, and Theorem 3.2.9, we have the following theorem concerning systems with constant matrices such as (3.2.22).

Theorem 3.2.10. *(Distinct eigenvalues) Let* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *be* n *distinct real eigenvalues of the matrix* A *of* (3.2.22)*, and let* $K_1, K_2, \ldots, K_n$ *be the corresponding eigenvectors. Then the general solution of* (3.2.22) *on the interval* $I = (-\infty, \infty)$ *is given by*

$$x(t) = c_1K_1e^{\lambda_1 t} + c_2K_2e^{\lambda_2 t}, \ldots, c_nK_ne^{\lambda_n t}$$

for constants $c_i, i = 1, 2, \ldots, n$.

Remark 3.2. If we let

$$\phi_i = K_ie^{\lambda_i t}, \quad i = 1, 2, \ldots, n,$$

then the fundamental matrix Φ of (3.2.22) is formed by taking its columns to be

$$\Phi = \big[\phi_1(t), \phi_2(t), \ldots, \phi_n(t)\big].$$

Copyright Elsevier 2022

Example 3.6. (*Distinct eigenvalues*) Consider the linear system

$$x' = \begin{pmatrix} -4 & 1 & 1 \\ 1 & 5 & -1 \\ 0 & 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Then the eigenpairs are given by

$$\lambda_1 = -3, K_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_2 = -4, K_2 = \begin{pmatrix} 10 \\ -1 \\ 1 \end{pmatrix}, \quad \lambda_3 = 5, K_3 = \begin{pmatrix} 1 \\ 8 \\ 1 \end{pmatrix},$$

and hence the general solution can be written as

$$x(t) = c_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^{-3t} + c_2 \begin{pmatrix} 10 \\ -1 \\ 1 \end{pmatrix} e^{-4t} + c_3 \begin{pmatrix} 1 \\ 8 \\ 1 \end{pmatrix} e^{5t}.$$

Moreover, the fundamental matrix $\Phi(t)$ is given by

$$\Phi(t) = \begin{pmatrix} e^{-3t} & 10e^{-4t} & e^{5t} \\ 0 & -e^{-4t} & 8e^{5t} \\ e^{-3t} & e^{-4t} & e^{5t} \end{pmatrix}. \tag{3.2.24}$$

Now suppose we impose that every solution satisfies the initial condition

$$x(t_0) := x_0 = \begin{pmatrix} x_{01} \\ x_{02} \\ x_{03} \end{pmatrix}.$$

Then by Theorem 3.2.7, the unique solution is

$$x(t) = \Phi(t)\Phi^{-1}(t_0)x_0,$$

where $\Phi(t)$ is given by (3.2.24).

Next, we discuss repeated roots. For example, the system

$$x' = \begin{pmatrix} 3 & -18 \\ 2 & -19 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{3.2.25}$$

has the repeated eigenvalue $\lambda_1 = \lambda_2 = -3$. If $K_1 = \binom{k_1}{k_2}$ is the corresponding eigenvector, then we have the two equations $6k_1 - 18k_2 = 0$, $2k_1 - 6k_2 = 0$, which are both equivalent to $k_1 = 3k_2$. As a consequence, we have the single eigenvector $K_1 = \binom{3}{1}$, and so one solution is given by

$$\phi_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} e^{-3t}.$$

Copyright Elsevier 2022

Since we are interested in finding the general solution, we need to examine the question of finding another solution.

In general, if m is a positive integer and $(\lambda - \lambda_1)^m$ is a factor of the characteristic equation $\det(A - \lambda I) = 0$, whereas $(\lambda - \lambda_1)^{m+1}$ is not a factor, then λ_1 is said to be an eigenvalue of multiplicity m. We discuss two such scenarios:

(a) For some $n \times n$ matrix A, it may be possible to find m linearly independent eigenvectors $K_1, K_2, \ldots, K_n$ corresponding to an eigenvalue λ_1 of multiplicity $m \leq n$. In this case the general solution of the system contains the linear combination

$$c_1 K_1 e^{\lambda_1 t} + c_2 K_2 e^{\lambda_2 t} + \ldots + c_n K_n e^{\lambda_n t}.$$

(b) If there is one eigenvector corresponding to an eigenvalue of multiplicity m, then we can find m linearly independent solutions of the form

$$\begin{aligned}
\phi_1 &= K_{11} e^{\lambda_1 t}, \\
\phi_2 &= K_{21} t e^{\lambda_1 t} + K_{22} e^{\lambda_1 t}, \\
&\vdots \\
\phi_m &= K_{m1} \frac{t^{m-1}}{(m-1)!} e^{\lambda_1 t} + K_{m2} \frac{t^{m-2}}{(m-2)!} e^{\lambda_1 t} + K_{mm} e^{\lambda_1 t},
\end{aligned}$$

where K_{ij} are column vectors, known as *generalized eigenvectors*. For an illustration of case (b), we suppose λ_1 is an eigenvalue of multiplicity two with only one corresponding eigenvector. To find the second eigenvector, in general, we assume a second solution of

$$x' = Ax$$

of the form

$$\phi_2(t) = K t e^{\lambda_1 t} + P e^{\lambda_1 t}, \tag{3.2.26}$$

where

$$P = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \text{ and } K = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}$$

are to be found. Differentiate $\phi_2(t)$ and substitute back into $x' = Ax$ to get

$$(AK - \lambda_1 K) t e^{\lambda_1 t} + (AP - \lambda_1 P - K) e^{\lambda_1 t} = 0.$$

Since this equation must hold for all t, we have

$$(A - \lambda_1 I) K = 0 \tag{3.2.27}$$

Copyright Elsevier 2022

and

$$(A - \lambda_1 I)P = K. \tag{3.2.28}$$

Eq. (3.2.27) implies that K must be an eigenvector of A associated with the eigenvalue λ_1, and hence we have one solution $\phi_1(t) = Ke^{\lambda_1 t}$. To find the second solution given by (3.2.26), we must solve for the vector P in (3.2.28). Now we go back to finding a second solution for (3.2.24). We already know K given by K_1. Let $P = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$. Then from (3.2.28) we have $(A + 3I)P = K$, which implies that $6p_1 - 18p_2 = 3$ or $2p_1 - 6p_2 = 1$. Since these two equations are equivalent, we may choose $p_1 = 1$ and find $p_2 = 1/6$. However, for simplicity, we choose $p_1 = 1/2$, so that $p_2 = 0$. Using (3.2.26), we find that

$$\phi_2(t) = \begin{pmatrix} 3 \\ 1 \end{pmatrix} te^{-3t} + \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} e^{-3t}.$$

Thus the general solution is

$$x(t) = c_1\phi_1(t) + c_2\phi_2(t) = c_1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} e^{-3t} + c_2\left(\begin{pmatrix} 3 \\ 1 \end{pmatrix} te^{-3t} + \begin{pmatrix} 1/2 \\ 0 \end{pmatrix} e^{-3t}\right).$$

Moreover, the fundamental matrix $\Phi(t)$ is given by

$$\Phi(t) = \begin{pmatrix} 3e^{-3t} & 3te^{-3t} + \frac{1}{2}e^{-3t} \\ e^{-3t} & te^{-3t} \end{pmatrix}.$$

3.2.3 Exponential matrix e^{At}

In the previous section we looked at the concept of a fundamental matrix $\Phi(t)$ for the linear system $x'(t) = A(t)x(t)$ and applied it only to constant matrices. In this section we study in detail such constant matrices; namely, we consider

$$x'(t) = Ax(t), \quad x(t_0) = x_0, \tag{3.2.29}$$

where A is an $n \times n$ constant matrix, and $x \in \mathbb{R}^n$. Note that

$$||Ax - Ay|| \le K||x - y||,$$

where $K > 0$ is constant, and $||\cdot||$ is the matrix norm of A defined by

$$||A|| = \sup_{x \in \mathbb{R}^n - \{0\}} \frac{|Ax|}{|x|}.$$

Then by Picard's theorem (Theorem 3.1.2) there exists a solution on the interval $|t - t_0| < h$ for $x_0 : [-h, h] \to \mathbb{R}^n$. We may define Picard's iteration by

$$x_{n+1}(t) = x_0 + \int_{t_0}^{t} Ax_n(s)ds, \; x_0(t) = x_0, \tag{3.2.30}$$

Copyright Elsevier 2022

with the understanding that

$$x_i = (\phi_1^i, \phi_2^i, \dots, \phi_n^i)^T, \ i = 1, 2, \dots, n,$$

where $\phi_j^i \in \mathbb{R}^n$, $j = 1, 2, \dots, n$. Setting $n = 0$ in (3.2.30), we get

$$x_1 = x_0 + \int_{t_0}^{t} Ax_0 ds = x_0 + (t - t_0)Ax_0.$$

Similarly, for $n = 1$, we have

$$\begin{aligned} x_2 &= x_0 + \int_{t_0}^{t} Ax_1(s)ds = x_0 + \int_{t_0}^{t} A(s - t_0)Ax_0 ds \\ &= x_0 + (t - t_0)Ax_0 + \frac{(t - t_0)^2}{2} A^2 x_0. \end{aligned}$$

Using an induction argument, we arrive at

$$\begin{aligned} x_n &= x_0 + (t - t_0)Ax_0 + \frac{(t - t_0)^2}{2} A^2 x_0 + \ldots + \frac{(t - t_0)^n}{n!} A^n x_0 \\ &= \sum_{k=0}^{n} \frac{(t - t_0)^k}{k!} A^k x_0. \end{aligned}$$

Letting $n \to \infty$, we have

$$x(t) = e^{A(t - t_0)} x_0, \tag{3.2.31}$$

where

$$e^{A(t - t_0)} = \sum_{n=0}^{\infty} \frac{(t - t_0)^n}{n!} A^n. \tag{3.2.32}$$

It readily follows from the ratio test that the infinite series in (3.2.32) converges absolutely for all $t - t_0 \in \mathbb{R}$. Thus $e^{A(t - t_0)}$ exists for all times. Note that the infinite series given by (3.2.32) can be differentiated term by term, and hence

$$\begin{aligned} \frac{d}{dt} e^{A(t - t_0)} &= A \sum_{n=0}^{\infty} n \frac{(t - t_0)^n}{n!} A^n \\ &= \sum_{n=1}^{\infty} \frac{(t - t_0)^{n-1}}{(n - 1)!} A^n \\ &= A \sum_{n=1}^{\infty} \frac{(t - t_0)^{n-1}}{(n - 1)!} A^{n-1} \end{aligned}$$

Copyright Elsevier 2022

$$
\begin{aligned}
&= A \sum_{k=0}^{\infty} \frac{(t-t_0)^k}{k!} A^k \\
&= A e^{A(t-t_0)}.
\end{aligned}
$$

This shows that if

$$x(t) = e^{A(t-t_0)} x_0,$$

then $x(t)$ satisfies $x' = Ax$. Moreover,

$$x(t) = e^{A(t-t_0)} x_0 = \Big[I + (t-t_0)A + \frac{(t-t_0)^2}{2} A^2 + \ldots + \frac{(t-t_0)^n}{n!} A^n + \ldots \Big] x_0,$$

and hence

$$x(t_0) = I x_0 = x_0.$$

We have the following definition.

Definition 3.2.8. Let A be an $n \times n$ constant matrix. Then we define the *exponential matrix* function by $e^{A(t-t_0)}$, which is the solution of $x' = Ax$, and $e^{A(0)} = I$ (identity matrix).

Thus we have already proved that

$$x(t) = e^{A(t-t_0)} x_0 \tag{3.2.33}$$

is the unique solution of

$$x'(t) = Ax(t), \quad x(t_0) = x_0,$$

for all $t \in \mathbb{R}$.

Example 3.7. Consider the system

$$x' = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad x(0) = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

Then the matrix A is diagonal and given by

$$A = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}.$$

Since A is a diagonal matrix, by induction we can easily show that

$$A^n = \begin{pmatrix} 3^n & 0 \\ 0 & (-1)^n \end{pmatrix}, \; n = 1, 2, \ldots.$$

Copyright Elsevier 2022

Hence

$$\begin{aligned} &I + tA + \frac{(tA)^2}{2} + \ldots + \frac{(tA)^n}{n!} \\ &\quad = \begin{pmatrix} 1 + \frac{3t}{1!} + \frac{(3t)^2}{2!} + \ldots + \frac{(3t)^n}{n!} & 0 \\ 0 & 1 + \frac{(-t)}{1!} + \frac{(-t)^2}{2!} + \ldots + \frac{(-t)^n}{n!} \end{pmatrix}. \end{aligned}$$

We deduce from the exponential function properties that the exponential matrix e^{At} is given by

$$e^{At} = \begin{pmatrix} e^{3t} & 0 \\ 0 & e^{-t} \end{pmatrix},$$

and the solution is

$$x(t) = e^{At}x_0 = \begin{pmatrix} e^{3t} & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

In the next theorem we list some properties of the exponential matrix. For simplicity, we start at the initial time zero, that is, we take $t_0 = 0$. We are concerned with

$$x' = Ax, \tag{3.2.34}$$

where A is an $n \times n$ constant matrix, and $x \in \mathbb{R}^n$.

Theorem 3.2.11. *Let A and B be $n \times n$ constant matrices. Then:*

(1) $\frac{d}{dt}e^{At} = Ae^{At}$, $t \in \mathbb{R}$;
(2) $\det(e^{At}) \neq 0$, $t \in \mathbb{R}$, *and* e^{At} *is a fundamental matrix solution for* (3.2.34);
(3) $e^{At} = I + At + \frac{1}{2}(At)^2 + \frac{1}{6}(At)^3 + \cdots + \frac{1}{k!}(At)^k + \cdots$;
(4) $e^{At}e^{As} = e^{A(t+s)}$, $t, s \in \mathbb{R}$;
(5) $\left(e^{At}\right)^{-1} = e^{-At}$;
(6) *If* $AB = BA$, *then* $e^{At}B = Be^{At}$;
(7) *If* $AB = BA$, *then* $e^{At}e^{Bt} = e^{(A+Bt)}$, $t \in \mathbb{R}$;
(8) *If* P *is a nonsingular matrix, then* $e^{PBP^{-1}} = Pe^{B}P^{-1}$.

Proof. The proof of (1) is already done. For (2), we already know that e^{At} solves (3.2.34). Moreover, at $t = 0$, $e^{At} = I$, and hence $\det(e^{A0}) = \det(I) \neq 0$. By the uniqueness of solution this must hold for all $t \in \mathbb{R}$. Now since e^{At} solves (3.2.34),

Copyright Elsevier 2022

we have that e^{At} is a fundamental matrix for (3.2.34). The proof of (3) is already done. For (4), we let $L(t) = e^{At}e^{As} - e^{A(t+s)}$, where $s \in \mathbb{R}$ is fixed, and $t \in \mathbb{R}$. Then

$$L'(t) = Ae^{At}e^{As} - Ae^{A(t+s)} = A\left(e^{At}e^{As} - e^{A(t+s)}\right)$$

or $L'(t) = AL(t)$, from which we conclude that $L(t)$ solves (3.2.34). On the other hand, $L(0) = e^{As} - e^{As} = 0$, and hence by the uniqueness of solutions we get that $L(t) = 0$ for all $t \in \mathbb{R}$. This completes the proof of (4).

By the proof of (4), we have that

$$e^{At}e^{-At} = e^{At-At} = e^{A(t+(-t))} = I.$$

This implies that

$$\left(e^{At}\right)^{-1} = e^{-At}, \ t \in \mathbb{R}. \qquad \square$$

Remark 3.3. In the case where A is an $n \times n$ constant matrix, the exponential matrix is equivalent to

$$e^{A(t-t_0)} = \Phi(t)\Phi^{-1}(t_0),$$

which is the principal matrix.

Next, we discuss an important topic that relies on the *Jordan canonical form*, which helps us find the exponential matrix e^{At}. The next theorem is well known in linear algebra, and you may consult Brauer and Nohel [17] for its proof.

Theorem 3.2.12. *(Jordan canonical form) For every $n \times n$ matrix A, we can construct a matrix C such that*

$$C^{-1}AC = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & j_k \end{pmatrix}, \tag{3.2.35}$$

where

$$J_k = \lambda_k I + N_k$$

with

$$N_k = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Copyright Elsevier 2022

Remark 3.4. (The simplest case) If A has n real and distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then the matrix $C^{-1}AC$ will be a diagonal matrix with entries λ_i, that is,

$$C^{-1}AC = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

In this case the matrix C is formed from the corresponding eigenvectors of the matrix A.

If (3.2.35) holds, then let

$$B = C^{-1}AC,$$

and hence

$$e^{tA} = e^{tCBC^{-1}} = Ce^{tB}C^{-1}.$$

In Section 3.2.1, we encountered different scenarios on the type and nature of the eigenvalues when computing the eigenvectors. The next result, called *Putzer algorithm*, allows us to find the fundamental matrix or, in particular, the exponential matrix e^{At} when the matrix is constant, regardless of the order or multiplicities of the eigenvalues. Putzer's algorithm depends on the next theorem, known as the Cayley–Hamilton theorem.

Theorem 3.2.13. *(Cayley–Hamilton) Every $n \times n$ constant matrix satisfies its characteristic equation*

$$p_A(\lambda) := \det(A - \lambda I) = 0.$$

In other words, if

$$p_A(\lambda) = (-1)^n \Big(\lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n\Big),$$

then

$$p_A(A) = (-1)^n \Big(A^n + c_1 A^{n-1} + \cdots + c_{n-1}A + c_n I\Big) = 0.$$

Theorem 3.2.14. *(Putzer algorithm) Suppose A is an $n \times n$ constant matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, which may come in any order or multiplicities. Suppose $\mu_1(t), \mu_2(t), \ldots, \mu_n(t)$ satisfy the corresponding initial value problems*

$$\mu_1' = \lambda_1\mu_1,\ \mu_1(0) = 1;\ \mu_j' = \lambda_j\mu_j + \mu_{j-1},\ \mu_j(0) = 0,\ j = 2, 3, \ldots, n. \tag{3.2.36}$$

Copyright Elsevier 2022

Let the matrices $P_0, P_1, \ldots, P_n$ *be defined recursively by*

$$P_0 = I, \quad P_j = \prod_{k=1}^{j}(A - \lambda_k I), \ j = 1, 2, \ldots, n. \tag{3.2.37}$$

Then

$$e^{At} = \sum_{j=0}^{n-1} \mu_{j+1}(t) P_j. \tag{3.2.38}$$

Proof. Let

$$\Phi(t) = \sum_{j=0}^{n-1} \mu_{j+1}(t) P_j.$$

Then

$$\begin{aligned}\Phi(0) &= \sum_{j=0}^{n-1} \mu_{j+1}(0) P_j \\ &= \mu_1(0)P_0(0) + \mu_2(0)P_1(0) + \ldots + \mu_n(0)P_{n-1}(0) \\ &= \mu_1(0)P_0(0) + 0 + \ldots + 0 \ \ \text{(by (3.2.36))} \\ &= \mu_1(0)I = I \ \ \text{(by (3.2.37))}.\end{aligned}$$

Next, we differentiate Φ with respect to t:

$$\begin{aligned}\Phi'(t) &= \sum_{j=0}^{n-1} \mu'_{j+1}(t) P_j \\ &= \lambda_1 \mu_1 P_0 + \sum_{j=1}^{n-1} \mu'_{j+1}(t) P_j, \ \ \text{(by (3.2.36))} \\ &= \lambda_1 \mu_1 P_0 + \sum_{j=1}^{n-1} (\lambda_{j+1}\mu_{j+1} + \mu_j) P_j, \ \ \text{(by (3.2.36))}.\end{aligned} \tag{3.2.39}$$

On the other hand,

$$A\Phi = \sum_{j=0}^{n-1} \mu_{j+1}(t) A P_j.$$

From (3.2.37) we have

$$P_j = \prod_{k=1}^{j}(A - \lambda_k I) = (A - \lambda_j I) \prod_{k=1}^{j-1}(A - \lambda_k I) = (A - \lambda_j I) P_{j-1}.$$

Copyright Elsevier 2022

This implies that

$$AP_j = P_{j+1} + \lambda_{j+1} P_j, \quad j = 1, 2, \ldots n-1.$$

Substituting AP_j into $A\Phi$ in the above expression yields

$$A\Phi = \sum_{j=0}^{n-1} \mu_{j+1}(t)[P_{j+1} + \lambda_{j+1} I P_j],$$

$$\begin{aligned} A\Phi = {} & \lambda_1 \mu_1(t) P_0 + (\lambda_2 \mu_2(t) + \mu_1(t)) P_1 + \ldots \\ & + (\lambda_n \mu_n(t) + \mu_{n-1}(t)) P_{n-1} + \mu_n(t) P_n. \end{aligned}$$

By the Cayley–Hamilton theorem $P_n = 0$, and hence the last expression of $A\Phi$ is the same as the right side of (3.2.39) so that we have $\Phi'(t) = A\Phi(t)$ with $\Phi(0) = I$. Thus expression (3.2.38) holds by the uniqueness of the solution. This completes the proof. ☐

Example 3.8. Consider the linear system

$$x' = \begin{pmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

satisfying the initial condition

$$x(0) := x_0 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

The matrix

$$A = \begin{pmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{pmatrix}$$

has the eigenvalues $3, 3, -3$. We may order them arbitrarily. Let, for example, $\lambda_1 = 3$, $\lambda_2 = 3$, and $\lambda_3 = -3$. Using (3.2.37) with $j = 1, 2$, we see that

$$P_0 = I, \quad P_1 = A - 3I, \quad P_2 = (A - 3I)^2.$$

Similarly, by (3.2.36)

$$\begin{aligned} \mu_1' &= 3\mu_1, \ \mu_1(0) = 1; \\ \mu_2' &= 3\mu_2 + \mu_1(t), \ \mu_2(0) = 0; \\ \mu_3' &= -3\mu_3 + \mu_2(t), \ \mu_3(0) = 0. \end{aligned}$$

Copyright Elsevier 2022

Solving the above initial value problems gives

$$\mu_1(t) = e^{3t}, \quad \mu_2(t) = te^{3t}, \quad \text{and} \quad \mu_3(t) = \frac{te^{3t}}{6} - \frac{e^{3t}}{36} + \frac{e^{-3t}}{36}.$$

By using (3.2.38) with $n = 3$, we obtain

$$e^{At} = \sum_{j=0}^{2} \mu_{j+1}(t) P_j = \mu_1 P_0 + \mu_2 P_1 + \mu_3 P_2,$$

and the unique solution is given by

$$x(t) = e^{At} x_0.$$

Example 3.9. Assume we have three masses m_1, m_2, and m_3 that are connected to each other by four springs with spring constants k_1, k_2, k_3, and k_4, respectively. As seen in Fig. 3.2, the first and last springs are attached to two fixed walls, and the masses slide along a frictionless horizontal surface. Displacement of the masses from their equilibrium positions is determined by the coordinates x, y, and z, each positive to the right. The motion of the masses depends strongly on the initial conditions, that is, by hand, any of the masses is displaced by pulling forward or backward, and we let them go. Next, we attempt to find the equations of motion. The springs operate according to Hooke's law: Force = spring constant multiplied by elongation. The springs restore after compression and extension. In addition, the first three springs are elongated by x, $y - x$, $z - y$, respectively. Then using Newton's second law, Force $=$ sum of the Hooke forces, we arrive at the system of second-order differential equations

$$\begin{aligned} m_1 x''(t) &= -k_1 x(t) + k_2[y(t) - x(t)], \\ m_2 y''(t) &= -k_2[y(t) - x(t)] + k_3[z(t) - y(t)], \\ m_3 z''(t) &= -k_3[z(t) - y(t)] - k_4 z(t). \end{aligned} \tag{3.2.40}$$

System (3.2.40) can be written as a second-order vector–matrix system

$$M\mathbf{x}''(t) = A\mathbf{x}(t),$$

where the *displacement* $\mathbf{x}$, *mass matrix* M, and *stiffness matrix* A are given by

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad M = \begin{pmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{pmatrix}, \quad A = \begin{pmatrix} -k_1 - k_2 & k_2 & 0 \\ k_2 & -k_2 - k_3 & k_3 \\ 0 & k_3 & -k_3 - k_4 \end{pmatrix}.$$

Copyright Elsevier 2022

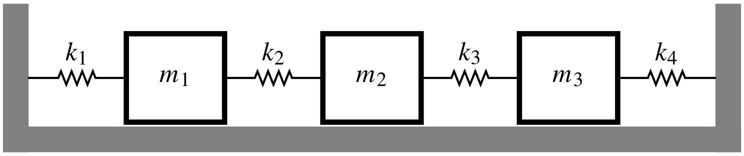


FIGURE 3.2

Three masses connected by four springs.

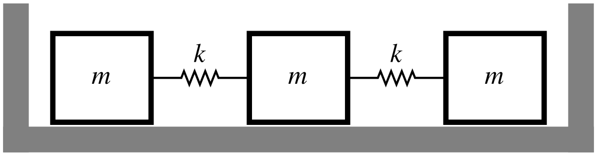


FIGURE 3.3

Three identical boxcars connected by identical springs.

Set $m_1 = m_2 = m_3 = 1$, $k_1 = 2$, $k_2 = 1$, $k_3 = 1$, and $k_4 = 2$ in the second-order vector–matrix system

$$M\mathbf{x}''(t) = A\mathbf{x}(t).$$

Then using Putzer's algorithm, we find the solution

$$\begin{aligned}\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \big(a_1 \cos(t) + b_1 \sin(t)\big) \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \\ &+ \big(a_2 \cos(\sqrt{3}t) + b_2 \sin(\sqrt{3}t)\big) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ &+ \big(a_3 \cos(2t) + b_3 \sin(2t)\big) \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix},\end{aligned}$$

where a_1, a_2, b_1, b_2, a_3, and b_3 are arbitrary constants.

Example 3.10. (Boxcars) A special case of mass-spring system is *boxcars*. In Fig. 3.3 we assume three identical boxcars on a level track connected by two identical springs. Excepts for the springs on fixed ends, this problem is the same as in Example 3.9. With this in mind, we have that $k_1 = k_4 = 0$ and $k_2 = k_3 = k$. In addition, $m_1 = m_2 = m_3 = m$. Consequently, we arrive at the system

$$M\mathbf{x}''(t) = A\mathbf{x}(t),$$

Copyright Elsevier 2022

where the *displacement* $\mathbf{x}$, *mass matrix* M, and *stiffness matrix* A are given by

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad M = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix}, \quad A = \begin{pmatrix} -k & k & 0 \\ k & -2k & k \\ 0 & k & -k \end{pmatrix}.$$

For illustration, we take $\frac{k}{m} = 1$. Then we have the second-order system

$$\mathbf{x}'' = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix} \mathbf{x},$$

which has the vector solution

$$\begin{aligned} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \left(a_1 + b_1 t\right) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &\quad + \left(a_2 \cos(t) + b_2 \sin(t)\right) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ &\quad + \left(a_3 \cos(\sqrt{3}t) + b_3 \sin(\sqrt{3}t)\right) \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \end{aligned}$$

where a_1, a_2, a_3, b_1, b_2, and b_3 are arbitrary constants.

3.3 $x' = A(t)x + g(t)$

Now we consider the non-homogeneous system

$$x' = A(t)x + g(t), \tag{3.3.1}$$

where $A(t)$ is an $n \times n$ matrix, and $g(t)$ is an $n \times 1$ vector, both continuous on some interval I. It was proven in Theorem 3.2.1 that system (3.3.1) has a unique solution on I if an initial condition is specified for $t_0 \in I$. We are interested in finding the *general solution* of the non-homogeneous system (3.3.1). With the non-homogeneous system we associate the complementary homogeneous system

$$x' = A(t)x. \tag{3.3.2}$$

Knowing the solution of (3.3.2) is enough to find the general solution of (3.3.1), and such a solution will be given by the *variation of parameters formula*.

Copyright Elsevier 2022

Definition 3.3.1. Let x_h be the homogeneous solution of (3.3.2), and let x_p be the particular solution of (3.3.1) on an interval I. Then the general solution $x : I \to \mathbb{R}^n$ of (3.3.1) on the interval I is given by

$$x(t) = x_h(t) + x_p(t).$$

We have seen that if Φ is the fundamental matrix of the homogeneous system (3.3.2), then $x_h = \Phi(t)k$ for some constant $n \times 1$ vector k.

Theorem 3.3.1. *Let $A(t)$ be an $n \times n$ matrix, and let $g(t)$ be an $n \times 1$ vector, both continuous on some interval I. Suppose Φ is the fundamental matrix of the homogeneous system (3.3.2). Then $x(t)$ is a solution of*

$$x' = A(t)x + g(t), \quad x(t_0) = x_0, \ t \geq t_0, \tag{3.3.3}$$

on I if and only if x satisfies

$$x(t) = \Phi(t)\Phi^{-1}(t_0)x_0 + \int_{t_0}^{t} \Phi(t)\Phi^{-1}(s)g(s)ds, \tag{3.3.4}$$

where $t_0 \in I$ and $x_0 \in \mathbb{R}^n$.

Proof. First, note that the uniqueness of the solution of (3.3.3) follows from Theorem 3.2.1. Let Φ be the fundamental matrix of the homogeneous system (3.3.2). Suppose x satisfies (3.3.4). Taking the derivative in t in (3.3.4) gives

$$\begin{aligned}
x'(t) &= \Phi'(t)\Phi^{-1}(t_0)x_0 + \Phi'(t)\int_{t_0}^{t} \Phi^{-1}(s)g(s)ds + \Phi(t)\Phi^{-1}(t)g(t) \\
&= A(t)\Phi(t)\Phi^{-1}(t_0)x_0 + A(t)\Phi(t)\int_{t_0}^{t} \Phi^{-1}(s)g(s)ds + g(t) \\
&= A(t)\Big[\Phi(t)\Phi^{-1}(t_0)x_0 + \Phi(t)\int_{t_0}^{t} \Phi^{-1}(s)g(s)ds\Big] + g(t) \\
&= A(t)x(t) + g(t).
\end{aligned}$$

Finally, replacing t by t_0 in (3.3.4) gives $x(t_0) = Ix_0 = x_0$. Now let $x(t)$ be the unique solution of (3.3.3), and let us show that it satisfies (3.3.4). As we have mentioned earlier, $x_h = \Phi(t)k$ for some constant $n \times 1$ vector k. To derive the particular solution x_p, we look for a particular solution of the form

$$x_p(t) = \Phi(t)u(t)$$

for some continuous function u on I. Differentiating x_p and plugging back into (3.3.3), we have

$$x_p' = \Phi'(t)u(t) + \Phi(t)u'(t)$$

Copyright Elsevier 2022

or

$$\Phi'(t)u(t) + \Phi(t)u'(t) = A(t)\Phi(t)u(t) + g(t).$$

Since

$$\Phi'(t) = A(t)\Phi(t),$$

the above expression reduces to

$$\Phi(t)u'(t) = g(t).$$

Multiplying the above expression by $\Phi^{-1}(t)$ from the left, we get $u'(t) = \Phi^{-1}(t)g(t)$ or $u(t) = \int_{t_0}^{t} \Phi^{-1}(s)g(s)ds$. Finally, the equality $x_p(t) = \Phi(t)u(t)$ gives

$$x_p(t) = \Phi(t)\int_{t_0}^{t} \Phi^{-1}(s)g(s)ds.$$

It follows that the general solution is

$$x(t) = x_h(t) + x_p(t) = \Phi(t)k + \int_{t_0}^{t} \Phi(t)\Phi^{-1}(s)g(s)ds. \tag{3.3.5}$$

Using the initial condition in (3.3.5), we get

$$x_0 = \Phi(t_0)k + 0$$

or

$$k = \Phi^{-1}(t_0)x_0.$$

A substitution of k into (3.3.5) gives (3.3.4). This completes the proof. ☐

Eq. (3.3.4) is called the *variation of parameters formula*.

Remark 3.5. In the case where A is an $n \times n$ constant matrix the exponential matrix is equivalent to

$$e^{A(t-t_0)} = \Phi(t)\Phi^{-1}(t_0),$$

which is the principal matrix. In this case, (3.3.4) reduces to or takes the form

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-s)}g(s)ds. \tag{3.3.6}$$

Example 3.11. Solve

$$x'(t) = \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix} x + \begin{pmatrix} -15te^{-2t} \\ -4te^{-2t} \end{pmatrix}, \quad x(0) = \begin{pmatrix} 7 \\ 3 \end{pmatrix}.$$

Copyright Elsevier 2022

Using Putzer's algorithm, we find that

$$e^{At} = \frac{1}{7}\begin{pmatrix} e^{-2t} + 6e^{5t} & -2e^{-2t} + 2e^{5t} \\ -3e^{-2t} + 3e^{5t} & 6e^{-2t} + e^{5t} \end{pmatrix}.$$

Note that $e^{A0} = I$. Using formula (3.3.6), we have

$$\begin{aligned} x(t) &= \frac{1}{7}\begin{pmatrix} e^{-2t} + 6e^{5t} & -2e^{-2t} + 2e^{5t} \\ -3e^{-2t} + 3e^{5t} & 6e^{-2t} + e^{5t} \end{pmatrix}\begin{pmatrix} 7 \\ 3 \end{pmatrix} \\ &\quad + \int_{t_0}^{t} \frac{1}{7}\begin{pmatrix} e^{-2(t-s)} + 6e^{5(t-s)} & -2e^{-2(t-s)} + 2e^{5(t-s)} \\ -3e^{-2(t-s)} + 3e^{5(t-s)} & 6e^{-2(t-s)} + e^{5(t-s)} \end{pmatrix}\begin{pmatrix} -15se^{-2s} \\ -4se^{-2s} \end{pmatrix} ds \\ &= \frac{1}{14}\begin{pmatrix} (6 + 28t - 7t^2)e^{-2t} + 92e^{5t} \\ (-4 + 14t + 21t^2)e^{-2t} + 46e^{5t} \end{pmatrix}. \end{aligned}$$

3.4 Discussion

When the matrix A is not constant and we wish to compute the fundamental matrix of the system

$$x' = A(t)x, \tag{3.4.1}$$

where $A(t)$ is an $n \times n$ matrix with continuous components on some interval I, then the Putzer algorithm will not help much. Suppose we are able to find two matrices Λ and S such that

$$A = S\Lambda S^{-1}. \tag{3.4.2}$$

Then we claim that also

$$e^{At} = Se^{\Lambda t}S^{-1} \tag{3.4.3}$$

for all t. To prove the claim, we use (3.2.32). As a consequence of (3.4.2), we have that $At = S(\Lambda t)S^{-1}$, and hence

$$(At)^k = \big(S(\Lambda t)S^{-1}\big)\big(S(\Lambda t)S^{-1}\big)\ldots\big(S(\Lambda t)S^{-1}\big).$$

Using (3.2.32) with $t_0 = 0$, we get

$$\begin{aligned} e^{At} &= I + At + \frac{1}{2}(At)^2 + \frac{1}{6}(At)^3 + \ldots + \frac{1}{k!}(At)^k + \ldots \\ &= I + S(\Lambda t)S^{-1} + \frac{1}{2}S(\Lambda t)^2S^{-1} + \frac{1}{6}S(\Lambda t)^3S^{-1} + \ldots + \frac{1}{k!}S(\Lambda t)^kS^{-1} + \ldots \end{aligned}$$

Copyright Elsevier 2022

$$
\begin{aligned}
&= S\big[I + \Lambda t + \frac{1}{2}(\Lambda t)^2 + \frac{1}{6}(\Lambda t)^3 + \ldots + \frac{1}{k!}(\Lambda t)^k + \ldots\big]S^{-1} \\
&= Se^{\Lambda t}S^{-1},
\end{aligned}
$$

as claimed. Thus we state the general theorem.

Theorem 3.4.1. *For every $n \times n$ matrix A with entries in complex numbers, we can find an invertible matrix S and an upper triangular matrix Λ such that (3.4.2) holds.*

Another point of discussion is that if we integrate (3.4.1) from t_0 to t, then we get $x(t) = e^{\int_{t_0}^t A(s)ds}$. Now let

$$\Phi(t) = e^{\int_{t_0}^t A(s)ds}. \tag{3.4.4}$$

For $\Phi(t)$ to be fundamental matrix solution, we must have

$$A(t)\Big(\int_{t_0}^t A(s)ds\Big) = \Big(\int_{t_0}^t A(s)ds\Big)A(t). \tag{3.4.5}$$

Let us see why. Let $J = \int_{t_0}^t A(s)ds$. Then

$$e^J = I + J + \frac{1}{2!}J^2 + \frac{1}{3!}J^3 + \ldots + \frac{1}{k!}J^k + \cdots,$$

and

$$
\begin{aligned}
\frac{d}{dt}e^{\int_{t_0}^t A(s)ds} &= \frac{d}{dt}\Big(I + \int_{t_0}^t A(s)ds + \frac{1}{2!}\Big(\int_{t_0}^t A(s)ds\Big)^2 \\
&+ \ldots + \frac{1}{(k-1)!}\Big(\int_{t_0}^t A(s)ds\Big)^{k-1}A(t) + \frac{1}{k!}\Big(\int_{t_0}^t A(s)ds\Big)^k + \ldots\Big) \\
&= A(t) + \int_{t_0}^t A(s)ds\ A(t) + \frac{1}{2!}\Big(\int_{t_0}^t A(s)ds\Big)^2 A(t) \\
&+ \ldots + \frac{1}{(k-1)!}\Big(\int_{t_0}^t A(s)ds\Big)^{k-1}A(t) + \frac{1}{k!}\Big(\int_{t_0}^t A(s)ds\Big)^k A(t) + \ldots \\
&= \big[I + \int_{t_0}^t A(s)ds + \frac{1}{2!}\Big(\int_{t_0}^t A(s)ds\Big)^2 + \ldots \\
&+ \frac{1}{(k-1)!}\Big(\int_{t_0}^t A(s)ds\Big)^{k-1}A(t) + \frac{1}{k!}\Big(\int_{t_0}^t A(s)ds\Big)^k + \ldots\big]A(t) \\
&\neq A(t)\big[I + \int_{t_0}^t A(s)ds + \frac{1}{2!}\Big(\int_{t_0}^t A(s)ds\Big)^2 \\
&+ \ldots + \frac{1}{(k-1)!}\Big(\int_{t_0}^t A(s)ds\Big)^{k-1}A(t) + \frac{1}{k!}\Big(\int_{t_0}^t A(s)ds\Big)^k + \ldots\big] \\
&= A(t)\Phi(t).
\end{aligned}
$$

Copyright Elsevier 2022

Thus if (3.4.5) holds, then $\Phi'(t) = A(t)\Phi(t)$. We have the following application of the above discussion.

Example 3.12. Find the fundamental matrix $\Phi(t)$ for the system

$$x' = \begin{pmatrix} 1 & t \\ t & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Let

$$A(t) = \begin{pmatrix} 1 & t \\ t & 1 \end{pmatrix}$$

and

$$\Phi(t) = e^{\int_0^t A(s)ds}.$$

Now we verify condition (3.4.5):

$$A(t)\int_{t_0}^t A(s)ds = \begin{pmatrix} 1 & t \\ t & 1 \end{pmatrix} \begin{pmatrix} t & t^2/2 \\ t^2/2 & t \end{pmatrix} = \begin{pmatrix} t + t^3/2 & 3t^2/2 \\ 3t^2/2 & t + t^3/2 \end{pmatrix}.$$

On the other hand,

$$\Big(\int_{t_0}^t A(s)ds\Big)A(t) = \begin{pmatrix} t & t^2/2 \\ t^2/2 & t \end{pmatrix} \begin{pmatrix} 1 & t \\ t & 1 \end{pmatrix} = \begin{pmatrix} t + t^3/2 & 3t^2/2 \\ 3t^2/2 & t + t^3/2 \end{pmatrix}.$$

The matrix $\int A(t)dt$ has the eigenvalues

$$\lambda_1 = t + \frac{t^2}{2} \text{ and } \lambda_2 = t - \frac{t^2}{2}.$$

Luckily, the corresponding eigenvectors are independent of t. The eigenpairs are

$$\lambda_1, K_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix};\ \lambda_2, K_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Let

$$H = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Then the inverse of H is

$$H^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Next, we try to diagonalize $\int_0^t A(s)ds$ by the transformation

$$J = H^{-1}\Big(\int_0^t A(s)ds\Big)H := \begin{pmatrix} t + \frac{t^2}{2} & 0 \\ 0 & t - \frac{t^2}{2} \end{pmatrix}.$$

Copyright Elsevier 2022

As consequence of (3.4.3), we get the exponential matrix of J,

$$e^J = \begin{pmatrix} e^{t+\frac{t^2}{2}} & 0 \\ 0 & e^{t-\frac{t^2}{2}} \end{pmatrix}.$$

We can now calculate the fundamental matrix $\Phi(t)$:

$$\begin{aligned}\Phi(t) = e^{\int_{t_0}^{t} A(s)ds} &= He^J H^{-1} \\ &= \frac{e^t}{2}\begin{pmatrix} e^{\frac{t^2}{2}} + e^{-\frac{t^2}{2}} & e^{\frac{t^2}{2}} - e^{-\frac{t^2}{2}} \\ e^{\frac{t^2}{2}} - e^{-\frac{t^2}{2}} & e^{\frac{t^2}{2}} + e^{-\frac{t^2}{2}} \end{pmatrix}.\end{aligned}$$

3.5 Exercises

Exercise 3.1. Find the region of existence of the following system and then find the maximal interval of existence:

$$\begin{aligned} x_1' &= -\frac{x_3}{x_2^2}, \\ x_2' &= 1, \\ x_3' &= \frac{x_1}{x_2^2}, \end{aligned}$$

with initial vector

$$\begin{pmatrix} x_1(\frac{1}{\pi}) \\ x_2(\frac{1}{\pi}) \\ x_3(\frac{1}{\pi}) \end{pmatrix} = \begin{pmatrix} -1 \\ \frac{1}{\pi} \\ 0 \end{pmatrix}.$$

Exercise 3.2. Show that $||A||_1$ is a norm, that is, verify all the items in Definition 3.2.2.

Exercise 3.3. Show that if A is an $n \times n$ invertible matrix, then

$$||A|| > 0, \text{ and } ||A^{-1}|| \geq \frac{1}{||A||}.$$

Exercise 3.4. Suppose an $n \times n$ matrix $A(t)$ and an $n \times 1$ vector $g(t)$ are continuous on $-\infty < t < \infty$, $\int_{-\infty}^{\infty} |A(t)|dt < \infty$, and $\int_{-\infty}^{\infty} |g(t)|dt < \infty$. Show that the solution $\phi(t)$ of $y'(t) = A(t)y(t) + g(t)$, exists for $-\infty < t < \infty$. In addition, compute the upper bound for $|\phi(t)|$ on $-\infty < t < \infty$.

Copyright Elsevier 2022

Exercise 3.5. In Example 3.4 keep the norm of the matrix A as a function of t and obtain a slightly better estimate of the error.

Exercise 3.6. Prove that every solution of

$$x'' - x' + t^2 x = t^4$$

exists for all $t \in \mathbb{R}$.

Exercise 3.7. Use the method of *energy function* to show that the solutions

$$x'' + \cos^2(x)x' - \sin^2(x) = 0$$

exist for all times, that is, the maximal interval of existence is $[0, \infty)$.

Exercise 3.8. Obtain two iterations of the 3 × 3 linear system

$$\begin{aligned} x_1'(t) &= 5x_1 - x_2, \\ x_2'(t) &= -5x_2 + 9x_3, \\ x_3'(t) &= -5x_1 - x_2, \end{aligned}$$

with the initial condition

$$\mathbf{x}(0) = \mathbf{x_0} = \begin{pmatrix} x_{01} \\ x_{02} \\ x_{03} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Exercise 3.9. Use Theorem 3.2.10 to solve, for all $t \geq t_0$, the linear system

$$x' = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

with initial condition

$$x(t_0) := x_0 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Exercise 3.10. Use Theorem 3.2.10 to solve the linear system with the initial condition

$$x' = \begin{pmatrix} 1 & -12 & -14 \\ 1 & 2 & -3 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad x(0) = x_0 := \begin{pmatrix} 4 \\ 6 \\ -7 \end{pmatrix}.$$

Exercise 3.11. **(a)** Let

$$x'(t) = Ax(t),$$

Copyright Elsevier 2022

where A is a constant $n \times n$ matrix. Suppose A has only one eigenvector K associated with an eigenvalue λ_1 of multiplicity three. We can find a second solution of the form (3.2.26), where the eigenvector P satisfies (3.2.28). Show that if we assume a third solution $\phi_3(t)$ of the form

$$\phi_3(t) = K\frac{t^2}{2}e^{\lambda_1 t} + Pte^{\lambda_1 t} + Qe^{\lambda_1 t},$$

then the vector Q is given by the equation

$$(A - \lambda_1 I)Q = P.$$

(b) Find the fundamental matrix for the linear system

$$x' = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Exercise 3.12. **(a)** Let

$$x'(t) = Ax(t),$$

where A is an $n \times n$ constant matrix with real entries. Let K be an eigenvector corresponding to the complex eigenvalue $\lambda_1 = \alpha + i\beta$ (with real α and β). Show that the corresponding two linearly independent solutions $\phi_1(t)$ and $\phi_2(t)$ are given by

$$\phi_1(t) = \Big(\mathrm{Re}(K)\cos(\beta t) - \mathrm{Im}(K)\sin(\beta t)\Big)e^{\alpha t}$$

and

$$\phi_2(t) = \Big(\mathrm{Im}(K)\cos(\beta t) + \mathrm{Re}(K)\sin(\beta t)\Big)e^{\alpha t}.$$

For example, if $K = \left(\begin{smallmatrix} 1 \\ 2i \end{smallmatrix}\right)$, then

$$K = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2i \end{pmatrix} = \mathrm{Re}(K) + \mathrm{Im}(K).$$

(b) Find the fundamental matrix for the linear system

$$x' = \begin{pmatrix} 1 & 2 \\ -1/2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Exercise 3.13. Let Φ be a fundamental matrix of $x' = Ax$, and let C be a nonsingular matrix with $AC = CA$. Prove that now $C\Phi$ is a fundamental matrix of $x' = Ax$. (See Theorem 3.2.5.)

Exercise 3.14. Prove parts (5), (6), and (7) of Theorem 3.2.11.

Copyright Elsevier 2022

Exercise 3.15. Finish the calculation in Example 3.8 and show that

$$e^{At} = \cosh(3t)I + \frac{1}{3}\sinh(3t)A.$$

Exercise 3.16. Use Theorem 3.2.14 to compute e^{At} for

(a)

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix},$$

(b)

$$A = \begin{pmatrix} 5 & -4 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 5 \end{pmatrix}.$$

Exercise 3.17. Use Theorem 3.2.14 to compute e^{At} for

$$x' = \begin{pmatrix} -1 & -2 & 2 \\ 3 & 5 & -3 \\ 3 & 4 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

satisfying the initial condition

$$x(0) = x_0 := \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Exercise 3.18. Consider two non-homogeneous systems $x' = Ax + g_1(t)$ and $x' = Ax + g_2(t)$ with solutions z_1 and z_2, respectively. Let $g(t) = g_1(t) + g_2(t)$. Show that $z_1 + z_2$ solves $x' = Ax + g(t)$. Here g_1, g_2 are $n \times 1$ vector functions that are continuous on $\mathbb{R}$.

Exercise 3.19. Use the variation of parameters formula to solve the given system subject to the indicated initial condition.

(a)

$$x'(t) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} x + \begin{pmatrix} e^t \\ e^{2t} \\ te^{3t} \end{pmatrix}, \quad x(0) = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

(b)

$$x'(t) = \begin{pmatrix} 3 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} x + \begin{pmatrix} 0 \\ t \\ 2e^t \end{pmatrix}, \quad x(0) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Copyright Elsevier 2022

Exercise 3.20. Imitate the proof that leads to the variation of parameters formula (1.5.2) for non-homogeneous scalar differential equations to derive (3.3.6) for the non-homogeneous system

$$x'(t) = Ax(t) + g(t), \quad x(t_0) = x_0,$$

where A is an $n \times n$ constant matrix, and g is continuous on some interval I.

Exercise 3.21. Find the fundamental matrix $\Phi(t)$ for the system

$$x' = \begin{pmatrix} -\frac{1}{t} & \frac{2}{t^2} \\ -1 & \frac{3}{t} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Exercise 3.22. Assume the fundamental matrix $\Phi(t)$ of a certain 2×2 differential system is given by

$$\Phi(t) = \begin{pmatrix} 2 & t \\ t & t^2 \end{pmatrix}.$$

Find the corresponding system of differential equations.

Exercise 3.23. Write the system

$$x_1' = -x_1 + x_2 e^{2t}, \quad x_2' = -x_2$$

in the form of

$$x'(t) = A(t)x(t).$$

Compute the eigenvalues of $A(t)$ and then find the solution of the system. Are solutions bounded or unbounded?

Copyright Elsevier 2022

CHAPTER

Stability of linear systems

4

We look at the stability of linear systems via the variation of parameters. The chapter also includes a nice section on Floquet theory with its application to Mathieu's equation.

4.1 Definitions and examples

Consider the scalar differential equation

$$x' = ax, \ t \geq 0, \tag{4.1.1}$$

for constant a. Then (4.1.1) has a solution $x(t) = Ce^{at}$ for every constant C. Note that $x(t) = 0$ is a solution of (4.1.1) since it satisfies both sides of the equation. It is clear that solutions go to infinity for $a > 0$ and decay to zero for $a < 0$ as $t \to \infty$. Thus the behavior of the solutions depends on the sign of the constant a. It seems that the solution $x = 0$ is an attractor asymptotically stable in the case $a < 0$, which we will discuss in detail in this chapter.

Notice that if $a = 0$, then solutions are constants. Thus if we start near the zero solution $x = 0$, then we stay near the zero solution, and hence the zero solution is stable but not asymptotically stable.

In the above example, we were able to find the solution. However, this is not the case for the nonlinear differential equation

$$x'(t) = \frac{x(x-1)(x+2)}{1+\sin^2(x)}, \ t \geq 0, \ x(0) = x_0. \tag{4.1.2}$$

Eq. (4.1.2) is nonlinear, and the true solution cannot be explicitly found. On the other hand, we notice that $x = -2, 0, 1$ are all *constant solutions* or *equilibrium solutions* of (4.1.2). However, we can use calculus to analyze those constant solutions. Note that the right side of (4.1.2) is continuous on $\mathbb{R}$, and hence a solution exists for all $t \geq 0$. Moreover, any solution passing through (t_0, x_0) is unique. For the next discussion, we refer the reader to Fig. 4.1.

(a) For $x < -2$, we have from (4.1.2) that $x' < 0$, and hence solutions grow negatively unbounded; in particular, they pull away from the constant solution -2.

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00010-3
Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

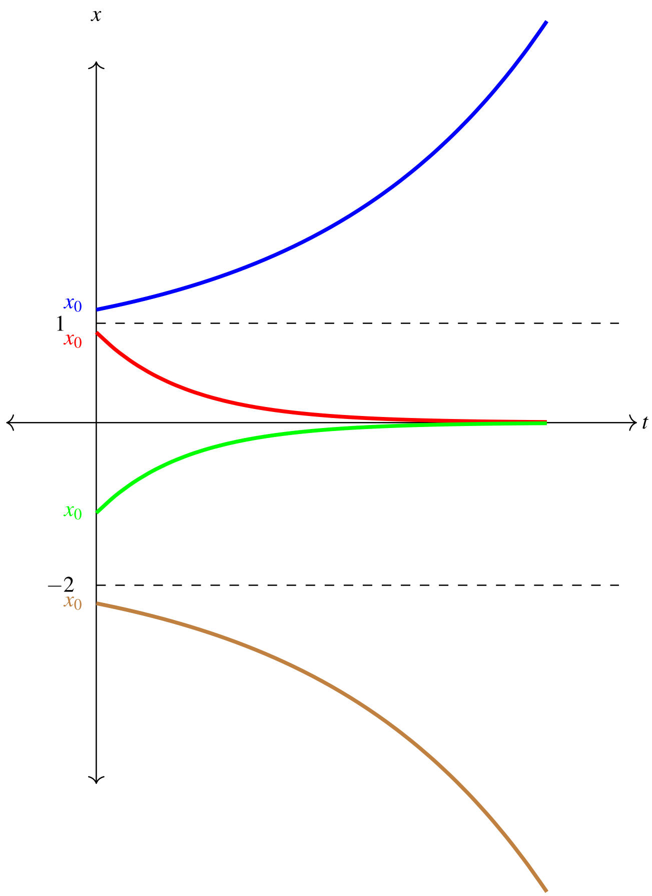


FIGURE 4.1

Diagram for the constant solutions.

- **(b)** For $-2 < x < 0$, we have from (4.1.2) that $x' > 0$, and hence solutions increase toward the constant solution $x = 0$ and away from the constant solution $x = -2$.
- **(c)** For $0 < x < 1$, we have from (4.1.2) that $x' < 0$, and hence solutions decrease toward the constant solution $x = 0$ and away from the constant solution $x = 1$. Note that in (b) and (c) solutions cannot cross the constant solution $x = 0$ due to the uniqueness. Finally,
- **(d)** For $x > 1$, we have from (4.1.2) that $x' > 0$, and hence solutions unboundedly increase and away from the constant solution $x = 1$.

Thus by the above discussion and Fig. 4.1, we see that the constant solution $x = 0$ is an attractor (asymptotically stable), and the other constant solutions $x = -2, 1$ are repellers (unstable), which we discuss later.

Copyright Elsevier 2022

Consider the system of ordinary differential equations

$$x' = f(t, x), \tag{4.1.3}$$

where $f \in C\big([0, \infty) \times D, \mathbb{R}^n\big)$, and $D \subset \mathbb{R}^n$ is open. We say that a vector $x^* \in \mathbb{R}^n$ is an *equilibrium*, or *constant solution*, or *equilibrium solution* of (4.1.3) if

$$f(t, x^*) = 0, \;\; 0 \le t < w,$$

for positive constant w. In practice, it is easier to talk about a *zero solution*, or $x = 0$. To do so, we translate every nonzero equilibrium point to zero by the change of variables

$$\tilde{x}(t) = x(t) - x^*,$$

and from (4.1.3) we have that

$$\tilde{x}'(t) = (x(t) - x^*)' = x'(t) = f(t, \tilde{x}(t) + x^*).$$

In this case, $\tilde{x}(t) \equiv 0$ is an equilibrium point of (4.1.3), which we may call the *zero solution* of (4.1.3). Thus, in most cases, we may require $f(t, 0) = 0$, $0 \le t < w$, when we talk about the zero solution.

Example 4.1. (*Lotka–Volterra predator–prey model*) Let $x(t)$ and $y(t)$ be the number of preys and predators at time t, respectively. To keep the model simple, we make the following assumptions:

- The predator species are dependent on a single prey species as its only food supply;
- The prey species have an unlimited food supply; and
- There is no threat to the prey other than the specific predator.

In the absence of predation, the prey population would grow at a natural rate:

$$\frac{dx}{dt} = ax(t), \;\; a > 0.$$

On the other hand, in the absence of prey, the predator population would decline at a natural rate:

$$\frac{dy}{dt} = -cy(t), \;\; c > 0.$$

The effects of predator eating prey is an interaction rate of decline $(-bxy, \; b > 0)$ in the prey population x and an interaction rate of growth $(dxy, \; d > 0)$ of predator population y. Hence we obtain the predator–prey model

$$\begin{aligned} \frac{dx}{dt} &= ax - bxy, \\ \frac{dy}{dt} &= -cy + dxy. \end{aligned} \tag{4.1.4}$$

Copyright Elsevier 2022

The Lotka–Volterra model consists of a system of linked differential equations that cannot be separated from each other and that cannot be solved in closed form. Thus we search for the equilibrium solutions by setting the right sides of both equations in (4.1.4) equal to zero, that is,

$$x(a - by) = 0, \quad y(-c + dx) = 0.$$

From the first equation we obtain $x = 0$, $y = a/b$, and from the second equation we get $y = 0$, $x = c/d$. Now if $x = 0$, then from the second equation we have that $-cy = 0$, and hence $y = 0$. Thus we have the equilibrium solution

$$x_1^* = (0, 0).$$

Similarly, if $y = a/b$, then $-\frac{c}{b}a + dx\frac{a}{b} = 0$, which implies that $x = c/d$. So the other equilibrium solution is

$$x_2^* = (c/d, a/b).$$

The first equilibrium solution x_1^* effectively represents the extinction of both species. If both populations are at 0, then they will continue to be so indefinitely. The second equilibrium solution x_2^* represents a fixed point at which both populations sustain their current non-zero numbers and do so indefinitely. The levels of the population at which this equilibrium is achieved depend on the chosen values of the parameters a, b, c, and d. We further discuss the stability of both equilibrium solutions x_1^* and x_2^* and hope that x_1^* is unstable. (See Fig. 4.2.)

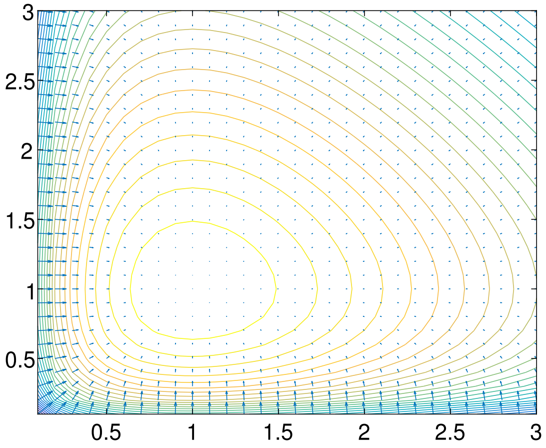


FIGURE 4.2

Prey–predator dynamics as described by the level curves. The arrows describe the velocity and direction of solutions for $a = b = c = d = 1$.

Copyright Elsevier 2022

Next, we consider (4.1.3) with the initial condition $x(t_0) = x_0$ with the assumption that $f(t, 0) = 0$.

Definition 4.1.1. The zero solution $x = 0$ of (4.1.3):

- **(a)** is *stable* (S) if for all $\varepsilon > 0$ and $t_0 \geq 0$, there is $\delta = \delta(t_0, \varepsilon) > 0$ such that $|x(t_0)| < \delta$ implies $|x(t, t_0, x_0)| < \varepsilon$,
- **(b)** is *uniformly stable* (US) if δ is independent of t_0,
- **(c)** is *unstable* if it is not stable,
- **(d)** is *asymptotically stable* (AS) if it is stable and $\lim\limits_{t\to\infty} |x(t, t_0, x_0)| = 0$.
- **(e)** is *uniformly asymptotically stable* (UAS) if it is US and there exists $\gamma > 0$ such that for each $\mu > 0$, there exists $T = T(\mu) > 0$ such that $|x(t_0)| < \gamma$, $t \geq t_0 + T$, implies $|x(t, t_0, x_0)| < \mu$.

There are various stability definitions that can be found in many of the referenced books. As far as this book is concerned, the stability definitions that we have listed are the most important ones. We furnish a few examples to illustrate the above definitions.

Example 4.2. In this example we show that the zero solution is US and AS but not UAS. We consider the initial value problem given by

$$(1-t)x'(t) = x, \; x(t_0) = x_0, \;\; t \geq t_0 > 1.$$

For simplicity, we may write our problem as

$$x' = -\frac{x}{t-1}, \; x(t_0) = x_0, \;\; t \geq t_0 > 1. \tag{4.1.5}$$

Separating the variables in (4.1.5), integrating from t_0 to t, and then applying the initial data, we arrive at the solution

$$x(t, t_0, x_0) = \frac{t_0 - 1}{t-1} x_0, \;\; t \geq t_0 > 1. \tag{4.1.6}$$

For any $\varepsilon > 0$, set $\delta = \varepsilon$, so that for $|x_0| < \delta$ and $t \geq t_0 > 1$, we have that

$$|x(t, t_0, x_0)| = \left|\frac{t_0-1}{t-1} x_0\right| = \left|\frac{t_0-1}{t-1}\right| |x_0| \leq |x_0| < \delta = \varepsilon.$$

This shows that the zero solution is US. Moreover,

$$\lim_{t\to\infty} |x(t, t_0, x_0)| = \lim_{t\to\infty} \left|\frac{t_0-1}{t-1}\right| |x_0| = 0,$$

and so the zero solution is AS. On the other hand, if we evaluate the solution $x(t, t_0, x_0) = \dfrac{t_0-1}{t-1} x_0$ at $t = t_0 + T$, then we get $x(t_0 + T, t_0, x_0) = \dfrac{t_0 - 1}{t_0 + T - 1} x_0 \to 1$ as $t_0 \to \infty$. This implies that $|x(t, t_0, x_0)| \not< \varepsilon$ for $t \geq t_0 + T$. We conclude that the zero solution is not UAS. So this example shows that US plus AS do not necessarily imply UAS.

Copyright Elsevier 2022

Example 4.3. The differential equation

$$x' = 0, \;\; x(t_0) = x_0,$$

has the constant solution $x(t) = x_0$, and hence the zero solution is US but not AS. On the other hand,

$$x' = x, \;\; x(t_0) = x_0,$$

has the solution $x(t) = x_0 e^{t-t_0}$, and its zero solution is unstable.

Example 4.4. The differential equation

$$x' = -x, \;\; x(t_0) = x_0,$$

has the solution

$$x(t, t_0, x_0) = x_0 e^{-(t-t_0)}, \;\; t \geq t_0 \geq 0,$$

and its zero solution is US and UAS. To see this, for any $\varepsilon > 0$, we set $\delta = \varepsilon$. Then for $|x_0| < \delta$, we have that $x(t, t_0, x_0)| = |x_0 e^{-(t-t_0)}| \leq |x_0| \leq \varepsilon = \delta$. So the zero solution is US. For the UAS, we must show that there exists $\gamma > 0$ such that for each $\mu > 0$, there exists $T = T(\mu) > 0$ such that $|x(t_0)| < \gamma$, $t \geq t_0 + T$, implies $|x(t, t_0, x_0)| < \mu$. Set $\delta = 1$ of the uniform stability, and let $\gamma = \delta = 1$. Let $T(\mu) > -\text{Ln}(\mu)$. Then for $t \geq t_0 + T(\mu)$,

$$|x(t, t_0, x_0)| = |x_0 e^{-(t-t_0)}| \leq e^{-T(\mu)} \leq e^{\text{Ln}(\mu)} = \mu.$$

This shows the zero solution is UAS.

4.2 $x' = A(t)x$

Now we are in a position to discuss the stability of the zero solution of the linear system

$$x'(t) = A(t)x(t), \;\; x(t_0) = x_0, \tag{4.2.1}$$

where $A(t)$ is an $n \times n$ matrix with continuous entries on $[0, \infty)$, using the concept of the fundamental matrix. We begin with the following theorem.

Theorem 4.2.1. *Let* $\Phi(t)$ *be the fundamental matrix of* (4.2.1)*. Then the zero solution of* (4.2.1) *is*

(a) *stable if and only if there exists a positive constant* M *such that*

$$|\Phi(t)| \leq M, \;\; t \geq 0;$$

(b) *asymptotically stable if and only if*

$$|\Phi(t)| \to 0 \;\; \textit{as} \;\; t \to \infty.$$

Copyright Elsevier 2022

Proof. (a) ($\Leftarrow$) Let $\Phi(t)$ be the fundamental matrix of (4.2.1). Then the unique solution x of (4.2.1) is given by $x(t) = \Phi(t)\Phi^{-1}(t_0)x_0$. Thus for any $\varepsilon > 0$ such that $|x_0| < \varepsilon$, set $\delta = \frac{\varepsilon}{|\Phi^{-1}(t_0)|M}$. Then

$$|x(t)| = |\Phi(t)\Phi^{-1}(t_0)x_0| \le |\Phi(t)||\Phi^{-1}(t_0)||x_0| \le M|\Phi^{-1}(t_0)|\delta = \varepsilon.$$

($\Rightarrow$) Set $\varepsilon = 1$ from the stability proof. Then

$$|x(t)| = |\Phi(t)\Phi^{-1}(t_0)x_0| < 1 \text{ for } t \ge t_0 \text{ if } |x_0| < \delta(1, t_0),$$

which implies that

$$|\Phi(t)\Phi^{-1}(t_0)| < \frac{1}{\delta(1, t_0)}.$$

Therefore

$$|\Phi(t)| = |\Phi(t)\Phi^{-1}(t_0)\Phi(t_0)| \le |\Phi(t)\Phi^{-1}(t_0)||\Phi(t_0)| \le \frac{1}{\delta(1, t_0)}|\Phi(t_0)| := M.$$

This completes the proof of (a).

Next, we prove (b). We already know that the zero solution is stable. Now

$$|x(t)| = |\Phi(t)\Phi^{-1}(t_0)x_0| \to 0 \text{ as } t \to \infty$$

if and only if

$$|\Phi(t)| \to 0 \text{ as } t \to \infty.$$

This completes the proof. ☐

Theorem 4.2.2. *Let $\Phi(t)$ be the fundamental matrix of* (4.2.1). *Then the zero solution of* (4.2.1) *is uniformly stable if and only if there exists a positive constant M such that*

$$|\Phi(t)\Phi^{-1}(s)| \le M, \quad t \ge s \ge 0. \tag{4.2.2}$$

Proof. ($\Leftarrow$) Let $\Phi(t)$ be the fundamental matrix of (4.2.1). Then the unique solution x of (4.2.1) is given by $x(t) = \Phi(t)\Phi^{-1}(s)x(s)$. Since (4.2.2) holds for all s such that $t \ge s \ge 0$, we have that

$$|x(t)| = |\Phi(t)\Phi^{-1}(s)x(s)| \le |\Phi(t)\Phi^{-1}(s)||x_0|.$$

Thus for any $\varepsilon > 0$, choosing $\delta = \frac{\varepsilon}{M}$, we get that for $|x_0| < \delta$,

$$|x(t)| \le |\Phi(t)||\Phi^{-1}(s)||x_0| \le M\delta = \varepsilon.$$

($\Rightarrow$) Set $\varepsilon = 1$ from the proof of uniform stability. Then

$$|x(t)| = |\Phi(t)\Phi^{-1}(s)x(s)| < 1 \text{ for } t \ge s \text{ if } |x(s)| < \delta(1),$$

Copyright Elsevier 2022

which implies that

$$|\Phi(t)\Phi^{-1}(s)| \le \frac{1}{\delta(1)} := M.$$

This completes the proof. □

The next theorem provides necessary and sufficient conditions for the UAS.

Theorem 4.2.3. *Let $\Phi(t)$ be the fundamental matrix of* (4.2.1). *Then the zero solution of* (4.2.1) *is uniformly asymptotically stable if and only if there exist positive constants M and β such that*

$$|\Phi(t)\Phi^{-1}(s)| \le Me^{-\beta(t-s)}, \ \ t \ge s \ge 0. \tag{4.2.3}$$

Proof. ($\Leftarrow$) Let $\Phi(t)$ be the fundamental matrix of (4.2.1). Then the unique solution x of (4.2.1) is given by $x(t) = \Phi(t)\Phi^{-1}(s)x(s)$. It is evident from (4.2.3) that the zero solution is US. For all $t \ge s \ge 0$, we have that

$$|x(t)| = |\Phi(t)\Phi^{-1}(s)x(s)| \le |\Phi(t)\Phi^{-1}(s)||x_0|.$$

As a result, we may set $\gamma = 1$. Now we take $T(\mu) = -\frac{1}{\beta}\mathrm{Ln}(\frac{\mu}{M})$, $0 < \mu < M$. Then for $|x_0| < 1$ and $t \ge t_0 + T(\mu)$, we have

$$|x(t)| = |\Phi(t)\Phi^{-1}(t_0)x_0| < Me^{-\beta(t-t_0)} = Me^{-\beta T(\mu)} \le \mu.$$

($\Rightarrow$) Since the zero solution is UAS, there exists $\gamma > 0$ such that for each $\mu > 0$, there exists $T = T(\mu) > 0$ such that $|x(t_0)| < \gamma$, $t \ge t_0 + T$, implies $|x(t, t_0, x_0)| < \mu$. Let $\alpha = \frac{\mu}{\gamma}$. Then $|x_0| \le \gamma$ and $t \ge t_0 + T(\mu)$ yield

$$|\Phi(t)\Phi^{-1}(t_0)x_0| \le \mu. \tag{4.2.4}$$

Therefore

$$|\Phi(t + T(\mu))\Phi^{-1}(t)| \le \alpha < 1, \ t \ge 0. \tag{4.2.5}$$

For $t \ge t_0$, there is a positive integer k such that

$$t_0 + kT(\mu) \le t \le t_0 + (k+1)T(\mu).$$

Now let $t_n = t_0 + (k+1)T(\mu)$, $n = 1, 2, \dots, k$. Then by (4.2.4) and (4.2.5)

$$\begin{aligned} |\Phi(t)\Phi^{-1}(t_0)| &= \left|\Phi(t)\Phi^{-1}(t_k)\Phi(t_k)\Phi^{-1}(t_{k-1})\dots\Phi(t_1)\Phi^{-1}(t_0)\right| \\ &\le \left|\Phi(t)\Phi^{-1}(t_k)\right| \prod_{n=0}^{k-1} \Phi(t_{n+1})\Phi^{-1}(t_n) \\ &\le M\alpha^k. \end{aligned}$$

Copyright Elsevier 2022

Let $\beta = -\frac{1}{T(\gamma)}\text{Ln}(\alpha) > 0$. Then

$$\begin{aligned}
|\Phi(t)\Phi^{-1}(t_0)| &\le \frac{M}{\alpha}\alpha^{k+1} \\
&= \frac{M}{\alpha}e^{-(k+1)\beta T(\gamma)} \\
&\le \frac{M}{\alpha}e^{-\beta(t-t_0)} \\
&:= Le^{-\beta(t-t_0)}.
\end{aligned}$$

This completes the proof. □

Example 4.5. The nonlinear system

$$x'(t) = \begin{pmatrix} 0 & 1 \\ \frac{-2}{t^2} & \frac{2}{t} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \; t > 0,$$

has the fundamental matrix

$$\Phi(t) = \begin{pmatrix} t^2 & t \\ 2t & 1 \end{pmatrix}, \; t > 0.$$

By Theorem 4.2.1, its zero solution is unstable, since there is no constant M such that $||\Phi(t)|| \le M$.

Example 4.6. The system

$$x'(t) = \begin{pmatrix} 0 & 1 \\ -2 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \; t > 0,$$

has the fundamental matrix

$$\Phi(t) = e^{-t}\begin{pmatrix} \cos(t) & \sin(t) \\ -(\cos(t) + \sin(t)) & \cos(t) - \sin(t) \end{pmatrix}, \; t > 0.$$

After some calculations, we can verify that for some positive constant K,

$$||\Phi(t)\Phi^{-1}(s)|| \le Ke^{-(t-s)}, \; t \ge s \ge 0,$$

and by Theorem 4.2.3 the zero solution is UAS.

4.3 Floquet theory

This section is concerned with homogeneous systems of the form

$$x'(t) = A(t)x(t), \tag{4.3.1}$$

Copyright Elsevier 2022

where the matrix $A(t)$ is periodic with period T (constant), that is,

$$A(t+T) = A(t) \text{ for all } t \in \mathbb{R}.$$

For the rest of this section, we assume that $A(t)$ is a periodic matrix with period T. Our aim is to find an expression for the fundamental matrix in this case. Unlike the case of a constant coefficient matrix, the fundamental matrix cannot be expressed as $\Phi(t) = (\phi_1, \dots, \phi_n)$ where the ϕ_i, $i = 1, 2, \dots, n$, satisfy

$$\phi_i(t) = k_i e^{\lambda_i t},$$

and (λ_i, K_i), $i = 1, 2, \dots n$, is eigenpairs for the constant matrix A. To reinforce this notion, we offer the following example.

Example 4.7. Consider the nonlinear system

$$x'(t) = \begin{pmatrix} \sin(t) & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ t \in \mathbb{R}.$$

Then the corresponding $A(t)$ matrix is periodic with period $T = 2\pi$. Moreover, $A(t)$ has the eigenpairs

$$\big(\sin(t), \begin{pmatrix} 1 \\ 0 \end{pmatrix}\big), \big(2, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\big).$$

Without thinking and not paying attention to the fact that A is not a constant matrix, using the obtained eigenpairs, we form the fundamental matrix

$$\Phi(t) = \begin{pmatrix} e^{\sin(t)} & 0 \\ 0 & e^{2t} \end{pmatrix}.$$

It is clear that the first column $\phi_1(t) = \begin{pmatrix} e^{\sin(t)} \\ 0 \end{pmatrix}$ is not a vector solution of the system since

$$\begin{pmatrix} \sin(t) & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} e^{\sin(t)} \\ 0 \end{pmatrix} = \begin{pmatrix} \sin(t) e^{\sin(t)} \\ 0 \end{pmatrix} \neq \phi_1'(t).$$

In other words,

$$A(t)\phi_1(t) \neq \phi_1'(t).$$

Hence we cannot use the eigenpairs to obtain the fundamental matrix for the periodic differential system.

Example 4.7 is a living proof of the need to develop a new approach to find the fundamental matrix of systems in the form (4.3.1). To motivate the subject, we consider the scalar differential equation

$$x'(t) = a(t)x(t), \tag{4.3.2}$$

Copyright Elsevier 2022

where $a : \mathbb{R} \to \mathbb{R}$ is continuous and periodic with minimal period T. Then every solution $\phi(t)$ of (4.3.2) has the form

$$\phi(t) = ce^{\int_0^t a(s)ds}. \tag{4.3.3}$$

Set the constant $c = 1$, and for a constant λ, we observe that

$$\int_t^{t+T} a(s)ds = \lambda T, \quad \text{a constant},$$

for all $t \in \mathbb{R}$, since $a(t+T) = a(t)$. Using (4.3.3) with $c = 1$, we have

$$\phi(t+T) = e^{\int_0^{t+T} a(s)ds} = e^{\int_0^t a(s)ds} e^{\int_t^{t+T} a(s)ds} = \phi(t)e^{\lambda t}. \tag{4.3.4}$$

The number λ in (4.3.4) has a profound effect on the behavior of the solution ϕ. At a glance, if ϕ is periodic with period T, that is, $\phi(t+T) = \phi(t)$, then (4.3.4) implies that

$$e^{\lambda t} = 1.$$

This holds if and only if $\lambda = \frac{2\pi i}{T}$. The number λ is called a *characteristic exponent* for (4.3.2), and $e^{\lambda t}$ is called the corresponding *characteristic multiplier*. More information about the form of the solutions can be revealed by letting the new function

$$p(t) = \phi(t)e^{-\lambda t}.$$

Then the function p is periodic with period T since

$$p(t+T) = \phi(t+T)e^{-\lambda(t+T)} = \phi(t)e^{\lambda T}e^{-\lambda(t+T)} = \phi(t)e^{-\lambda t} = p(t).$$

Thus any solution of (4.3.2) can be written in the form

$$\phi(t) = p(t)e^{\lambda t}. \tag{4.3.5}$$

Eq. (4.3.5) is very significant in the sense that it allows us to study the behavior of all solutions on $-\infty < t < +\infty$. As noted above, ϕ is periodic with period T if and only if $\lambda = \frac{2n\pi i}{T}$. Similarly, $\phi(t) \to 0$ as $t \to \infty$ and becomes unbounded as $t \to -\infty$ if and only if $\operatorname{Re}\lambda < 0$.

Now we go back to the periodic system (4.3.1).

Theorem 4.3.1. *Let $A(t+T) = A(t)$ for all $t \in \mathbb{R}$. If $\Phi(t)$ is a fundamental matrix of (4.3.1), then so is $\Phi(t+T)$, and there exists a nonsingular constant matrix B such that*

(a) $\Phi(t+T) = \Phi(t)B$ *for all* $t \in \mathbb{R}$*;*

(b) $\det\big(B\big) = e^{\int_0^t \operatorname{tr}\big(A(s)\big)ds}$.

Copyright Elsevier 2022

Proof. (a) Set $Z(t) = \Phi(t+T)$. Then

$$Z'(t) = \Phi'(t+T) = A(t+T)\Phi(t+T) = A(t)\Phi(t+T) = Z(t)\Phi(t+T).$$

Moreover, $\det\big(Z(t)\big) = \det\big(\Phi(t+T)\big) \neq 0$, since $\Phi(t)$ is a fundamental matrix. This proves that $\Phi(t+T)$ is also a fundamental matrix. The second part of (a) follows from Theorem 3.2.5. This completes the proof of (a). For the proof of (b), we let $W(t) = \det\big(\Phi(t)\big)$. From Theorem 3.2.3 we have that

$$W(t) = W(t_0)e^{\int_{t_0}^{t} \mathrm{tr}\big(A(s)\big)ds},$$

and hence

$$\begin{aligned} W(t+T) &= W(t_0)e^{\int_{t_0}^{t+T} \mathrm{tr}(A(s))ds} \\ &= W(t_0)e^{\left(\int_{t_0}^{t} \mathrm{tr}\big(A(s)\big)ds + \int_{t}^{t+T} \mathrm{tr}\big(A(s)\big)ds\right)} \\ &= W(t_0)e^{\int_{t_0}^{t} \mathrm{tr}\big(A(s)\big)ds} \cdot e^{\int_{t}^{t+T} \mathrm{tr}\big(A(s)\big)ds} \\ &= W(t)e^{\int_{0}^{T} \mathrm{tr}\big(A(s)\big)ds}. \end{aligned} \tag{4.3.6}$$

In addition,

$$\Phi(t+T) = \Phi(t)B$$

or

$$\det\big(\Phi(t+T)\big) = \det\big(\Phi(t)\big)\det\big(B\big).$$

Thus

$$W(t+T) = W(t)\det(B\big). \tag{4.3.7}$$

A comparison of (4.3.6) and (4.3.7) gives

$$\det(B\big) = e^{\int_{0}^{T} \mathrm{tr}\big(A(s)\big)ds}.$$

□

Remark 4.1. Note that since the matrix B is constant, we can uniquely determine it by evaluating both sides of (a) at $t = 0$:

$$B = \Phi^{-1}(0)\Phi(T).$$

We have the following definition of Floquet multipliers.

Definition 4.3.1. Let $\Phi(t)$ be the fundamental matrix for the Floquet system (4.3.1). Then the eigenvalues $\rho_1, \rho_2, \ldots, \rho_n$ of

Copyright Elsevier 2022

$$B := \Phi^{-1}(0)\Phi(T)$$

are called the *Floquet multipliers* of the Floquet system (4.3.1).

The *Floquet exponents* are $\mu_1, \mu_2, \dots, \mu_n$ satisfying

$$\rho_1 = e^{\mu_1 T}, \rho_2 = e^{\mu_2 T}, \dots, \rho_n = e^{\mu_n T}. \tag{4.3.8}$$

Note that μ_j, $j = 1, 2, \dots, n$, may be complex.

Remark 4.2. **(i)** If $\Phi(0) = I$, then from item (b) of Theorem 4.3.1 it follows that

$$\det\big(B\big) = \prod_{j=1}^{n} \rho_j = e^{\int_0^T \operatorname{tr}\left(A(s)\right)ds}.$$

(ii) Since the trace is the sum of the eigenvalues, we also have

$$\operatorname{tr}(B) = \sum_{j=1}^{n} \rho_j.$$

(iii) The exponential exponents are not unique since if $\rho_j = e^{\mu_j T}$, then $\rho_j = e^{(\mu_j + \frac{2\pi i}{T})T}$.

(iv) The Floquet multipliers ρ_j are intrinsic parameters of system (4.3.1) and do not depend on the choice of the fundamental matrix $\Phi(t)$.

Proof (of (iv)*).* Suppose $\Psi(t)$ is another fundamental matrix. Then by Theorem 4.3.1, there is a nonsingular matrix B such that

$$\Psi(t+T) = \Psi(t)B.$$

Now by Theorem 3.2.5

$$\Psi(t) = \Phi(t)C,$$

where C is a nonsingular matrix. Thus

$$\Psi(t+T) = \Phi(t+T)C$$

or

$$\Psi(t)B = (\Phi(t)B)C.$$

This gives

$$\Phi(t)CB = \Phi(t)BC, \quad \text{or} \quad CB = BC.$$

As a consequence,

$$CBC^{-1} = B,$$

so the eigenvalues of B and C are the same. ☐

Copyright Elsevier 2022

Theorem 4.3.2. *Let ρ be the Floquet multiplier of* (4.3.1) *with the corresponding Floquet exponent μ. Then there exists a solution $x(t)$ of* (4.3.1) *such that*

- **(a)** $x(t+T)=\rho x(t)$;
- **(b)** *There exists a continuously differentiable periodic $n \times 1$ vector function p with period T such that $x(t)=e^{\mu t}p(t)$.*

Proof. Let ρ be an eigenvalue of the matrix B with corresponding eigenvector b. Let $\Phi(t)$ be a fundamental matrix of the Floquet system (4.3.1) such that $x(t)=\Phi(t)b$. Then

$$\begin{aligned} x(t+T) &= \Phi(t+T)b \\ &= \Phi(t)Bb \\ &= \rho\Phi(t)b \\ &= \rho x(t). \end{aligned}$$

This proves (a). As for (b), we take $p(t)=x(t)e^{-\mu t}$. Then

$$\begin{aligned} p(t+T) &= x(t+T)e^{-\mu(t+T)} \\ &= \rho x(t)e^{-\mu(t+T)} \\ &= \rho e^{-\mu T}x(t)e^{-\mu t} \\ &= x(t)e^{-\mu t} \\ &= p(t), \end{aligned}$$

where we have used $\rho=e^{\mu T}$. This shows that $p(t)$ is periodic with period T. As a result, we have that $x(t)=e^{\mu t}p(t)$. This completes the proof of (b) and the theorem. ☐

Remark 4.3. Suppose a Floquet multiplier ρ is complex. Since ρ is an eigenvalue of the matrix B, then so is its conjugate $\bar{\rho}$. Then we have two solutions, x and $\bar{x}$. Let the Floquet exponent $\mu=\zeta+i\eta$, and consider the periodic complex vector $p(t)=q(t)+ir(t)$. We may write the corresponding two solutions as

$$x(t)=e^{\mu t}p(t) \ \text{ and } \ x(t)=e^{\bar{\mu}t}\bar{p}(t).$$

Explicitly, we have

$$\begin{aligned} x(t) &= e^{(\zeta+i\eta)t}(q(t)+ir(t)) \\ &= e^{\zeta t}\Big[\big(q(t)\cos(\eta t)-r(t)\sin(\eta t)\big)+i\big(r(t)\cos(\eta t)+q(t)\sin(\eta t)\big)\Big] \end{aligned}$$

and

$$\begin{aligned} \bar{x}(t) &= e^{(\zeta-i\eta)t}(q(t)-ir(t)) \\ &= e^{\zeta t}\Big[\big(q(t)\cos(\eta t)-r(t)\sin(\eta t)\big)-i\big(r(t)\cos(\eta t)+q(t)\sin(\eta t)\big)\Big]. \end{aligned}$$

Copyright Elsevier 2022

As a consequence, we may write the linearly independent real solutions

$$x_R = \mathrm{Re}\Big[e^{\mu t} p(t)\Big] = e^{\zeta t}\Big[q(t)\cos(\eta t) - r(t)\sin(\eta t)\Big], \tag{4.3.9}$$

$$x_I = \mathrm{Im}\Big[e^{\mu t} p(t)\Big] = e^{\zeta t}\Big[q(t)\sin(\eta t) + r(t)\cos(\eta t)\Big]. \tag{4.3.10}$$

Lemma 4.1. *(Logarithmic matrix) If B is an $n \times n$ matrix, then there exists an $n \times n$ (complex) matrix C such that $B = e^C$.*

Proof. Let λ_j, $j = 1, 2, \ldots, m$, with $m \le n$ denote the eigenvalues of B. We already know that

$$e^{P^{-1}CP} = P^{-1}e^C P.$$

Therefore we may assume that P is in Jordan canonical form, that is,

$$P = diag\big(J_1, J_2, \ldots, J_n\big)$$

with

$$J_j = \lambda_j + N_j,\ \lambda_j \neq 0,$$

where J_j and N_j are defined in Theorem 3.2.12. Suppose that

$$B = diag\big(B_1, B_2, \ldots, B_n\big),$$

where

$$B_j = \begin{pmatrix} \lambda_j & 1 & & \\ & \lambda_j & \ddots & \\ & & & 1 \\ & & & \lambda_j \end{pmatrix},\ j = 1, 2, \ldots m.$$

Let $l_j \times l_j$ be the size of $B_j = \lambda_j I_j + N_j$, where $N^{l_j} = 0$ (zero matrix) for all $k \ge m$. Thus $\sum_{k=1}^{m} l_k = n$, and so we have constructed a matrix B of size $n \times n$. Thus if we can prove that for every B_j, there exists a matrix C_j such that $B_j = e^{C_j}$, then $B = e^C$. Define

$$C_j := (\mathrm{Ln}(\Lambda_j) I_j + L_j,$$

where

$$L_j := -\sum_{k=1}^{m} \frac{(-N_j)^k}{k\lambda_j^k} \equiv \mathrm{Ln}(I_j + \frac{N_j}{\lambda_j}).$$

Recall that at $x = 0$,

$$\mathrm{Ln}(1 + x) = \sum_{k=1}^{\infty} \frac{(-x)^k}{k},\ \ |x| < 1,$$

Copyright Elsevier 2022

which implies that

$$1 + x = e^{\mathrm{Ln}(1+x)} = 1 + (x - x^2/2 + x^3/3 + \dots) + \frac{1}{2!}(x - x^2/2 \\ + x^3/3 + \dots)^2 + \frac{1}{3!}(x - x^2/2 + x^3/3 + \dots)^3 + \dots.$$

Since $N^{l_j} = 0$ (zero matrix) for all $k \geq m$, we must have

$$e^{L_j} = I_j + \frac{N_j}{\lambda_j}$$

or

$$B_j = e^{\mathrm{Ln}(\lambda_j) I_j + L_j},$$

where we have loosely used the fact that $B_j = e^{\mathrm{Ln}(B_j)}$. This completes the proof. □

Now we are ready to state and prove the Floquet theorem.

Theorem 4.3.3. *(Floquet theorem) Every fundamental matrix of* (4.3.1) *has the form*

$$\Phi(t) = P(t)e^{Ct},$$

where $P(t)$ is a T-periodic matrix, and C is a constant $n \times n$ matrix.

Proof. Let $\Phi(t)$ be a fundamental matrix of (4.3.1). Then

$$\begin{aligned}(\Phi^{-1}(t)\Phi(t+T))' &= (\Phi^{-1}(t))'\Phi(t+T) + \Phi^{-1}(t)\Phi'(t+T) \\ &= -\Phi^{-1}(t)A(t)\Phi(t+T) + \Phi^{-1}(t)A(t)\Phi(t+T) \\ &= 0.\end{aligned}$$

Thus

$$\Phi^{-1}(t)\Phi(t+T) = B \;\; \text{(constant matrix)},$$

or

$$\Phi(t+T) = \Phi(t)B.$$

Since $\Phi^{-1}(t)$ and $\Phi(t+T)$ are nonsingular, B is also nonsingular. Thus by Lemma 4.1 $B = e^{CT}$ for some constant matrix C. Left to show that $P(t)$ is periodic with period T. We define $P(t) = \Phi(t)e^{-Ct}$. Then

$$\begin{aligned}P(t+T) &= \Phi(t+T)e^{-C(t+T)} \\ &= \Phi(t)e^{CT}e^{-C(t+T)} \\ &= \Phi(t)e^{-Ct} \\ &= P(t).\end{aligned}$$

This completes the proof. □

Copyright Elsevier 2022

Example 4.8. Let us find the Floquet multipliers and exponents of

$$x'(t) = \begin{pmatrix} -1 & 1 \\ 0 & 1+\cos(t) - \frac{\sin(t)}{2+\cos(t)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

We notice that the corresponding coefficient matrix is periodic with period 2π. We easily solve

$$x_2' = \left(1 + \cos(t) - \frac{\sin(t)}{2+\cos(t)}\right) x_2$$

using separation of variables and obtain

$$x_2 = c_2 e^{t+\sin(t)}(2+\cos(t)).$$

Substituting x_2 into $x_1' = -x_1 + x_2$ and then using the variation of parameters formula, we obtain

$$x_1 = c_2 e^{t+\sin(t)} + c_1 e^{-t}.$$

Setting $c_1 = c_2 = 1$, we obtain the fundamental matrix

$$\Phi(t) = \begin{pmatrix} e^{t+\sin(t)} & e^{-t} \\ e^{t+\sin(t)}(2+\cos(t)) & 0 \end{pmatrix}.$$

From Definition 4.3.1, we have

$$B := \Phi^{-1}(0)\Phi(T) = \Phi^{-1}(0)\Phi(2\pi) = \begin{pmatrix} 0 & \frac{1}{3} \\ 1 & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} e^{2\pi} & e^{-2\pi} \\ 3e^{2\pi} & 0 \end{pmatrix} = \begin{pmatrix} e^{2\pi} & 0 \\ 0 & e^{-2\pi} \end{pmatrix}.$$

Then the eigenvalues of B, $\rho_1 = e^{2\pi}$ and $\rho_2 = e^{-2\pi}$ are the Floquet multipliers with Floquet exponents $\mu_1 = 2\pi$ and $\mu_2 = -2\pi$.

Example 4.9. Find the Floquet multiplier and exponent of

$$x'(t) = (2 - \cos(t))x(t).$$

Using separation of variables, we find the solution

$$x(t) = Ke^{2t - \sin(t)}$$

for some constant K. Set $K = 1$. Then the Floquet multiplier is $\rho = x^{-1}(0)x(2\pi) = 1e^{4\pi}$, and the Floquet exponent is $\mu = 4\pi$.

One of the main purposes of Floquet theory is to study the stability of the zero solution of (4.3.1). It is evident from Theorem 4.3.2 that the behavior of a solution depends on the Floquet exponent μ. In fact, from Theorem 4.3.2 we have that

$$x(t) = e^{\mu t} p(t),$$

Copyright Elsevier 2022

where $p(t)$ is a periodic $n \times 1$ vector. Let x_j be the jth component of the vector solution. Then from Theorem 4.3.2 we have that

$$x_j(t) = \rho_j x_j(t)$$

and

$$x_j(t + NT) = \rho_j^N x_j(t) \to 0 \text{ as } N \to \infty.$$

Now we can easily see that each Floquet multiplier falls into one of the following categories:

(a) If $|\rho| < 1$, then $\mathrm{Re}(\mu) < 0$ and so $x(t) \to 0$ as $t \to \infty$;
(b) If $\rho = \pm 1$, then $\mathrm{Re}(\mu) = 0$, and so the solution is periodic with period T;
(c) If $|\rho| > 1$, then $\mathrm{Re}(\mu) > 0$, and so $x(t) \to \infty$ as $t \to \infty$.

We have the following general theorem concerning the stability of the zero solution of (4.3.1).

Theorem 4.3.4. *The zero solution of* (4.3.1) *is*

(i) *stable if and only if the Floquet multipliers ρ satisfy $|\rho| \leq 1$ and there is a complete set of eigenvectors for any multiplier of modulus 1;*
(ii) *asymptotically stable if and only if $|\rho| < 1$ for every ρ.*

Proof. Let $\Phi(t)$ be the fundamental matrix of (4.3.1). Then by the Floquet theorem, we have that $\Phi(t) = P(t)e^{Ct}$. Then the solutions of (4.3.1) are given by $x(t) = P(t)e^{Ct}x_0$, where $x_0 \in \mathbb{R}^n$. We make two claims:
(1) the stability of the zero solution of (4.3.1) is the same as for

$$y'(t) = Cy(t), \tag{4.3.11}$$

which has solutions of the form $y(t) = e^{Ct}x_0$. To see this, let M be a positive constant. Since $P(t)$ is continuous and periodic, we have that $||P(t)|| \leq M$ for all $t \in \mathbb{R}$. Thus $|x(t)| = |P(t)e^{Ct}x_0| \leq M||e^{Ct}x_0|| = M|y(t)|$, and this completes the proof of the first claim.
(2) $y(t) = P^{-1}(t)x(t)$ is a solution of $y'(t) = Cy(t)$. To see this, let $y(t) = P^{-1}(t)x(t)$. Then

$$\begin{aligned} y(t) &= P^{-1}(t)\Phi(t)x_0 \\ &= P^{-1}(t)P(t)e^{Ct}x_0 \\ &= e^{Ct}x_0, \end{aligned}$$

which is a solution of $y'(t) = Cy(t)$. This completes the proof of the second claim.

Next, we address the proof of stability.

If $|\rho| < 1$ for all Floquet multipliers, then $\mathrm{Re}(\mu) < 0$ for all Floquet exponents, and the zero solution is AS for (4.3.11) and hence for (4.3.1).

Copyright Elsevier 2022

Now we assume that $|\rho| \le 1$ for all ρ and there are no generalized eigenvectors (see (3.2.26)) of e^{CT} corresponding to Floquet multipliers with $|\rho| = 1$. Then $\mathrm{Re}(\mu) \le 0$ for all eigenvalues of C, and there are no generalized eigenvectors of C corresponding to eigenvalues with $\mathrm{Re}(\mu) = 0$. As a consequence, we have that $|e^{Ct}|$ is uniformly bounded. Thus the zero solution is stable for (4.3.11) and hence for (4.3.1).

If $\rho > 1$ for some ρ or if there is a generalized eigenvector of e^{CT} with $\rho = 1$, then either C has an eigenvalue with $\mathrm{Re}(\mu) > 1$ or C has a generalized eigenvector with $\mathrm{Re}(\mu) = 0$. In both cases, $|y(t)| \to \infty$ as $t \to \infty$. This implies that the zero solution of (4.3.11) is unstable. On the other hand, since $P(t)$ is continuous and periodic, we have that $P^{-1}(t)$ is also continuous and periodic. Thus

$$y(NT) = P^{-1}(NT)x(NT)x_0$$

for a positive integer N. Therefore

$$\infty = \lim_{N\to\infty} |y(NT)| \le ||P^{-1}(NT)|| \lim_{N\to\infty} |x(NT)|,$$

and the zero solution is also unstable for (4.3.1). This completes the proof. □

We provide the following example.

Example 4.10. Consider the periodic system with period π

$$x'(t) = \begin{pmatrix} -1 + \frac{1}{2}\sin^2(t) & -1 - \frac{1}{2}\cos(t)\sin(t) \\ 1 - \frac{1}{2}\cos(t)\sin(t) & -1 + \frac{1}{2}\cos^2(t) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

By inspection we found the fundamental matrix to be

$$\Phi(t) = \begin{pmatrix} -e^{-t/2}\sin(t) & e^{-t}\cos(t) \\ e^{-t/2}\cos(t) & e^{-t}\sin(t) \end{pmatrix}.$$

We encourage the reader to verify that

$$\Phi'(t) = A(t)\Phi(t).$$

Then

$$\begin{aligned} B &:= \Phi^{-1}(0)\Phi(T) = \Phi^{-1}(0)\Phi(\pi) \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -e^{-\pi} \\ -e^{-\pi/2} & 0 \end{pmatrix} = \begin{pmatrix} -e^{-\pi/2} & 0 \\ 0 & -e^{-\pi} \end{pmatrix}, \end{aligned}$$

and thus the eigenvalues of B are given by $\rho_1 = -e^{-\pi/2}$ and $\rho_2 = -e^{-\pi}$ with $|\rho_i| < 1, i = 1, 2$. By Theorem 4.3.4 the origin $(0, 0)$ is asymptotically stable.

Copyright Elsevier 2022

4.3.1 Mathieu's equation

We begin by considering the 2×2 real entries and constant matrix

$$D = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Let $L(\lambda) = \det(D - \lambda I)$. Then $L(\lambda)$ can be easily computed to be

$$L(\lambda) = \lambda^2 - (a+d)\lambda + ad - bc = \lambda^2 - \operatorname{tr}(D)\lambda + \det(D). \tag{4.3.12}$$

On the other hand, if λ_1 and λ_2 are the eigenvalues of D, not necessarily distinct, then

$$L(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2. \tag{4.3.13}$$

A quick comparison of (4.3.12) with (4.3.13) gives

$$\operatorname{tr}(D) = \lambda_1 + \lambda_2 \tag{4.3.14}$$

and

$$\det(D) = \lambda_1\lambda_2. \tag{4.3.15}$$

Next, we consider Mathieu's equation

$$x''(t) + \big(\alpha + \beta\cos(t)\big)x(t) = 0 \tag{4.3.16}$$

with parameters α and β. Eq. (4.3.16) was discovered by Émile Léonard Mathieu, who encountered it while studying vibrating elliptical drumheads. It has applications in many fields of the physical sciences, such as optics, quantum mechanics, and general relativity. The equation tends to occur in problems involving periodic motion or in the analysis of partial differential equations and boundary value problems possessing elliptic symmetry. Using the transformation $x_1 = x$, $x_2 = x'$, we write (4.3.16) in the system form

$$x'(t) = \begin{pmatrix} 0 & 1 \\ -\alpha - \beta\cos(t) & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

It is clear that the system is a Floquet system with a period of 2π. Let $\Phi(t)$ be a fundamental matrix. Then the Floquet multipliers ρ_1 and ρ_2 are the eigenvalues of the matrix

$$B = \Phi^{-1}(0)\Phi(T).$$

Since $\operatorname{tr}(A) = 0$, where

$$A(t) = \begin{pmatrix} 0 & 1 \\ -\alpha - \beta\cos(t) & 0 \end{pmatrix},$$

Copyright Elsevier 2022

by Theorem 4.3.1, we have that

$$\det\big(B\big) = \prod_{j=1}^{n} \rho_j = e^{\int_0^T \operatorname{tr}\big(A(s)\big)ds} = e^0 = 1.$$

Thus by (4.3.15),

$$\rho_1\rho_2 = 1.$$

As consequence of (4.3.12)–(4.3.15), the eigenvalues or Floquet multipliers must satisfy the quadratic equation $\rho^2 - (\rho_1 + \rho_2)\rho + 1 = 0$. Letting $\phi := \phi(\alpha, \beta) = \rho_1 + \rho_2$, we arrive at the characteristic equation

$$\rho^2 - \phi\rho + 1 = 0. \tag{4.3.17}$$

Eq. (4.3.17) has the roots

$$\rho_{1,2} = \frac{1}{2}\big(\phi \pm \sqrt{\phi^2 - 4}\big).$$

Despite the fact that we cannot solve for ϕ, we can still deduce some information about the solutions by considering particular values for ϕ. Below we consider several cases that depend on assigned values of ϕ.

1. Case 1: $\phi > 2$.
   Since $\rho_{1,2} = \frac{1}{2}\big(\phi \pm \sqrt{\phi^2 - 4}\big)$, we have $\rho_1 > 1$, and since $\rho_1\rho_2 = 1$, we have $\rho_1 > 1 > \rho_2 > 0$. As $\rho_2 = \frac{1}{\rho_1}$, we conclude that $\mu_2 = -\mu_1$. Thus according to Theorem 4.3.2, the solution must be of the form

   $$x(t) = c_1 e^{\mu_1 t} p_1(t) + c_2 e^{-\mu_1 t} p_2(t),$$

   where $p_1(t)$ and $p_2(t)$ are both periodic with period 2π. Moreover, the zero solution is unstable, and solutions are unbounded.
2. Case 2: $\phi = 2$.
   Then $\rho_1 = \rho_2 = 1$. In this case, Theorem 4.3.2 guarantees one solution of the form $e^{\mu t}p(t)$. On the other hand, if B has two corresponding linearly independent eigenvectors, then we can find two linearly independent 2π-periodic functions $p_1(t)$ and $p_2(t)$, so that the general solution has the form $x(t) = c_1 p_1(t) + c_2 p_2(t)$, since $\rho = 1$ implies $\mu = 0$. Now if B has only one eigenvector, then we end up with a solution of the form $p_1(t)$ and another of the form $tp_1(t) + p_2(t)$. This is due to the fact that the Jordan canonical form of the matrix B is $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, and the solution takes the form

   $$\Phi(t) = P(t)e^{\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} t} = P(t)\begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} = \begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix}\begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}.$$

Copyright Elsevier 2022

In conclusion, we have a solution of the form

$$x(t) = (c_1 + c_2 t) p_1(t) + c_2 p_2(t),$$

which is unbounded, and, moreover, the zero solution is unstable.

3. Case 3: $\phi = -2$.
Then $\rho_1 = \rho_2 = -1$. This is similar to Case 3. We know that $\rho = e^{\mu T} = -1$, which implies that $\mu = \frac{i\pi}{T}$, since

$$e^{\mu T} = e^{\frac{i\pi}{T} T} = e^{i\pi} = -1.$$

The solution is given by

$$x(t) = (c_1 + c_2 t) p_1(t) e^{\frac{i\pi}{T} t} + c_2 p_2(t) e^{\frac{i\pi}{T} t},$$

which is a combination of a periodic solution with period 4π and a nonperiodic solution. Again, solutions are unbounded, and the zero solution is unstable.

4. Case 4: $-2 < \phi < 2$.
Without loss of generality, we may define η by $\phi = 2\cos(\eta T)$, where $0 < \eta T < \pi$, so that

$$\begin{aligned}\rho &= \frac{1}{2}\left(\phi \pm \sqrt{\phi^2 - 4}\right)\\ &= \cos(\eta T) \pm i\ \sin(\eta T)\\ &= e^{\pm i\eta T}.\end{aligned}$$

Since $\zeta = 0$, Eqs. (4.3.9) and (4.3.10) reduce to

$$\begin{aligned}x_R &= \mathrm{Re}\left[e^{i\eta t} p(t)\right] = q(t)\cos(\eta t) - r(t)\sin(\eta t),\\ x_I &= \mathrm{Im}\left[e^{i\eta t} p(t)\right] = q(t)\sin(\eta t) + r(t)\cos(\eta t),\end{aligned}$$

where $p(t) = q(t) + i\ r(t)$ is periodic with period 2π. With this in mind, the solution takes the form

$$x(t) = c_1\Big(q(t)\cos(\eta t) - r(t)\sin(\eta t)\Big) + c_2\Big(q(t)\sin(\eta t) + r(t)\cos(\eta t)\Big).$$

We easily see that the zero solution is stable and solutions are bounded, but it is not periodic in general as there are two frequencies 2π and η. In such a case, the solution is said to be pseudoperiodic. Next, we examine the period of η. Now $e^{i\eta t}$ has the period $\bar{T} := \frac{2\pi}{\eta}$. Since $0 < 2\cos(\eta T) < \pi$, or, $\phi \neq \pm 2$, we must have $\eta T \neq j\pi$, $j = 1, 2, \dots$, or $\dfrac{2\pi T}{\bar{T}} \neq j\pi$, which implies that $\dfrac{2T}{j} \neq \bar{T}$. This gives $\bar{T} \neq 2T, T, \dfrac{2}{3}T, \dots.$

Copyright Elsevier 2022

Note that in order that $\bar{T} = nT$, we must have

$$\eta = \frac{2\pi}{nT}$$

for $n \neq 1, 2$.

4.3.2 Applications to Mathieu's equation

4.3.2.1 Inverted pendulum

An application of Mathieu's equation is the inverted pendulum depicted in Fig. 4.3. We assume that the pendulum is frictionless and has a massless rod of length l. If F_{pivot} is the force acting on the pivot, τ_{net} is the net force to external torque, and I is the moment of inertia, then by Newton's second law, we have

$$F_{\text{pivot}} = ma = m\frac{d^2y}{dt^2}, \tag{4.3.18}$$

$$\tau_{\text{net}} = I\frac{d^2\theta}{dt^2} = ml^2\frac{d^2\theta}{dt^2}. \tag{4.3.19}$$

If F is the force acting on the particle and τ is the torque, then

$$\tau = rF\sin(\theta),$$

where r is the distance from the axis of rotation to the particle. So the gravitational torque is

$$\tau_{\text{grav}} = rF_{\text{grav}}\sin(\theta) = mgl\sin(\theta).$$

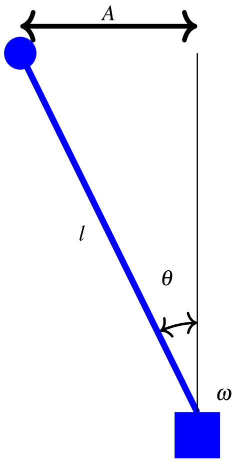


FIGURE 4.3

Inverted pendulum plotted in xy-coordinates.

Copyright Elsevier 2022

Since the harmonic motion of the pendulum can be expressed as $y(t) = A\cos(\omega t)$, Eq. (4.3.18) becomes

$$F_{\text{pivot}} = ma = m\omega^2 A\cos(\omega t).$$

Now the torque exerted by the pivot, τ_{pivot} is given by

$$\tau_{\text{pivot}} = r F_{\text{pivot}}\ \sin(\theta) = -ml\omega^2 A\cos(\omega t)\sin(\theta).$$

Thus the total torque $\tau_{\text{net}} = \tau_{\text{grav}} + \tau_{\text{pivot}}$. As a result, expression (4.3.19) yields the expression

$$ml^2\frac{d^2\theta}{dt^2} = mgl\sin(\theta) - ml\omega^2 A\cos(\omega t)\sin(\theta)$$

or

$$\frac{d^2\theta}{dt^2} + \big(\frac{g}{l} - \frac{\omega^2 A}{l}\cos(t)\big)\sin(\theta),$$

setting $\sin(\theta) \approx \theta$ for small oscillations, the above expression reduces to

$$\theta'' + \big(\alpha + \beta\cos(t)\big) = 0$$

with $\alpha = \dfrac{g}{l}$, $\beta = -\dfrac{\omega^2 A}{l}$, and period $T = 2\pi$.

4.4 Exercises

Exercise 4.1. Discuss the flow of solutions around the equilibrium solutions as we did in Fig. 4.1 for the following systems:

(a)

$$x'(t) = (1 - x^2)xe^x,\ \ x(0) = x_0 \neq 0;$$

(b)

$$x'(t) = x^3 + x^2 + 6x,\ \ x(0) = x_0 \neq 0.$$

Exercise 4.2. Discuss the stability as we did in Example 4.2 for the following IVPs:

(a)

$$(t^2 + 9)x'(t) + tx = 0,\ \ x(0) = x_0 \neq 0;$$

(b)

$$(1 + t)x'(t) = -x,\ x(t_0) = x_0,\ \ t \geq t_0 > -1.$$

Copyright Elsevier 2022

Exercise 4.3. Show that the zero solution of

$$x'(t) = -x^3, \;\; x(t_0) = x_0 \neq 0,$$

is UAS.

Exercise 4.4. The solution of

$$x'(t) = a(t)x(t), \; x(t_0) = x_0 \neq 0, \; t \geq t_0 \geq 0, \tag{4.4.1}$$

is given by

$$x(t) = x_0 e^{\int_{t_0}^{t} a(s)ds},$$

where $a(t)$ is continuous for $t \geq 0$. Show that the zero solution of (4.4.1) is

(a) stable if and only if there exists a positive constant M such that

$$e^{\int_{t_0}^{t} a(s)ds} < M, \;\; t \geq 0;$$

(b) asymptotically stable if and only if

$$\int_0^t a(s)ds \to -\infty \;\text{ as }\; t \to \infty;$$

(c) uniformly stable if and only if there exists a positive constant M such that $\displaystyle\int_{t_1}^{t_2} a(s)ds \leq M$ for all $0 < t_1 < t_2 < \infty$;

(d) uniformly asymptotically stable if and only if there exist positive constants M and α such that $\displaystyle\int_{t_0}^{t} a(s)ds \leq -\alpha(t - t_0) + M$.

Exercise 4.5. Find the fundamental matrix and study the stability of the zero solution of the given system for $t > 0$:

(a)

$$x'(t) = \begin{pmatrix} 0 & 1 \\ -\frac{1}{2t^2} & -\frac{5}{2t} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

(b)

$$x'(t) = \begin{pmatrix} 0 & 1 \\ -\frac{8}{t^2} & -\frac{10}{t} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Exercise 4.6. Find the Floquet multipliers and exponents of

$$x'(t) = \begin{pmatrix} \frac{\cos(t)+\sin(t)}{2+\sin(t)-\cos(t)} & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Copyright Elsevier 2022

Exercise 4.7. Find the Floquet multipliers and exponents of

$$x'(t) = \begin{pmatrix} 1 + \frac{\cos(t)}{2+\sin(t)} & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Exercise 4.8. Find the Floquet multipliers and exponents of

(a)

$$x'(t) = \cos(3t)x(t);$$

(b)

$$x'(t) = \Big(\sin(3t) + \cos(3t)\Big)x(t);$$

(c)

$$x'(t) = \Big(-1 + \cos(2t)\Big)x(t).$$

Exercise 4.9. Prove that $B_j = e^{\mathrm{Ln}(B_j)}$ in the proof of Lemma 4.1.

Exercise 4.10. Carry out a similar discussion as in Section 4.3.1 for the second-order Hill equation

$$x''(t) + a(t)x(t) = 0,$$

where the function a is continuous and periodic with period T.

Exercise 4.11. Find the logarithm of the rotation matrix

$$A = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

Exercise 4.12. Show that all solutions of (4.2.1) are bounded if and only if its zero solution is stable.

Copyright Elsevier 2022

CHAPTER

5

Qualitative analysis of linear systems

This chapter is exclusively devoted to the study of the stability of linear systems, near-linear systems, perturbed systems, and autonomous systems in the plane.

5.1 Preliminary theorems

In Section 4.2 of Chapter 4, we discussed the stability, uniform stability, asymptotic stability, and uniform asymptotic stability of the zero solution of the non-autonomous linear system

$$x' = A(t)x$$

by imposing conditions on its fundamental and principal matrices. In this section, we give explicit criteria that satisfy the conditions that were imposed only for the linear system with constant coefficients

$$x' = Ax, \tag{5.1.1}$$

where A is an $n \times n$ matrix with constant entries. We begin with the following lemma.

Lemma 5.1. *Let*

$$r(t) = t^m e^{-\alpha t}, \ t \geq 0,$$

where m is a non-negative integer, and α is a positive constant. Then there exists a positive constant D depending on m and α such that $r(t) \leq D$.

Proof. If $m = 0$ then $r(t) \leq 1 := D$. On the other hand, if $m > 0$, then by repeated application of L'Hôpital's rule we have $\lim_{t\to\infty} t^m e^{-\alpha t} = 0$. Moreover, since $r(0) = 0$, $r(t)$ is continuous, and $r(t) \geq 0$, we have that $r(t) \leq D$ for some positive constant D. This completes the proof. □

Definition 5.1.1. Let A be a square matrix. We define the *spectrum* of A as

$$\sigma(A) = \{\lambda : \det(A - \lambda I) = 0\}.$$

Thus the inequality

$$Re\big(\sigma(A)\big) < \zeta$$

for some real number ζ means that all the eigenvalues of A have real parts less than ζ.

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00011-5
Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

The next theorem is concerned with eigenvalues of multiplicity m.

Theorem 5.1.1. *Let* λ *be an eigenvalue of a matrix* A *with* $Re\big(\sigma(A)\big) < \zeta$. *Then for any integer* $m \geq 0$, *there exists a positive constant* D *such that*

$$t^m e^{\lambda t} \leq D e^{\zeta t} \text{ for all } t \geq 0.$$

Proof. Let $\lambda = \alpha + i\beta$ with $\alpha, \beta \in \mathbb{R}$. If $\beta = 0$, that is, λ is real, then since $Re\big(\sigma(A)\big) < \zeta$, we have that $\alpha - \zeta < 0$. So by Lemma 5.1

$$t^m e^{(\alpha-\zeta)t} \leq D \text{ for all } t \geq 0.$$

Multiplying this inequality by $e^{\zeta t}$, we obtain

$$t^m e^{\lambda t} = t^m e^{\alpha t} \leq D e^{\zeta t} \text{ for all } t \geq 0.$$

Now if $\beta \neq 0$, that is, λ is complex, then

$$|t^m e^{\lambda t}| = |t^m e^{(\alpha+i\beta)t}| = t^m e^{\alpha t} \leq D e^{\zeta t} \text{ for all } t \geq 0,$$

since $\alpha = Re\big(\sigma(A)\big) < \zeta$. This completes the proof. □

Theorem 5.1.2. *Let* λ *be an eigenvalue of a matrix* A *with* $Re\big(\sigma(A)\big) < \zeta$. *Then for some positive constant* D, *any solution* x *of* (5.1.1) *satisfies*

$$|x(t)| \leq D e^{\zeta t} \text{ for all } t \geq 0.$$

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be eigenvalues of A. We can form a set of fundamental solution, $x_1, x_2, \ldots, x_n$ and find an integer $N \geq 0$ such that each solution is of the form $x_j = t^m e^{\lambda_N t}$, $N = 1, 2, \ldots, k$, for integer m. Then by Theorem 5.1.1, we have that $|x_j(t)| \leq V_j e^{\zeta t}$ for positive constants V_j. If x is an arbitrary solution of (5.1.1), then it can be written in the form

$$x(t) = c_1 x_1(t) + c_2 x_2(t) + \ldots + c_n x_n(t).$$

Then for $t \geq 0$,

$$\begin{aligned}
|x(t)| &\leq |c_1||x_1(t)| + |c_2||x_2(t)| + \ldots + |c_n||x_n(t)| \\
&\leq |c_1| V_1 e^{\zeta t} + |c_2| V_2 e^{\zeta t} + \ldots + |c_n| V_n e^{\zeta t} \\
&= \Big(|c_1| V_1 + |c_2| V_2 + \ldots + |c_n| V_n\Big) e^{\zeta t}.
\end{aligned}$$

This completes the proof. □

The next theorem plays an important role in the study of stability, and its proof depends on the Putzer algorithm, Theorem 3.2.14.

Copyright Elsevier 2022

Theorem 5.1.3. *Let λ_i, $i = 1, 2, \ldots, k$, be the distinct eigenvalues of A, where λ_i has multiplicity n_i, and $n_1 + n_2 + \ldots + n_k = n$. If $Re\big(\sigma(A)\big) < \zeta$, then there is a positive constant D such that*

$$e^{At} \leq De^{\zeta t} \text{ for all } t \geq 0.$$

Proof. From Theorem 3.2.14 we know that

$$e^{At} = \sum_{j=0}^{n-1} \mu_{j+1}(t) P_j, \tag{5.1.2}$$

where $\mu_1(t), \mu_2(t), \ldots, \mu_n(t)$ satisfy the corresponding initial value problems

$$\mu_1' = \lambda_1 \mu_1,\ \mu_1(0) = 1;\ \mu_j' = \lambda_j \mu_j + \mu_{j-1}(t),\ \mu_j(0) = 0,\ j = 2, 3, \ldots, n,$$

and the matrices $P_0, P_1, \ldots, P_n$ are given by

$$P_0 = I,\quad P_j = \prod_{k=1}^{j} (A - \lambda_j I),\ j = 1, 2, \ldots, n.$$

Then we may easily verify that

$$\mu_1(t) = e^{\lambda_1 t},\ \ \mu_{j+1}(t) = e^{\lambda_{j+1} t} \int_0^t e^{-\lambda_{j+1} t} \mu_j(s) ds$$

for $j = 1, 2, \ldots, n-1$. Without loss of generality, we may choose the order in $\{\lambda_j\}$ so that the real parts do not decrease, that is, if $\lambda_j = \rho_j + i\sigma_j$, then $\rho_j \leq \rho_{j+1}$. We claim that

$$|\mu_j(t)| \leq \frac{t^{j-1}}{(j-1)!} e^{\rho_j t} \text{ for } t > 0,\ j = 1, 2, \ldots, n. \tag{5.1.3}$$

Obviously, (5.1.3) holds for $j = 1$, since $\mu_1(t) = e^{\lambda_1 t}$. Thus let it hold for j. Now from the above integral defining $\mu_{j+1}(t)$ we have

$$\mu_{j+1}(t) = e^{\rho_{j+1} t} \int_0^t e^{-\rho_{j+1} s} \frac{s^{j-1}}{(j-1)!} e^{\rho_j s} ds.$$

Since $e^{(\rho_j - \rho_{j+1})s} \leq 1$, the above integral reduces to

$$\mu_{j+1}(t) = e^{\rho_{j+1} t} \frac{1}{(j-1)!} \int_0^t s^{j-1} ds = \frac{t^j}{j!} e^{\rho_{j+1} t}.$$

By our setup, we have ρ^n is the largest of the real parts, and hence

$$\mu_{j+1}(t) \leq \frac{t^j}{j!} e^{\rho_n t}.$$

Copyright Elsevier 2022

For any exponent j and any $\rho^* > 0$, we have

$$t^j < he^{\rho^* t}$$

for all $t \geq 0$ and some positive constant h. Thus

$$\mu_{j+1}(t) \leq he^{(\rho^*+\rho_n)t}, \quad t \geq 0. \tag{5.1.4}$$

Now substitute (5.1.3) into (5.1.2) and then take the matrix norm to get

$$||e^{At}|| \leq \sum_{j=0}^{n-1} |\mu_{j+1}(t)|||P_j|| \leq \Big[h\sum_{j=0}^{n-1} ||P_j||\Big]e^{\zeta t} := De^{\zeta t}. \tag{5.1.5}$$

This completes the proof. □

As a direct consequence of Theorem 5.1.3, we have the following lemma.

Lemma 5.2. *Suppose all the eigenvalues of A have negative real parts. Then there are positive constants h, K, and ρ such that*

$$\mu_{j+1}(t) \leq he^{-\rho t} \ \text{ and } \ ||e^{At}|| \leq Ke^{-\rho t}, \ t \geq 0. \tag{5.1.6}$$

Theorem 5.1.4. **(i)** *Assume that every eigenvalue of the matrix A has a nonpositive real part and that the eigenvalues with zero real parts are simple. Then there is a positive constant D such that*

$$||e^{At}|| \leq D, \ t \geq 0.$$

Moreover, if x is a solution of (5.1.1), then it is bounded, and the zero solution is stable.

(ii) *If $Re(\sigma(A)) < 0$, then the zero solution is AS.*

(iii) *If there is an eigenvalue of A with positive real part, then the zero solution is unstable.*

Proof. If $\lambda = i\beta$ is a simple pure imaginary, then it contributes a bounded function of the form $e^{i\beta t}$ to a fundamental set of solutions. Thus we get a fundamental set of solutions with all elements bounded on $[0, \infty)$. Since e^{At} is formed from such a set, an argument similar to the proof of Theorem 5.1.3 shows that

$$||e^{At}|| \leq D$$

for some positive constant D. Now if x is a solution of (5.1.1) with $x(t_0) = x_0$, then

$$x(t) = e^{A(t-t_0)}x_0 \ \text{ for } \ t \geq t_0.$$

Let $\varepsilon > 0$ and set $\delta = \varepsilon/D$. Then for $|x_0| < \delta$, we have

$$|x(t)| = ||e^{A(t-t_0)}|||x_0| \leq D\delta = \varepsilon.$$

Copyright Elsevier 2022

This proves that the solution is bounded, and the zero solution is stable. The proof of (i) is complete.

As for (ii), the zero solution is stable by (i) since $Re(\sigma(A)) < 0$. Thus we may choose a real number $\zeta < 0$ such $Re(\sigma(A)) < \zeta$. By Theorem 5.1.3 there is a positive constant D such that $||e^{At}|| \leq De^{\zeta t}$ for all $t \geq 0$. Hence, for any initial condition $x(0) = x_0$, the solution x of (5.1.1) satisfies

$$|x(t, 0, x_0)| = |e^{At}x_0| \leq D|x_0|e^{\zeta t} \to 0 \text{ as } t \to \infty.$$

This completes the proof of (ii). ☐

Remark 5.1. Consider two matrices

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

It is clear that each matrix has an eigenvalue 0 with multiplicity 2 (not simple). The zero solution of the corresponding system $x' = Ax$ is stable, whereas the zero solution of $x' = Bx$ is unstable. Therefore careful examination of the stability is required when dealing with eigenvalues with zero real parts of multiplicity higher than one.

Theorem 5.1.5. *If all the eigenvalues of A are distinct with $Re\big(\sigma(A)\big) < -\zeta$, where $\zeta > 0$, then the zero solution of* (5.1.1) *is UAS.*

Proof. By Lemma 5.2, we have that

$$||e^{At}|| \leq Ke^{-\zeta t}, \ t \geq 0,$$

and hence

$$||e^{A(t-s)}|| \leq Ke^{-\zeta(t-s)}, \ t \geq s \geq 0.$$

Now if x is a solution of (5.1.1) with $x(s) = x_0$, then

$$x(t) = e^{A(t-s)}x_0 \ \text{ for } \ t \geq s,$$

and the results follow from Theorem 4.2.3, since

$$e^{A(t-s)} = \Phi(t)\Phi^{-1}(s).$$

This completes the proof. ☐

5.2 Near-constant systems

In this section, we examine when solutions of a linear non-autonomous system behave like solutions of an autonomous linear system near zero. We have already seen that the stability of zero solution of the system

$$x'(t) = A(t)x(t)$$

Copyright Elsevier 2022

does not depend on the sign of the eigenvalues of the matrix $A(t)$. For example, the system

$$x'(t) = \begin{pmatrix} -1+\frac{3}{2}\cos^2(t) & 1-\frac{3}{2}\cos(t)\sin(t) \\ -1-\frac{3}{2}\cos(t)\sin(t) & -1+\frac{3}{2}\sin^2(t) \end{pmatrix} x(t) \tag{5.2.1}$$

has the eigenvalues

$$\lambda = \frac{-1 \pm i\sqrt{7}}{4}$$

and the principal matrix

$$\Psi(t) = \begin{pmatrix} e^{t/2}\cos(t) & e^{-t}\sin(t) \\ -e^{t/2}\sin(t) & e^{-t}\cos(t) \end{pmatrix}.$$

This tells us that solutions are unbounded and the zero solution is unstable. Thus the question that we wish to examine is that if, in some sense, the matrix $A(t)$ is close to a constant matrix A, then how closely do solutions of the non-autonomous system

$$x'(t) = A(t)x(t) \tag{5.2.2}$$

behave like solutions of the autonomous system

$$x'(t) = Ax(t)? \tag{5.2.3}$$

We assume that both matrices are $n \times n$ and $A(t)$ has continuous entries on some interval that includes the initial time t_0.

Theorem 5.2.1. *Assume that every eigenvalue of the matrix A has a nonpositive real part and that the eigenvalues with zero real parts are simple. If*

$$\int_0^\infty ||A(t) - A||dt \le E \tag{5.2.4}$$

for some positive constant E, then solutions of (5.2.2) *are bounded, and its zero solution is uniformly stable.*

Proof. By Theorem 5.1.4 there is a positive constant D such that

$$e^{At} \le D.$$

Rewrite (5.2.2) as

$$x'(t) = Ax(t) + \big(A(t) - A\big)x(t) \tag{5.2.5}$$

and set $g(t) = \big(A(t) - A\big)x(t)$. If x is a solution of (5.2.2) with $x(t_0) = x_0$, then it satisfies (5.2.5), and by the variation of parameters formula (3.3.6) we have that

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^t e^{A(t-s)}\big(A(s) - A\big)x(s)ds. \tag{5.2.6}$$

Copyright Elsevier 2022

Taking the appropriate norms of both sides of (5.2.6) and estimating give, for $t \geq t_0 \geq 0$,

$$\begin{aligned} |x(t)| &= ||e^{A(t-t_0)}|||x_0| + \int_{t_0}^{t} ||e^{A(t-s)}||\,||A(s)-A||\,|x(s)|ds \\ &\leq D|x_0| + \int_{t_0}^{t} D||A(s)-A||\,|x(s)|ds \\ &\leq D|x_0|e^{\int_{t_0}^{t} D||A(s)-A||\,ds} \quad \text{(by Gronwall's inequality)} \qquad (5.2.7) \\ &\leq D|x_0|e^{DE}. \qquad (5.2.8) \end{aligned}$$

It follows that solutions are bounded and the zero solution is uniformly stable by choosing $\delta = \frac{\varepsilon}{De^{DE}}$ for any $\varepsilon > 0$ with $|x_0| < \delta$. □

Theorem 5.2.2. *Assume that all the eigenvalues of the matrix A of system (5.2.3) are distinct with negative real parts and there is a positive number (to be identified later in the proof) θ such that*

$$||A(t)-A|| < \theta \tag{5.2.9}$$

for all $t \geq t_0 \geq 0$. Then the zero solution of (5.2.2) is uniformly asymptotically stable.

Proof. Since $Re(\sigma(A)) < 0$, we may choose a real number $\zeta > 0$ such $Re(\sigma(A)) < -\zeta$. By Lemma 5.2, there is a positive constant K such that $||e^{At}|| \leq Ke^{-\zeta t}$ for all $t \geq 0$. If x is a solution of (5.2.2) with $x(t_0) = x_0$, then it satisfies (5.2.5), and by the variation of parameters formula (3.3.6), we have that

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-s)}\big(A(s)-A\big)x(s)ds.$$

Let

$$0 < \theta < \frac{\zeta}{K}. \tag{5.2.10}$$

Taking the appropriate norms of both sides gives, for $t \geq 0$,

$$\begin{aligned} |x(t)| &= ||e^{A(t-t_0)}|||x_0| + \int_{t_0}^{t} ||e^{A(t-s)}||\,||A(s)-A||\,|x(s)|ds \\ &\leq Ke^{-\zeta(t-t_0)}|x_0| + \int_{t_0}^{t} Ke^{-\zeta(t-s)}\,\theta\,|x(s)|ds \end{aligned}$$

or

$$|x(t)|e^{\zeta(t-t_0)} \leq K|x_0| + \int_{t_0}^{t} Ke^{\zeta(s-t_0)}\,\theta\,|x(s)|ds.$$

Copyright Elsevier 2022

Let $z(t) = e^{\zeta(t-t_0)}|x(t)|$. Then the above inequality reduces to

$$z(t) \le K|x_0| + k\theta \int_{t_0}^{t} z(s)ds.$$

An application of Gronwall's inequality leads to

$$z(t) \le k|x_0|e^{k\theta(t-t_0)} \text{ for all } t \ge t_0 \ge 0,$$

and, consequently,

$$x(t)| \le k|x_0|e^{(k\theta-\zeta)(t-t_0)} \text{ for all } t \ge t_0 \ge 0,$$

from which we arrive at the uniform stability and uniform asymptotic stability of the zero solution by imitating the proof of Theorem 4.2.3. This completes the proof. ☐

Remark 5.2. There is no standard way for forming (5.2.5). It depends on what we want from the eigenvalues of A and conditions (5.2.4) and (5.2.9).

Example 5.1. We show that the zero solution of the non-autonomous system

$$x'(t) = \begin{pmatrix} \frac{e^{-t}}{3} & 1 - \frac{te^{-t}}{6} \\ -2 + \frac{1}{16(1+t)^2} & -3 + \frac{e^{-t}}{16} \end{pmatrix} x \tag{5.2.11}$$

is uniformly asymptotically stable. We write the system in the form

$$\begin{aligned} x'(t) &= \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} x + \left[\begin{pmatrix} \frac{e^{-t}}{3} & 1 - \frac{te^{-t}}{6} \\ -2 + \frac{1}{16(1+t)^2} & -3 + \frac{e^{-t}}{16} \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} \right] x \\ &= Ax + \big(A(t) - A\big)x. \end{aligned}$$

Now the eigenvalues of A are -2 and -1. We easily verify that

$$e^{A(t-t_0)} = \begin{pmatrix} e^{-(t-t_0)} & e^{-3(t-t_0)} \\ -e^{-(t-t_0)} & -3e^{-(t-t_0)} \end{pmatrix}, \ t \ge t_0 \ge 0,$$

and

$$||e^{A(t-t_0)}||_\infty \le 4e^{-(t-t_0)}, \ \ t \ge t_0 \ge 0.$$

Moreover,

$$||A(t) - A||_\infty = \max\{\frac{1}{3} + \frac{1}{6e}, \frac{1}{16} + \frac{1}{16}\} = \frac{1}{8} \text{ for all } t \ge t_0 \ge 0,$$

and condition (5.2.10) is satisfied with $K = 4$, $\zeta = 1$, and $\theta = 2$. Thus by Theorem 5.2.2 the zero solution of (5.2.11) is uniformly asymptotically stable.

Copyright Elsevier 2022

5.3 Perturbed linear systems

Suppose we know about the stability of the zero solution of the linear system

$$x'(t) = A(t)x(t), \ t \geq 0. \tag{5.3.1}$$

The question is, what can we say about the stability of the zero solution of the *perturbed* linear system

$$x'(t) = A(t)x(t) + B(t)x(t), \ t \geq 0, \tag{5.3.2}$$

where A and B are $n \times n$ continuous matrices on the interval $[0, \infty)$? In other words, under what conditions on $B(t)$ is the perturbed system (5.3.2) stable? It turns out that if, in some sense, the norm of the matrix B is small, then a type of stability of (5.3.2) can be deduced. Note that in (5.3.2), $B(t)$ is known as the *perturbation term*.

Theorem 5.3.1. *Assume that the zero solution of* (5.3.1) *is uniformly sable and that*

$$\int_0^\infty ||B(t)||dt \leq E \tag{5.3.3}$$

for some positive constant E. Then the zero solution of (5.3.2) *is uniformly stable.*

Proof. Let $\Phi(t)$ be the fundamental matrix of (5.3.1). Due to the US of the zero solution, for some positive constant K, we have

$$||\Phi(t)\Phi^{-1}(s)|| \leq K.$$

By the variation of parameters formula (3.3.4), we have that

$$x(t) = \Phi(t)\Phi^{-1}(t_0)x_0 + \int_{t_0}^t \Phi(t)\Phi^{-1}(s)B(s)ds$$

or

$$\begin{aligned} |x(t)| &= ||\Phi(t)\Phi^{-1}(t_0)|||x_0| + \int_{t_0}^t ||\Phi(t)\Phi^{-1}(s)||\,||B(s)||\,|x(s)|ds \\ &\leq K|x_0| + \int_{t_0}^t K||B(s)||\,|x(s)|ds \\ &\leq K|x_0|e^{\int_{t_0}^t K||B(s)||\,ds} \quad \text{(by Gronwall's inequality)} \\ &\leq K|x_0|e^{KE}, \end{aligned} \tag{5.3.4}$$

which implies the uniform stability. This completes the proof. □

The next example shows that the integrability condition (5.3.3) for $||B(t)||$ is necessary.

Copyright Elsevier 2022

Example 5.2. Letting

$$x_1' = x_2,$$

the second-order differential equation

$$x'' + x = 0$$

can be written in the matrix form

$$x'(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

with

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Thus its zero solution is uniformly stable. On the other hand, Theorem 5.3.1 says nothing about the perturbed system

$$x'' - \frac{2}{t}x' + x = 0, \ t > 0,$$

which is equivalent to

$$x'(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 & 0 \\ 0 & \frac{2}{t} \end{pmatrix} x.$$

Here $B(t) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{2}{t} \end{pmatrix}$ is the perturbed matrix in the sense that $|B(t)| = \frac{2}{t} \to 0$ as $t \to \infty$. Moreover,

$$\int_{t^*}^{\infty} ||B(t)||dt = \int_{t^*}^{\infty} \frac{2}{t} dt = \infty$$

for any $t^* > 0$, and hence $B(t)$ does not satisfy condition (5.3.3).

Note that the perturbed system has two linearly independent solutions

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \sin(t) - t\cos(t) \\ t\sin(t) \end{pmatrix}, \ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos(t) + t\sin(t) \\ t\cos(t) \end{pmatrix},$$

and consequently its solutions are unbounded, and as a result, its zero solution is unstable.

Theorem 5.3.2. *Assume that the zero solution of* (5.3.1) *is uniformly asymptotically stable, that is, if* $\Phi(t)$ *is the fundamental matrix of* (5.3.1)*, then by* (4.2.3) *there are positive constants* K *and* ξ *such that*

$$||\Phi(t)\Phi^{-1}(s)|| \leq K e^{\xi(t-s)}, \ t \geq s \geq 0.$$

Copyright Elsevier 2022

Suppose

$$\int_0^{\infty} ||B(t)||dt \le \gamma(t-t_0)+\beta,\ \ t \ge t_0 \ge 0, \tag{5.3.5}$$

for positive constants γ and β. Then the zero solution of (5.3.2) *is uniformly asymptotically stable, provided that*

$$\gamma < \frac{\xi}{K}.$$

Proof. Imitate the proof of Theorem 5.3.1 and get

$$|x(t)|e^{\xi(t-t_0)} \le K|x_0| + \int_{t_0}^{t} Ke^{\xi(s-t_0)}\, ||B(s)||\, |x(s)|ds.$$

Using Gronwall's inequality along with condition (5.3.5), we arrive at

$$|x(t)| \le Ke^{K\beta}|x_0|e^{K(\gamma-\frac{\xi}{K})(t-t_0)}. \tag{5.3.6}$$

The rest of the proof follows along the lines of Theorem 5.2.2 using inequality (5.3.6). This completes the proof. ☐

5.4 Autonomous systems in the plane

In most applications, the corresponding modeling systems can be written as two-dimensional autonomous systems in the form

$$x' = P(x,y),\ \ y' = Q(x,y), \tag{5.4.1}$$

where P and Q are continuous on some subset of $\mathbb{R}^2$. Recall that the system is called *autonomous* if P and Q do not depend explicitly on t. System (5.4.1) can be written in the form $z' = f(z)$, where

$$z = \begin{pmatrix} x \\ y \end{pmatrix} \text{ and } f(z) = \begin{pmatrix} P(z) \\ Q(z) \end{pmatrix}.$$

Thus the second-order differential equation

$$x'' = g(x,x')$$

can be written in the form (5.4.1) as the system

$$x' = y,\ \ y' = g(x,y).$$

We begin with the following lemma. We denote the solution of (5.4.1) by the pair $(x(t), y(t)),\ t \in (a,b)$.

Copyright Elsevier 2022

Lemma 5.3. *Let* $(x(t), y(t))$ *be a solution of* (5.4.1). *Then for any constant* k*, the pair* $(x(t+k), y(t+k))$ *is also a solution of* (5.4.1).

Proof. Let $x_1 = x(t+k)$ and $x_2 = y(t+k)$. Then by the chain rule we have

$$\begin{aligned} x_1' &= x'(t+k) = P(x(t+k), y(t+k)) = P(x_1, x_2), \\ x_2' &= y'(t+k) = Q(x(t+k), y(t+k)) = Q(x_1, x_2), \end{aligned}$$

which imply that the pair $(x_1(t), x_2(t))$ is a solution of (5.4.1) for $t \in (a-k, b-k)$. This completes the proof. ☐

The results of Lemma 5.3 do not hold for non-autonomous systems. To see this, we consider the system

$$x' = ty, \ \ y' = y.$$

It readily follows that

$$x(t) = (t-1)e^t, \ \ y(t) = e^t$$

is a solution of the system. However, for any nonzero constant k, we have $x(t+k) = (t+k-1)e^{t+k}$ and $x'(t+k) = (t+k)e^{t+k} \neq ty(t)$.

Definition 5.4.1. The solution $(x(t), y(t))$ of (5.4.1) as t varies describes parametrically a curve in the plane, which we call a *trajectory* (or orbit).

Theorem 5.4.1. *Different trajectories of* (5.4.1) *cannot intersect.*

Proof. Let C_1 and C_2 be two trajectories represented by $(x_1(t), y_1(t))$ and $(x_2(t), y_2(t))$, respectively, with a common point (x_0, y_0). Then there are two different times (otherwise, uniqueness is violated), say t_1 and t_2, such that

$$(x_0, y_0) = (x_1(t_1), y_1(t_1)) = (x_2(t_2), y_2(t_2)).$$

By Lemma 5.3

$$x(t) = x_1(t + t_1 - t_2), \ \ y(t) = y_1(t + t_1 - t_2)$$

is a solution. Since $(x(t_2), y(t_2)) = (x_0, y_0)$, we have that $x(t)$ and $y(t)$ must agree respectively with $x_2(t)$ and $y_2(t)$ due to the uniqueness. Thus C_1 and C_2 must coincide. This completes the proof. ☐

Remark 5.3. Be aware that a trajectory is a curve represented parametrically by one or more solutions. Thus $x(t), y(t)$ and $x(t+k), y(t+k)$ for $k \neq 0$ represent distinct solutions but the same trajectory.

Example 5.3. The damped harmonic motion represented by the second-order differential equation

$$u'' + 2u' + 2u = 0$$

Copyright Elsevier 2022

can be easily transformed into the planar autonomous system

$$\begin{aligned} x' &= y, \\ y' &= -2x - 2y. \end{aligned}$$

The eigenvalues of the system are found to be $\lambda = -1 \pm i$. Hence the solution is

$$\begin{aligned} x &= e^{-t}\left(c_1 \cos(t) + c_2 \sin(t)\right), \\ y &= e^{-t}\left(-c_1 \sin(t) + c_2 \cos(t)\right). \end{aligned}$$

To describe the trajectories, we find that

$$x^2 + y^2 = (c_1^2 + c_2^2)e^{-2t},$$

which describes a family of spirals.

We begin a detailed study of the linear plane autonomous system

$$X' = AX, \tag{5.4.2}$$

where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is a nonsingular matrix, and

$$X = \begin{pmatrix} x \\ y \end{pmatrix}.$$

The characteristic equation for this system is

$$\lambda^2 - (a+d)\lambda + ad - bc = 0.$$

By introducing

$$\begin{aligned} p &= a + d, \\ q &= ad - bc, \end{aligned}$$

the characteristic equation becomes

$$\lambda^2 - p\lambda + q = 0.$$

Let λ_1 and λ_2 be the roots of this equation,

$$\lambda_{1,2} = \frac{1}{2}\left[p \pm \sqrt{p^2 - 4q}\right]. \tag{5.4.3}$$

Copyright Elsevier 2022

It is obvious that the signs of $\lambda_{1,2}$ and hence the stability of the zero solution depend on the discriminant

$$\Delta = p^2 - 4q = (\mathrm{tr}(A))^2 - 4\det(A).$$

Based on Theorem 5.1.4, we can now state the following stability theorem of the zero solution of (5.4.2).

Theorem 5.4.2. *The zero solution of* (5.4.2) *is*

a) asymptotically stable if any of the followings occurs:
i)

$$\Delta > 0,\ q > 0,\ \textit{and}\ p < 0,$$

ii)

$$\Delta = 0,\ p < 0,$$

iii)

$$\Delta < 0,\ p < 0;$$

b) stable if

$$\Delta < 0,\ p = 0;$$

c) unstable if any of the followings occurs:
i)

$$\Delta > 0,\ q > 0,\ \textit{and}\ p > 0,$$

ii)

$$\Delta = 0,\ p > 0,$$

iii)

$$\Delta < 0,\ p > 0.$$

Remark 5.4. In the case where $p = 0$ and $q = 0$, further analysis is required to determine the stability of the zero solution.

Note that according to Theorem 5.4.2, the zero solution $(0, 0)$ of the system in Example 5.3 is asymptotically stable.

Next, we consider the behavior of the trajectories of system (5.4.2) for various cases of the eigenvalues. In what follows, we use the matrix transformation

$$Y = BX,\ \ \det(B) \neq 0,$$

so that the essential behavior near the critical point $(0, 0)$ remains unchanged.

Copyright Elsevier 2022

1. *Real and distinct roots.* If we let

$$B = \begin{pmatrix} c & \lambda_1 - a \\ c & \lambda_2 - a \end{pmatrix},$$

then system (5.4.2) transforms into the system

$$\begin{aligned} x' &= \lambda_1 x, \\ y' &= \lambda_2 y, \end{aligned}$$

where, for simplicity, x and y are again used as the new coordinates. For some arbitrary constants c_1 and c_2, the solutions of the system are

$$x(t) = c_1 e^{\lambda_1 t}, \quad y(t) = c_2 e^{\lambda_2 t}. \tag{5.4.4}$$

Eliminating the time t in (5.4.4) gives

$$y = c x^{\frac{\lambda_2}{\lambda_1}}, \tag{5.4.5}$$

where c is an arbitrary constant. When λ_1 and λ_2 have the same sign, then Eq. (5.4.5) represent parabolic curves that are tangent to the origin as shown in Fig. 5.1(a). The origin is called a *proper node*. We have the following three subcases.
1.(a) If both λ_1, λ_2 are negative, then the origin is asymptotically stable. In this case, we have a *stable node*.
1.(b) If λ_1, λ_2 are positive, then the origin is unstable. In this case, we have an *unstable node*.
1.(c) If λ_1, λ_2 have opposite signs, then (5.4.5) represents hyperbolic curves as shown in Fig. 5.1(b). The origin is unstable and is called a *saddle point*.

2. *Real and equal roots.* In this case, $\Delta = (a - d)^2 + 4bc = 0$, which implies that

$$\lambda_1 = \lambda_2 = \frac{(a + d)}{2} := \lambda.$$

2. (a) If b or $c = 0$ and $a = d$, then system (5.4.2) reduces to

$$\begin{aligned} x' &= \lambda x, \\ y' &= \lambda y. \end{aligned}$$

As before, solving the system and eliminating t lead to the trajectories that are lines through the origin given by

$$y = \frac{c_2}{c_1} x.$$

Thus, if $\lambda < 0$, then the origin is asymptotically stable and is a *proper node*, as shown in Fig. 5.1(c). If $\lambda > 0$, then the origin is unstable.

Copyright Elsevier 2022

(b) This is a more complicated case. If we let

$$B = \begin{pmatrix} \frac{a-d}{2b} & 1 \\ \frac{1}{b} & 0 \end{pmatrix}, \; b \neq 0,$$

then system (5.4.2) transforms into the system

$$\begin{aligned} x' &= \lambda x, \\ y' &= x + \lambda y. \end{aligned}$$

For some arbitrary constants c_1 and c_2, the solutions of the system are

$$x(t) = c_1 e^{\lambda t}, \;\; y(t) = (c_1 t + c_2) e^{\lambda t}. \tag{5.4.6}$$

The trajectories are shown in Fig. 5.1(d), and the origin is an *improper node*; it is asymptotically stable when $\lambda < 0$ and unstable when $\lambda > 0$.

3. *Complex roots.* In this case, we let

$$\lambda_1 = \alpha + i\beta \;\text{ and }\; \lambda_2 = \alpha - i\beta$$

and choose the matrix

$$B = \begin{pmatrix} c & \alpha - a \\ 0 & \beta \end{pmatrix}$$

so that system (5.4.2) transforms into

$$\begin{aligned} x' &= \alpha x - \beta y, \\ y' &= \beta x + \alpha y. \end{aligned}$$

3. (a) For $\alpha = 0$, this system becomes

$$\begin{aligned} x' &= -\beta y, \\ y' &= \beta x, \end{aligned}$$

with solutions

$$x(t) = c_1 \cos(\beta t) + c_2 \sin(\beta t), \;\; y(t) = c_1 \sin(\beta t) + c_2 \cos(\beta t)$$

for arbitrary constants c_1 and c_2. It is easy to verify that the trajectories are circles given by

$$x^2 + y^2 = c_1^2 + c_2^2$$

and are shown in Fig. 5.1(e). In this case the origin is stable but not asymptotically stable.

Copyright Elsevier 2022

3. (b) For $\alpha \neq 0$, the above system has the solutions

$$x(t) = e^{\alpha t}(c_1 \cos(\beta t) + c_2 \sin(\beta t)), \quad y(t) = e^{\alpha t}(c_1 \sin(\beta t) + c_2 \cos(\beta t))$$

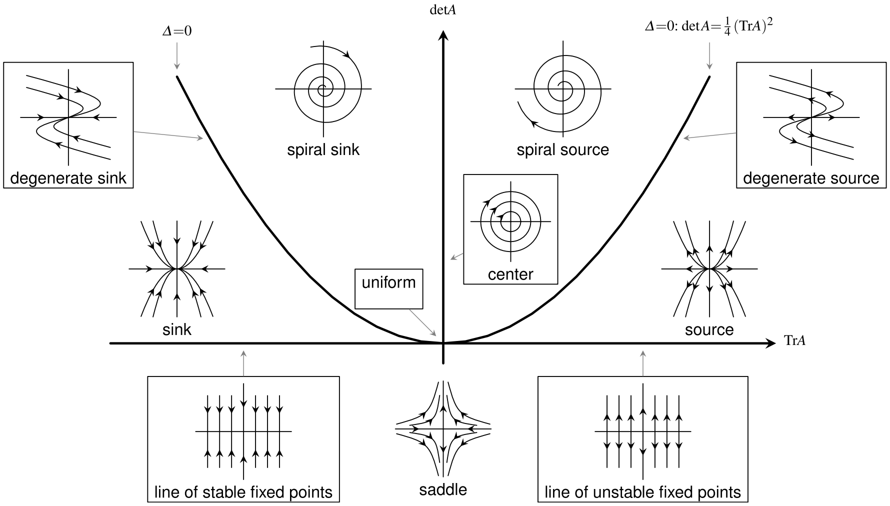
for arbitrary constants c_1 and c_2. The trajectories are spirals given by

$$x^2 + y^2 = (c_1^2 + c_2^2)e^{2\alpha t}$$

and are shown in Fig. 5.1(f). In this case the origin is a *focal point*, which is asymptotically stable when $\alpha < 0$ and unstable when $\alpha > 0$.

The solutions of a planar system of linear differential equations can be classified according to the trace and the determinant of the coefficient matrix. This diagram schematically shows the different types of solutions.

Poincaré Diagram: Classification of Phase Portraits in the $(\det A, \mathrm{Tr}A)$-plane


$\det A$
$\Delta{=}0$
$\Delta{=}0$: $\det A{=}\frac{1}{4}(\mathrm{Tr}A)^2$
degenerate sink
spiral sink
spiral source
degenerate source
center
uniform
sink
source
$\mathrm{Tr}A$
line of stable fixed points
saddle
line of unstable fixed points

5.5 Hamiltonian and gradient systems

Consider the pendulum with friction discussed in Example 3.5:

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -\frac{g}{L}\sin(x_1) - \frac{b}{m}x_2. \end{aligned}$$

Define the *energy function* $H : [t_0, \alpha) \times \mathbb{R}^2 \to [0, \infty)$ by

$$H(x_1, x_2) = \frac{1}{2}x_2^2(t) + \frac{g}{L}\Big(1 - \cos(x_1(t))\Big).$$

Copyright Elsevier 2022

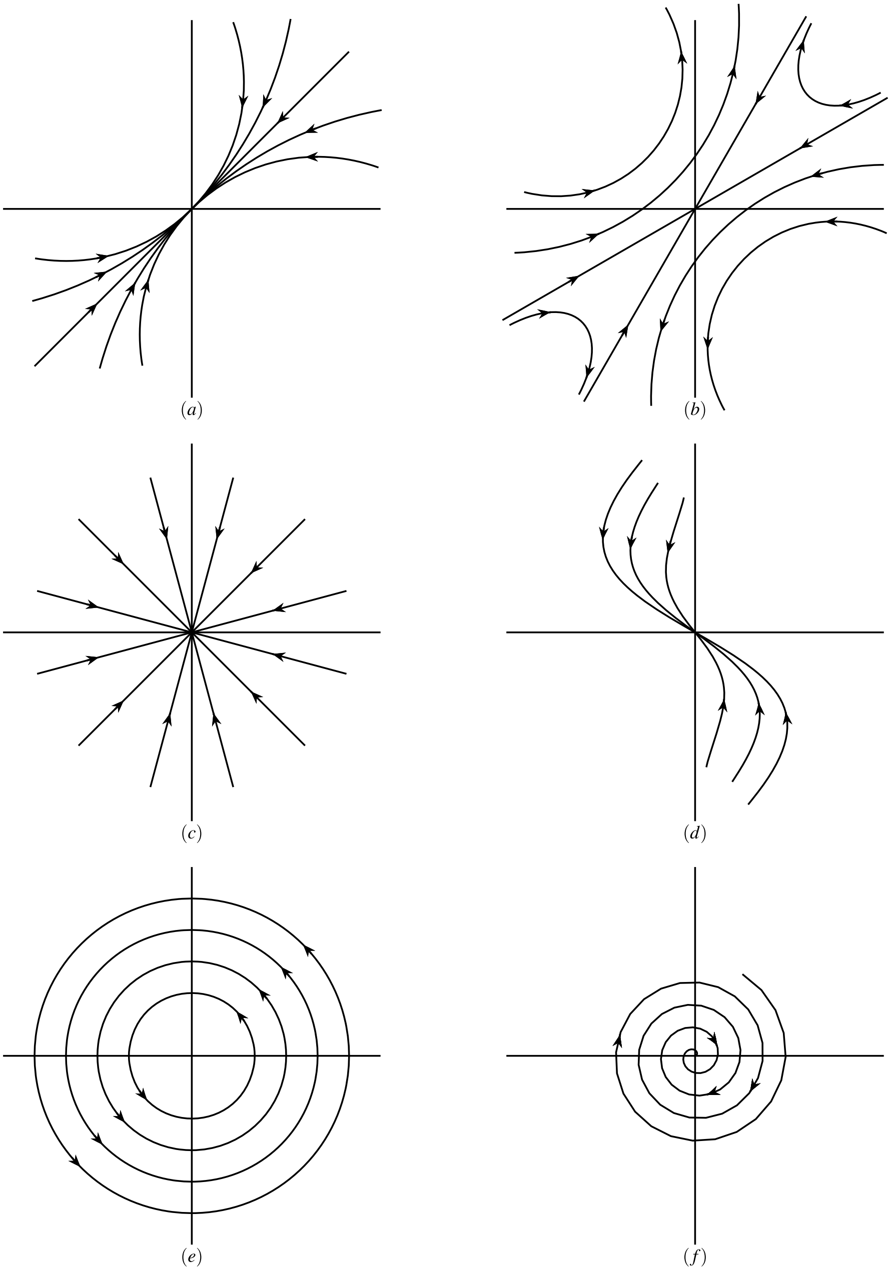


FIGURE 5.1

A phase portrait gallery.

Copyright Elsevier 2022

Then along the solutions of (3.2.7) we have that

$$\begin{aligned}\frac{d}{dt}H(x_1,x_2) &= x_2\,x_2' + \frac{g}{L}\sin(x_1)\,x_1'\\ &= x_2\Big(-\frac{g}{L}\sin(x_1) - \frac{b}{m}x_2\Big) + \frac{g}{L}\sin(x_1)x_2\\ &= -\frac{b}{m}x_2^2.\end{aligned}$$

If the friction $b > 0$, then $\frac{d}{dt}H(t) < 0$, and the *energy is dissipated*. On the other hand, if $b = 0$, then the energy is *conserved*, and theoretically, this solution lies on a *level curve*

$$H(x_1(t), x_2(t)) = \text{constant} := c$$

of the function $H(x_1, x_2)$ with variables x_1 and x_2. Thus, in the case $b > 0$, this solution will not stay on a level curve. Hence centers are possible for (3.2.5).

When the energy is conserved, we obtain an example of a *Hamiltonian system*. In our situation the system

$$\begin{cases} x_1' = x_2, \\ x_2' = -\frac{g}{L}\sin(x_1) \end{cases} \tag{5.5.1}$$

can be written in the form

$$\begin{aligned} x_1' = x_2 &= \frac{\partial H}{\partial x_2},\\ x_2' = -\frac{g}{L}\sin(x_1) &= -\frac{\partial H}{\partial x_1}.\end{aligned}$$

Since the Hamiltonian function conserves the energy, it is constant along solutions. The level curves are given by

$$\frac{1}{2}x_2^2(t) - \frac{g}{L}\cos(x_1(t)) = c.$$

System (5.5.1) has the equilibrium solutions $(\pm n\pi, 0)$, $n = 0, 1, \ldots$. It will be shown in Chapter 6, when there is no damping or friction, that is, $b = 0$, then $(\pm n\pi, 0)$ are centers for $n = 2, 4, 6, \ldots$ and unstable saddles for $n = 1, 3, 5, \ldots$. Phase portraits and level curves are shown in Fig. 5.2 for different values of c.

Thus we have the following definition.

Definition 5.5.1. A planar system of differential equations is called a *Hamiltonian system* if there exists a differentiable function $H(x_1, x_2) : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\begin{cases} x_1'(t) = \frac{\partial H}{\partial x_2}, \\ x_2'(t) = -\frac{\partial H}{\partial x_1}. \end{cases} \tag{5.5.2}$$

Such a function $H(x_1, x_2)$ is called a *Hamiltonian function* of system (5.5.2).

Copyright Elsevier 2022

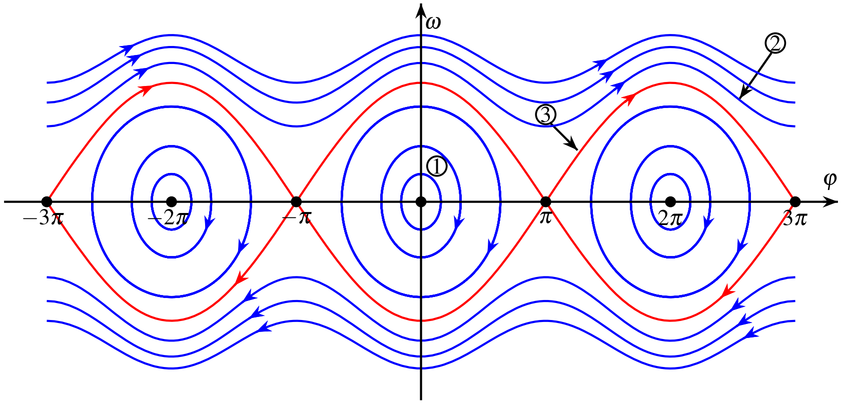


FIGURE 5.2

Phase portraits and level curves for (5.5.1) for different values of c.

Consider the second-order differential equation of the form

$$x'' - f(x)x' - g(x) = 0.$$

In system form, we have

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= g(x_1) + f(x_1)x_2. \end{aligned}$$

Let

$$H(x_1, x_2) = \frac{1}{2}x_2^2(t) + \left(-\int_0^{x_1} g(s)ds\right),$$

where $\frac{1}{2}(x')^2$ corresponds to the *kinetic energy*, and $-\int_0^{x(t)} g(s)ds$ represents the *potential energy*. Then it is easy to verify that

$$\begin{aligned} x_1' &= x_2 = \frac{\partial H}{\partial x_2}, \\ x_2' &= g(x_1) + f(x_1)x_2 = -\frac{\partial H}{\partial x_1}, \end{aligned}$$

only if $f(x) = 0$. We conclude that the *Newtonian system*

$$x'' = g(x)$$

with g continuously differentiable on $\mathbb{R}$ is a Hamiltonian system.

Copyright Elsevier 2022

For the Hamiltonian system (5.5.2), the energy is always conserved. To see this, let $(x_1(t), x_2(t))$ be any solution of (5.5.2). Then

$$\begin{aligned} \frac{d}{dt}H(x_1(t), x_2(t)) &= \frac{\partial H}{\partial x_1}x_1' + \frac{\partial H}{\partial x_2}x_2' \\ &= \frac{\partial H}{\partial x_1}\frac{\partial H}{\partial x_2} + \frac{\partial H}{\partial x_2}\left(-\frac{\partial H}{\partial x_1}\right) = 0. \end{aligned} \tag{5.5.3}$$

That is why a Hamiltonian function is sometimes called an *energy function*. In applications, the Hamiltonian function is defined by

$$H(x, y) = K(x, y) + V(x, y),$$

where K is the kinetic energy, and V is the potential energy.

Next, we establish conditions for the planar system

$$x' = f(x, y), \quad y' = g(x, y) \tag{5.5.4}$$

to be a Hamiltonian system. If (5.5.4) is a Hamiltonian system, then there is a function H satisfying

$$\frac{\partial H}{\partial y} = f(x, y) \text{ and } -\frac{\partial H}{\partial x} = g(x, y).$$

Assuming that the function H has continuous second-order partial derivatives, we arrive at

$$\frac{\partial f}{\partial x} = \frac{\partial^2 H}{\partial x \partial y} = \frac{\partial^2 H}{\partial y \partial x} = -\frac{\partial g}{\partial y}. \tag{5.5.5}$$

Thus we have shown that condition (5.5.5) is a sufficient condition for system (5.5.4) to be a Hamiltonian system.

The system

$$\begin{aligned} x' &= x + y^2, \\ y' &= -y - x^2 \end{aligned}$$

is Hamiltonian since $\frac{\partial f}{\partial x} = 1 = -\frac{\partial g}{\partial y} = -(-1)$.

The next theorem provides necessary and sufficient conditions for the existence of Hamiltonian function.

Theorem 5.5.1. *The planar system* (5.5.4) *is a Hamiltonian system if and only if*

$$\frac{\partial f}{\partial x} = -\frac{\partial g}{\partial y}. \tag{5.5.6}$$

Proof. We already know that if system (5.5.4) is Hamiltonian, then (5.5.6) is true. It remains to show that if (5.5.4) holds, then there is $H(x, y)$ satisfying

$$x'(t) = \frac{\partial H}{\partial y} = f(x, y)$$

Copyright Elsevier 2022

and

$$y'(t) = -\frac{\partial H}{\partial x} = g(x, y).$$

Let us begin with $\frac{\partial H}{\partial x} = -g(x, y)$. Then a direct integration with respect to x keeping y fixed gives

$$H(x, y) = -\int g(x, y)dx + C_1(y),$$

where the constant of integration C_1 may depend on y since H is a function of two variables x and y. Once we determine C_1, we can obtain H. Differentiating the last expression in H with respect to y and setting it equal to $f(x, y)$ leads to

$$\frac{\partial H}{\partial y} = \frac{\partial}{\partial y}\Big(-\int g(x, y)dx + C_1(y)\Big) = f(x, y) \text{ by definition}$$

or

$$C_1'(y) = \frac{\partial}{\partial y}\int g(x, y)dx + f(x, y).$$

Then integrating the last expression with respect to y gives the desired function $C_1(y)$, that is,

$$C_1(y) = \int \Big(\frac{\partial}{\partial y}\int g(x, y)dx\Big)dy + \int f(x, y)dy.$$

The proof will be accomplished if we can show $C_1(y)$ does not depend on x. Equivalently, we can show that the right side of the above expression is independent of x:

$$\begin{aligned}
&\frac{\partial}{\partial x}\Big[\int f(x, y)dy + \int \Big(\frac{\partial}{\partial y}\int g(x, y)dx\Big)dy\Big] \\
&= \int \frac{\partial f}{\partial x}dy + \int \frac{\partial}{\partial x}\Big(\frac{\partial}{\partial y}\int g(x, y)dx\Big)dy \\
&= \int \frac{\partial f}{\partial x}dy + \int \frac{\partial}{\partial y}\Big(\frac{\partial}{\partial x}\int g(x, y)dx\Big)dy \\
&= \int \frac{\partial f}{\partial x}dy + \int \frac{\partial g}{\partial y}dy \\
&= \int \frac{\partial f}{\partial x}dy - \int \frac{\partial f}{\partial x}dy = 0.
\end{aligned}$$

This completes the proof. □

Example 5.4. We already know that the system

$$x' = x + y^2, \;\; y' = -y - x^2$$

Copyright Elsevier 2022

is Hamiltonian since $\frac{\partial f}{\partial x} = 1 = -\frac{\partial g}{\partial y} = -(-1)$. Let us now try to find the Hamiltonian function H. It follows from Theorem 5.5.1 that

$$H(x,y) = -\int g(x,y)dx + C_1(y) = -\int (-y-x^2)dx + c_1(y) = xy + \frac{x^3}{3} + C_1(y),$$

where

$$C_1'(y) = \frac{\partial}{\partial y}\int g(x,y)dx + f(x,y) = \frac{\partial}{\partial y}\Big(\int (-y-x^2)dx\Big) + x + y^2.$$

It follows that $C_1(y) = \frac{y^3}{3}$. Finally,

$$H(x,y) = xy + \frac{x^3}{3} + \frac{y^3}{3}.$$

Next, we try to explore the stability of an equilibrium solution of a Hamiltonian system. We have seen that the autonomous system (5.5.4) is called Hamiltonian or conservative if there exists a Hamiltonian function $H : \Omega \to \mathbb{R}$ that is not constant on any open set in Ω, but it follows from (5.5.3) that it is constant on solutions (orbits).

Theorem 5.5.2. *Suppose* (5.5.4) *is a Hamiltonian system. Then none of its equilibrium solutions can be asymptotically stable.*

Proof. Suppose q is an asymptotically stable equilibrium solution or a point of the Hamiltonian system given by (5.5.4). Then there is a neighborhood Λ of q such that for $p \in \Lambda$, $\phi(t,p) \to q$ as $t \to \infty$, where $\phi(t,p)$ is an orbit or trajectory through p. However, if the system is assumed conservative with Hamiltonian function H, then H is constant on orbits, so $H(p) = H(\phi(t,p)) = H(q)$ for each $p \in \Lambda$. This implies that H is constant on an open set, which is a contradiction. This completes the proof. □

When a Hamiltonian system has an equilibrium point at which the integral has a minimum, then an inference of stability can be made.

Definition 5.5.2. A function $E : \Omega \to \mathbb{R}$ is said to have a *strong minimum* at q if there is a neighborhood Λ of q such that $E(x) > E(q)$ for every $x \in \Lambda$ except for $x = q$.

The name *strong* is implied from the strict inequality.

Theorem 5.5.3. *Suppose* (5.5.4) *is a Hamiltonian system with equilibrium solution q and Hamiltonian function H. If H has a strong minimum at q, then q is stable.*

Proof. Shift the equilibrium solution q to the origin $(0,0)$. Let Λ be a neighborhood of $(0,0)$. Let $\varepsilon > 0$ be small enough such that the set $A = \{d : ||d|| = \varepsilon\} \subset \Lambda$. Set

$$V(d) = H(d) - H(0).$$

Copyright Elsevier 2022

Denote the minimum of V on the set A by V_ε. Now we chose $\delta > 0$ with $\delta < \varepsilon$, so that we have $V(d) < \varepsilon$ if $||d|| < \delta$. Then we have $||d(t)|| < \varepsilon$ for all $t > 0$. This implies that the origin is stable. This completes the proof. ☐

Next, we state parallel results concerning strong maximums and saddle points.

Definition 5.5.3. A function $E : \Omega \to \mathbb{R}$ is said to have a *strong maximum* at q if there is a neighborhood Λ of q such that $E(q) > E(x)$ for every $x \in \Lambda$ except for $x = q$.

The name *strong* is implied from the strict inequality.

Theorem 5.5.4. *Suppose* (5.5.4) *is a Hamiltonian system with equilibrium solution q and Hamiltonian function H. If H has a strong maximum at q, then q is an unstable saddle.*

Remark 5.5. Note that if (x_0, y_0) is a critical point of (5.5.4), then it will also be a critical point of the Hamiltonian function H. We can use the second derivative test again to determine whether or not $H(x, y)$ has a saddle at (x_0, y_0) or a maximum or a minimum there. The Jacobian at the critical point is

$$J(x_0, y_0) = \begin{pmatrix} H_{xy}(x_0, y_0) & H_{yy}(x_0, y_0) \\ -H_{xx}(x_0, y_0) & -H_{xy}(x_0, y_0) \end{pmatrix}.$$

Furthermore,

$$\det J(x_0, y_0) = -\Big(H_{xy}(x_0, y_0)\Big)^2 + H_{xx}(x_0, y_0)H_{yy}(x_0, y_0).$$

Recall from Section 4.3.1 that if λ_1 and λ_2 are eigenvalues of $J(x_0, y_0)$, not necessarily distinct, then $\det J(x_0, y_0) = \lambda_1\lambda_2$. Thus if we assume that none of the eigenvalues of the $J(x_0, y_0)$ is zero, then we have $\det J(x_0, y_0) \neq 0$.

If $\det J(x_0, y_0) > 0$, which is the same as saying that the function H has a strict minimum or maximum at (x_0, y_0), then (x_0, y_0) is a stable center for (5.5.4) if both eigenvalues are negative. This is the case since by Theorem 5.5.2 no critical point of a Hamiltonian system can be asymptotically stable. If the both eigenvalues are positive, then (x_0, y_0) is unstable for (5.5.4).

If $\det(x_0, y_0) < 0$, the equilibrium point (x_0, y_0) is unstable saddle for (5.5.4).

Consider the Hamiltonian system given in Example 5.4. We have already found the Hamiltonian function

$$H(x, y) = xy + \frac{x^3}{3} + \frac{y^3}{3}.$$

The two equilibrium solutions are $(0, 0)$ and $(-1, -1)$. Define

$$D(x, y) = H_{xx}H_{yy} - \Big(H_{xy}\Big)^2.$$

Copyright Elsevier 2022

Then

$$D(x, y) = 4xy - 1.$$

Now $D(0,0) = -1$, and $(0,0)$ is an unstable saddle. On the other hand, $D(-1-1) = 3 > 0$, and since $H_{xx}(-1,-1) = -2 < 0$, the Hamiltonian function has a strong maximum at the equilibrium solution $(-1,-1)$. Thus by Theorem 5.5.4 $(-1,-1)$ is unstable.

Next, we discuss dissipative systems that we call *gradient systems*.

Definition 5.5.4. A planar system of differential equations is called *gradient* if there exists a differentiable function $V(x, y) : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\begin{cases} x'(t) = -\frac{\partial V}{\partial x}, \\ x_2'(t) = -\frac{\partial V}{\partial y}. \end{cases} \tag{5.5.7}$$

The function $V(x, y)$ is called the *gradient function* of system (5.5.7).

The name *gradient* comes from the fact that system (5.5.7) can be written in the form

$$X' = -\nabla V,$$

where

$$X = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \nabla V = \begin{pmatrix} \frac{\partial V}{\partial x} \\ \frac{\partial V}{\partial y} \end{pmatrix}.$$

For the gradient system (5.5.7), the energy is dissipated. To see this, assume that V is an energy function for system (5.5.7). Then if $(x(t), y(t))$ is any solution of (5.5.7), then

$$\begin{aligned} \frac{d}{dt} V(x(t), y(t)) &= \frac{\partial V}{\partial x} x' + \frac{\partial V}{\partial y} y' \\ &= -\left[\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2\right] < 0, \end{aligned} \tag{5.5.8}$$

provided that $\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 \neq 0$.

Next, we establish conditions for the planar system

$$x' = f(x, y), \quad y' = g(x, y) \tag{5.5.9}$$

to be a gradient system. If (5.5.9) is a gradient system, then there is a function V satisfying

$$-\frac{\partial V}{\partial x} = f(x, y) \text{ and } -\frac{\partial V}{\partial y} = g(x, y).$$

Copyright Elsevier 2022

Assuming that the function V has continuous second-order partial derivatives, we arrive at

$$\frac{\partial f}{\partial y} = -\frac{\partial^2 V}{\partial x \partial y} = -\frac{\partial^2 V}{\partial y \partial x} = \frac{\partial g}{\partial x}. \tag{5.5.10}$$

Thus we have shown that condition (5.5.10) is a sufficient condition for system (5.5.9) to be a gradient system. The system of differential equations

$$x' = -2xy, \quad y' = -(x^2 - 1)$$

is gradient since $\frac{\partial f}{\partial y} = -2x = \frac{\partial g}{\partial x}$. Following parallel steps as in Theorem 5.5.1, we easily find the gradient function

$$V(x, y) = x^2 y - y.$$

Note that

$$\frac{dV}{dt} = 2xyx' + (x^2 - 1)y' = -[4x^2y^2 + (x^2 - 1)^2] = -[V_x^2 + V_y^2] < 0$$

for $V_x^2 + V_y^2 \neq 0$. As in the case of a Hamiltonian system, if (x_0, y_0) is a critical point of (5.5.9), then it will also be a critical point of the gradient function V. Now we state parallel theorems to Theorems 5.5.3 and 5.5.4.

Theorem 5.5.5. *Suppose* (5.5.9) *is a gradient system with equilibrium solution q and gradient function H. If H has a strong minimum at q, then q is stable.*

Theorem 5.5.6. *Suppose* (5.5.9) *is a gradient system with equilibrium solution q and gradient function H. If H has a strong maximum at q, then q is an unstable saddle.*

Example 5.5. Consider the system

$$x' = -8x^3 + 2x, \quad y' = -2y + 2,$$

which is gradient with gradient function

$$V(x, y) = 2x^4 + y^2 - x^2 - 2y.$$

The system has three equilibrium solutions given by $(-1/2, 1)$, $(0, 1)$, $(1/2, 1)$. Using the second derivative test, we find that V has a strong minimum at $(-1/2, 1)$ and $(1/2, 1)$. Thus by Theorem 5.5.5 the points are stable for the system. The gradient function V has neither a maximum nor a minimum at $(0, 1)$ and hence nothing can be concluded about the stability of $(0, 1)$. We will see in Chapter 6 that $(0, 1)$ is unstable.

Copyright Elsevier 2022

5.6 Exercises

Exercise 5.1. Assume that every eigenvalue of an $n \times n$ constant matrix A has a non-positive real part and that the eigenvalues with zero real part are simple. Suppose that $g : [0, \infty) \to \mathbb{R}^n$ is continuous with

$$\int_0^t |g(s)|ds < \infty.$$

Then show that every solution of

$$x'(t) = Ax(t) + g(t)$$

is bounded on $[0, \infty)$.

Exercise 5.2. Prove (iii) of Theorem 5.1.4.

Exercise 5.3. Examine the stability of the zero solution of the given system.

(a)

$$x'(t) = \begin{pmatrix} 0 & 2 & 1 \\ -2 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} x;$$

(b)

$$x'(t) = \begin{pmatrix} 1 & 2 & 1 \\ 6 & -1 & 0 \\ -1 & -2 & -1 \end{pmatrix} x;$$

(c)

$$x'(t) = \begin{pmatrix} 6 & -1 \\ 5 & 2 \end{pmatrix} x.$$

Exercise 5.4. Show that the zero solution of the non-autonomous system

$$x'(t) = \begin{pmatrix} -1 + te^{-t} & 2 + e^{-t} \\ 2 + \frac{1}{(1+t)^2} & -4 + (1+t)^{\frac{-1}{2}} \end{pmatrix} x$$

is stable.

Exercise 5.5. Assume all the eigenvalues of the matrix A of system (5.2.3) have negative real parts and there is a positive number θ such that $||A(t) - A|| < \theta$ for t is sufficiently large. Show that solutions of (5.2.2) are bounded and its zero solution is asymptotically stable.

Copyright Elsevier 2022

Exercise 5.6. Let A be a constant matrix with $Re(\rho(A)) < 0$. Choose constants $K > 0$ and $\alpha < 0$ such that

$$||e^{At}|| \le Ke^{\alpha t}, \ \ t \ge 0.$$

Let C be a continuous matrix-valued function such that for positive constant β, we have

$$||C(t)|| \le \frac{\beta}{K}, \ \ t \ge 0.$$

Show that the zero solution of

$$x' = [A + C(t)]x(t)$$

is

1. stable if $\beta = -\alpha$,
2. unstable if $\beta > -\alpha$, and
3. asymptotically stable if $0 < \beta < -\alpha$.

Exercise 5.7. a) Let A be a constant matrix with $Re(\rho(A)) < 0$. Let C be a continuous matrix-valued function on $[0, \infty)$. Show that if

$$\int_0^\infty ||C(t)||dt < \infty,$$

then the zero solution of

$$x' = [A + C(t)]x(t)$$

is asymptotically stable.

b) Use the results of part a) to show that the zero solution of

$$u''(t) + \left(\frac{1}{t^2+1} + 1\right)u'(t) + \left(e^{-t} + 1\right)u(t) = 0$$

is asymptotically stable for $t \ge 0$.

Exercise 5.8. Give an example as an application of Theorem 5.3.2.

Exercise 5.9. Show that all trajectories of the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

are periodic and determine their periods.

Copyright Elsevier 2022

Exercise 5.10. Describe the type of stability of the origin of each system and sketch the trajectories.
(a) $x' = x$, $y' = 2x + 2y$;
(b) $x' = -x$, $y' = x - y$;
(c) $x' = -x + 2y$, $y' = x - y$;
(d) $x' = -x + y$, $y' = 2x$;
(e) $x' = 2x - 8y$, $y' = x - 2y$;
(f) $x' = -3x + 2y$, $y' = -2x$.

Exercise 5.11. Consider the damped harmonic motion represented by the second-order differential equation

$$mx'' + ax' + kx = 0,$$

where m, a, and k are positive constants. By changing the equation into a system discuss the nature and stability of the origin in the following cases: $a = 0$; $a^2 - 4mk = 0$; $a^2 - 4mk < 0$; $a^2 - 4mk > 0$. Interpret the results physically.

Exercise 5.12. Decide whether each system is Hamiltonian or not; in the case it is, find the Hamiltonian function. Next, find the equilibrium solutions (if any) and classify their stabilities when possible.
(a) $x'' + x + x^3$;
(b) $\theta'' + \sin(\theta) = 0$;
(c) $x'' + x^2 - 1$;
(d) $x'' + x^2 + 1$.

Exercise 5.13. Decide whether each system is gradient or not; in the case it is, find the gradient function. Next, find the equilibrium solutions (if any) and classify their stabilities when possible.
(a) $x' = -4x(x - 1)(x - 2)$, $y' = -2y$;
(b) $x' = 1 - 2x$, $y' = -7 - 3y$;
(c) $x' = 3 - 2y^2x$, $y' = -2yx^2 - 4$;
(d) $x' = 4x^3 + 4xy$, $y' = 2x^2 + 2y - 1$.

Copyright Elsevier 2022

CHAPTER

Nonlinear systems

6

In general, nonlinear differential equations are important in applications, and their solutions are not known explicitly. Sometimes these equations may be linearized by an expansion process in which nonlinear terms are discarded. When nonlinear terms make vital contributions to the solution, this cannot be done, but sometimes it is enough to retain a few "small" ones. In this chapter, we look for different ways to obtain vital information regarding the behavior of solutions.

6.1 Bifurcations in scalar systems

In this section, we limit our discussion to the scalar differential equation

$$x' = f(x), \tag{6.1.1}$$

where $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable. We assume that x^* is an equilibrium solution of (6.1.1). We always assume that the equilibria are isolated, that is, if x^* is an equilibrium, then there is an open interval containing x^* but no other equilibria.

Fig. 6.1 shows a generic plot of $f(x)$ for $x^* = a, b, c$. The equilibrium solution a is an *attractor* (stable) since arrows to its left and right point toward it. The equilibrium solution b is a *repeller* (unstable) since the arrow to its left and the arrow to its right point away from it. Finally, the equilibrium solution c is *semi-stable* since the arrow to its left points toward it and the one to its right points away from it. In the case of the equilibrium solution a, we may say that it is *locally asymptotically stable* with respect to small perturbation or deviation from the equilibrium solution. To better illustrate the notion, consider the scenario of a population of foxes in a protected refuge that is in a locally asymptotically stable state a. Suppose a weather event caused a small number of the foxes to die. Local asymptotic stability means that the system will eventually return to the original state a. However, if the number of dead foxes is big, then the perturbation is not small, and in this event, there is no guarantee that the population of foxes will return to the original state a. In the case of the population return to state a for all perturbations, no matter how large, the state a is called *globally asymptotically stable*. Such topics will be considered in detail in Chapter 7.

Now we examine the dynamics of (6.1.1) in the neighborhood of its equilibrium solution x^*. If (6.1.1) is in an equilibrium state x^* and $\eta(t)$ represents a small pertur-

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00012-7
Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

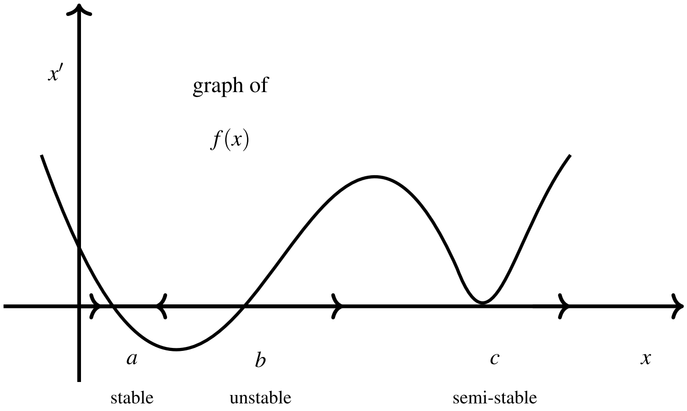


FIGURE 6.1

Generic phase line diagram.

bation from that state, then $x(t) = x^* + \eta(t)$. It follows that $x'(t) = \eta'(t)$. Thus using Taylor series expansion about x^*, we have

$$\begin{aligned} x'(t) = \eta'(t) = f(x^* + \eta) \\ &= f(x^*) + \frac{f'(x^*)}{1!}(x^* + \eta - x^*) + \frac{f''(x^*)}{2!}(x^* + \eta - x^*)^2 + \cdots \\ &= f'(x^*)\eta + \frac{f''(x^*)}{2!}\eta^2 + \cdots \\ &= f'(x^*)\eta + O(\eta^2) + \cdots . \end{aligned}$$

If $f'(x^*) \neq 0$, then the term $|f'(x^*)\eta| \gg |\frac{f''(x^*)}{2!}\eta^2|$. Thus we may neglect the term $O(\eta^2)$ and higher-order terms for that same reason. Then we arrive at the linearization of the system about the equilibrium state x^* given by

$$\eta'(t) = s\eta, \tag{6.1.2}$$

where $s = f'(x^*)$ is the slope of $f(x)$ at x^*. Eq. (6.1.2) has the solution

$$\eta(t) = ce^{st} \tag{6.1.3}$$

for some constant c. It is clear from (6.1.3) that if $s > 0$, then solutions grow exponentially, and the equilibrium solution x^* is unstable. On the other hand, if $s < 0$, then the solutions decay exponentially to zero, and the equilibrium solution x^* is asymptotically stable and hence is stable. If $s = 0$, then nothing can be said about the stability. Note that different stabilities may be obtained for different systems when

Copyright Elsevier 2022

$s = 0$. To see this, we consider the scalar differential equations

$$x' = x^2,\ x' = -x^2,\ x' = x^3,\ \text{ and }\ x' = -x^3.$$

All they share the equilibrium solution $x^* = 0$ and $s = f'(0) = 0$. Consider the first equation with $f(x) = x^2$. Then $f'(x) = 2x$, which is negative for $x < 0$ and positive for $x > 0$. Thus by the previous discussion the arrow to its left points toward $x^* = 0$, and the one to it right points away from it. It follows that $x^* = 0$ is semi-stable. By similar arguments we find that $x^* = 0$ is semi-stable for the second equation, unstable for the third equation, and stable for the fourth equation.

Example 6.1. If $x(t)$ represents a population at time t, $r > 0$ is the growth rate, and $K > 0$ is the carrying capacity, then the logistics model of the population growth is given by

$$x' = rx(1 - \frac{x}{K}). \tag{6.1.4}$$

Setting the right hand of (6.1.4) equal to zero, we arrive at two equilibrium solutions

$$x^* = 0,\ K.$$

Let $f(x) = rx(1 - \frac{x}{K})$. Then $f'(x) = r - \dfrac{2r}{K}x$. It follows from (6.1.2) that

$$s = f'(0) = r\ \text{ and }\ s = f'(K) = -r.$$

We conclude that $x^* = 0$ is unstable and $x^* = K$ is stable.

Although the dynamics of scalar differential equations is simple, since both solutions converge to a constant or become unbounded as $t \to \infty$, they can have some interesting behaviors if solutions depend on some embedded parameters. To be specific, when the parameters are varied and pass some critical values, the systems may experience some abrupt changes or undergo some qualitative changes. These qualitative changes are called *bifurcations*, and the parameter values at which the bifurcations occur are called *bifurcation points*. We are interested in scalar differential equations of the form

$$x' = f(x, \mu), \tag{6.1.5}$$

where f is continuous in both arguments, and $\mu \in \mathbb{R}$ is a parameter.

Thus the system is said to undergo a bifurcation as μ crosses some critical value μ_0 if, in some way, the nature of the solutions changes qualitatively between the regions $\mu < \mu_0$ and $\mu > \mu_0$.

Three most important one-dimensional equilibrium bifurcations are described locally by the following differential equations:

$$\begin{aligned} x' &= \mu - x^2, && \text{saddle-node;} \\ x' &= \mu x - x^2, && \text{transcritical; and} \\ x' &= \mu x - x^3, && \text{pitchfork.} \end{aligned}$$

Copyright Elsevier 2022

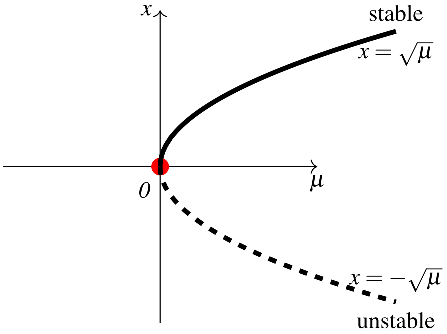


FIGURE 6.2

Bifurcation diagram determined by $x' = \mu - x^2$.

We will study each of these in more detail. We begin by considering *saddle node bifurcation*, which is the basic mechanism for creation and destruction of fixed points. As parameter varies, two fixed points move toward each other, collide, and mutually annihilate, that is, the saddle-node bifurcation results in fixed points being created or destroyed.

Example 6.2. Consider the scalar differential equation with parameter μ

$$x' = \mu - x^2. \tag{6.1.6}$$

Clearly, there are no equilibrium solutions for $\mu < 0$, and $x^* = \pm\sqrt{\mu}$ for $\mu > 0$. It is clear that $s = f'(\sqrt{\mu}) = -2\sqrt{\mu} < 0$ and $s = f'(-\sqrt{\mu}) = 2\sqrt{\mu} > 0$, which imply that the solid bold branch $x = \sqrt{\mu}$ is stable and the dashed branch $x = -\sqrt{\mu}$ is unstable. The bifurcation diagram is shown in Fig. 6.2, and since there is an exchange of stability at $\mu = 0$, we have a bifurcation at $\mu = 0$. This type of bifurcation is known by the names *saddle node bifurcation*, *fold bifurcation, turning point*, and *blue-sky bifurcation*. Note that at the bifurcation point $\mu = 0$, $s = f'(0) = 0$, and hence the linearization term disappears when equilibrium solutions coalesce.

Remark 6.1. In (6.1.6) of Example 6.2, if we change the term $-x^2$ to x^2, then

$$x' = \mu + x^2.$$

There are no equilibrium solutions for $\mu > 0$, and the curve $y = \mu + x^2$ does not intersect the x-axis. As $\mu \searrow 0$, the graph $y = \mu + x^2$ intersects the origin tangentially, and we have one equilibrium solution. As μ becomes negative, two equilibrium solutions appear at $x = \pm\sqrt{-\mu}$. Then, using the stability indicator s, we see that $x = -\sqrt{-\mu}$ is stable and $x = \sqrt{-\mu}$ is unstable. We conclude there is a *saddle node bifurcation* at $(0, 0)$ as indicated in the bifurcation diagram shown in Fig. 6.3.

Next, we discuss a class of bifurcations that occur as stability changes when going from one equilibrium solution to another. This type of bifurcation is called *transcritical bifurcation*.

Copyright Elsevier 2022

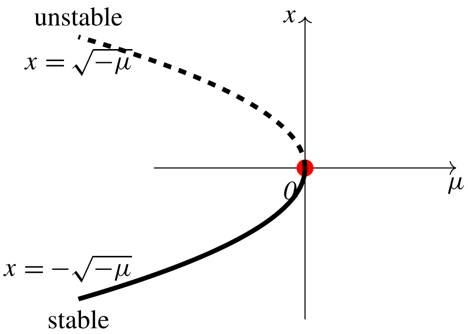
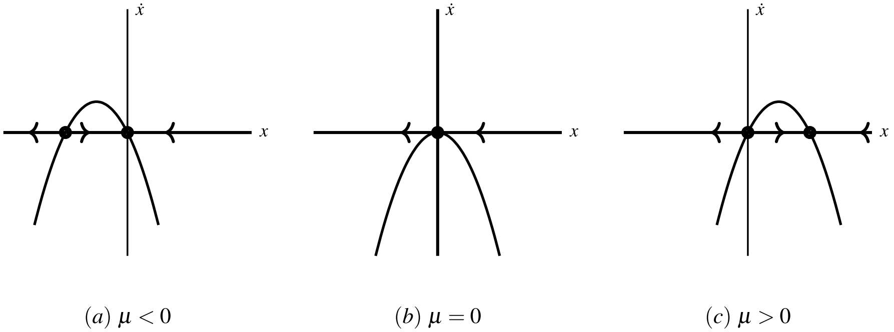


FIGURE 6.3

Bifurcation diagram determined by $x' = \mu + x^2$.



(a) $\mu < 0$ (b) $\mu = 0$ (c) $\mu > 0$

FIGURE 6.4

Phase portraits for different values of μ.

Example 6.3. To illustrate the idea, we consider the population growth model with carrying capacity

$$x' = \mu x - x^2.$$

The system has the equilibrium solutions $x^* = 0, \mu$. Using $s = f'(x^*)$, it follows that $x^* = 0$ is stable if $\mu < 0$ and unstable if $\mu > 0$. Similarly, $x^* = \mu$ is stable for $\mu > 0$ and unstable for $\mu < 0$. Fig. 6.4 shows the phase portraits for different values of μ. Also, Fig. 6.5 shows the bifurcation diagram. The stable branches are shown by bold solid lines, and the unstable branches are shown by dashed lines.

Next, we discuss a type of bifurcation called *pitchfork bifurcation*. The name comes from the shape of the bifurcation diagram that resembles a pitchfork. This type of bifurcation is common in problems that have a *symmetry*, for example, the buckling beam. In the case of pitchfork bifurcation the equilibrium solutions tend to appear and disappear symmetrically, creating a type of bifurcations that are different

Copyright Elsevier 2022

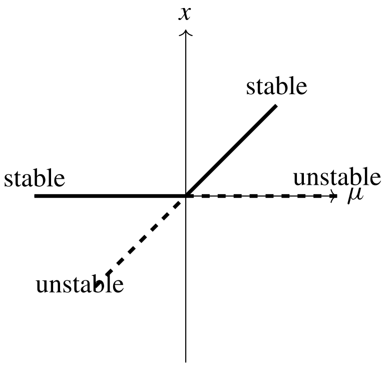


FIGURE 6.5

Bifurcation diagram that shows $(\mu, x) = (0, 0)$ is a transcritical bifurcation point.

from those that we have already discussed. Pitchfork bifurcations can come in one of two types, in the *supercritical bifurcation* and in the *subcritical bifurcation*.

Supercritical bifurcation is typified by equations of the form

$$x' = \mu x - x^3.$$

For $\mu < 0$, the origin is the only equilibrium solution, and it is stable. On the other hand, for $\mu > 0$, the origin is unstable, and the other two equilibrium solutions $\pm\sqrt{\mu}$ are stable.

Similarly, subcritical bifurcation is typified by equations of the form

$$x' = \mu x + x^3.$$

For $\mu < 0$, the origin is stable, and the other two equilibrium solutions $\pm\sqrt{\mu}$ are unstable. On the other hand, for $\mu > 0$, the origin is the only equilibrium solution, and it is unstable.

Example 6.4. We consider the scalar differential equation

$$x' = \mu x - x^3.$$

It is clear that $x^* = 0$ is an equilibrium solution regardless of the sign of μ and $x^* = \pm\sqrt{\mu}$ only if $\mu \geq 0$. Note that for any fixed value of μ, the function $y = \mu x - x^3$ is an odd function in x. Hence critical points will appear and disappear symmetrically with respect to the origin, or the system has the left–right symmetry. Using $s = f'(x^*) = \mu - 3(x^*)^2$, it follows $x^* = 0$ is stable for $\mu \leq 0$ and unstable for $\mu \geq 0$. In addition, $s = f(\pm\sqrt{\mu}) = -2\mu < 0$ implies that the equilibrium solutions $\pm\sqrt{\mu}$ are stable. The bifurcation diagram is displayed in Fig. 6.6.

Fig. 6.6 indicates that two stable equilibrium solutions or branches are created at the bifurcation point $(0, 0)$ and still exist after the bifurcation. A similar remark can be drawn regarding the subcritical bifurcation.

Copyright Elsevier 2022

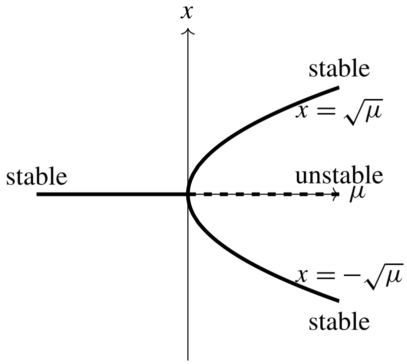


FIGURE 6.6

Bifurcation diagram determined for $x' = \mu x - x^3$; supercritical bifurcation at $(0, 0)$.

The next theorem is the scalar version of the *implicit function theorem*.

Theorem 6.1.1. *For scalar system* (6.1.5)*, suppose that* $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *with continuous first partial derivatives. Suppose* x_0 *is an equilibrium solution at* μ_0*, meaning that* $f(x_0, \mu_0) = 0$*. Assume that*

$$f(x_0, \mu_0) = 0, \quad \frac{\partial f}{\partial x}(x_0, \mu_0) \neq 0.$$

Then there exist $\delta, \varepsilon > 0$ *and a* C^1*-function*

$$\tilde{x} : (\mu_0 - \varepsilon, \mu_0 + \varepsilon) \to \mathbb{R}$$

such that $x = \tilde{x}$ *is the unique solution of*

$$f(x, \mu) = 0$$

with $|x - x_0| < \delta$ *and* $|\mu - \mu_0| < \varepsilon$*.*

For bifurcation theory, we have the following important result.

Corollary 6.1. *Suppose that* $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *has continuous first partial derivatives. Suppose* x_0 *is an equilibrium solution at* μ_0 *for* (6.1.5)*. Then a necessary condition for a solution* (x_0, μ_0) *of* (6.1.5) *to be a bifurcation point of equilibria is that*

$$\frac{\partial f}{\partial x}(x_0, \mu_0) = 0. \tag{6.1.7}$$

The next theorem provides necessary conditions that guarantee the different types of bifurcation points we have discussed so far.

Copyright Elsevier 2022

Theorem 6.1.2. *Suppose that* $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *is continuous in both arguments and has continuous first and second partial derivatives. Suppose that at* (x_0, μ_0)*,* f *satisfies the necessary bifurcation conditions*

$$f(x_0, \mu_0) = 0, \quad \frac{\partial f}{\partial x}(x_0, \mu_0) = 0.$$

(1) If

$$\frac{\partial f}{\partial \mu}(x_0, \mu_0) \neq 0, \quad \frac{\partial^2 f}{\partial x^2}(x_0, \mu_0) \neq 0,$$

then a saddle node bifurcation occurs at (x_0, μ_0).
(2) If

$$\frac{\partial f}{\partial \mu}(x_0, \mu_0) = 0, \quad \frac{\partial^2 f}{\partial x \partial \mu}(x_0, \mu_0) \neq 0, \quad \frac{\partial^2 f}{\partial x^2}(x_0, \mu_0) \neq 0,$$

then a transcritical bifurcation occurs at (x_0, μ_0).
(3) If

$$\frac{\partial f}{\partial \mu}(x_0, \mu_0) = 0, \quad \frac{\partial^2 f}{\partial x^2}(x_0, \mu_0) = 0,$$

$$\frac{\partial^2 f}{\partial x \partial \mu}(x_0, \mu_0) \neq 0, \quad \frac{\partial^3 f}{\partial x^3}(x_0, \mu_0) \neq 0,$$

then a pitchfork bifurcation occurs at (x_0, μ_0)*. In addition,*
(i) if $\frac{\partial^3 f}{\partial x^3}(x_0, \mu_0) < 0$*, then there is a supercritical bifurcation at* (x_0, μ_0)*, and*
(ii) if $\frac{\partial^3 f}{\partial x^3}(x_0, \mu_0) > 0$*, then there is a subcritical bifurcation at* (x_0, μ_0).

Example 6.5. We consider the scalar system

$$x' = \mu x - e^x.$$

Then its bifurcation points must satisfy

$$\mu x - e^x = 0.$$

Plotting the graphs $y = \mu x$ and $y = e^x$, we see that the line $y = x$ is tangent to the curve at $(x, y) = (1, e)$. In other words, if we set $f(x, \mu) = \mu x - e^x$, then $f(1, e) = 0$. This says that $x_0 = 1$ is an equilibrium solution for $\mu = e$. Moreover, $\frac{\partial f}{\partial x}(1, e) = 0$. In addition, $\frac{\partial f}{\partial \mu}(x_0, \mu_0) = 1 \neq 0$, and $\frac{\partial^2 f}{\partial x^2}(x_0, \mu_0) = -e \neq 0$, and so by item 1 of Theorem 6.1.2 there is a saddle node bifurcation at $(1, e)$.

Remark 6.2. Suppose bifurcation for (6.1.5) occurs at (x_0, μ_0). We may translate to the origin $(0, 0)$ by the change of variables $x \to x - x_0$, $\mu \to \mu - \mu_0$. Moreover, if

Copyright Elsevier 2022

$x = \tilde{x}(\mu)$ is a solution branch, then $x \to x - \tilde{x}(\mu)$ maps the branch to $x = 0$. Suppose we have a saddle node bifurcation. For $(0, 0)$ to be a bifurcation point, we must have

$$f(x_0, \mu_0) = 0, \quad \frac{\partial f}{\partial x}(x_0, \mu_0) = 0.$$

Further, suppose that

$$\frac{\partial f}{\partial \mu}(0, 0) = a \neq 0, \quad \frac{\partial^2 f}{\partial x^2}(0, 0) = b \neq 0.$$

Then if in the Taylor expansion of $f(x, \mu)$, we neglect quadratic terms in μ and cubic terms in x, then near $(x, \mu) = (0, 0)$, we obtain

$$\begin{aligned} f(x, \mu) &= \mu \frac{\partial f}{\partial \mu}(0, 0) + x \frac{\partial f}{\partial x}(0, 0) + \frac{x^2}{2} \frac{\partial^2 f}{\partial x^2}(0, 0) + \cdots \\ &= a\mu + bx^2 + \cdots . \end{aligned} \tag{6.1.8}$$

The form $a\mu + bx^2$ in (6.1.8) is called a *normal form* or *representative* for scalar saddle node bifurcations. Thus the original system may be approximated near the origin by

$$x' = a\mu + bx^2.$$

The signs of a and b determine the type of bifurcations and the shape of the bifurcation diagrams as shown in Example 6.2 and Remark 6.1.

We end the section with the following application.

Example 6.6. (*Bucking of a rod*) The potential energy of the system of two rigid rods of length L connected by a torsional spring with spring constant k is given by the function

$$R(x) = \frac{1}{2} k x^2 + 2\lambda L \big(\cos(x) - 1\big),$$

where λ is the strength of compressive force, and x is the angle of the rod to the horizontal axis. The term $\frac{1}{2}kx^2$ is the energy required to compress the spring by the angle $2x$, and $2\lambda L\big(\cos(x) - 1\big)$ is the work done on the system by the external force. Equilibrium solutions satisfy $R'(x) = 0$ or

$$x - \mu \sin(x) = 0, \quad \mu = \frac{2\lambda L}{k},$$

where μ is a dimensionless force parameter. The above equation has the unbuckled equilibrium solution $x = 0$. Set

$$f(x, \mu) = R'(x) = \mu \sin(x) - x.$$

Copyright Elsevier 2022

Then the necessary conditions for an equilibrium bifurcation are

$$\frac{\partial f}{\partial x} = \mu - 1 = 0,$$

which occurs at $\mu = 1$. By finding the Taylor series of f at $(0, 1)$ we arrive at

$$f(x, \mu) = (\mu - 1)x - \frac{1}{6}x^3 + \cdots .$$

We can see from Example 6.4 or verifying the conditions of item 3 of Theorem 6.1.2 that there is a pitchfork bifurcation at $(x, \mu) = (0, 1)$. The bifurcation equilibria near this point are given for $0 < \mu - 1 \ll 1$ by

$$x = \sqrt{6(\mu - 1)} + \cdots .$$

This can also be seen by plotting the graphs $y = x$ and $y = \mu \sin(x)$.

This simple one-dimensional problem is an illustration for the buckling of an elastic beam, one of the first bifurcation problems which was originally studied by Euler in 1757.

For the rest of this chapter, if $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, then $|x| = \sum_{i=1}^{n} |x_i|$, and for an $n \times n$ matrix, $A = a_{ij}$, $|A| = \sum_{i,j=1}^{n} |a_{ij}|$.

6.2 Stability of systems by linearization

Consider the general system of ordinary differential equations

$$\begin{aligned} x_1' &= f_1(x_1, \ldots, x_n), \\ x_2' &= f_2(x_1, \ldots, x_n), \\ &\vdots \\ x_n' &= f_n(x_1, \ldots, x_n). \end{aligned}$$

Using the vector notations

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Copyright Elsevier 2022

and

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix},$$

the above system can be written in the vector form

$$x' = f(x), \tag{6.2.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ with continuously differentiable components. We assume that the vector $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ is a fixed point of (6.2.1).

Define the linear part of f at x^* by $J(x^*)$, where

$$J(x^*) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^*) & \frac{\partial f_1}{\partial x_2}(x^*) & \cdots & \frac{\partial f_1}{\partial x_n}(x^*) \\ \frac{\partial f_2}{\partial x_1}(x^*) & \frac{\partial f_2}{\partial x_2}(x^*) & \cdots & \frac{\partial f_2}{\partial x_n}(x^*) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x^*) & \frac{\partial f_n}{\partial x_2}(x^*) & \cdots & \frac{\partial f_n}{\partial x_n}(x^*) \end{pmatrix}, \tag{6.2.2}$$

which is called the *Jacobian* of f at x^*. Since $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, Taylor's theorem for functions of several variables says that

$$f(x) = J(x^*)(x - x^*) + g(x) \ \ \text{(we have used } f(x^*) = 0\text{)}, \tag{6.2.3}$$

where g is a function that is small in the neighborhood of x^* in the sense that

$$\lim_{x \to x^*} \frac{|g(x)|}{|x - x^*|} = 0. \tag{6.2.4}$$

Note that since $f(x^*) = 0$, (6.2.3) implies that $g(x^*) = 0$. Our hope is to prove that if $Re(\rho(J(x^*))) < 0$ or the origin for

$$y' = J(x^*)y$$

is asymptotically stable, then the equilibrium point x^* of (6.2.1) is asymptotically stable. Note that by introducing the transformation $y = x - x^*$ we may assume without loss of generality that the fixed point is at the origin. Thus we have the following general theorem.

Theorem 6.2.1. *Assume that* $Re(\rho(A)) < 0$ *where* A *is an* $n \times n$ *constant matrix. Let* U *be an open set of* $\mathbb{R} \times \mathbb{R}^n$ *that contains* $[0, \infty) \times B_r(0)$, *where* $B_r(0) = \{x \in \mathbb{R}^n : |x - 0| \le r\}$ *for positive constant* r. *Assume that* $g : U \to \mathbb{R}^n$ *is continuous in both arguments and Lipschitz in the second argument such that*

$$\lim_{x \to 0} \frac{|g(t, x)|}{|x|} = 0 \ \textit{uniformly for} \ \ t \in [0, \infty) \tag{6.2.5}$$

Copyright Elsevier 2022

with $g(t,0)=0$. *Then the origin for the perturbed nonlinear system*

$$x' = Ax + g(t,x) \tag{6.2.6}$$

is asymptotically stable.

Proof. Since $Re(\rho(A)) < 0$, we can find constants $K > 0$ and $\xi < 0$ such that

$$||e^{At}|| \le Ke^{\xi t}, \ \ t > 0.$$

Let $x(t)$ be a solution of (6.2.6) on some interval (a,b) containing 0 with $x(0)=x_0$. Then the variation of parameters formula (3.3.6) yields

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-s)}g(s,x(s))ds. \tag{6.2.7}$$

Taking the appropriate norms of both sides of (6.2.7) and estimating give, for $t \ge 0$,

$$|x(t)| = K|x_0|e^{\xi t} + \int_0^t Ke^{\xi(t-s)}|g(s,x(s))|ds, \ \ t \in [0,b).$$

Multiplying both sides by $e^{-\xi t}$, we arrive at the expression

$$|x(t)|e^{-\xi t} = K|x_0| + \int_0^t Ke^{-\xi s}|g(s,x(s))|ds, \ \ t \in [0,b).$$

Chose $\eta > 0$ such that $\eta K < -\xi$. Thus by (6.2.5) we have that

$$|g(t,x)| \le \eta|x| \ \text{ for } (t,x) \in [0,\infty) \times \bar{B}_r(0).$$

Set $\delta^* = \delta/(2K+1) < \delta$. Suppose that $|x_0| < \delta^*$, and let (a,b) be the maximal interval of existence with $x(0)=x_0$.

Let $[0,c] \subseteq (a,b)$ be an interval such that $|x(t)| \le \delta$ for all $t \in [0,c]$. This is possible due to the continuity of x. By the above inequality of g we have that

$$|g(t,x)| \le \eta|x| \ \text{ for } t \in [0,c].$$

Then we have

$$|x(t)|e^{-\xi t} = K|x_0| + \int_0^t Ke^{-\xi s}\eta|x(s)|ds, \ \ t \in [0,c].$$

Use Gronwall's inequality and then evaluate the integral to get

$$\begin{aligned} |x(t)|e^{-\xi t} &\le K^2|x_0|\eta\frac{1}{K\eta}[e^{K\eta t}-1] \\ &= K|x_0|e^{K\eta t}. \end{aligned}$$

Copyright Elsevier 2022

Since $|x_0| < \delta^*$, $K|x_0| < \delta$, and so $|x(t)|e^{-\xi t} \le \delta e^{K\eta t}$. It follows that

$$|x(t)| \le \delta e^{(\xi + K\eta)t}, \quad t \in [0, c], \tag{6.2.8}$$

where $\xi + K\eta < 0$.

Of course, we only showed that inequality (6.2.8) holds for $|x(t)| \le \delta$ on $[0, c]$. It is left as an exercise to show that the inequality holds on $[0, \infty)$. This completes the proof. □

Remark 6.3. (1) If the origin is stable for the linear system $x' = Ax$ but not asymptotically stable, then it is not necessary that the origin is stable for the nonlinear system (6.2.6).
(2) If the origin is unstable for the linear system $x' = Ax$, then it is unstable for the nonlinear system (6.2.6).

Remark 6.4. Let $D \subset \mathbb{R}^n$ containing the origin. Consider the autonomous system

$$x' = f(x), \tag{6.2.9}$$

where $f : D \to \mathbb{R}^n$ with continuously differentiable components and $f(0) = 0$. Let $A = J(0)$ be the Jacobian matrix of f at $x = 0$. Rewrite (6.2.9) as

$$x' = Ax + g(x), \tag{6.2.10}$$

where $g(x) = f(x) - Ax$. The differentiability of f implies $g(x) = o(|x|)$ as $x \to \infty$. Therefore $g(x)$ satisfies condition (6.2.5).

The next theorem provides results concerning the UAS of the zero solution. Its proof is left as an exercise.

Theorem 6.2.2. *Assume that $Re(\rho(A)) < 0$, where A is an $n \times n$ constant matrix (or the zero solution of the linear part is AS). Assume that a continuous function $g(t, x)$ satisfies Lipschitz or local Lipschitz condition in the second argument for $x \in D$. Suppose that*

$$\lim_{x \to 0} \frac{|g(t, x)|}{|x|} = 0 \text{ uniformly for } t \in [0, \infty) \tag{6.2.11}$$

with $g(t, 0) = 0$. Then the origin for the perturbed nonlinear system

$$x' = Ax + g(t, x)$$

is uniformly asymptotically stable.

We have the following example.

Example 6.7. The zero solution of the nonlinear non-autonomous system

$$x' = -x + \frac{t}{t+1}x^2$$

Copyright Elsevier 2022

is UAS. It is obvious that the zero solution of the linear part $x' = -x$ is AS. We must show that condition (6.2.11) holds. Let $g(t,x) = \frac{t}{t+1}x^2$. Then g is continuous, and for any open subset $D = (\alpha, \beta) \subset \mathbb{R}$ containing the origin, we have

$$|g(t,x) - g(t,y)| \le 2\beta|x-y| \text{ for all } \ (t,x),(t,y) \in [0,\infty) \times D.$$

Moreover,

$$\lim_{x\to 0} \frac{|g(t,x)|}{|x|} \le \lim_{x\to 0} \frac{x^2}{|x|} = 0 \text{ uniformly for } t \in [0,\infty).$$

Thus by Theorem 6.2.2 the zero solution is UAS.

The next example shows the needs for a different approach when the eigenvalues of the matrix A fail to have negative real parts.

Example 6.8. Consider the planar autonomous system

$$x' = -y + x\sqrt{x^2+y^2}, \ \ y' = x + y\sqrt{x^2+y^2}.$$

It is clear that (0, 0) is the only equilibrium point. We rewrite the system in the form

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x\sqrt{x^2+y^2} \\ y\sqrt{x^2+y^2} \end{pmatrix}.$$

The eigenvalues of

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

are $\pm i$. Thus Theorem 6.2.1 does not apply. Instead, we use the polar coordinates. Namely,

$$x = r\cos(\theta), \ y = r\sin(\theta), \ \text{ where } r^2 = x^2 + y^2,$$

and $r = r(t)$ and $\theta = \theta(t) = \tan^{-1}(\frac{y}{x})$. Thus

$$\frac{d}{dt}(x^2+y^2) = \frac{d}{dt}r^2$$

or

$$xx' + yy' = rr'.$$

A substitution of x' and y' yields

$$\begin{aligned} rr' &= x(-y + x\sqrt{x^2+y^2}) + y(x + y\sqrt{x^2+y^2}) \\ &= (x^2+y^2)\sqrt{x^2+y^2} \\ &= r^2 r. \end{aligned}$$

Copyright Elsevier 2022

Thus we have the first-order differential equation

$$r' = r^2, \tag{6.2.12}$$

which implies that the function r is increasing. Thus trajectories are spiraling away from the origin, and hence the origin is unstable. In addition,

$$\theta' = \frac{xy' - x'y}{x^2 + y^2} = 1.$$

Choosing $\theta(0) = 0$, we have $\theta(t) = t$, which indicates trajectories spiral away from the origin counterclockwise. Note that using separation of variables, (6.2.12) can be easily solved to find $r(t) = \frac{1}{k-t}$ for some suitable constant k. Notice that $r(t) \to \infty$ as $t \to k^-$.

Example 6.9. In Example 3.5, we considered a pendulum with friction and arrived at the planar autonomous system

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -\frac{g}{L}\sin(x_1) - \frac{b}{m}x_2. \end{aligned} \tag{6.2.13}$$

It is clear that system (6.2.13) has infinitely many equilibrium points $(n\pi, 0)$, $n = 0, 1, \dots,$ which correspond to vertical positions with zero velocity. We will only analyze two equilibrium points $(0, 0)$ and $(\pi, 0)$. For the equilibrium point $(0, 0)$, we write system (6.2.13) in the form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -\frac{g}{L} & -\frac{b}{m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{g}{L} \begin{pmatrix} 0 \\ x_1 - \sin(x_1) \end{pmatrix}.$$

The linear part has the eigenvalues

$$\lambda_1 = -\frac{b}{2m} + \frac{1}{2}\Big(\frac{b^2}{m^2} - 4\frac{g}{L}\Big)^{1/2}$$

and

$$\lambda_1 = -\frac{b}{2m} - \frac{1}{2}\Big(\frac{b^2}{m^2} - 4\frac{g}{L}\Big)^{1/2}.$$

If $\frac{b^2}{m^2} - 4\frac{g}{L} < 0$ (when damping is weak), then the eigenvalues are complex with negative real parts, and the origin is a stable spiral. If $\frac{b^2}{m^2} - 4\frac{g}{L} \geq 0$ (when the damping is strong), that is,

$$\frac{b^2}{m^2} > 4\frac{g}{L},$$

Copyright Elsevier 2022

we have that $\lambda_2 \leq \lambda_1 < 0$, and hence the equilibrium solution $(0,0)$ for system (6.2.13) is asymptotically stable by Theorem 6.2.1.

The physical and geometrical interpretation for $(x_1, x_2') = (0,0)$ being a stable spiral corresponds to the pendulum hanging vertically downward ($x_1 = 0$) with zero velocity ($x_2' = 0$), and thus, when the damping is weak, a perturbation will cause the pendulum to spiral and approach the vertically downward position.

For the equilibrium point $(\pi, 0)$, we use the transformation

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \pi \\ 0 \end{pmatrix}$$

in (6.2.13), obtaining

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} y_2 \\ \frac{g}{L}\sin(y_1) - \frac{b}{m}y_2 \end{pmatrix}.$$

To examine this system near $(0,0)$, we write the above new nonlinear system in y in the form

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ \frac{g}{L} & -\frac{b}{m} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \frac{g}{L} \begin{pmatrix} 0 \\ y_1 - \sin(y_1) \end{pmatrix}.$$

The linear parts have the eigenvalues

$$\lambda_1 = -\frac{b}{2m} + \Big(\frac{b^2}{m^2} + 4\frac{g}{L}\Big)^{1/2},$$
$$\lambda_1 = -\frac{b}{2m} - \Big(\frac{b^2}{m^2} + 4\frac{g}{L}\Big)^{1/2}.$$

It is clear that $\lambda_2 < 0 < \lambda_1$ and the equilibrium point $(0,0)$ and hence $(\pi, 0)$ are unstable.

Note that we could calculate the Jacobian given by (6.2.2) at $(\pi, 0)$ and obtain

$$J((\pi,0)) = \begin{pmatrix} 0 & 1 \\ \frac{g}{L} & -\frac{b}{m} \end{pmatrix},$$

instead of using transformation.

If we compute $J((\pm n\pi, 0))$, then we will see that $(\pm n\pi, 0)$ are stable for $n = 2, 4, \ldots$ and unstable for $n = 1, 3, \ldots$.

Example 6.10. (*Lotka–Volterra predator–prey model*) In Example 4.1 we considered the Lotka–Volterra system

$$\begin{aligned} \frac{dx}{dt} &= ax - bxy, \\ \frac{dy}{dt} &= -cy + dxy \end{aligned} \tag{6.2.14}$$

Copyright Elsevier 2022

and found that the equilibrium solutions are

$$x_1^* = (0, 0) \text{ and } x_2^* = (c/d, a/b).$$

The first equilibrium solution x_1^* effectively represents the extinction of both species. If both populations are at 0, then they will continue to be so indefinitely. The second equilibrium solution x_2^* represents a fixed point at which both populations sustain their current, non-zero numbers, and do so indefinitely. The levels of population at which this equilibrium is achieved depend on the chosen values of the parameters a, b, c, and d. Now we will further discuss the stability of both equilibrium solutions x_1^* and x_2^* and hope that x_1^* is unstable and x_2^* is stable. The Jacobian matrix is

$$J(x, y) = \begin{pmatrix} a - by & -bx \\ dy & -c + dx \end{pmatrix}.$$

For $x_1^* = (0, 0)$, we have

$$J(0, 0) = \begin{pmatrix} a & 0 \\ 0 & -c \end{pmatrix}$$

with eigenvalues $a > 0$ and $-c < 0$. Thus the equilibrium point $(0, 0)$ is a saddle and hence unstable.

For $x_2^* = (c/d, a/b)$,

$$J(c/d, a/b) = \begin{pmatrix} 0 & -\frac{bc}{d} \\ \frac{da}{b} & 0 \end{pmatrix},$$

which has purely imaginary eigenvalues. Hence linearization is inconclusive. Note that the use of polar coordinates will not work here in the sense that it is hard, if not impossible, to obtain any information from it. We look at the vector fields surrounding the equilibrium point by dividing the region of interest into four quadrants as depicted in Fig. 6.1. We do so by drawing vertical and horizontal lines through the equilibrium point $(c/d, a/b)$. Those lines are called the x and y *nullclines*. We begin by writing system (6.2.14) in the form

$$x' = x(a - by), \quad y' = y(-c + dx). \tag{6.2.15}$$

The x nullclines for system (6.2.15) where $x' = 0$ are $x = 0$ and $y = \frac{a}{b}$. Trajectories cross these two lines vertically. Similarly, trajectories cross the y nullclines where $y' = 0$ horizontally. We see that the equilibria are the intersections of the x and y nullclines. On the ray to the left of the equilibrium and below the line $y = a/b$, we have $x < \frac{c}{d}$ and $y = \frac{a}{b}$. We know that the vector field is vertical, and so we only need to check the sign of y'. From (6.2.15) we have

$$y' = y(-c + dx) = \frac{a}{b}(-c + dx) < \frac{a}{b}(-c + d\frac{c}{d}) = 0.$$

Copyright Elsevier 2022

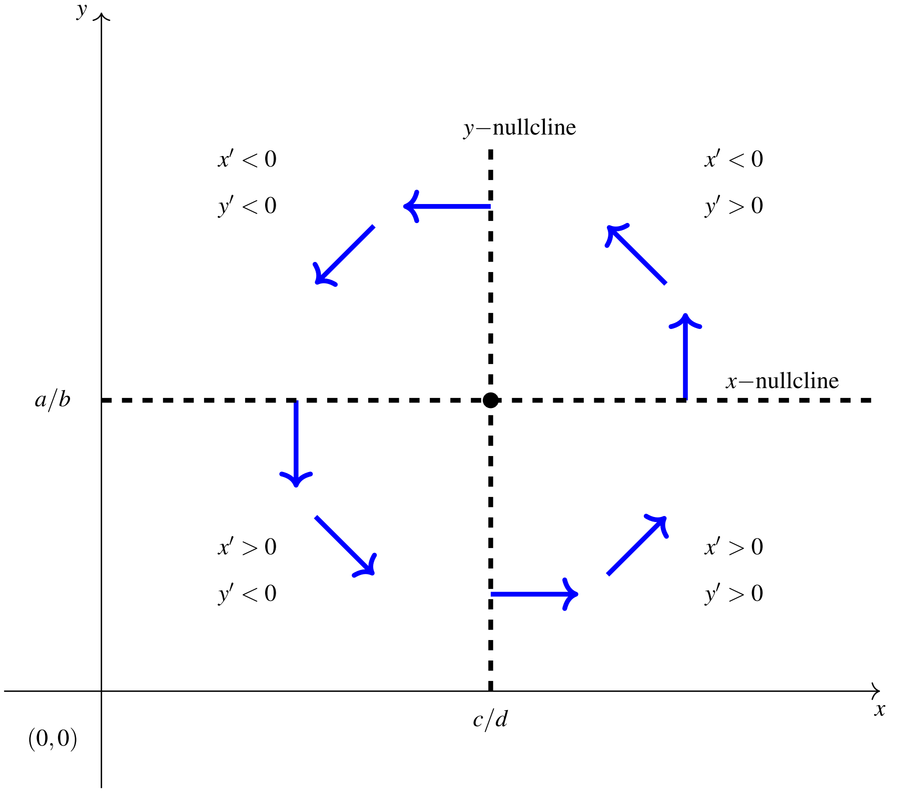


FIGURE 6.7

Nullclines and vector field.

Thus in that region, we have that $x' > 0$ and $y' < 0$. Similarly, if we consider the region to the left of the vertical line $x = \frac{c}{d}$ and above the horizontal line $y = c/d$, then we have that $x = \frac{c}{d}$ and $y > \frac{a}{b}$. Again, the vector field is horizontal, and we only need to check the sign of x'. From (6.2.15) we have

$$x' = x(a - by) < \frac{c}{d}(a - b\frac{a}{b}) = 0.$$

Thus in that particular region, we have $x' < 0$, $y' < 0$. This task can be easily performed for the other two regions. The complete vector field around the equilibrium is displayed in Fig. 6.7. To obtain the trajectories, we write the system in the form

$$\frac{dy}{dx} = \frac{y(-c + d\,x)}{x(a - by)}.$$

Copyright Elsevier 2022

After rearranging the terms, we get

$$\frac{(a - by)}{y} dy = \frac{(-c + d\,x)}{x} dx.$$

Simplifying and then integrating both sides yield

$$a\mathrm{Ln}(y) - by = -c\mathrm{Ln}(x) + d\,x + K \text{ for some constant } K.$$

Exponentiating both sides gives

$$y^a e^{-by} = x^{-c} e^{d\,x} e^K.$$

We consider the y nullcline where the value of x is fixed at c/d. We choose a specific trajectory by assigning a fixed value for K. Then the right side of the above relation is positive β, so that

$$y^a = \beta e^{by}. \tag{6.2.16}$$

The left and right sides of (6.2.16) are the power and exponential functions, respectively. Therefore if we plot them on the same graph, then there can be at most two crossings, or Eq. (6.2.16) can have at most two solutions for y. Thus along the y nullcline, there can be at most two crossings. That means that an orbit or trajectory cannot spiral into or out from an equilibrium point. Otherwise, many values of y would be required, which is not the case here. This determines that the equilibrium point $(c/d, a/b)$ is encircled with periodic trajectories, and hence it is stable. Note that the periodic trajectories encircling the equilibrium point $(c/d, a/b)$ indicate that a small perturbation from the equilibrium solution puts the populations on a periodic orbit that stays near the equilibrium and the system never returns to equilibrium. A phase diagram is shown in Fig. 6.8.

6.3 An SIR epidemic model

Here is the perfect place to apply what we have studied so far to an SIR model, where

$S = S(t)$ is the number of susceptible individuals (not yet having the disease but capable of getting it)
$I = I(t)$ is the number of infected individuals, and
$R = R(t)$ is the number of recovered or removed individuals.

We assume that the total population is constant and everyone is susceptible. Then

$$S + I + R = N \text{ (constant)},$$

that is,

$$S + I + R = 1 \text{ (by normalization)}.$$

Copyright Elsevier 2022

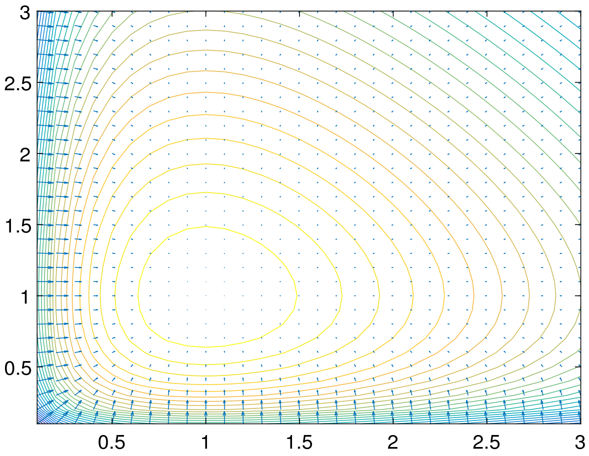


FIGURE 6.8

Prey–predator dynamics as described by the level curves. The arrows describe the velocity and direction of solutions. We chose $a = b = c = d = 1$.

Let

$\mu =$ birth rate $=$ death rate (because the number of people is constant),
$\gamma =$ removal rate, and
$\lambda =$ infection rate.

We assume that all newborns are susceptible and arrive at the epidemic model

$$\begin{cases} S'(t) = (S + I + R)\mu - \mu S - (SI)\lambda, \\ I'(t) = -\mu I - \gamma I + \lambda SI, \\ R'(t) = -\mu R + \gamma I. \end{cases}$$

We have $(S + I + R)\mu = \mu$, since the population is constant and normalized at 1. The term of μS represents the death of susceptible, and the negative implies loss or removal from being susceptible. On the other hand, the term $(SI)\lambda$ denotes those who got infected, and the minus represents that they are infected and not susceptible anymore. Similarly, the term μI represents those who die from the infection, and the minus indicates a loss for those who are infected already. In addition, the term γI indicates those who are removed from the population due to beating the epidemic or got vaccinated. Finally, the term μR denotes those who are removed due to death. Keep in mind that *minus* means loss from a certain class for S, I, and R, and *plus* implies gain for the same classes. Thus by only analyzing $S \geq 0$ and $I \geq 0$ we arrive at the 2×2-planar system

$$\begin{cases} S' = \mu - \mu S - \lambda SI, \\ I' = -(\mu + \gamma)I + \lambda SI. \end{cases} \tag{6.3.1}$$

Copyright Elsevier 2022

Note that system (6.3.1) is nonlinear and $(0,0)$ is not an equilibrium solution. Also, the system is valid in the region $0 \le S + I \le 1$ since we are neglecting R. It is important to note that the system does not allow those who recover from the epidemic to become susceptible again. Suppose $S(t_0) = S_0$ and $I(t_0) = I_0$. Since the right side of (6.3.1) is continuous with continuous first partial derivatives, we know that the solutions $S(t, t_0, S_0)$ and $I(t, t_0, I_0)$ exist and are unique. Moreover, $S(t, t_0, S_0)$ and $I(t, t_0, S_0)$ are continuous and differentiable with respect to all variables. Next, we examine how big the interval on which the solutions exist. Due to our constraints, we are interested in the triangular region of existence

$$T = \{(S, I) : S \ge 0, I \ge 0,\ S + I < 1\}.$$

Along $S = 0$, we have $S' = \mu > 0$, $0 \le I < 1$, and $I' = -(\mu + \gamma)I < 0$ if $I > 0$. Moreover,

$$\frac{dI}{dS} = \frac{I'}{S'} < 0.$$

Thus any trajectory that starts on the I-axis moves downward inside T in the positive direction of S.

Now, along $I = 0$, we have $S' = \mu(\mu - S) > 0$, $0 \le I < 1$, and $I' = 0$. Thus any trajectory that starts on the S-axis moves toward the point $(1, 0)$, at which the line $S + I = 1$ intersects the S-axis. Note that $(1, 0)$ is an equilibrium solution.

Along the line $S + I = 1$, we have

$$(S + I)' = S' + I' = \mu - \mu(S + I) - \gamma I,$$

and since $S + I = 1$, this expression reduces to

$$(S + I)' = -\gamma I < 0.$$

This tells us that any trajectory starting at the line $S + I = 1$ moves to the left of the line toward the inside of the triangle T. We conclude that the triangular region T is positively invariant, that is, if $(S_0, I_0) \in T$, then $S(t, t_0, S_0),\ I(t, t_0, I_0) \in T$ for all $t \ge t_0$. Thus the maximum interval of existence for solutions in T is $[t_0, \infty)$. From now on we refer to the triangular region T as the feasible region. Next, we search for the equilibrium solutions (S^*, I^*) in the feasible region. To do so, we must solve for S and I by setting the right sides of (6.3.1) to zero, that is, $S' = 0$, and $I' = 0$:

$$\mu - \mu S - \lambda SI = 0, \text{ and } -(\mu + \gamma)I + \lambda SI = 0.$$

From the second equation we obtain

$$\big(-\mu - \gamma + \lambda S\big)I = 0.$$

Thus if $I \ne 0$, then we arrive at

$$S^* = \frac{\mu + \gamma}{\lambda}.$$

Copyright Elsevier 2022

Clearly, $S^* \geq 0$, and in order for S^* to be in the feasible region, we must have

$$\mu + \gamma < \lambda.$$

Substituting $S^* = \frac{\mu+\gamma}{\lambda}$ into the equation $\mu - \mu S^* - \lambda S^* I^* = 0$, we arrive at

$$I^* = \frac{\mu}{\mu+\gamma} - \frac{\mu}{\lambda}.$$

It is clear $I^* > 0$, since

$$I^* = \frac{\mu}{\mu+\gamma} - \frac{\mu}{\lambda} > \frac{\mu}{\lambda} - \frac{\mu}{\lambda} = 0.$$

Next, we show that (S^*, I^*), belongs to the feasible region:

$$\begin{aligned} 0 < S^* + I^* &= \frac{\mu+\gamma}{\lambda} + \frac{\mu}{\mu+\gamma} - \frac{\mu}{\lambda} \\ &= \frac{\gamma}{\lambda} + \frac{\mu}{\mu+\gamma} \\ &< \frac{\gamma}{\mu+\gamma} + \frac{\mu}{\mu+\gamma} \\ &= 1. \end{aligned}$$

Thus the two equilibrium solutions $(1, 0)$ and (S^*, I^*) belong to the feasible region T. Next, we use Theorem 6.2.1 to study the stability of the two equilibrium solutions. We do not want solutions to approach $(1, 0)$ and stay there. Instead, we rather have that solutions approach (S^*, I^*), that is, we hope for $(1, 0)$ to be unstable.

We begin with $(1, 0)$. The Jacobian matrix is

$$J(S, I) = \begin{pmatrix} -\mu - \lambda I & -\lambda S \\ \lambda I & -(\mu+\gamma) + \lambda S \end{pmatrix}.$$

Thus

$$J(1, 0) = \begin{pmatrix} -\mu & -\lambda \\ 0 & -(\mu+\gamma) + \lambda \end{pmatrix}$$

with two eigenvalues

$$d_1 = -\mu \text{ and } d_2 = -(\mu+\gamma) + \lambda.$$

Clearly, $d_1 < 0$, and $d_2 > 0$ as long as $\mu + \gamma < \lambda$. Thus by Theorem 6.2.1 the equilibrium solution $(1, 0)$ is unstable. Recall that if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Copyright Elsevier 2022

is a nonsingular matrix, then the characteristic equation is

$$\lambda^2 - (a+d)\lambda + ad - bc = 0.$$

By introducing

$$\begin{aligned} p &= a+d, \\ q &= ad - bc, \end{aligned}$$

the characteristic equation becomes

$$\lambda^2 - p\lambda + q = 0.$$

The roots of this equation are

$$\lambda_{1,2} = \frac{1}{2}\left[p \pm \sqrt{p^2 - 4q}\right]. \tag{6.3.2}$$

Now the Jacobian at (S^*, I^*) is

$$J(S^*, I^*) = \begin{pmatrix} -\mu - \lambda I^* & -\lambda S^* \\ \lambda I^* & 0 \end{pmatrix}.$$

Using (6.3.2), we arrive at the two eigenvalues

$$d_{1,2} = \frac{-(\mu + \lambda I^*) \pm \sqrt{(\mu + \lambda I^*)^2 - 4\lambda^2 S^* I^*}}{2}$$

with negative real parts. Hence by Theorem 6.2.1 the equilibrium solution (S^*, I^*) is asymptotically stable.

6.4 Limit cycle

A central issue in the theory of nonlinear systems in differential equations is to determine the existence of closed orbits or limit cycles. Those closed orbits or limit cycles represent periodic solutions of the system. We consider the two-dimensional autonomous system

$$x' = P(x, y), \;\; y' = Q(x, y), \tag{6.4.1}$$

where P and Q are continuous on some subset of $\mathbb{R}^2$. Systems given by (6.4.1) may admit solutions that commonly tend to a limiting finite periodic solution. This limiting closed curve in the phase plane is called a *limit cycle*.

Definition 6.4.1. A limit cycle of system (6.4.1) is a closed curve, and no other solution that is a closed curve exists in its neighborhood.

Copyright Elsevier 2022

Note that every neighboring trajectory spirals and approaches the limit cycle from the inside or from the outside as $t \to \infty$ or as $t \to -\infty$.

Definition 6.4.2. If every neighboring trajectory approaches a limit cycle as $t \to \infty$ or as $t \to -\infty$, then the limit cycle is said to be *stable*.

It is hard, if not impossible, to determine if a nonlinear system admits a limit cycle. However, the next result, known as Bendixson–Dulac criteria should provide a criterion for the *nonexistence* of limit cycles or closed trajectories of the system given by (6.4.1).

Theorem 6.4.1. *(Bendixson–Dulac theorem) Suppose the first partial derivatives of $P(x, y)$ and $Q(x, y)$ are continuous on some connected subset U in $\mathbb{R}^2$. If*

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}$$

is not identically zero and has the same sign in any open set of U, then system (6.4.1) has no closed trajectory in U.

Proof. Suppose the contrary, that is, there is a closed curve C in U. Then by Green's theorem we have

$$\int_C P(x, y)dy - Q(x, y)dx = \int\int_R (\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y})dxdy, \qquad (6.4.2)$$

where R is the region enclosed by C. Let $x = x(t)$, $y = y(t)$ be a parameterization of C. Then

$$\int_C P(x, y)dy - Q(x, y)dx = \int_0^T (P\frac{dy}{dt} - Q\frac{dx}{dt})dt,$$

where T is the period of C. Using (6.4.1), we arrive at

$$\int_C P(x, y)dy - Q(x, y)dx = \int_0^T (PQ - QP)dt = 0.$$

Thus from (6.4.2) we have

$$\int\int_R (\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y})dxdy = 0,$$

which can only hold if $\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}$ is zero or changes sign, which is a contradiction. Thus C cannot be close, and the system does not admit a limit cycle. This completes the proof. □

Example 6.11. Let g and h be differentiable functions. Consider the two-dimensional system

$$x' = g(y), \quad y' = (1 + x^2)y + h(x).$$

Copyright Elsevier 2022

Then

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} = 1 + x^2 > 0, \quad x \in \mathbb{R},$$

and by Theorem 6.4.1 the system does not admit a periodic solution.

The next theorem generalizes Theorem 6.4.1. For its proof, we refer the reader to Exercise 6.20.

Theorem 6.4.2. *(Generalization of Bendixson–Dulac theorem) Suppose the first partial derivatives of* $P(x, y)$*,* $Q(x, y)$*, and* $h(x, y)$ *are continuous on some connected subset* U *in* $\mathbb{R}^2$*. If*

$$\frac{\partial (hP)}{\partial x} + \frac{\partial (hQ)}{\partial y}$$

is not identically zero and of one sign in any open set of U *in* $\mathbb{R}^2$*, then system* (6.4.1) *cannot have a closed trajectory in* U*.*

We have the following example as an application.

Example 6.12. For constants a, b, c, k, consider the two-dimensional system

$$x' = (y - a)x, \quad y' = bx + cy + ky^2.$$

Then

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} = y - a + c + 2ky,$$

and hence Theorem 6.4.1 is not applicable. Let us try Theorem 6.4.2 with $h(x, y) = x^{-(2k+1)}$. Then

$$\frac{\partial (hP)}{\partial x} + \frac{\partial (hQ)}{\partial y} = (c + 2ka)x^{-(2k+1)} \neq 0$$

if $c + 2ka \neq 0$, and by Theorem 6.4.2 the system does not admit a periodic solution in the half-plane $x > 0$ or $x < 0$.

In fact, we can do better. If $x = 0$, then starting from any point on the y-axis, the equation has a solution that lies on the y-axis. Thus we conclude that the equation has no closed cycle or trajectory and hence a periodic solution in $\mathbb{R}^2$, since otherwise the uniqueness will be violated.

So far we have discussed results concerning the nonexistence of periodic solutions, or periodic trajectories, or limit cycles. Next, we look at the existence of limit cycles.

Theorem 6.4.3. *A closed orbit or limit cycle of* (6.4.1) *surrounds at least one equilibrium point of the system. In other words, if no equilibrium point of the system exists in a region, then there can be no periodic orbits in that region.*

Copyright Elsevier 2022

Let (0, 0) be the only equilibrium solution of (6.4.1). Moreover, suppose any solution of (6.4.1) that starts near the origin will spiral away from the origin. Since there are no equilibrium solutions, our first thought might be that all solutions of (6.4.1) have trajectories that spiral out to infinity. However, the next theorem, due to Poincaré and Bendixson, presents us a different scenario. It shows, under certain conditions, the existence of a limit cycle. Its proof requires high-level arguments, and we omit it.

Theorem 6.4.4. *(Poincaré–Bendixson) Suppose R is a closed bounded region of the plane enclosed by two simple curves μ_1 and μ_2. If R is free of equilibrium solutions of* (6.4.1) *and there is a trajectory C confined to R, then either C is a limit cycle, or it spirals toward a limit cycle.*

The main point of the theorem is that if we find one solution that exists for all t large enough (that is, we let t go to infinity) and stays within a bounded region, then we have found either a periodic orbit or a solution that spirals toward a limit cycle; that is, in the long term, the behavior will be very close to a periodic function. We should take the theorem more as a qualitative statement rather than something to help us in computations. In practice, it is hard to find solutions and therefore hard to show rigorously that they exist at all times. Another caveat to consider the theorem is that it only works in two dimensions. In three dimensions and higher, there is simply too much room.

As an application, we provide the following example.

Example 6.13.

$$x' = y + x[1 - x^2 - y^2], \;\; y' = -x + y[1 - x^2 - y^2]. \tag{6.4.3}$$

It is clear that (0, 0) is the only equilibrium solution of (6.4.3). Let us pass to polar coordinates. Namely,

$$x = r\cos(\theta), \;\; y = r\sin(\theta), \;\; \text{where } r^2 = x^2 + y^2,$$

and $r = r(t)$ and $\theta = \theta(t) = \tan^{-1}(\frac{y}{x})$. Thus

$$\frac{d}{dt}(x^2 + y^2) = \frac{d}{dt}r^2$$

or

$$xx' + yy' = rr'.$$

A substitution of x' and y' yields

$$r' = r(1 - r^2),$$

$$\theta' = \frac{xy' - x'y}{x^2 + y^2} = 1.$$

Copyright Elsevier 2022

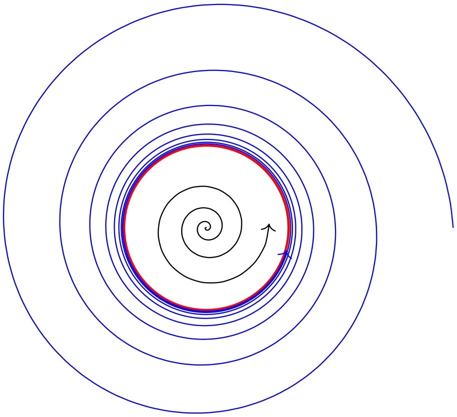


FIGURE 6.9

Limit cycle.

For constants c and k, the solutions are given by

$$r(t) = \frac{1}{\sqrt{1 + ce^{-2t}}},\ c > -1,\ \theta(t) = -t + k,\ t \geq 0.$$

For $c = 0$, we have that $r = 0$ and $(x(t), y(t)) = (\cos(t), -\sin(t))$ is a periodic solution with trajectory given by the unit circle in red (light gray in print version) in Fig. 6.9. For $c > 0$, we have $0 < r(t) < 1$, and as t increases, $r \nearrow 1$, and the corresponding solution approaches the unit circle from inside of the circle in a spiral way as $t \to \infty$, as indicated in Fig. 6.9. For $-1 < c < 0$, we have $r(t) > 1$, and as t increases, $r \searrow 1$, and the corresponding solution approaches the unit circle from outside of the circle in a spiral way as $t \to \infty$, as indicated in Fig. 6.9. Therefore the unit circle $x^2 + y^2 = 1$, given parametrically by $(x(t), y(t)) = (\cos(t), -\sin(t))$ is a *limit cycle* for (6.4.3), and it is asymptotically stable. Note that $(0, 0)$ is the only equilibrium solution inside the unit circle. As for the construction of the region R, we let $\mu_1 < \mu < \mu_2$, where $0 < \mu_1 < 1$ and $1 < \mu_2 < \infty$, and define the annulus region

$$R = \{(x, y) : \mu_1 < x^2 + y^2 = 1 < \mu_2\}.$$

Clearly, R contains no equilibrium solution, and by Theorem 6.4.4, there is a periodic cycle in R.

6.5 Lotka–Volterra competition model

If two populations $x_1(t)$ and $x_2(t)$ were growing logistically and not affecting each other, their growth can be described by two uncoupled logistic equations

Copyright Elsevier 2022

$$\begin{cases} x_1'(t) = r_1 x_1(t)\left(1 - \frac{x_1(t)}{K_1}\right), \\ x_2'(t) = r_2 x_2(t)\left(1 - \frac{x_2(t)}{K_2}\right), \\ x_1(0) > 0,\ x_2(0) > 0, \end{cases}$$

where $r_i, i = 1, 2$, are the growth rates (birth–death), and $K_i, i = 1, 2$, are positive constants. If the two populations were competing for the same resources, then the growth rate of each population is affected by the other. In that case the model becomes

$$\begin{cases} x_1'(t) = r_1 x_1(t)\left(1 - \frac{x_1(t)}{K_1} - \lambda_1 x_2(t)\right), \\ x_2'(t) = r_2 x_2(t)\left(1 - \frac{x_2(t)}{K_2} - \lambda_2 x_1(t)\right), \\ x_1(0) > 0,\ x_2(0) > 0, \end{cases}$$

where $\lambda_i, i = 1, 2$, are positive constants. To simplify the calculations, we convert the system by using non-dimensional variables by measuring x_1 in units of K_1, x_2 in units of K_2, and time in units of $\frac{1}{r_1}$. We let

$$x_1(t) = K_1 x(t),\ x_2(t) = K_2 y(t),\ \text{ and } t = \frac{s}{r_1}.$$

By the chain rules, it follows that

$$K_1 \frac{dx}{ds} = \frac{dx_1}{dt}\frac{dt}{ds} = \frac{dx_1}{dt}\frac{1}{r_1}$$

or

$$x_1'(t) = r_1 K_1 \frac{dx}{ds}.$$

Substituting this back into the first equation of the model, we arrive at

$$r_1 K_1 \frac{dx}{ds} = r_1 K_1 x(s)\left(1 - x(s) - \lambda_1 K_2 y(s)\right),$$

from which we get

$$\frac{dx}{ds} = x(s)\left(1 - x(s) - \lambda_1 K_2 y(s)\right).$$

Similarly, we obtain

$$\frac{dy}{ds} = r y(s)\left(1 - y(s) - \lambda_2 K_1 x(s)\right),$$

where $r = \frac{r_2}{r_1}$.

Copyright Elsevier 2022

Interchanging s with t and letting $\mu_1 = \lambda_1 K_2$ and $\mu_2 = \lambda_2 K_1$, we arrive at the dimensionless Lotka–Volterra competition model

$$\begin{cases} x'(t) = x(t)\big(1 - x(t) - \mu_1 y(t)\big), \\ y'(t) = ry(t)\big(1 - y(t) - \mu_2 x(t)\big), \\ x(0) = x_0 > 0,\ y(0) = y_0 > 0. \end{cases} \tag{6.5.1}$$

Next, we show that solutions of (6.5.1) are positive and remain bounded for all $t > 0$.

Proposition 6.1. *If $x(t)$, $y(t)$ is a solution pair of (6.5.1) with positive initial conditions x_0 and y_0, respectively, then they are positive and bounded for all $t > 0$.*

Proof. Let $x(t)$, $y(t)$ be a solution of (6.5.1) with positive initial conditions x_0 and y_0. Dividing the first equation of (6.5.1) by $x(t)$ and the second equation by $y(t)$, respectively, and integrating, we get the integral relations

$$x(t) = x_0 e^{\int_0^t \left(1 - x(s) - \mu_1 y(s)\right) ds}$$

and

$$y(t) = y_0 e^{r \int_0^t \left(1 - y(s) - \mu_2 s(s)\right) ds}.$$

We see that if the initial conditions are positive, then the components of the solution are positive. It is left to show that the solutions remain bounded for all $t > 0$. Since x and y are positive, we have that

$$x(t)\big(1 - x(t) - \mu_1 y(t)\big) \le x(t)\big(1 - x(t)\big)$$

and

$$ry(t)\big(1 - y(t) - \mu_2 x(t)\big) \le ry(t)\big(1 - y(t)\big).$$

It follows that

$$x'(t) \le x(t)\big(1 - x(t)\big) \text{ and } y'(t) \le ry(t)\big(1 - y(t)\big).$$

From now on, we write x for $x(t)$ and y for $y(t)$. Consider the differential equation

$$z' = z(1 - z),\ z(0) = x(0) = x_0.$$

Separating the variables and then integrating, we get $\int \frac{dz}{z(1-z)} = \int dt$. Using partial fractions on the left side integral, we obtain $\mathrm{Ln}|\frac{z}{1-z}| = t + C$. Taking the exponential on both sides, solving for z, and then applying the initial condition, we arrive at the unique solution

$$z(t) = \frac{x_0 e^t}{1 + x_0(e^t - 1)} \to 1 \text{ as } t \to \infty.$$

Copyright Elsevier 2022

Since z is positive and has a limit, we conclude that z is bounded for all $t > 0$. Moreover,

$$z(t) \leq 1 \text{ as } t \to \infty.$$

From this discussion we arrive at

$$x(t) \leq z(t) \leq 1.$$

Similar results can be obtained for $y(t)$. This concludes the proof. □

Setting the right sides of (6.5.1) equal to zero, we obtain the equilibrium solutions

$$E_0 = (0, 0), \ E_1 = (0, 1), \ E_2 = (1, 0),$$

and

$$E_3, \text{ which comprises of the solutions (if they exist) of } x + \mu_1 y = 1, \ \mu_2 x + y = 1.$$

Next, we use Theorem 6.2.1 to analyze the stability of the equilibrium solutions. The Jacobian matrix is

$$J(x, y) = \begin{pmatrix} 1 - 2x - \mu_1 y & -\mu_1 x \\ -r\mu_2 y & r(1 - 2x - \mu_2 x) \end{pmatrix}.$$

For $E_0 = (0, 0)$, we have

$$J(0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix}$$

with eigenvalues $\lambda = 1, r$, and hence by Theorem 6.2.1 and Remark 6.3 E_0 is unstable or repeller. At E_1,

$$J(0, 1) = \begin{pmatrix} 1 - \mu_1 & 0 \\ -r\mu_2 & -r \end{pmatrix}$$

with eigenvalues $\lambda = 1 - \mu_1$ and $-r$, and hence by Theorem 6.2.1 and Remark 6.3 E_1 is either a saddle or asymptotically stable. Similarly, at E_2,

$$J(1, 0) = \begin{pmatrix} -1 & -\mu_1 \\ 0 & r(1 - \mu_2) \end{pmatrix}$$

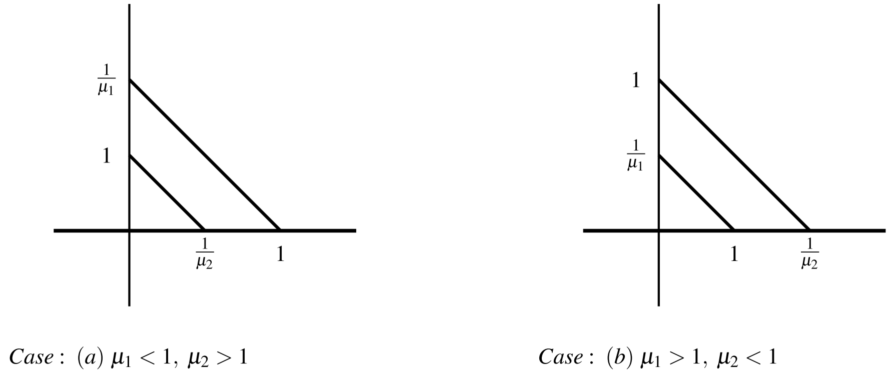
with eigenvalues $\lambda = -1$ and $r(1 - \mu_2)$, and hence by Theorem 6.2.1 and Remark 6.3 E_2 is either a saddle or asymptotically stable. For, E_3 we solve the system

$$x + \mu_1 y = 1, \ \mu_2 x + y = 1.$$

After some calculations, we obtain

$$E_3 = (\frac{1 - \mu_1}{1 - \mu_1\mu_2}, \frac{1 - \mu_2}{1 - \mu_1\mu_2}).$$

Copyright Elsevier 2022



Case: (*a*) $\mu_1 < 1,\ \mu_2 > 1$ *Case*: (*b*) $\mu_1 > 1,\ \mu_2 < 1$

FIGURE 6.10

No equilibrium points in the first quadrant.

The stability of the equilibrium solution of E_3 is important to biologists since it shows the coexistence of both species. We will analyze the stability by considering different cases that are based on the eigenvalues.

Case (a): $\mu_1 < 1, \mu_2 > 1$**.**
In this case the lines $x + \mu_1 y = 1$ and $\mu_2 x + y = 1$ do not intersect in the first quadrant, and it follows that E_3 is not an equilibrium solution as depicted in Fig. 6.10. In addition, in Case (a), E_2 is asymptotically stable, and E_1 is a saddle point. Since the interior of the first quadrant does not contain any equilibrium solution, it follows from Theorem 6.4.3 that there is no periodic solution in the first quadrant. Recall that by Theorem 6.4.3 a closed orbit or limit cycle surrounds at least one equilibrium point. Thus by Theorem 6.4.3 the omega limit set of any trajectory is the equilibrium solution E_2, and hence E_2 is globally asymptotically stable; see Fig. 6.11. The biological meaning is that population x wins the competition. In other words, population y becomes instinct.

Case (b): $\mu_1 > 1, \mu_2 < 1$**.**
In this case the roles of E_1 and E_2 are reversed, but still only one competitor survives.

Next, we consider the cases where E_3 is in the first quadrant, that is, E_3 is an equilibrium solution.

It turned out that E_3 is an equilibrium solution in two cases.

Case (c): $\mu_1 < 1, \mu_2 < 1$ and **Case (d):** $\mu_1 > 1, \mu_2 > 1$**.**
Next, we make use of Theorem 6.4.2 (generalization of the Bendixson–Dulac theorem) to study the possibilities of the existence of a periodic solution that surrounds E_3. Let $h(x, y) = \frac{1}{xy}$. Then h is a continuously differentiable function in $G = \{(x, y) : x > 0, y > 0\}$ with

$$\frac{\partial}{\partial x}\Big[\frac{1}{xy}x\big(1 - x - \mu_1 y\big)\Big] = -\frac{1}{y}$$

Copyright Elsevier 2022

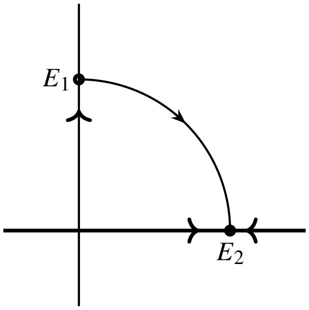


FIGURE 6.11

E_2 is globally asymptotically stable.

and

$$\frac{\partial}{\partial y}\Big[\frac{1}{xy}ry\big(1-y-\mu_2x\big)\Big]=-\frac{r}{x}.$$

Thus

$$\frac{\partial}{\partial x}\Big[\frac{1}{xy}x\big(1-x-\mu_1y\big)\Big]+\frac{\partial}{\partial y}\Big[\frac{1}{xy}ry\big(1-y-\mu_2x\big)\Big]=-(\frac{1}{y}+\frac{r}{x}),$$

which is negative in G. Therefore there is no periodic solution in the first quadrant in both case (c) and case (d).

Let (x^*,y^*) be a solution of $x+\mu_1y=1$, $\mu_2x+y=1$, that is, $(x^*,y^*)=(\frac{1-\mu_1}{1-\mu_1\mu_2},\frac{1-\mu_2}{1-\mu_1\mu_2})$. Then

$$\begin{aligned}J(x^*,y^*)&=\begin{pmatrix}1-2x^*-\mu_1y^* & -\mu_1x^*\\ -r\mu_2y^* & r(1-2y^*-\mu_2x^*)\end{pmatrix}\\&=\begin{pmatrix}1-2x^*-\mu_1\frac{1-\mu_2}{1-\mu_1\mu_2} & -\mu_1x^*\\ -r\mu_2y^* & r(1-2y^*-\mu_2\frac{1-\mu_1}{1-\mu_1\mu_2})\end{pmatrix}\\&=\begin{pmatrix}-x^* & -\mu_1x^*\\ -r\mu_2y^* & -ry^*\end{pmatrix},\end{aligned}$$

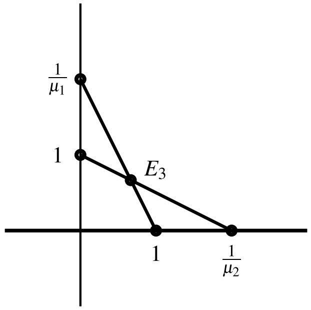
and $\operatorname{tra}(J(x^*,y^*))=-x^*-ry^*<0$ for $x^*,y^*>0$ implies that the sum of the eigenvalues is negative. Moreover,

$$\det(J(x^*,y^*))=rx^*y^*-r\mu_1\mu_2x^*y^*=rx^*y^*(1-\mu_1\mu_2).$$

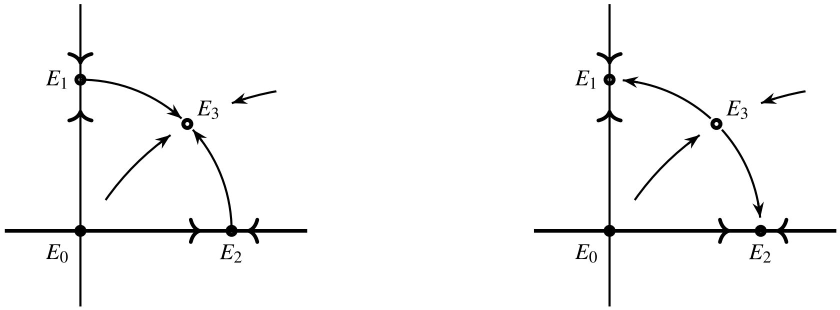
As

$$\det(J(x^*,y^*)-\lambda I)=\lambda^2-\operatorname{tra}(J(x^*,y^*))\lambda+\det(J(x^*,y^*)),$$

Copyright Elsevier 2022



Case(c): $\mu_1 < 1,\ \mu_2 < 1$ *Case*(d): $\mu_1 > 1,\ \mu_2 > 1$



(e): $\mu_1 < 1,\ \mu_2 < 1,\ 1-\mu_1\mu_2 > 0$ (f): $\mu_1 > 1,\ \mu_2 > 1, 1-\mu_1\mu_2 < 0$

FIGURE 6.12

Coexistence.

we have that

$$\lambda^2 + (x^* + ry^*)\lambda + rx^*y^*(1 - \mu_1\mu_2) = 0.$$

It follows that the eigenvalues are given by

$$\lambda = \frac{-x^* - ry^* \pm \sqrt{(x^* + ry^*)^2 - 4rx^*y^*(1 - \mu_1\mu_2)}}{2}.$$

So E_3 is asymptotically stable if $1 - \mu_1\mu_2 > 0$ and a saddle if $1 - \mu_1\mu_2 < 0$. Fig. 6.12(d) and (f).

If $\mu_1\mu_2 = 1$, then the two lines are parallel or coincident. If coincident, then there is a line of equilibrium solutions, and if parallel, then cases (a) and (b) apply. In case (c), $1 - \mu_1 > 0$ and $1 - \mu_2 > 0$, so E_1 and E_2 are saddle unstable points. Since $\mu_1\mu_2 < 1$, E_3 is an attractor, that is, all solutions tend to E_3. In that case the two populations coexist. See Fig. 6.12(d).

Copyright Elsevier 2022

In case (d), $\mu_1 - 1 > 0$ and $\mu_2 - 1 > 0$, and hence E_1 and E_2 are asymptotically stable. Also, in case (d) $\mu_1\mu_2 > 1$, and so E_3 is a saddle point. See Fig. 6.12(f). Here the initial conditions determine the winner.

6.6 Bifurcation in planar systems

We consider the two-dimensional autonomous system

$$x' = P(x, y, \mu),\;\; y' = Q(x, y, \mu) \tag{6.6.1}$$

depending upon a real parameter μ. We assume that P and Q are continuous in all arguments on some subset of $\mathbb{R}^2$. As μ changes values or signs, the behavior of solutions changes and may result in bifurcation points. For example, as we move across a particular equilibrium solution, the stability may change from stable to unstable, and vice versa. We begin by considering the following example.

Example 6.14. Consider the two-dimensional linear system

$$x' = x + \mu y,\;\; y' = x - y$$

for parameter μ. The matrix

$$\begin{pmatrix} 1 & \mu \\ 1 & -1 \end{pmatrix}$$

has the eigenvalues

$$\lambda = \pm\sqrt{1+\mu}.$$

If $\mu > -1$, then the eigenvalues are real with opposite signs, and therefore the origin $(0, 0)$ is a saddle point (unstable). For the particular value $\mu = -1$, the system becomes

$$x' = x - y,\;\; y' = x - y$$

and results in the whole line $y = x$ of fixed points. If $\mu < -1$, then the eigenvalues are purely imaginary, and the origin $(0, 0)$ is a center (stable). As $\mu > -1$ decreases to $\mu < -1$, the nature of the solution changes at $\mu = -1$; as a result, the equilibrium solution $(0, 0)$ changes from an unstable saddle to a stable center as μ passes the critical value -1. Thus we have a *bifurcation* at $\mu = -1$.

Example 6.15. Consider the planar autonomous system

$$x' = -y - x(x^2 + y^2 - \mu),\;\; y' = x - y(x^2 + y^2 - \mu).$$

It is clear that $(0, 0)$ is the only equilibrium point. We rewrite the system in the form

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} \mu & -1 \\ 1 & \mu \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -x(x^2+y^2) \\ -y(x^2+y^2) \end{pmatrix}.$$

Copyright Elsevier 2022

The eigenvalues of

$$A = \begin{pmatrix} \mu & -1 \\ 1 & \mu \end{pmatrix}$$

are $\mu \pm i$.

(1) If $\mu < 0$, then the trajectories spiral toward the origin, and the origin is asymptotically stable by Theorem 6.2.1.
(2) If $\mu > 0$, then the origin is unstable for the nonlinear system.
(3) If $\mu = 0$, then Theorem 6.2.1 says nothing about the stability or instability about the origin of the nonlinear system.

Next we examine in more detail the trajectories of the system by using polar coordinates as we did in Example 6.8. By doing so we arrive at

$$r' = r(\mu - r^2) \quad \text{and} \quad \theta' = 1. \tag{6.6.2}$$

Separating the variables and using partial fractions we arrive at

$$\int \frac{dr}{r(r - \sqrt{\mu})(r + \sqrt{\mu})} = t + k \quad \text{for some constant } k.$$

This leads to the solution

$$r(t) = \frac{\sqrt{\mu}}{\sqrt{1 + ce^{-2\mu t}}}. \tag{6.6.3}$$

Similarly, if we choose $\theta(0) = 0$, then we have $\theta(t) = t$.

Assume that $\mu > 0$ and consider the following three cases.

(3a) When $c = 0$, we have from (6.6.3) $r(t) = \sqrt{\mu}$, which represents a periodic solution with radius $\sqrt{\mu}$.
(3b) If $c < 0$, then (6.6.3) represents an orbit or trajectory that spirals counterclockwise toward the circle $r(t) = \sqrt{\mu}$ from the outside.
(3c) If $c > 0$, then (6.6.3) represents an orbit or trajectory that spirals counterclockwise toward the circle $r(t) = \sqrt{\mu}$ from the inside.

Let $0 < \mu_1 < \mu < \mu_2$ and define the annulus region

$$R = \{(x, y) : \mu_1 < x^2 + y^2 < \mu_2\}.$$

Clearly, R contains no equilibrium solution, and then cases (3a)–(3c) prove that the circle given by $r^2(t) = \mu$ or $x^2 + y^2 = \mu$ is a *stable limit cycle* and the origin is unstable.

Next, we examine the stability of the origin when $\mu = 0$.

In this case (6.6.2) gives

$$r' = -r^3,$$

Copyright Elsevier 2022

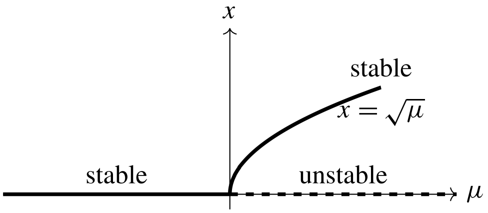


FIGURE 6.13

Bifurcation diagram for Example 6.15.

which has the solution

$$r(t) = \frac{1}{\sqrt{2t+K}}$$

for a suitable constant K. Clearly, $r(t) \to 0$ as $t \to \infty$, and hence the zero solution (or the origin) is asymptotically stable (spirally stable).

Finally, for $\mu < 0$, from (6.6.2) we have

$$r' = r(r^2 + l^2) \text{ and } \theta' = 1,$$

where $l^2 = -\mu > 0$. Separating the variables and using partial fractions, we arrive at

$$\int \frac{dr}{r(r^2+l^2)} = -dt \text{ for some constant } k.$$

This leads to the solution

$$r(t) = \frac{cl^2 e^{-2l^2 t}}{1 - cl^2 e^{-2l^2 t}} \text{ for some constant } c.$$

Clearly, $r(t) \to 0$, as $t \to \infty$, and hence the zero solution (or the origin) is asymptotically stable (spirally stable). In conclusion, for $\mu \leq 0$, the origin $(0, 0)$ is stable, and for $\mu > 0$, (0,0) is unstable, and $r = \sqrt{\mu}$ is stable. Thus as μ passes through the critical value $\mu = 0$ from negative to positive, the origin changes from stable spiral to unstable spiral, and there appears a new limit cycle that is a periodic solution. This type of bifurcation is common and known as *Hopf bifurcation*.

The bifurcation diagram is shown in Fig. 6.13.

The next theorem provides a necessary condition for the existence of Hopf bifurcation. The proof is complicated, and hence we omit it.

Theorem 6.6.1. *Consider the two-dimensional autonomous system*

$$x' = P(x, y, \mu), \ y' = Q(x, y, \mu) \tag{6.6.4}$$

Copyright Elsevier 2022

depending upon a real parameter μ, where P and Q are differentiable in the neighborhood of $(0, 0, 0)$. Assume that for small μ,

$$P(0, 0, \mu) = Q(0, 0, \mu) = 0. \tag{6.6.5}$$

Let the Jacobian

$$J(0, 0, 0) = \begin{pmatrix} P_x(0, 0, 0) & P_y(0, 0, 0) \\ Q_x(0, 0, 0) & Q_y(0, 0, 0) \end{pmatrix} = \begin{pmatrix} 0 & -w \\ w & 0 \end{pmatrix} \text{ for some } w \neq 0. \tag{6.6.6}$$

Let

$$\begin{aligned} a = {} & \frac{1}{16}\big(P_{xxx} + Q_{xxy} + P_{xyy} + Q_{yyy}\big) \\ & + \frac{1}{16w}\big(P_{xy}(P_{xx} + P_{yy}) - Q_{xy}(Q_{xx} + Q_{yy}) - P_{xx}Q_{xx} + P_{yy}Q_{yy}\big) \end{aligned} \tag{6.6.7}$$

evaluated at $(0, 0, 0)$ with $a \neq 0$. Suppose that

$$a\big(P_{\mu x} + Q_{\mu y}\big) > 0. \tag{6.6.8}$$

Then for sufficiently small $\mu < 0$, there is a unique periodic solution with amplitude $A(\mu)$ such that

$$\lim_{\mu \to 0^-} \frac{A(\mu)}{\sqrt{|\mu|}} = c > 0.$$

If

$$a\big(P_{\mu x} + Q_{\mu y}\big) < 0, \tag{6.6.9}$$

then the periodic solution exists for $\mu > 0$. The stability of the periodic solution is determined by the sign of

$$P_{\mu x} + Q_{\mu y} \tag{6.6.10}$$

and is the opposite of the stability of the zero solution in the range of μ where the periodic solution exists.

Example 6.16. As an application of Theorem 6.6.1, we consider the system given in Example 6.14. It readily follows that the Jacobian

$$J(0, 0, 0) = \begin{pmatrix} \mu & -1 \\ 1 & \mu \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Thus condition (6.6.6) is satisfied with $w = 1$. After simple calculations, it follows that (6.6.7) evaluated at $(0, 0, 0)$ implies that $a = -1$. Moreover,

$$a\big(P_{\mu x} + Q_{\mu y}\big) = -1(1 + 1) < 0.$$

Thus by Theorem 6.6.1, the system possesses a periodic solution for $\mu > 0$ that is stable (spirally stable) since the origin is unstable in the region where μ is positive.

Copyright Elsevier 2022

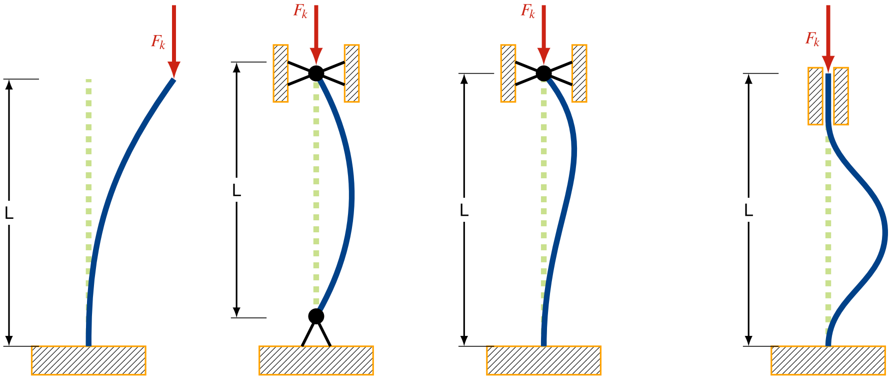


FIGURE 6.14

Buckling beam.

We end this section with an application to Euler's buckling beam.

Example 6.17. (Euler's buckling beam) One of the most famous applications to bifurcations is the so-called Euler's buckling beam. Consider different scenarios of beams in Fig. 6.14, where excessive pressure is applied to the beam. The beam can support a small force or pressure, but as the force is increased, the beam will go through a qualitative change in its vertical standing position. At one point, as we keep increasing the force, the beam position will get compromised and results in a so-called *buckling,* that is, there will be a critical value at which if the pressure is increased beyond it, then the beam will buckle. Fig. 6.14 provides different situations for such buckling beams along with supported ends. As for modeling, suppose we have a beam of length one unit and the beam is placed along the horizontal axis with one end supported by a fixed wall and a force $\mathbb{F} \geq 0$ applied along its axis. Let $\delta > 0$ be the viscous damping, and let $K > 0$ be the stiffness. The small deflection of the beam is denoted by $v(x, t)$, where x denotes the horizontal axes. After transforming the model from a partial differential equation to a second-order differential equation, it was shown in [27] that the amplitude u in variable time t satisfies the second-order differential equation

$$u''(t) + \delta u'(t) - \pi^2(\mathbb{F} - \pi^2)u(t) + \frac{1}{2}K\pi^4 u^3(t) = 0. \tag{6.6.11}$$

To better study the differential equation, we put it in the system form

$$\begin{cases} x' = y, \\ y' = \pi^2(\mathbb{F} - \pi^2)x - \delta y - \frac{1}{2}K\pi^4 x^3, \end{cases} \tag{6.6.12}$$

where the force $\mathbb{F}$ is the only parameter. The system has three equilibrium solutions

Copyright Elsevier 2022

$$(0,0),\ \left(\sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}},0\right),\ \text{and}\ \left(-\sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}},0\right).$$

Note that the equilibrium solutions

$$\left(\sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}},0\right)\ \text{and}\ \left(-\sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}},0\right)$$

are only valid for $\mathbb{F} > \pi^2$.

Now

$$J(0,0) = \begin{pmatrix} 0 & 1 \\ \pi^2(\mathbb{F}-\pi^2) & -\delta \end{pmatrix}$$

with eigenvalues

$$\lambda = \frac{-\delta \pm \sqrt{\delta^2 + 4\pi^2(\mathbb{F}-\pi^2)}}{2},$$

and, as a consequence, $(0,0)$ is stable if $\mathbb{F} < \pi^2$ and unstable if $\mathbb{F} > \pi^2$. The physical implication is that the beam will buckle if $\mathbb{F} > \pi^2$.

Now we consider the equilibrium solution $(\sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}},0)$ for $\mathbb{F} > \pi^2$. We translate it to the origin with transformation

$$x_1 = x - \sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}},\ y_1 = 0$$

and arrive at the linearized matrix

$$J(0,0) = \begin{pmatrix} 0 & 1 \\ -2\pi^2(\mathbb{F}-\pi^2) & -\delta \end{pmatrix}$$

with eigenvalues

$$\lambda = \frac{-\delta \pm \sqrt{\delta^2 - 8\pi^2(\mathbb{F}-\pi^2)}}{2}.$$

It follows that $(0,0)$ is stable if $\mathbb{F} > \pi^2$ and that the equilibrium solution $(\sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}},0)$ is stable for $\mathbb{F} > \pi^2$. The physical implication is that the beam will buckle if $\mathbb{F} > \pi^2$.

Similarly, we can easily show that the equilibrium solution $(-\sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}},0)$ is also stable for $\mathbb{F} > \pi^2$. Thus we have a pitchfork bifurcation at $\mathbb{F} = \pi^2$. The bifurcation diagram in Fig. 6.15 shows that bending is symmetric and can take place in the negative direction as well.

Copyright Elsevier 2022

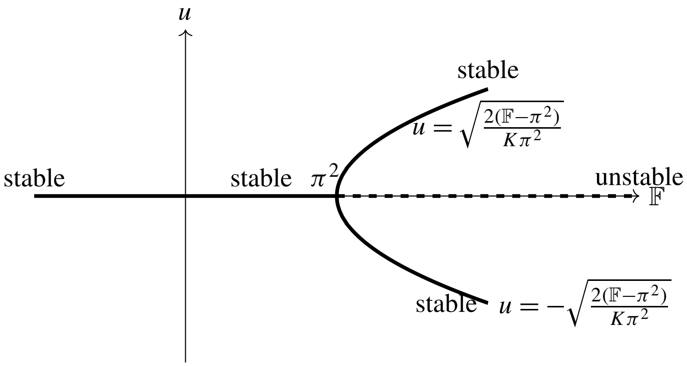


FIGURE 6.15

Bifurcation diagram for $u = \pm\sqrt{\frac{2(\mathbb{F}-\pi^2)}{K\pi^2}}$.

6.7 Manifolds and Hartman–Grobman theorem

In the study of dynamical systems and differential equations, the stable manifold theorem is an important result about the structure of the set of orbits approaching a given hyperbolic fixed point. Stable manifolds are directly related to the important Hartman–Grobman theorem, which states that the flow generated by a smooth vector field in a neighborhood of a hyperbolic equilibrium point is topologically conjugate with the flow generated by its linearization. Hartman's counterexample shows that, in general, the conjugacy cannot be taken to be C^1. However, the stable manifold theorem tells us that there are important structures for the two flows that can be matched up by smooth changes of variables.

Let $f : E \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be C^1 and consider the autonomous nonlinear differential equation

$$x' = f(x). \tag{6.7.1}$$

Definition 6.7.1. An equilibrium point x_0 of (6.7.1) is called *hyperbolic* if none of the eigenvalues of the matrix $Df(x_0)$ has a zero real part, where throughout this section, we adopt the notation

$$Df(x_0) = J(x_0).$$

For example, the system

$$x' = y, \quad y' = -x - x^3 - 3y$$

has $(0, 0)$ as its only equilibrium solution. The linearization at $(0, 0)$ is

$$Df(0,0) = \begin{pmatrix} 0 & 1 \\ -1 & -3 \end{pmatrix},$$

Copyright Elsevier 2022

which has the eigenvalues

$$\lambda = \frac{-2 \pm \sqrt{5}}{2}.$$

Thus (0, 0) is a hyperbolic equilibrium solution since none of the eigenvalues has the zero real part.

Our aim is to show that the local behavior of the nonlinear system (6.7.1) near a hyperbolic equilibrium point x_0 is qualitatively determined by the behavior of the linear system

$$x' = Ax, \tag{6.7.2}$$

where the matrix $A = Df(x_0)$. Now we define the flow or trajectory of (6.7.1) that suits our notations for this section. Denote the maximal interval of existence (α, β) of the solution $\phi(t, x_0)$ of the IVP

$$x' = f(x), \;\; x(0) = x^* \tag{6.7.3}$$

by $I(x^*)$ since α or β usually depend on x^*. If $x^* \in E$ and $t \in I(x^*)$, then the set of mappings ϕ_t defined by

$$\phi_t(x^*) = \phi(t, x^*)$$

is called the flow of (6.7.1). We define the set $\Omega \subseteq \mathbb{R} \times E$ as

$$\Omega = \{(t, x^*) \in \mathbb{R} \times E : \; t \in I(x^*)\}.$$

We leave it an exercise to show that the set Ω is an open subset of $\mathbb{R} \times E$ and $\phi \in C^1(\Omega)$.

Theorem 6.7.1. *Let* $E \subseteq \mathbb{R}^n$ *be open with* $f \in C^1(E)$*, and let* $\phi_t : E \to E$ *be the flow of* (6.7.1) *defined for all* $t \in \mathbb{R}$.
(i) The set $S \subseteq E$ *is said to be invariant with respect to the flow* ϕ_t *if* $\phi_t(S) \subseteq S$ *for all* $t \in \mathbb{R}$.
(ii) The set E *is said to be positively (or negatively) invariant with respect to the flow* ϕ_t *if* $\phi_t(S) \subseteq S$ *for all* $t > 0$ *(or* $t < 0$).

We furnish the following example.

Example 6.18. Consider the nonlinear system of differential equations

$$x' = \begin{pmatrix} x_1 + x_2^2 \\ -x_2 \end{pmatrix}$$

with

$$x(0) = \begin{pmatrix} a \\ b \end{pmatrix} = x^*.$$

Copyright Elsevier 2022

Now $x_2' = -x_2$ has the solution $x_2 = be^{-t}$. Substituting this solution of x_2 into the first differential equation and using the variation of parameters formula, we arrive at $x_1 = ae^t + \frac{1}{3}b^2(e^t - e^{-2t})$. Thus

$$\phi_t(x^*) = \begin{pmatrix} ae^t + \frac{1}{3}b^2(e^t - e^{-2t}) \\ be^{-t} \end{pmatrix}.$$

Next, we show that the set

$$S = \{x \in \mathbb{R}^2 : x_1 = -x_2^2/3\}$$

is invariant under the flow ϕ_t. Let $x^* \in S$. Then $a = -b^2/3$, and it follows that

$$\phi_t(x^*) = \begin{pmatrix} -\frac{1}{3}b^2e^{-2t} \\ be^{-t} \end{pmatrix} \in S.$$

Therefore $\phi_t(S) \subseteq S$ for all $t \in \mathbb{R}$. We will see later that the set S will be called the *stable manifold* for system (6.7.1).

We are ready to formally define stable and unstable manifolds.

Consider the linear system given by (6.7.2) and suppose the eigenvalues of A and their corresponding generalized eigenvectors are $\lambda_j = \alpha_j + i\beta_j$ and $k_j + i\, d_j$ for $j = 1, 2, \ldots, n$. Let those vectors be the columns of the matrix P. Then $P^{-1}AP = J$, where J is the Jordan form of A. Then the sets

$$E^s := Span\{k_j, d_j : \alpha_j < 0\},$$

$$E^u := Span\{k_j, d_j : \alpha_j > 0\},$$

$$E^c := Span\{k_j, d_j : \alpha_j = 0\}$$

are the *stable, unstable*, and *center* subspaces for system (6.7.2).

Definition 6.7.2. Suppose that x_0 is a hyperbolic equilibrium point of (6.7.1).
(i) The *stable manifold* of x_0 is the set

$$W^s(x_0) = \{x^* \in E : \lim_{t\uparrow\infty} \phi_t(x^*) = x_0\}.$$

(ii) The *unstable manifold* of x_0 is the set

$$W^u(x_0) = \{x^* \in E : \lim_{t\downarrow-\infty} \phi_t(x^*) = x_0\}.$$

We have the following example.

Copyright Elsevier 2022

Example 6.19. Consider the system given in Example 6.18. Then the eigenvalues of $A = Df(0,0)$ are $\lambda = \pm 1$. If $\lambda = 1$, then its corresponding eigenvector is given by $v_1 = (1,0)^T$. Similarly, $v_2 = (0,1)^T$ is the corresponding eigenvector for $\lambda = -1$. Thus

$$E^u = \{(x,y) \in \mathbb{R}^2 : y = 0\}$$

and

$$E^s = \{(x,y) \in \mathbb{R}^2 : x = 0\}.$$

Now $\lim_{t\uparrow\infty} \phi_t(x^*) = x_0$ if and only if $a + b^2/3 = 0$. Thus the stable manifold is

$$W^s(x_0) = \{x^* \in \mathbb{R}^2 : a = -b^2/3\}.$$

Similarly, $\lim_{t\downarrow-\infty} \phi_t(x^*) = (0,0)$ if and only if $b = 0$. Thus the unstable manifold is

$$W^u(0,0) = \{x^* \in \mathbb{R}^2 : b = 0\}.$$

6.7.1 The stable manifold theorem

The stable manifold theorem that we are about to state and give a sketch of its proof states that (6.7.1) has stable and unstable manifolds W^s and W^u that are tangent at x_0 to the stable and unstable subspaces E^s and E^u of the linearized system (6.7.2), where x_0 is a hyperbolic equilibrium solution of (6.7.1). Furthermore, W^s and W^u are of the same dimensions as E^s and E^u. Besides, W^s and W^u are positively and negatively invariant under ϕ_t, respectively, and satisfy

$$\lim_{t\uparrow\infty} \phi_t(x^*) = x_0 \quad \text{for all } \; x^* \in W^s,$$

$$\lim_{t\downarrow-\infty} \phi_t(x^*) = x_0 \quad \text{for all } \; x^* \in W^u.$$

We furnish the following example that should serve as warm up.

Example 6.20. Let $g \in C^1(E)$, where $E \subseteq \mathbb{R}^2$ is an open set containing $(0,0)$. Consider the nonlinear system

$$\begin{cases} x' = -x, \\ y' = y + g(x), \end{cases}$$

where we assume that $g(0) := g(0,0) = 0$ and that the function g may contain linear terms. The only equilibrium solution of the system is the origin and the Jacobian is

$$Df(0,0) = \begin{pmatrix} -1 & 0 \\ g'(0) & 1 \end{pmatrix}.$$

Copyright Elsevier 2022

Then the eigenvalues of $A = Df(0,0)$ are $\lambda = \pm 1$, and the zero equilibrium is hyperbolic. If $\lambda = 1$, then its corresponding eigenvector is given by $v_1 = (0,1)^T$. Similarly, $v_2 = (2, -g'(0))^T$ is the corresponding eigenvector for $\lambda = -1$. Thus

$$E^u = \{(x,y) \in \mathbb{R}^2 : x = 0\}$$

and

$$E^s = \{(x,y) \in \mathbb{R}^2 : y = -\frac{1}{2}g'(0)x\}.$$

To find the stable and unstable manifolds, we directly solve the system. Assume an initial point $x(0) = a$, $y(0) = b$ for $a, b \in \mathbb{R}$. The first equation in the system has the solution

$$x(t) = ae^{-t}.$$

Substituting x back into the second equation and using the variation of parameters formula, we arrive at

$$y(t) = be^t + e^t \int_0^t e^{-s} g(ae^{-s}) ds.$$

Making the substitution $u = e^{-s}$, we arrive at the flow

$$\phi_t(c) = \begin{pmatrix} ae^{-t} \\ be^t + e^t \int_{e^{-t}}^1 g(au)du \end{pmatrix}.$$

If $c = (a,b)^T \in \mathbb{R}^2$ is in the unstable manifold $W^u(0,0)$, then $\lim_{t\downarrow -\infty} \phi_t(c) \to 0$. Then from the first component of ϕ_t it follows that $\lim_{t\downarrow -\infty} ae^{-t} = 0$ if and only if $a = 0$. Substituting $a = 0$ into the second component of ϕ_t and taking the limit, that is, $\lim_{t\uparrow\infty} \left(be^1 + e^t \int_{e^{-t}}^t g(0)ds\right) = \lim_{t\downarrow -\infty} be^t = 0$ (since $g(0) = 0$), it follows that

$$W^u(0,0) = \{(x,y) \in \mathbb{R}^2 : x = 0\}.$$

Similarly, if $c = (a,b)^T \in \mathbb{R}^2$ is in the stable manifold $W^s(0,0)$, then $\lim_{t\uparrow\infty} \phi_t(c) \to 0$. Then from the first component of ϕ_t it follows that $\lim_{t\uparrow\infty} ae^{-t} = 0$ for any a. From the second component of ϕ_t we have

$$\lim_{t\uparrow\infty} e^t \left(b + e^t \int_{e^{-t}}^t g(0)ds\right) = 0. \tag{6.7.4}$$

To that effect, we claim that

$$b = -\int_0^1 g(au)du.$$

Copyright Elsevier 2022

To see that this satisfies (6.7.4), we first notice that

$$\int_0^1 g(au)du = \int_0^{e^{-t}} g(au)du + \int_{e^{-t}}^1 g(au)du.$$

Then it follows

$$\lim_{t\uparrow\infty} e^t \Big(\int_{e^{-t}}^1 g(au)du - \int_0^1 g(au)du \Big) = - \lim_{t\uparrow\infty} e^t \int_0^{e^{-t}} g(au)du.$$

We must show that the term $e^t \int_0^{e^{-t}} g(au)du$ is small for all $t > T$ for sufficiently large T. Since g is continuous with $g(0) = 0$, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|g(au)| < \varepsilon$ for $|au| < \delta$. Moreover, as $0 \le u \le e^{-t}$, we have $|au| = |a||u| \le |a|e^{-t}$. Thus we can choose $T > 0$ sufficiently large so that $|a|e^{-t} < \delta$ for $t > T$. Therefore $\left| e^t \int_0^{e^{-t}} g(au)du \right| < \varepsilon$. As this can be done for any $\varepsilon > 0$, we conclude that (6.7.4) holds. It follows that the stable manifold is

$$W^s(0, 0) = \{(x, y(x)) \in \mathbb{R}^2 : y(x) = - \int_0^1 g(xu)du\}.$$

Note that we could easily use L'Hospital's rule on $\lim_{t\uparrow\infty} e^t \int_0^{e^{-t}} g(xu)du$ to show that (6.7.4) holds. Next, we show that W^s is tangent to E^s:

$$\frac{dy}{dx}|_{x=0} = - \int_0^1 g'(xu)udu\Big|_{x=0}.$$

Making the substitution $v = xu$, this expression becomes

$$\frac{dy}{dx}|_{x=0} = - \frac{1}{x^2} \int_0^x g'(v)vdv\Big|_{x=0}.$$

Using L'Hospital's rule, we find the slope at the origin to be

$$- \frac{1}{x^2} \int_0^x g'(v)vdv\Big|_{x=0} = - \lim_{x\to 0} \frac{1}{x^2} \int_0^x g'(v)vdv = - \lim_{x\to 0} \frac{g'(x)x}{2x} = - \frac{1}{2} g'(0).$$

Thus W^s is tangent to E^s at the origin, and, moreover, $W^s \in C^1$.

Let $g(x) = xe^x$ in Example 6.20. Then $g(0) = 0$ and $g \in C^1$. Furthermore,

$$E^s = \{(x, y) \in \mathbb{R}^2 : y = - \frac{1}{2} g'(0)x\} = \{(x, y) \in \mathbb{R}^2 : y = - \frac{1}{2} x\},$$

and

$$y(x) = - \int_0^1 g(xu)du = \int_0^1 xue^{xu}du = -e^x + \frac{1}{x}e^x - \frac{1}{x}.$$

Copyright Elsevier 2022

Substitute the Maclaurin series for e^x into the last expression and simplify to get

$$y = \frac{-x}{2}\left(1 + \frac{2x}{3} + \frac{x^2}{3} + \cdots\right).$$

Thus

$$W^s(0,0) = \left\{(x, y(x)) \in \mathbb{R}^2 : y(x) = \frac{-x}{2}\left(1 + \frac{2x}{3} + \frac{x^2}{3}\right)\right\}.$$

In Example 6.20, we were able to compute the solutions to find W^s and W^u. This is not the case for most nonlinear systems. The next theorem affirms the existence of such manifolds in the neighborhoods of hyperbolic equilibrium solution, regardless of whether solutions can be found or not. We have seen that the stable manifold of a hyperbolic critical point is the collection of all states that asymptotically approach that fixed point as time goes to infinity. Knowing the stable and unstable manifolds of a fixed point can be useful for simple purposes, for example, for knowing which initial conditions of a system of ordinary differential equations lead to which asymptotic behavior. It is also very useful for spotting "exotic" behavior in dynamical systems.

As mentioned before, if the equilibrium solution is not the origin, then we use transformation so that it coincides with the origin. Next, we define the differentiability of manifolds.

Definition 6.7.3. (Homeomorphism) Let K_1 and K_2 be two subsets of a metric space X. Then K_1 and K_2 are said to be *homeomorphic* if there is a continuous one-to-one map $g : K_1 \to K_2$ such that g^{-1} is also continuous.

Such two sets are called *homeomorphic* or topologically equivalent if there is a homeomorphism of K_1 onto K_2.

Definition 6.7.4. (Differentiable manifold) An n-dimensional differentiable manifold M is a connected metric space with an open covering $M = \cup_\alpha U_\alpha$ such that

(1) for all α, $\cup_\alpha$ is homeomorphic to the open unit ball $B = \{x \in \mathbb{R}^n : |x| < 1\}$ in $\mathbb{R}^n$ centered at the origin, that is, for all α, there exists a homeomorphism $g_\alpha : U_\alpha \to B$ of U_α onto B, and

(2) if $U_\alpha \cap U_\beta \neq \emptyset$ and $g_\alpha : U_\alpha \to B$, $g_\beta : U_\beta \to B$ are homeomorphisms, then $g_\alpha\big(U_\alpha \cap U_\beta\big)$ and $g_\beta\big(U_\alpha \cap U_\beta\big)$ are subsets of $\mathbb{R}^n$, the map

$$g = g_\alpha \circ g_\beta^{-1} : g_\beta\big(U_\alpha \cap U_\beta\big) \to g_\alpha\big(U_\alpha \cap U_\beta\big)$$

is differentiable (or of class C^n), and for all $x \in g_\beta\big(U_\alpha \cap U_\beta\big)$, the Jacobian determinant is not zero, that is, $\det \frac{\partial g}{\partial x} \neq 0$.

For complete proof of the next theorem, we refer to [36].

Theorem 6.7.2. *Suppose* $f(0) = 0$ *and let* $E \subseteq \mathbb{R}^n$ *be an open subset containing the origin with* $f \in C^1(E)$. *Let* $\phi_t : E \to E$ *be the flow of* (6.7.1) *defined for all* $t \in \mathbb{R}$. *Let*

$$x' = Ax := Df(0)x. \tag{6.7.5}$$

Copyright Elsevier 2022

Suppose the linearized matrix $Df(0)$ has k eigenvalues with negative real parts and $n-k$ eigenvalues with positive real parts. Then there exists a k-dimensional differentiable manifold W^s that is tangent to the stable subspace E^s of (6.7.5)*, and there exists an $(n-k)$-dimensional differentiable manifold W^u that is tangent to the unstable space E^u of the linearized system* (6.7.5)*. Moreover,*

$$\lim_{t\uparrow\infty}\phi_t(p)=0 \ \textit{ for any } \ p\in W^s,$$

and

$$\lim_{t\downarrow-\infty}\phi_t(p)=0 \ \textit{ for any } \ p\in W^u.$$

Proof. (We only provide a sketch of the proof) This sketch of the proof is geared toward giving guidelines on how to find the stable and unstable manifolds.

Since $f\in C^1(E)$ and $f(0)=0$, we may write system (6.7.1) as

$$x'=Ax+F(x), \tag{6.7.6}$$

where $A=Df(0)$ and $F(x)=f(x)-Ax$ with $F\in C^1(E)$, $F(0)=0$ and $DF(0)=0$. There is an invertible $n\times n$ matrix C such that

$$B=C^{-1}AC=\begin{pmatrix}P & 0\\ 0 & Q\end{pmatrix},$$

where the $k\times k$ matrix P has the eigenvalues λ_i, $i=1,2,\dots,k$, each with negative real part, and the $(n-k)\times(n-k)$ matrix Q has the eigenvalues λ_i, $i=k+1,k+2,\dots,n$, each with a positive real part. Make the change of variables $y=C^{-1}x$. Then $x'=Ax+F(x)$ becomes $Cy'=ACy+F(Cy)$. Multiplying both sides from the left by C^{-1} gives $y'=C^{-1}ACy+C^{-1}F(Cy)$. Set $G(y)=C^{-1}F(Cy)$ to arrive at

$$y'=\begin{pmatrix}P & 0\\ 0 & Q\end{pmatrix}y+G(y). \tag{6.7.7}$$

Since each step we have made is reversible, solving system (6.7.7) is equivalent to solving the original system (6.7.6). Next, we labor to find stable and unstable manifolds. Define

$$U(t)=\begin{pmatrix}e^{Pt} & 0\\ 0 & 0\end{pmatrix} \quad\text{and}\quad V(t)=\begin{pmatrix}0 & 0\\ 0 & e^{Qt}\end{pmatrix},$$

so that

$$e^{\begin{pmatrix}P & 0\\ 0 & Q\end{pmatrix}t}=U(t)+V(t).$$

For $t\in\mathbb{R}$ and $a\in\mathbb{R}^n$, define the integral equation

$$u(t,a)=U(t)a+\int_0^t U(t-s)G(u(s,a))ds$$

Copyright Elsevier 2022

$$-\int_t^\infty V(t-s)G(u(s,a))ds. \tag{6.7.8}$$

Perko showed the integral equation (6.7.8) satisfies (6.7.7) and $\lim_{t\to\infty} u(t,a)=0$. Define the iterative scheme for computing the solution:

$$\begin{aligned}
u(t,a) &= 0,\\
u^{(k+1)}(t,a) &= U(t)a + \int_0^t U(t-s)G(u^{(k)}(s,a))ds\\
&\quad -\int_t^\infty V(t-s)G(u^{(k)}(s,a))ds.
\end{aligned}$$

Next, we give some insight on why this integral equation given by (6.7.8) is chosen.

In general, using the variation of parameters formula, the solution of (6.7.7) is

$$u(t,a) = \begin{pmatrix} e^{Pt} & 0 \\ 0 & e^{Qt} \end{pmatrix} a + \int_0^t \begin{pmatrix} e^{P(t-s)} & 0 \\ 0 & e^{Q(t-s)} \end{pmatrix} G(u(s,a))ds.$$

Keep in mind that the goal is to remove the parts that blow up as $t \to \infty$. Thus we try to separate convergent and non-convergent parts from the above integral equation. Below we explain how this is done:

$$\begin{aligned}
u(t,a) &= \begin{pmatrix} e^{Pt} & 0 \\ 0 & e^{Qt} \end{pmatrix} a + \int_0^t \begin{pmatrix} e^{P(t-s)} & 0 \\ 0 & e^{Q(t-s)} \end{pmatrix} G(u(s,a))ds\\
&= \begin{pmatrix} e^{Pt} & 0 \\ 0 & 0 \end{pmatrix} a + \begin{pmatrix} 0 & 0 \\ 0 & e^{Qt} \end{pmatrix} a + \int_0^t \begin{pmatrix} e^{P(t-s)} & 0 \\ 0 & 0 \end{pmatrix} G(u(s,a))ds\\
&\quad + \int_0^t \begin{pmatrix} 0 & 0 \\ 0 & e^{Q(t-s)} \end{pmatrix} G(u(s,a))ds\\
&= \begin{pmatrix} e^{Pt} & 0 \\ 0 & 0 \end{pmatrix} a + \begin{pmatrix} 0 & 0 \\ 0 & e^{Qt} \end{pmatrix} a + \int_0^t \begin{pmatrix} e^{P(t-s)} & 0 \\ 0 & 0 \end{pmatrix} G(u(s,a))ds\\
&\quad + \int_0^\infty \begin{pmatrix} 0 & 0 \\ 0 & e^{Q(t-s)} \end{pmatrix} G(u(s,a))ds\\
&\quad - \int_t^\infty \begin{pmatrix} 0 & 0 \\ 0 & e^{Q(t-s)} \end{pmatrix} G(u(s,a))ds \text{ (remove terms that blow up to get)}
\end{aligned}$$

Copyright Elsevier 2022

$$
\begin{aligned}
&= \begin{pmatrix} e^{Pt} & 0 \\ 0 & 0 \end{pmatrix} a + \int_0^t \begin{pmatrix} e^{P(t-s)} & 0 \\ 0 & 0 \end{pmatrix} G(u(s,a))ds \\
&- \int_t^\infty \begin{pmatrix} 0 & 0 \\ 0 & e^{Q(t-s)} \end{pmatrix} G(u(s,a))ds \\
&= U(t)a + \int_0^t U(t-s)G(u(s,a))ds - \int_t^\infty V(t-s)G(u(s,a))ds.
\end{aligned}
$$

It is evident that the last $n-k$ components of a do not enter the computation, and so we take them to be zero. As a consequence, let $u(t,a)$ be given by (6.7.8), and let us examine what that implies for the initial conditions $u(0,a)$. First, we notice that

$$
\begin{aligned}
&u_j(0,a) = a_j, \quad j = 1, 2, \dots, k, \\
&u_j(0,a) = -\Big(\int_0^\infty V(-s)G(u(s,a))ds \Big)_j, \quad j = k+1, \dots, n.
\end{aligned}
$$

Hence the last $n-k$ components of the initial conditions satisfy

$$a_j = \psi_j(a_1, \dots a_k) := u_j(0, a_1, \dots, a_k, 0 \dots, 0), \quad j = k+1, \dots, n.$$

Finally, the stable manifold $W^s(0)$ is defined by

$$W^s = \{(y_1, \dots, y_n) : y_j = \psi_j(y_1, \dots, y_k), \quad j = k+1, \dots, n\}.$$

Here is the iterative scheme for calculating an approximation to W^s:

(1) Calculate the approximate solution $u^{(m)}(t,a)$.

(2) For each $j = k+1, \dots, n$, the function $\psi_j(a_1, \dots, a_k)$ is given by the jth component of $u^{(m)}(0,a)$.

To find the unstable manifold W^u, do the same by replacing t by $-t$, that is,

$$y' = -By - G(y).$$

Then the stable manifold for this system will be the unstable manifold for (6.7.7). During the process, we must replace the vector y by the vector $(y_{k+1}, \dots, y_n, y_1, \dots, y_k)$ to determine the $(n-k)$-dimensional manifold W^u by the above process. □

We end this section with the following popular example.

Example 6.21. Consider the planar nonlinear system

$$
\begin{cases}
x_1' = -x_1 - x_2^2, \\
y' = x_1^2 + x_2
\end{cases}
$$

with the only equilibrium solution $(0,0)$. Then

Copyright Elsevier 2022

$$A = Df(0) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The eigenvalues of A are $\lambda = 1, -1$ with corresponding eigenvectors $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, respectively. Thus the matrix C is given by $C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. It follows that

$$A = B = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad F(x) = G(x) = \begin{pmatrix} -x_2^2 \\ x_1^2 \end{pmatrix},$$

and

$$U(t) = \begin{pmatrix} e^{-t} & 0 \\ 0 & 0 \end{pmatrix}, \quad V(t) = \begin{pmatrix} 0 & 0 \\ 0 & e^t \end{pmatrix}.$$

Since $n = 2$ and $k = 1$, the last components $n - k$ of the n vector a are zero. So we take $a = \begin{pmatrix} a_1 \\ 0 \end{pmatrix}$. Our goal is to find the first three successive approximations $u^{(1)}(t, a)$, $u^{(2)}(t, a)$, and $u^{(3)}(t, a)$ and use $u^{(3)}(t, a)$ to approximate the function ψ_2 describing the stable manifold $W^s : x_2 = \psi_2(x_1)$. Define the iterate of the function $u(t, a)$ using (6.7.8)

$$u^{(k+1)}(t, a) = \begin{pmatrix} e^{-t} & 0 \\ 0 & 0 \end{pmatrix} a + \int_0^t \begin{pmatrix} e^{-(t-s)} & 0 \\ 0 & 0 \end{pmatrix} G(u^{(k)}(s, a))ds$$
$$- \int_t^\infty \begin{pmatrix} 0 & 0 \\ 0 & e^{(t-s)} \end{pmatrix} G(u^{(k)}(s, a))ds,$$

where

$$u^{(0)}(t, a) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$u^{(1)}(t, a) = U(t)a = \begin{pmatrix} e^{-t}a_1 \\ 0 \end{pmatrix},$$

$$u^{(2)}(t, a) = \begin{pmatrix} e^{-t} & 0 \\ 0 & 0 \end{pmatrix} a + \int_0^t \begin{pmatrix} e^{-(t-s)} & 0 \\ 0 & 0 \end{pmatrix} G(u^{(1)}(s, a))ds$$
$$- \int_t^\infty \begin{pmatrix} 0 & 0 \\ 0 & e^{(t-s)} \end{pmatrix} G(u^{(1)}(s, a))ds$$
$$= \begin{pmatrix} e^{-t}a_1 \\ 0 \end{pmatrix} - \int_t^\infty \begin{pmatrix} 0 \\ e^{(t-s)}e^{-2s}a_1^2 \end{pmatrix} ds = \begin{pmatrix} e^{-t}a_1 \\ -\frac{1}{3}e^{-2t}a_1^2 \end{pmatrix},$$

Copyright Elsevier 2022

$$
\begin{aligned}
u^{(3)}(t,a) &= \begin{pmatrix} e^{-t} & 0 \\ 0 & 0 \end{pmatrix} a - \int_0^t \begin{pmatrix} e^{-(t-s)} & 0 \\ 0 & 0 \end{pmatrix} G(u^{(2)}(s,a))ds \\
&\quad - \int_t^\infty \begin{pmatrix} 0 & 0 \\ 0 & e^{(t-s)} \end{pmatrix} G(u^{(2)}(s,a))ds \\
&= \begin{pmatrix} e^{-t}a_1 \\ 0 \end{pmatrix} - \int_0^t \begin{pmatrix} \frac{e^{-(t-s)}e^{42s}a_1^4}{9} \\ 0 \end{pmatrix} ds - \int_t^\infty \begin{pmatrix} 0 \\ e^{(t-s)}e^{-2s}a_1^2 \end{pmatrix} ds \\
&= \begin{pmatrix} e^{-t}a_1 + \frac{1}{27}(e^{-4t} - e^{-t})a_1^4 \\ -\frac{1}{3}e^{-2t}a_1^2 \end{pmatrix}.
\end{aligned}
$$

If we approximate $u(t,a)$ by $u^{(3)}(t,a)$, then

$$
u(t,a) = \begin{pmatrix} e^{-t}a_1 + \frac{1}{27}(e^{-4t} - e^{-t})a_1^4 \\ -\frac{1}{3}e^{-2t}a_1^2 \end{pmatrix} = \begin{pmatrix} u_1(t,a_1,0) \\ u_2(t,a_1,0) \end{pmatrix},
$$

and

$$
\psi_1(a_1) = u_1(0,a_1,0),\ \psi_2(a_1) = u_2(0,a_1,0)
$$

are approximated by

$$
\psi_1(a_1) = a_1, \quad \psi_2(a_1) = -\frac{1}{3}a_1^2.
$$

Thus the stable manifold is approximated by

$$
W^s = \{(y_1,y_2) \in \mathbb{R}^2 : y_2 = -\frac{1}{3}y_1^2\}.
$$

However, since the matrix $C = I$ for this example, the x and y spaces are the same. It follows that the stable manifold is given by

$$
W^s = \{(x_1,x_2) \in \mathbb{R}^2 : x_2 = -\frac{1}{3}x_1^2\}.
$$

The unstable manifold W^u can be found by applying the same procedure to the system with $t \to -t$ and x_1 and x_2 interchanged. Thus the stable manifold for the resulting system will be the unstable manifold for the original system. With this in mind, we arrive at the unstable manifold

$$
W^u = \{(x_1,x_2) \in \mathbb{R}^2 : x_1 = -\frac{1}{3}x_2^2\}.
$$

Remark 6.5. We used $u^{(3)}(t,a)$ to approximate $u(t,a)$, and so

$$
u(t,a) \approx \begin{pmatrix} e^{-t}a_1 + \frac{1}{27}(e^{-4t} - e^{-t})a_1^4 \\ -\frac{1}{3}e^{-2t}a_1^2 \end{pmatrix} \approx \begin{pmatrix} e^{-t}a_1 \\ -\frac{1}{3}e^{-2t}a_1^2 \end{pmatrix},
$$

Copyright Elsevier 2022

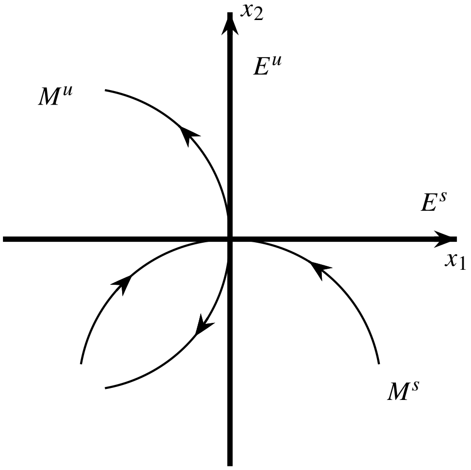


FIGURE 6.16

Local approximations of M^s and M^u of Example 6.21.

when $|a| = |(a_1, 0)^T|$ is small, or when a_1 is small. Thus the stable manifold can be approximated by

$$\begin{aligned}\psi_2(a_1) &= -\Big(\int_0^\infty V(-s)G(u(s,a))ds\Big)_2 \\ &= -\Big(\int_0^\infty \begin{pmatrix} 0 & 0 \\ 0 & e^{-s} \end{pmatrix} \begin{pmatrix} -\frac{1}{9}(e^{-4s} - e^{-t})a_1^4 \\ e^{-2s}a_1^2 \end{pmatrix} ds\Big)_2 \\ &= -\int_0^\infty e^{-s}e^{-2s}a_1^2 ds = -\frac{1}{3}a_1^2.\end{aligned}$$

Fig. 6.16 shows the stable and unstable manifolds in the neighborhood of (0, 0).

6.7.2 Global manifolds

The stable and unstable manifolds of Section 6.7.1 are local since they are defined in a small neighborhood of the origin. For that reason, they are referred to as the *local* stable and unstable manifolds at the origin. In this section, we define the *global manifolds* around the origin.

Definition 6.7.5. Let ϕ_t be the flow of (6.7.1). The *global stable* and *unstable manifolds* of (6.7.1) at the origin are defined by

$$G^s(0) = \cup_{t \le 0}\phi_t(W^s)$$

Copyright Elsevier 2022

and

$$G^u(0) = \cup_{t\geq 0}\phi_t(W^s),$$

respectively.

The global stable and unstable manifolds are unique and invariant with respect to the flow. Furthermore, for all $x \in G^s(0)$, $\lim_{t\to\infty}\phi_t(x) = 0$, and for all $x \in G^u(0)$, $\lim_{t\to-\infty}\phi_t(x) = 0$.

For the next theorem, we define a neighborhood around zero. Let $\delta > 0$. Define the neighborhood of the origin by

$$N_\delta(0) = \{x \in \mathbb{R}^n : ||x|| < \delta\}.$$

Theorem 6.7.3. *Assume that the hypotheses of Theorem 6.7.2 hold. Suppose the linearized matrix $Df(0)$ has k eigenvalues with negative real parts and $n-k$ eigenvalues with positive real parts. Set*

$$\mathrm{Re}(\lambda_j) < -\alpha < 0 < \beta < \mathrm{Re}(\lambda_m)$$

for $j = 1, \ldots, k$ and $m = k+1, \ldots, n$.

Then, given $\varepsilon > 0$, there exists $\delta > 0$ such that if $p \in N_\delta(0) \cap W^s$, then

$$|\phi_t(p)| \leq \varepsilon e^{-\alpha t}$$

for all $t \geq 0$. In a similar fashion, if $p \in N_\delta(0) \cap W^u$, then

$$|\phi_t(p)| \leq \varepsilon e^{\beta t}$$

for all $t \leq 0$.

Theorem 6.7.3 states that if solutions start in the stable manifold W^s, then they approach the origin exponentially. This is displayed in Example 6.19.

The center manifold theorem gives a detailed description of the dynamic of (6.7.1) in a neighborhood of a hyperbolic equilibrium once the center manifold W^c is determined.

6.7.3 Center manifold

Theorem 6.7.4. *Suppose $f(0) = 0$ and let $E \subseteq \mathbb{R}^n$ be an open subset containing the origin with $f \in C^1(E)$. Let $\phi_t : E \to E$ be the flow of (6.7.1) defined for all $t \in \mathbb{R}$. Let*

$$x' = Ax := Df(0)x. \tag{6.7.9}$$

Suppose the linearized matrix $Df(0)$ has k eigenvalues with negative real parts, j eigenvalues with positive real parts, and $m = n-k-j$ eigenvalues with zero real parts. Then there exists an m-dimensional differentiable manifold $W^c(0)$ that is

Copyright Elsevier 2022

tangent to the center subspace E^s of (6.7.9)*, a j-dimensional differentiable manifold $W^u(0)$ that is tangent to the unstable space E^u of the linearized system* (6.7.9)*, and an m-dimensional differentiable manifold W^c that is tangent to the center space E^c of the linearized system* (6.7.9)*. Moreover,*

$$\lim_{t\uparrow\infty}\phi_t(p)=0 \ \textit{ for any } \ p\in W^s,$$

$$\lim_{t\downarrow-\infty}\phi_t(p)=0 \ \textit{ for any } \ p\in W^u,$$

$$\lim_{t\downarrow-\infty}\phi_t(p)=0 \ \textit{ for any } \ p\in W^c.$$

In other words, the manifolds W^s, W^u, and W^c are invariant under the flow ϕ_t of (6.7.1).

Example 6.22. Consider the planar nonlinear system

$$\begin{cases} x_1' = x_1^2, \\ x_2' = -x_2 \end{cases}$$

with the only equilibrium solution $(0,0)$. Linearization around the equilibrium solution gives

$$A = Df(0) = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

The eigenvalues of A are $\lambda = 0, -1$ with corresponding eigenvectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Thus the stable subspace is $E^s = \{(x_1,x_2)\in\mathbb{R}^2 : x_1 = 0\}$, and the center subspace is $E^c = \{(x_1,x_2)\in\mathbb{R}^2 : x_2 = 0\}$. For constants c_1 and c_2, the system has the solution

$$x_1(t) = \frac{c_1}{1-c_1 t}, \quad x_2(t) = c_2 e^{-t}.$$

Note that if $c_1 = 0$, then $x_1 = 0$. Moreover, to get a better understanding of the dynamics of the solution, we eliminate t from both equations and arrive at the following relation between x_1 and x_2:

$$x_2 = c_2 e^{-\frac{1}{c_1}+\frac{1}{x_1}}.$$

Since $x_1 = 0$ for $c_1 = 0$, we have that the plane curves

$$x_2 = \Psi(x_1,c_1,c_2) = \begin{cases} c_2 e^{-\frac{1}{c_1}+\frac{1}{x_1}} & \text{for } x_1 < 0, \\ 0 & \text{for } x_1 = 0 \end{cases} \qquad (6.7.10)$$

Copyright Elsevier 2022

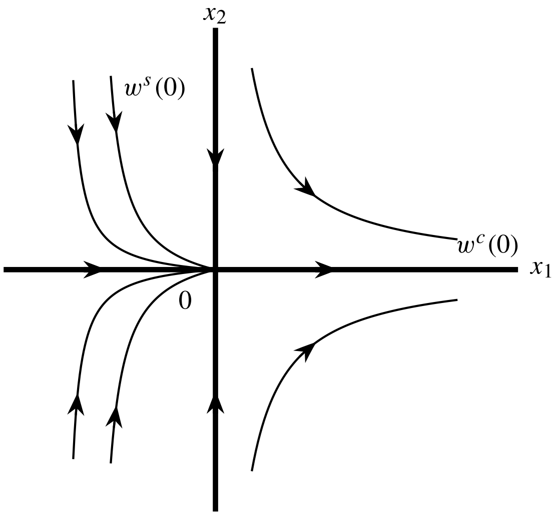


FIGURE 6.17

Center manifold.

satisfy

$$\Psi(0, c_1, c_2) = 0, \quad \frac{\partial \Psi}{\partial x_1}(0, c_1, c_2) = 0.$$

Now $c_1 < 0$ if and only if $x_1 < 0$ by $x_1(t) = \frac{c_1}{1-c_1 t}$ for all $t \geq 0$. It follows that when $x_1 < 0$ is sufficiently small, (6.7.10) is a local center manifold that is tangent to E^c at the origin. However, since (6.7.10) holds for any $c_1 < 0$ and c_2, there are infinitely many center manifolds, and all have the same orientations as seen in Fig. 6.17.

6.7.4 Center manifold and reduced systems

Let $E \subseteq \mathbb{R}^n$ contain the origin, and let $f : E \to \mathbb{R}^n$ be C^2. Consider the autonomous nonlinear differential equation

$$x' = f(x). \tag{6.7.11}$$

Since $f \in C^1(E)$ and $f(0) = 0$, we may write system (6.7.11) as

$$x' = Ax + F(x), \tag{6.7.12}$$

where $A = Df(0)$ and $F(x) = f(x) - Ax$ with $F \in C^1(E)$, $F(0) = 0$, and $DF(0) = 0$. There is an $n \times n$ invertible matrix C such that

$$C^{-1}AC = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix},$$

Copyright Elsevier 2022

where the $k \times k$ matrix P has eigenvalues λ_i, $i = 1, 2, \dots, k$, each with zero real part, and the $(n-k) \times (n-k)$ matrix Q has the eigenvalues λ_i, $i = k+1, k+2, \dots, n$, each with negative real part. Make the change of variables

$$\begin{pmatrix} y \\ z \end{pmatrix} = C^{-1}x, \quad y \in \mathbb{R}^k, \quad z \in \mathbb{R}^{n-k}$$

and transform (6.7.12) into the form

$$\begin{aligned} y' &= Px + F_1(y, z), \\ z' &= Qz + F_2(y, z), \end{aligned} \tag{6.7.13}$$

where

$$F_i(0,0) = 0, \ \frac{\partial F_i}{\partial y}(0,0) = \frac{\partial F_i}{\partial z}(0,0), \quad i = 1, 2.$$

If $z = h(y)$ is an invariant manifold for (6.7.13) with

$$h(0) = 0 \ \text{ and } \ \frac{\partial h}{\partial y}(0) = 0, \tag{6.7.14}$$

then we define the local *center manifold*

$$W^c = \{(y, z) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : y = h(y), \ \ ||y|| < \delta\}$$

for small $\delta > 0$. This way, W^c is tangent to E^c at the origin. We have the following theorem.

Theorem 6.7.5. *For system* (6.7.13)*, there exist* $\delta > 0$ *and* $h \in C^1$ *with* $h(y)$ *defined for all* $||y|| < \delta$ *such that* $z = h(y)$ *is a center manifold.*

On the center manifold W^c, we get the *reduced system*

$$y' = Px + F_1(y, h(y)),$$

and z' can be calculated on the center manifold in two ways: directly from the z' equation above or by differentiating $z = h(y)$, that is,

$$z' = Qz + F_2(y, h(y)) \ \text{ and } \ z' = \frac{d}{dt}h(y) = Dh(y)y' = Dh(y)[Px + F_1(y, h(y))],$$

where $Dh(y)$ is the matrix formed by partial derivatives of $h(y)$. Note that in one dimension, $Dh(y) = h'(y)$. Expanding h as a Taylor series (noting that the constant and linear terms vanish), the two equations for z' provide two different polynomials, and the coefficients of different monomials can be equated to determine the coefficients of the Taylor expansion. We end the section with the following two examples.

Copyright Elsevier 2022

Example 6.23. Consider the planar nonlinear system

$$\begin{cases} x' = xy, \\ y' = -x^2 - y \end{cases}$$

with the only equilibrium solution $(0, 0)$. Linearization around the equilibrium solution gives

$$A = Df(0) = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

The eigenvalues of A are $\lambda = 0, -1$ with corresponding eigenvectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Thus the stable subspace is $E^s = \{(x, y) \in \mathbb{R}^2 : x = 0\}$, and the center subspace is $E^c = \{(x, y) \in \mathbb{R}^2 : y = 0\}$. Since the matrix A is already in normal form, no coordinate transformation is needed. We consider a center manifold of the form

$$y = h(x) = ax^2 + bx^3 + cx^4 + \mathscr{O}(x^5). \tag{6.7.15}$$

There is no constant term and no linear term in $h(x)$ since the center manifold passes through the origin and should be tangent to E^c. Now we calculate the constants. We have

$$y' = -x^2 - y = -x^2 - \left(ax^2 + bx^3 + cx^4\right) + \mathscr{O}(x^5). \tag{6.7.16}$$

Next, we differentiate (6.7.15) with respect to t. Using the chain rules, we get $y' = \dfrac{dh(x)}{dt} = x'h'(x)$ or

$$y' = x'h'(x) = xyh'(x) = x\left(ax^2 + bx^3 + cx^4 + \ldots\right)\left(2ax + 3bx^2 + 4cx^3 + \ldots\right). \tag{6.7.17}$$

Setting the right side of (6.7.17) equal to (6.7.16) and equating the coefficients at x^2, x^3, and x^4 give

$$a = -1, \quad b = 0, \quad \text{and} \quad c = -2.$$

Thus the center manifold is given by

$$W^c = \{(x, y) \in \mathbb{R}^2 : y = -x^2 - 2x^4\},$$

and the dynamics on W^c is

$$x' = xh(x) = -x^3 - 2x^5 + \mathscr{O}(x^7).$$

Now for $x > 0$, we have $x' < 0$ and $x' > 0$ if $x < 0$. Thus the origin is stable, and the solutions are stable nodes as shown in the graph below, but the motion onto the center manifold in the y-direction is much faster than the motion on the center manifold, leading to a phase portrait shown in Fig. 6.18.

Copyright Elsevier 2022

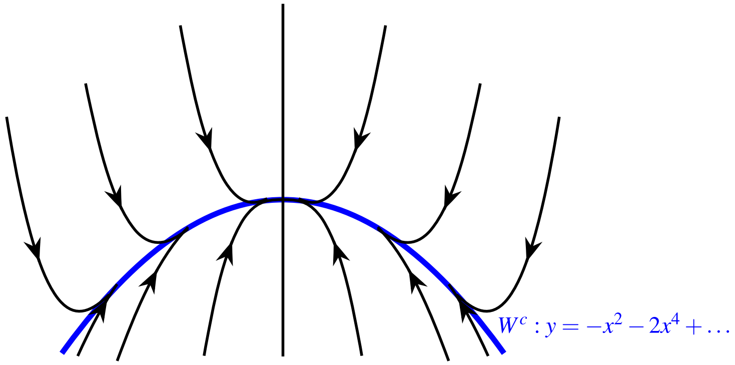


FIGURE 6.18

Exponential collapse of trajectories onto the center manifold and then slow convergence toward $(0, 0)$ on the center manifold.

The next example shows the need for a transformation.

Example 6.24. Consider the planar nonlinear system

$$\begin{cases} x' = y - x + xy, \\ y' = x - y - x^2 \end{cases} \tag{6.7.18}$$

with the only equilibrium solution $(0, 0)$. Linearization around the equilibrium solution gives

$$A = Df(0) = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}.$$

The eigenvalues of A are $\lambda = 0, -2$ with corresponding eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Thus the stable subspace is $E^s = \{(x, y) \in \mathbb{R}^2 : y = -x\}$, and the center subspace is $E^c = \{(x, y) \in \mathbb{R}^2 : y = x\}$.

Let $C = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Then $C^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Make the change of variable from x, y coordinates to u, v coordinates by

$$\begin{pmatrix} u \\ v \end{pmatrix} = C^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x + y \\ y - x \end{pmatrix}$$

or

$$\begin{pmatrix} x \\ y \end{pmatrix} = C \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u - v \\ u + v \end{pmatrix}.$$

Copyright Elsevier 2022

Under these transformations, system (6.7.18) in the u, v coordinates becomes

$$u' = uv - v^2, \quad v' = -2v + uv - u^2.$$

The new system is now in the normal form since

$$Df(0) = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}.$$

For similar reasons as in the above example, we define the center manifold $h(u)$ by

$$v(u) = h(u) = au^2 + bu^3 + cu^4 + \cdots .$$

Then using the chain rule, we arrive at

$$\begin{aligned} v' &= \left[2au + 3b^2u^2 + 4cu^3 + \dots\right]u' \\ &= \left[2au + 3b^2u^2 + 4cu^3 + \dots\right] \\ &\times \left[u\left(au^2 + bu^3 + cu^4 + \dots\right) - \left(au^2 + bu^3 + cu^4 + \dots\right)^2\right] \\ &= 2a^2u^4 + (5ab - 2a^3)u^5 + \cdots . \end{aligned} \tag{6.7.19}$$

On the other hand, using the second equation of the system with $v = h(u)$ gives

$$\begin{aligned} v' &= -2v + uv - u^2 \\ &= -2\left(au^2 + bu^3 + cu^4 + \dots\right) + u\left(au^2 + bu^3 + cu^4 + \dots\right) - u^2 \\ &= (2a + 1)u^2 + (2b - a)u^3 + (3c - b)u^4 + \cdots . \end{aligned} \tag{6.7.20}$$

Setting (6.7.19) equal to (6.7.20) and then equating the coefficients at u^2, u^3, and u^4 give

$$a = -\frac{1}{2}, \quad b = -\frac{1}{4}, \quad c = -\frac{3}{8},$$

and it follows that the center manifold in uv coordinates is

$$W^c = \{(u, v) \in \mathbb{R}^2 : v = -\frac{1}{2}u^2 - \frac{1}{4}u^3 - \frac{3}{8}u^4 + \dots\}.$$

The dynamics on the center manifold is

$$u' = uv - v^2 = -\frac{1}{2}u^3 - \frac{1}{4}u^4 - \frac{3}{8}u^5 + \cdots ,$$

which is stable for small u. Finally, the center manifold W^c in the original coordinates x, y is approximated from replacing u with $x + y$ and v with $y - x$:

$$W^c = \left\{(x, y) \in \mathbb{R}^2 : y - x = -\frac{1}{4}(x + y)^2 - \frac{1}{16}(x + y)^3 - \frac{3}{64}(x + y)^4 + \dots\right\}.$$

Copyright Elsevier 2022

6.7.5 Hartman–Grobman theorem

Consider the nonlinear autonomous system

$$x' = f(x), \tag{6.7.21}$$

where $E \subseteq \mathbb{R}^n$, and $f : E \to \mathbb{R}^n$ is a C^2 function. The *Hartman–Grobman theorem* states that near a hyperbolic equilibrium solution x_0, the nonlinear system given by (6.7.21) has the same qualitative structures as the linearized system

$$x' = Ax, \tag{6.7.22}$$

where $A = Df(x_0)$. In this section, we assume that x_0 is the origin, $E \subseteq \mathbb{R}^n$ contains the origin, and $f(0) = 0$. Recall from Definition 6.7.3 that two sets K_1 and K_2 are said to be *homeomorphic* if there is a continuous one-to-one map $g : K_1 \to K_2$ such that g^{-1} is also continuous.

In our case the two autonomous systems (6.7.21) and (6.7.22) are said to be topologically equivalent in a neighborhood of the origin or to have the same qualitative structure near the origin if there is a homeomorphism H mapping an open set U containing the origin onto an open set V containing the origin that maps trajectories of (6.7.21) in U onto trajectories (6.7.22) in V and preserves the orientation. If the homeomorphism H preserves parameterization by time t, then (6.7.21) and (6.7.22) are said to be topologically conjugate in the neighborhood of the origin.

We have the following example.

Example 6.25. Consider the two linear systems $x' = Ax$ and $y' = By$ with

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

Define

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}; \text{ then } R^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

It is easy to verify that $B = RAR^{-1}$. Let $y = H(x) = Rx$ or $x = R^{-1}y$. Then it follows that

$$y' = RAR^{-1}y = By.$$

Suppose $x(0) = x_0$. Then $x(t) = e^{At}x_0$ is the solution of $x' = Ax$, $x(0) = x_0$. Similarly, if $y(0) = y_0$, then $y = e^{Bt}y_0$ with $y_0 = Rx_0$.

Then by (3.4.3) it follows that

$$y(t) = H(x(t)) = Rx(t) = Re^{At}x_0 = Re^{At}R^{-1}y_0 = e^{Bt}y_0 = e^{Bt}Rx_0$$

is the solution of $x' = Bx$. In other words, H maps the trajectories of $x' = Ax$ onto trajectories of $x' = Bx$ and preserves the parameterization by t since

$$He^{At} = e^{Bt}H.$$

Copyright Elsevier 2022

Thus H is a homeomorphism, and $x' = Ax$ and $y' = By$ are topologically conjugate.

Next, we introduce the Hartman–Grobman theorem, proved independently by Grobman in 1959 and Hartman in 1960. We only offer an outline of the proof, which should serve as guidelines on how to find the homeomorphism H. For complete proof, we refer the reader to [26] or [33]. In plain language the theorem says that if the linearization matrix has no zero or purely imaginary eigenvalues, then the phase portrait for the nonlinear system near an equilibrium point is similar to that of its linearization.

Theorem 6.7.6. *(Hartman–Grobman theorem) Let $f(0) = 0$, and let $E \subseteq \mathbb{R}^n$ be an open subset containing the origin with $f \in C^1(E)$. Let ϕ_t be the flow of (6.7.21). Suppose the linearized matrix $Df(0)$ has no eigenvalue with zero real part (the origin is hyperbolic). Then there is a homeomorphism H mapping an open set U containing the origin onto an open set V containing the origin such that for each $x_0 \in U$, there is an open interval $I_0 \subseteq \mathbb{R}$ containing zero such that for all $x_0 \in U$ and $t \in I_0$, we have*

$$H \circ \phi_t(x_0) = e^{At} H(x_0).$$

In other words, H maps trajectories of (6.7.21) in U onto trajectories (6.7.22) in V and preserves the orientation and parameterization by time t.

Proof. (Outline) Suppose that

$$A = Df(0) = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix},$$

where all the eigenvalues of the matrix P have negative real parts and all the eigenvalues of the matrix Q have positive real parts. Let

$$x_0 = \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \in \mathbb{R}^n,$$

let ϕ_t be the flow of (6.7.21) and write the solution

$$x(t, x_0) = \phi_t(x_0) = \begin{pmatrix} y(t, y_0, z_0) \\ z(t, y_0, z_0) \end{pmatrix}$$

for y_0 and z_0 in E^s and E^u, respectively.

Define locally ($t = 1$) the transformation

$$T^1 x_0 = \begin{pmatrix} e^P y_0 + Y(x_0) \\ e^Q z_0 + Z(x_0) \end{pmatrix}.$$

Now for $x \in \mathbb{R}^n$, define

$$H_0(x) = \begin{pmatrix} \Phi(x) \\ \Psi(x) \end{pmatrix} = \begin{pmatrix} \Phi(y, z) \\ \Psi(y, z) \end{pmatrix}.$$

Copyright Elsevier 2022

Then

$$H_0 \circ T^1 = e^A \circ H_0,$$

which component-wise is equivalent to

$$\begin{aligned} e^P \Phi(x) &= \Phi(e^P y + Y(x), e^Q z + Z(x)), \\ e^Q \Psi(x) &= \Phi(e^P y + Y(x), e^Q z + Z(x)). \end{aligned} \tag{6.7.23}$$

Next, we define a continuous successive approximation from the second equation by

$$\begin{cases} \Psi_0(x) = z, \\ \Psi_{k+1}(x) = e^{-Q}\Phi_k(e^P y + Y(x), e^Q z + Z(x)), \ k = 0, 1, \dots. \end{cases}$$

Then $\Psi_k(x)$ is a Cauchy sequence of continuous functions, which converges uniformly as $k \to \infty$ to a continuous function $\Psi(x)$. We write the first equation in (6.7.23) by

$$e^{-P}\Phi(x) = \Phi(e^{-P} y + Y_1(x), e^{-Q} z + Z_1(x)), \tag{6.7.24}$$

where the functions Y_1 and Z_1 are defined by the inverse of T,

$$T^{-1}(y, z) = \begin{pmatrix} e^{-P} y + Y_1(y, z) \\ e^{-Q} z + Z_1(y, z) \end{pmatrix}.$$

With this setup, Eq. (6.7.24) can be solved by the method of successive approximations exactly as above with $\Phi_0(x) = y$. Finally, define the homeomorphism

$$H = \int_0^1 e^{-As} H_0 T^s ds = \int_0^1 e^{-As} H_0 \circ \phi_s(x, y) ds,$$

which satisfies

$$H \circ \phi_t(x, y) = e^{At} H(x, y).$$ □

Remark 6.6. (1) From the computational aspect, Theorem 6.7.6 is difficult to use since it requires the computation of the flow, which is not possible in most nonlinear systems.
(2) However, conceptually, it suffices to know that such a homeomorphism exists. This allows us to understand the dynamics of the nonlinear system near a hyperbolic equilibrium solution by simply looking at the eigenvalues of the linear system $x' = Ax$, $A = Df(0)$, obtained from linearizing the nonlinear system near the origin.

In the next example, we consider a system for which the flows can be computed.

Example 6.26. Consider

$$\begin{cases} y' = -y, \\ z' = z + y^2 \end{cases}$$

Copyright Elsevier 2022

with

$$x(0)=\begin{pmatrix}y_0\\ z_0\end{pmatrix},$$

$$Df(0)=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix},$$

and hence $e^P=e^{-1}$, and $e^Q=e$. The equation $y'=-y$ has the solution $y=y_0e^{-t}$. Substituting this solution into the second differential equation and using the variation of parameters formula, we arrive at

$$z=z_0e^t+\frac{1}{3}y_0^2(e^t-e^{-2t}).$$

Thus

$$\phi_t(x_0)=\begin{pmatrix}y_0e^{-t}\\ z_0e^t+\frac{1}{3}y_0^2(e^t-e^{-2t})\end{pmatrix},$$

and at $t=1$, $Y(x)=0$, $Z(x)=\frac{1}{3}y_0^2(e-e^{-2})$, where $x=(y,z)$. As a consequence,

$$e^Q\Psi(x)=\Phi(e^Py+Y(x),e^Qz+Z(x)),$$

which results in

$$e\Psi(x)=\Phi(e^{-1}y,ez+\frac{1}{3}y_0^2(e-e^{-2})).$$

Therefore the successive approximations are given by

$$\Psi_0(x)=z,\quad x=(y,z)\in\mathbb{R}^2,$$

$$\begin{aligned}\Psi_1(x)&=e^{-1}\Psi_0(e^{-1}y,ez+\frac{1}{3}y^2(e-e^{-2}))\\ &=e^{-1}[ez+\frac{1}{3}y^2(e-e^{-2})]\\ &=z+\frac{1}{3}(1-e^{-3})y^2,\end{aligned}$$

$$\begin{aligned}\Psi_2(x)&=e^{-1}\Psi_1(e^{-1}y,ez+\frac{1}{3}(e-e^{-2})y^2)\\ &=e^{-1}[(ez+\frac{1}{3}(e-e^{-2})y^2+\frac{1}{3}(1-e^{-3})e^{-2}y^2]\\ &=z+\frac{1}{3}(1-e^{-3})(y^2+e^{-3}y^2)\\ &=z+\frac{1}{3}(1-e^{-3})(1+e^{-3})y^2,\end{aligned}$$

Copyright Elsevier 2022

$$\begin{aligned}
\Psi_3(x) &= e^{-1}\Psi_2(e^{-1}y, ez + \frac{1}{3}(e - e^{-2})y^2) \\
&= e^{-1}[(ez + \frac{1}{3}(e - e^{-2})y^2 + \frac{1}{3}(1 - e^{-3})(1 + e^{-3})e^{-2}y^2] \\
&= z + \frac{1}{3}(1 - e^{-3})(y^2 + e^{-3}y^2) \\
&= z + \frac{1}{3}(1 - e^{-3})(1 + e^{-3} + e^{-6})y^2, \\
&\dots
\end{aligned}$$

$$\begin{aligned}
\Psi_m(x) &= z + \frac{1}{3}(1 - e^{-3})(1 + e^{-3} + e^{-6} + \ldots + (e^{-3})^{m-1})y^2 \\
&= z + \frac{1}{3}(1 - e^{-3})y^2 \sum_{n=1}^{m} (\frac{1}{e^3})^{n-1} \to z + \frac{1}{3}(1 - e^{-3})\frac{1}{1 - e^{-3}}y^2 \\
&= z + \frac{1}{3}y^2 \text{ as } m \to \infty.
\end{aligned}$$

Thus as $m \to \infty$,

$$\Psi_m(y,z) \to \Psi(y,z) = z + \frac{1}{3}y^2$$

uniformly for all $(y,z) \in \mathbb{R}^2$. Now the functions Y_1 and Z_1 are defined by the inverse of T,

$$T^{-1}(y,z) = \begin{pmatrix} e^{-P}y + Y_1(y,z) \\ e^{-Q}z + Z_1(y,z) \end{pmatrix}.$$

However,

$$T^{-1}(y,z) = \begin{pmatrix} ey \\ e^{-1}z - e\frac{e^3-1}{3e^2}y^2 \end{pmatrix},$$

which implies that $Y_1(y,z) = 0$ and $Z_1(y,z) = -ery^2$, where $r = \frac{e^3-1}{3e^2}$. To fully determine $\Phi(y,z)$, we use (6.7.24):

$$e^{-P}\Phi(x) = \Phi(e^{-P}y + Y_1(x), e^{-Q}z + Z_1(x))$$

or

$$e\Phi(x) = \Phi(ey, e^{-1}z - ery^2).$$

We define the successive approximations by

$$\begin{gathered}
\Phi_0(y,z) = y, \\
\Phi_{k+1}(y,z) = e^{-1}\Phi_k(ey, e^{-1}z - ery^2), \ k = 0, 1, \ldots,
\end{gathered}$$

Copyright Elsevier 2022

and get

$$\Phi(y,z)=y,\ \ (y,z)\in\mathbb{R}^2.$$

Thus

$$H_0(y,z)=\begin{pmatrix} y \\ z+\frac{1}{3}y^2 \end{pmatrix},$$

and the homeomorphism H is given by

$$\begin{aligned} H(y,z) &= \int_0^1 e^{-As}H_0T^s ds \\ &= \int_0^1 \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix} H_0[\begin{pmatrix} ye^{-s} \\ ze^s+\frac{1}{3}y^2(e^s-e^{-2s}) \end{pmatrix}]ds \\ &= \int_0^1 \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix} \begin{pmatrix} ye^{-s} \\ ze^s+\frac{1}{3}y^2(e^s-e^{-2s})+\frac{1}{3}y^2e^{-2s} \end{pmatrix} ds \\ &= \begin{pmatrix} y \\ z+\frac{1}{3}y^2 \end{pmatrix}. \end{aligned}$$

Note that H_0T^s is taken as the composition or H_0 is acting on T^s. Let's dig deeper and see what subset gets mapped onto the stable subspace E^s. Looking at the corresponding eigenvector of $\lambda=-1$, we see that $E^s=\{(y,z)\in\mathbb{R}^2 : z=0\}$. We are interested in finding the subset such that

$$H(y,z)=\begin{pmatrix} y \\ z+\frac{1}{3}y^2 \end{pmatrix}\in E^s.$$

It readily follows that the subset is

$$\{(y,z)\in\mathbb{R}^2 : z=-\frac{1}{3}y^2\},$$

which is the stable manifold W^s by simply taking $t\to\infty$ in the flow. In other words, the homeomorphism H maps W^s onto E^s. In addition, H maps $W^u=\{(y,z)\in\mathbb{R}^2 : y=0\}$ onto $E^u=\{(y,z)\in\mathbb{R}^2 : y=0\}$. Solving the linearized system $y'=-y$, $z'=z$, we get $y=c_1e^{-t}$ and $z=c_2e^t$ or $z=\frac{c}{y}$ for constants c_1, c_2, and c. Thus

$$H(y,z)=H(y,\frac{c}{y}),$$

from which it follows that

$$(y,z+\frac{1}{3}y^2)=(y,\frac{c}{y}).$$

Copyright Elsevier 2022

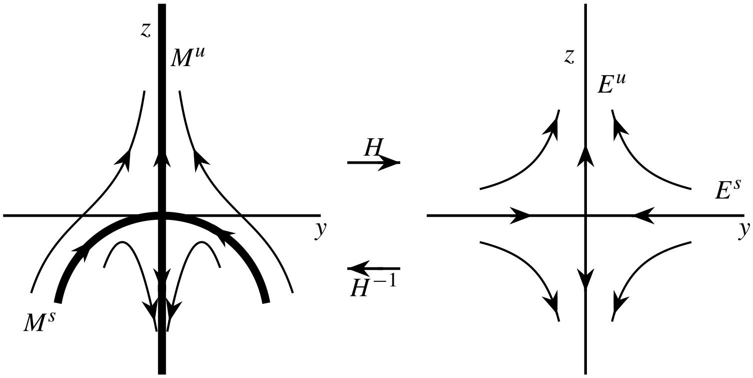


FIGURE 6.19

Under the homeomorphism H, the phase portrait for the nonlinear system near the origin is similar to that of its linearization.

Thus

$$z = \frac{c}{y} - \frac{1}{3}y^2.$$

We just showed that, under H, the curve $z = \frac{c}{y} - \frac{1}{3}y^2$ of the nonlinear system gets mapped onto $z = \frac{c}{y}$ of the linearized system. (See Fig. 6.19.)

6.8 Exercises

Exercise 6.1. For the following equations, find the equilibrium solutions, sketch the phase line, and determine the type of stability of all the equilibria.
(a) $x' = x^2(4 - x)$.
(b) $x' = 2x(1 - x) - \frac{1}{2}x$.
(c) $x' = (4 - x)(2 - x)^3$.

Exercise 6.2. Find the equilibrium solutions in terms of the parameter μ and determine their stability. Show that the origin is a turning point and sketch a bifurcation diagram and label all branches as stable or unstable. When possible, identify the type of bifurcation.
(a) $x' = \mu x(1 - x)$.
(b) $x' = (x^2 - \mu)(1 - x)$.

Exercise 6.3. Find the equilibrium solutions, determine their stability, and show that the origin is a turning point and sketch the bifurcation diagram.

$$x' = \mu - x^2.$$

Copyright Elsevier 2022

Exercise 6.4. For the following equations, show that each of them undergoes a transcritical bifurcation and find the bifurcation value. Sketch the bifurcation diagram.
(a) $x' = \mu x + 2x^2$.
(b) $x' = (2+\mu)x - 5x^2$.

Exercise 6.5. Find the equilibrium solution x_0 at μ_0 and use Theorem 6.1.2 to determine the type of bifurcation at (x_0, μ_0) for the scalar system

$$x' = \mu + x - e^x + 1.$$

Exercise 6.6. Show that

$$x' = \mu x + x^3$$

undergoes a pitchfork bifurcation at $\mu = 0$ and sketch the bifurcation diagram.

Exercise 6.7. Complete the proof of Theorem 6.2.1, that is, show that inequality (6.2.8) holds on $[0, \infty)$.

Exercise 6.8. Suppose the origin is uniformly stable for the linear system $x' = A(t)x$, where $A(t)$ is an $n \times n$ matrix with continuous entries. Suppose that $F : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous with

$$|F(t,x)| \le \gamma(t)|x| \text{ and } \int_0^\infty \gamma(t)dt < \infty,$$

where the function $\gamma(t) > 0$ is continuous. Show that if $F(t, 0) = 0$, then the zero solution of

$$x' = A(t)x + F(t,x),$$

is uniformly stable.

Exercise 6.9. Investigate the stability of each of the equilibrium points for the given systems:
(a) $x' = x - 3y + 2xy$, $y' = 4x - 6y - xy$.
(b) $x' = 6x - 5y + x^2$, $y' = 2x - y + y^2$.
(c) $x' = x - 2y + 3xy$, $y' = 2x - 3y - x^2 - y^2$.
(d) $x' = x - y$, $y' = x^2 - y$.
(e) $x' = y^2 - 1$, $y' = x^3 - y$.
(f) $x' = 8x - y^2$, $y' = -y + x^2$.

Exercise 6.10. Consider the system

$$x' = -x + hy,$$
$$y' = x - y.$$

Show that

Copyright Elsevier 2022

(1) $(0, 0)$ is asymptotically stable if $h = 0$,
(2) $(0, 0)$ is asymptotically stable spiral if $h < 0$, and
(3) $(0, 0)$ asymptotically stable node if $0 < h < 1$.

Exercise 6.11. Study the stability of the origin for the system

$$\begin{aligned} x' &= y + hx(x^2 + y^2), \\ y' &= -x + hy(x^2 + y^2). \end{aligned}$$

(Hint: use polar coordinates.)

Exercise 6.12. Study the stability of the origin for the system

$$\begin{aligned} x' &= -2x - y + x^2 y, \\ y' &= 4x + \mu y - y^2 \end{aligned}$$

for $\mu \neq 2$.

Exercise 6.13. Use polar coordinates to study the stability of the origin for the system

$$x' = -y - x\sqrt{x^2 + y^2}, \quad y' = x - y\sqrt{x^2 + y^2}.$$

Exercise 6.14. Show that the system

$$x' = y + \frac{x(1 - x^2 - y^2)}{\sqrt{x^2 + y^2}}, \quad y' = -x + \frac{y(1 - x^2 - y^2)}{\sqrt{x^2 + y^2}}$$

has a limit cycle.

Exercise 6.15. Show that the system

$$x' = y + x(1 - x^2 - y^2), \quad y' = -x + y(1 - x^2 - y^2)$$

has a limit cycle.

Exercise 6.16. Consider the system

$$\begin{aligned} x' &= 4x + 4y - x(x^2 + y^2), \\ y' &= -4x + 4y - y(x^2 + y^2). \end{aligned}$$

(a) Show that there is a closed path in the region $1 \leq r \leq 3$, where $r^2 = x^2 + y^2$.
(b) Find the general solution.

Exercise 6.17. Show that $x'' + x' + f(x) = 0$ has no periodic solution.

Exercise 6.18. Determine the region on which there is no periodic solution for

$$x'' + \beta(x^2 - 1)x' + x = 0.$$

Copyright Elsevier 2022

Exercise 6.19. Consider the predator–prey model

$$x' = x((1-x)-y), \quad y' = y(x-a),\ a>0,$$

where $x(t)$ and $y(t)$ are the numbers of preys and predators at time t, respectively.
(a) Sketch the nullclines and indicate the direction field.
(b) Find and analyze the equilibrium points.
(c) What can you say for each of the cases $a>1$, $a=1/2$, and $0<a<1$?

Exercise 6.20. (a) Prove Theorem 6.4.2.
(b) Use (a) to show that the system

$$x' = x(A-ax+by), \quad y' = y(B-cy+dx),$$

where a, b, c, d, A, and B are constants with $a,\ c>0$, has no closed trajectories or periodic orbits.
(Hint: let $h(x,y)=\frac{1}{xy}$.)

Exercise 6.21. Use polar coordinates to study the stability of the origin for the system

$$x' = -x - \frac{y}{\mathrm{Ln}(\sqrt{x^2+y^2})}, \quad y' = -y + \frac{x}{\mathrm{Ln}(\sqrt{x^2+y^2})}.$$

Exercise 6.22. Consider the nonlinear system

$$x' = y + \frac{1}{4}x[1-2(x^2+y^2)], \quad y' = -x + \frac{1}{2}y[1-(x^2+y^2)].$$

(a) Show that $(0,0)$ is the only critical point.
(b) Show, using polar coordinates (r,θ), that

$$rr' = \frac{1}{4}r^2[1+\sin^2(\theta)] - \frac{1}{2}r^4.$$

(c) Find the region R of Theorem 6.4.4.

Exercise 6.23. Consider the nonlinear system

$$x' = x-y-x^3, \quad y' = x+y-y^3.$$

Given that $(0,0)$ is the only critical point, use polar coordinates to show that there is a limit cycle in the region

$$R = \{(r,\theta) : 1<r<\sqrt{2}\}.$$

Exercise 6.24. Use Theorem 6.6.1 to show that the system

$$x' = x\mu - y - x^3 - xy^2 - y, \quad y' = x + y\mu - y^3 - x^2y$$

undergoes a Hopf bifurcation, that is, the system has a stable periodic solution with radius $\sqrt{\mu}$.

Copyright Elsevier 2022

Exercise 6.25. Decide if Theorem 6.6.1 applies to the van der Pol equation given in system form by

$$x' = y, \quad y' = -x - \mu(x^2 - 1)y.$$

Exercise 6.26. By solving each of the systems below find E^s, E^u, M^s, and M^u.
(a) $x' = -2x$, $y' = 3x + y$.
(b) $x' = 2x - y$, $y' = -y$.

Exercise 6.27. By solving the following system find E^s, E^u, M^s, and M^u:

$$x_1' = -x_1, \quad x_2' = x_2 + x_1^2, \quad x_3' = 0.$$

Exercise 6.28. Use the method of Example 6.21 to find M^s and M^u for

$$x_1' = -x_1, \quad x_2' = 2x_2 + x_1^2.$$

Exercise 6.29. Use the method of Example 6.21 to find M^s and M^u for

$$x_1' = -x_1 - x_2^3, \quad x_2' = x_2 + x_1^3.$$

(Hint: Find the first four successive approximations to $u^{(1)}(t,a)$, $u^{(2)}(t,a)$, $u^{(3)}(t,a)$, $u^{(4)}(t,a)$.)

Exercise 6.30. In Example 6.23, find the center manifold using

$$y = h(x) = ax^2 + bx^3 + cx^4 + dx^5 + ex^6 + \mathcal{O}(x^7).$$

Exercise 6.31. Find the center manifold W^c of the system

$$\begin{cases} x' = x^2 y, \\ y' = -x^2 - y. \end{cases}$$

Exercise 6.32. Find the center manifold W^c of the system

$$\begin{cases} x' = x + 2y - xy, \\ y' = 2x + 4y - y^2. \end{cases}$$

Exercise 6.33. For Example 6.26, verify that

$$e^{At} H(y,z) = H \circ \phi_t(x,y).$$

Exercise 6.34. Find the homeomorphism H for the nonlinear system

$$\begin{cases} y' = y + z^2, \\ z' = -z. \end{cases}$$

Copyright Elsevier 2022

Exercise 6.35. Approximate the homeomorphism H, W^s, and W^u for the nonlinear system

$$\begin{cases} y_1' = -y_1, \\ y_2' = -y_2^2 + y_1^3, \\ z' = z + y_1^3. \end{cases}$$

Exercise 6.36. Consider the two linear systems $x' = Ax$ and $y' = By$ with

$$A = \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}.$$

Find the homeomorphism H so that $e^{tA}H = He^{tB}$.

Copyright Elsevier 2022

CHAPTER 7

Lyapunov functions

Lyapunov functions are named after Aleksandr Mikhailovich Lyapunov, a Russian mathematician, who defended the thesis *The General Problem of Stability of Motion* at Kharkiv University in 1892. Lyapunov was a pioneer in creating the global approach to the analysis of the stability of nonlinear dynamical systems. His work, initially published in Russian and then translated to French, received little attention for many years. The current use of Lyapunov functions has proved that Lyapunov was ahead of his time in developing his theory that is being used in many different areas of sciences. In theory Lyapunov functions are scalar functions that may be used to prove the stability of equilibrium of a given dynamical system. Since the inception of Lyapunov functions, their successful usage has been extended to integral equations, integro-differential equations, and functional differential equations.

7.1 Lyapunov method

The existence of Lyapunov functions is a necessary and sufficient condition for stability in differential equations. There are no concrete procedures on how to find Lyapunov functions, but in some cases, the construction of Lyapunov functions is known. In the particular case of homogeneous autonomous systems with constant coefficients the Lyapunov function can be found as a quadratic form. As we will see later, Lyapunov functions allow us to study the stability of nonlinear dynamical systems when the equilibrium solution of an associated linear differential equation is a center.

Let D be an open set in $\mathbb{R}^n$ containing 0, and let t_0 be any initial time. Define

$$V : [t_0, \infty) \times D \to [0, \infty),$$

where V is any differentiable *scalar* function. If $\psi : [t_0, \infty) \times D \to D$ is any differentiable function, then $V(t) := V(t, \psi(t))$ is a scalar function of it, and using the chain rule, we can compute its derivative

$$V'(t) = \frac{\partial V}{\partial x_1}\psi_1'(t) + \ldots + \frac{\partial V}{\partial x_n}\psi_n'(t) + \frac{\partial V}{\partial t}.$$

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00013-9

Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

Let $f : [t_0, \infty) \times D \to \mathbb{R}^n$ with $f(t, 0) = 0$. Assume the existence of the unknown solution $x : [t_0, \infty) \to D$ for all $(t^*, x_0) \in [t_0, \infty) \times D$ and consider the system

$$x' = f(t, x), \tag{7.1.1}$$

where

$$f(t, x) = \begin{pmatrix} f_1(t, x) \\ f_2(t, x) \\ \vdots \\ f_n(t, x) \end{pmatrix}.$$

Then it follows from the above argument that

$$\begin{aligned} V'(t, x(t)) = {} & \frac{\partial V(t, x(t))}{\partial x_1} f_1(t, x(t)) + \dots \\ & + \frac{\partial V(t, x(t))}{\partial x_n} f_n(t, x(t)) + \frac{\partial V}{\partial t}. \end{aligned} \tag{7.1.2}$$

Thus expression (7.1.2) defines the derivative of the function $V(t, x)$ along the unknown solutions of (7.1.1). Note that if $\nabla V = grad\ V = (\frac{\partial V}{\partial x_1}, \dots, \frac{\partial V}{\partial x_n})$, then expression (7.1.2) takes the form

$$V'(t, x(t)) = grad\ V \cdot f(t, x) + \frac{\partial V}{\partial t},$$

where "·" represents the dot product.

Suppose we are able to find V such that for any $(t, x) \in [t_0, \infty) \times D$, we have

$$\frac{\partial V(t, x(t))}{\partial x_1} f_1(t, x(t)) + \dots + \frac{\partial V(t, x(t))}{\partial x_n} f_n(t, x(t)) + \frac{\partial V}{\partial t} \le 0.$$

Then along the solutions of (7.1.1), we have $V'(t, \psi(t)) \le 0$. Since V is nonincreasing along the solutions and V is increasing in $|x|$, with some work, we are able to get the boundedness of solutions and stability of equilibrium solutions.

The discussion that we gave above regarding the derivative of the function V is not rigor, and more precise definitions and conditions must be imposed. We start with the following definition.

Definition 7.1.1. Let $D \subset \mathbb{R}^n$ be an open set containing the origin, and let $V(t, x) : \mathbb{R}^+ \times D \to R$ be a given function. Then we define $W = \mathbb{R}^+ \times D$ and

$$V'_{(7.1.1)}(t, x(t)) = \limsup_{h \to 0^+} \frac{V(t + h, x + hf(t, x)) - V(t, x)}{h}, \ (t, x) \in W,$$

where f is the right-hand side function of (7.1.1).

Copyright Elsevier 2022

Let $x(t)$ be a solution of (7.1.1) and define

$$V'(t,x(t)) = \limsup_{h\to 0^+} \frac{V(t+h, x(t+h)) - V(t,x(t))}{h}.$$

Definition 7.1.2. Suppose $V(t,x) : \mathbb{R}^+ \times D \to \mathbb{R}^+$ is Lipschitz in x (uniformly in $t \in \mathbb{R}^+$), that is, there exists $L > 0$ such that for all $t \in \mathbb{R}^+$,

$$|V(t,x_1) - V(t,x_2)| \le L\|x_1 - x_2\|, \quad x_1, x_2 \in D.$$

We further assume that $V(t,x)$ is continuous in t and Lipschitz in x (uniformly in t) with Lipschitz constant $L > 0$. We have the following lemma.

Lemma 7.1. *Let $x(t)$ be a solution of* (7.1.1). *Then*

$$V'(t,x(t)) = V'_{(7.1.1)}(t,x(t)). \tag{7.1.3}$$

Moreover, if $Q : W \to [0,\infty)$ is continuous such that

$$V'(t,x(t)) \le -Q(t,x(t)), \tag{7.1.4}$$

then

$$V(t_2,x(t_2)) - V(t_1,x(t_1)) \le -\int_{t_1}^{t_2} Q(t,x(t))dt,\ 0 \le t_1 \le t_2. \tag{7.1.5}$$

Proof. First, we note that since V is a Lipschitz function, we have

$$-L||x_1 - x_2|| \le V(s,x_1) - V(s,x_2) \le L||x_1 - x_2||,\ s \ge 0,\ \ x_1,x_2 \in D.$$

Let $x(t)$ be a solution of (7.1.1). Then

$$\begin{aligned}
&V(t+h, x(t+h)) - V(t,x(t)) \\
&= V(t+h, x(t+h)) - V(t+h, x+hf(t,x)) \\
&+ V(t+h, x+hf(t,x)) - V(t,x(t)).
\end{aligned}$$

Divide the above expression by h, take $\limsup_{h\to 0^+}$, and then apply the Lipschitz condition to get

$$\begin{aligned}
&\limsup_{h\to 0^+} \frac{V(t+h, x(t+h)) - V(t,x(t))}{h} \\
&\le \limsup_{h\to 0^+} \frac{V(t+h, x+hf(t,x)) - V(t,x(t))}{h} \\
&+ L \lim_{h\to 0^+} \frac{\|x(t+h) - x(t) - hf(t,x(t))\|}{h}
\end{aligned}$$

Copyright Elsevier 2022

$$= \limsup_{h\to 0^+} \frac{V(t+h, x+hf(t,x)) - V(t,x(t))}{h}$$
$$+ L\|x' - f(t,x)\|,$$

which gives

$$V'(t,x(t)) \le \limsup_{h\to 0^+} \frac{V(t+h, x+hf(t,x)) - V(t,x)}{h} = V'_{(7.1.1)}(t,x(t)). \quad (7.1.6)$$

In a similar way, but reversing the use of Lipschitz inequality, we get

$$\limsup_{h\to 0^+} \frac{V(t+h, x(t+h)) - V(t,x(t))}{h}$$
$$\ge \limsup_{h\to 0^+} \frac{V(t+h, x+hf(t,x)) - V(t,x(t))}{h}$$
$$- L \lim_{h\to 0^+} \frac{\|x(t+h) - x(t) - hf(t,x(t))\|}{h}$$
$$= \limsup_{h\to 0^+} \frac{V(t+h, x+hf(t,x)) - V(t,x(t))}{h}$$
$$- L\|x' - f(t,x)\|,$$

which gives

$$V'(t,x(t)) \ge \limsup_{h\to 0^+} \frac{V(t+h, x+hf(t,x)) - V(t,x)}{h} = V'_{(7.1.1)}(t,x(t)). \quad (7.1.7)$$

This shows that

$$V'(t,x(t)) = V'_{(7.1.1)}(t,x(t)).$$

As for (7.1.5), we integrate both sides of (7.1.4) from t_1 to t_2. This completes the proof. □

Remark 7.1. Lemma 7.1 states that the derivative of V with respect to (7.1.1) is the derivative of V along the solution of (7.1.1) and is given by (7.1.2).

The situation is considerably simplified when V is independent of t. In this case (7.1.2) simplifies to

$$V'(x(t)) = \frac{\partial V(x(t))}{\partial x_1} f_1(t,x(t)) + \cdots + \frac{\partial V(x(t))}{\partial x_n} f_n(t,x(t)). \quad (7.1.8)$$

Recall that if x^* is an equilibrium point of (7.1.1), then the change of variables $x = y + x^*$ translates every nonzero equilibrium solution to the origin.

Copyright Elsevier 2022

Definition 7.1.3. Let D be an open set in $\mathbb{R}^n$ containing 0. A continuous autonomous function $V : D \to [0, \infty)$ is *positive definite* if

$$V(0) = 0 \text{ and } V(x) > 0 \text{ for } x \neq 0.$$

The function V is said to be *negative definite* if $-V$ is positive definite.

Definition 7.1.4. Let D be an open set in $\mathbb{R}^n$ containing 0. Let

$$V : D \to [0, \infty)$$

have continuous first partial derivatives. If V is positive definite and

$$V'(x(t)) = \frac{\partial V(x(t))}{\partial x_1} f_1(t, x(t)) + \cdots + \frac{\partial V(x(t))}{\partial x_n} f_n(t, x(t)) \leq 0$$

for $(t, x) \in [t_0, \infty) \times D$, and $x \neq 0$, then V is called a *Lyapunov function* for system (7.1.1).

If the inequality is strict, that is, $V'(x(t)) < 0$, then V is said to be a *strict Lyapunov function*.

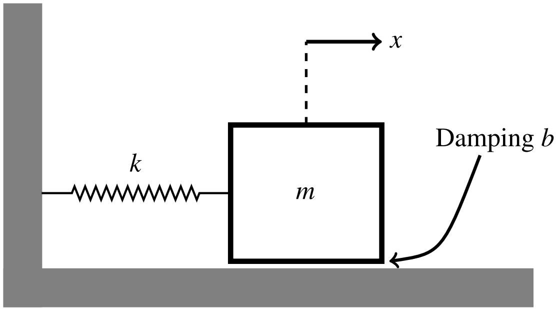


FIGURE 7.1

Damping harmonic motion.

Consider the damped harmonic motion depicted in Fig. 7.1, where m, k are positive constants and $a \geq 0$. If x is the displacement of the mass from its equilibrium position, then applying Newton's second law to all forces acting on the system, we arrive at the second-order differential equation

$$mx'' + bx' + kx = 0.$$

We introduce the transformation $x_1 = x$ and $x_2 = x'$ and arrive at the system

$$x_1' = x_2, \quad x_2' = -\frac{k}{m} x_1 - \frac{b}{m} x_2. \tag{7.1.9}$$

Copyright Elsevier 2022

Clearly, if $b = 0$, then linearization gives no information regarding the stability of the origin. Next, we try to construct a Lyapunov function that overcomes such difficulties. Since we are dealing with a physical system, we expect that the system energy will decrease over time to zero. Thus a logical choice for a Lyapunov function would be the sum of *kinetic energy* and *spring potential energy*, that is,

$$V(x) = \frac{1}{2}m(x')^2 + \frac{1}{2}kx^2.$$

Clearly, V is positive definite, and

$$\begin{aligned} V' &= mx'x'' + kxx' \\ &= x'\big(mx'' + kx\big) \\ &= x'(-bx') \\ &= -b(x')^2. \end{aligned}$$

Later on, we show that this is enough to demonstrate that the origin is stable. We make the following definitions regarding stability.

Definition 7.1.5.

$$x' = f(t, x), \tag{7.1.10}$$

where $x(t) \in \mathbb{R}^n$. Assume that f is continuous and locally Lipschitz with respect to x. Let $t \to x(t, t_0, x_0)$ denote the maximal defined solution of (7.1.10) satisfying the initial condition $x(t_0) = x_0$. Let

$$\phi : [t_0, \infty) \to \mathbb{R}^n$$

be a solution of the differential equation (7.1.10).

(a) We say that the solution ϕ is *stable* (S) on $[t_0, \infty)$ if for each $\varepsilon > 0$, there is $\delta = \delta(t_0, \varepsilon) > 0$ such that whenever $|\phi(t_0) - x_0| < \delta$, the solution $x(t, t_0, x_0)$ is defined for all $t \in [t_0, \infty)$, and

$$|\phi(t) - x(t, t_0, x_0)| < \varepsilon \ \text{ for all } \ t \geq t_0.$$

(b) We say that the solution ϕ is *uniformly stable* (US) on $[t_0, \infty)$ if it is stable and given $\varepsilon > 0$ as before, δ in (a) is independent of t_0 and such that whenever $|\phi(t_0) - x_0| < \delta$, the solution $x(t, t_0, x_0)$ is defined for all $t \in [t_0, \infty)$, and

$$|\phi(t) - x(t, t_0, x_0)| < \varepsilon \ \text{ for all } \ t \geq t_0.$$

(c) We say that the solution ϕ is *asymptotically stable* (AS) on $[t_0, \infty)$ if it is stable and given $\varepsilon > 0$ as before, there is $\delta_1 < \delta$ such that whenever $|\phi(t_0) - x_0| < \delta_1$, we have $\lim\limits_{t\to\infty} |\phi(t) - x(t, t_0, x_0)| = 0$.

Copyright Elsevier 2022

- **(d)** If ϕ is not stable, then we say that it is *unstable*, that is, there is some $\varepsilon > 0$ such that for every $\delta > 0$, there is some point x_0 with $|\phi(t_0) - x_0| < \delta$ such that $|\phi(t_1) - x(t_1, t_0, x_0)| \geq \varepsilon$ for some $t_1 \in [t_0, \infty)$.
- **(e)** We say that the solution ϕ is *uniformly asymptotically stable* (UAS) on $[t_0, \infty)$ if it is US and there exists $\gamma > 0$ such that for each $\mu > 0$, there exists $T = T(\mu) > 0$ such that $|\phi(t_0) - x_0| < \gamma$, $t \geq t_0 + T$, implies $|x(t, t_0, x_0) - \phi(t)| < \mu$.

Since any equilibrium solution can be translated to the origin, we have the equivalent definitions regarding the zero solution $x = 0$.

Let $f : [t_0, \infty) \times D \to \mathbb{R}^n$ with $D \subset \mathbb{R}^n$ containing the origin. Assume the existence of the unknown solution $x : [t_0, \infty) \to D$ on each $(t^*, x_0) \in [t_0, \infty) \times D$ and consider system (7.1.1) with the initial condition $x(t_0) = x_0$ and $f(t, 0) = 0$.

Definition 7.1.6. The zero solution ($x = 0$) of (7.1.1):

- **(a)** is *stable* (S) if for all $\varepsilon > 0$ and $t_0 \geq 0$, there is $\delta = \delta(t_0, \varepsilon) > 0$ such that $|x(t_0)| < \delta$ implies $|x(t, t_0, x_0)| < \varepsilon$,
- **(b)** is *uniformly stable* (US) if δ in (a) is independent of t_0,
- **(c)** is *unstable* if it is not stable,
- **(d)** is *asymptotically stable* (AS) if it is stable and $\lim_{t \to \infty} |x(t, t_0, x_0)| = 0$, and
- **(e)** is *uniformly asymptotically stable* (UAS) if it is US and there exists $\gamma > 0$ such that for each $\mu > 0$, there exists $T = T(\mu) > 0$ such that $|x(t_0)| < \gamma$, $t \geq t_0 + T$, implies $|x(t, t_0, x_0)| < \mu$.

Thus $x = 0$ is stable if solutions starting near $x = 0$ do not wander too far from $x = 0$ in future time. Now $x = 0$ is asymptotically stable if it is stable, and solutions near x_0 approach 0 as $t \to \infty$. The next result is useful in constructing positive definite or negative definite functions and is stated in the following theorem.

Theorem 7.1.1. *For constants a, b, and c, the function*

$$V(x, y) = ax^2 + bxy + cy^2$$

is positive definite if and only if

$$a > 0 \text{ and } 4ac - b^2 > 0$$

and is negative definite if and only if

$$a < 0 \text{ and } 4ac - b^2 > 0.$$

Example 7.1. Consider the planar system

$$x' = -x - xy^2,$$
$$y' = -y - yx^2.$$

Set

$$V(x, y) = ax^2 + bxy + cy^2.$$

Copyright Elsevier 2022

Computing the derivative of V along the solutions of the system, we obtain

$$V'(x,y) = -2[2a(x^2 + x^2y^2) + b(2xy + xy^3 + yx^3) + 2c(y^2 + x^2y^2)].$$

If we choose $b = 0$, $a > 0$, and $c > 0$, then V' is negative definite, and V is positive definite by Theorem 7.1.1.

7.1.1 Stability of autonomous systems

Let D be an open subset of $\mathbb{R}^n$ containing 0, and let $f : D \to \mathbb{R}^n$ with $f(0) = 0$. Assume the existence of the unknown solution $x : [t_0, \infty) \to D$ for each $(t^*, x_0) \in [t_0, \infty) \times D$ and consider the system

$$x' = f(x),\ x(0) = x_0,\ t \geq 0, \tag{7.1.11}$$

where

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix}.$$

Note that Definitions 7.1.3, 7.1.4, and 7.1.6 carry over to the autonomous system (7.1.11). We make the following definitions regarding an open ball so that we may state the stability notion in terms of such balls.

Definition 7.1.7. Let $x^* \in \mathbb{R}^n$ and $r > 0$. Then $B_r(x^*)$ denotes the open ball centered at x^* with radius r, that is,

$$B_r(x^*) = \{x \in \mathbb{R}^n : ||x - x^*|| < r\}.$$

Definition 7.1.8. The equilibrium point x^* of (7.1.11):

- **(a)** is *stable* (S) if for each ball $B_\varepsilon(x^*)$, there is a ball $B_\delta(x^*)$ ($\delta \leq \varepsilon$) such that if $x \in B_\delta(x^*)$, then $x(t, x_0)$ remains in the ball $B_\varepsilon(x^*)$ for $t \geq 0$,
- **(b)** is *asymptotically stable* (AS) if it is *stable* and there is a ball $B_c(x^*)$ such that for each $x \in B_c(x^*)$, $x(t, x_0) \to 0$ as $t \to \infty$.

We warm up with the following example.

Example 7.2. The origin $(0, 0)$ is the only equilibrium solution and is a node for the system

$$x' = -x,$$
$$y' = -y.$$

We apply Definition 7.1.8. The solution is easily computed and is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c_1 e^{-t} \\ c_2 e^{-t} \end{pmatrix}$$

Copyright Elsevier 2022

for $x(0) = c_1$ and $y(0) = c_2$. For all $\varepsilon > 0$, let $\delta = \varepsilon > 0$. Then for all $c \in B_\delta(0) = \{(x, y) \in \mathbb{R}^2 : \sqrt{x^2 + y^2} < \delta\}$ and $t > 0$, we have

$$||(x, y)|| = \sqrt{x^2(t) + y^2(t)} = \sqrt{(c_1^2 + c_2^2)e^{-2t}} \leq \delta = \varepsilon.$$

It follows that

$$(x(t, 0, c_1), y(t, 0, c_2)) \in B_\delta(0),$$

and hence $(0, 0)$ is stable. Moreover,

$$(x(t, 0, c_1), y(t, 0, c_2)) \to (0, 0) \text{ as } t \to \infty$$

shows that $(0, 0)$ is AS.

Theorem 7.1.2. *Let $V(x)$ be a Lyapunov function in D, that is, $V : D \to [0, \infty)$,*

$$V(0) = 0, \quad V(x) > 0, \text{ and } V' \leq 0 \text{ for } x \neq 0.$$

Then the zero solution of (7.1.11) *is stable.*

Proof. Pick $\varepsilon > 0$ such that the ball $B_\varepsilon(0)$ is in D. Let S_ε be the boundary of $B_\varepsilon(0)$. Due to the continuity of V at the origin and $V(0) = 0$, there must be a point on S_ε where V attains its minimum, which we denote by $m = \min_{x \in S_\varepsilon} V(x) > 0$. Let $B_\delta(0)$ be another ball in D. Since $V(0) = 0$ and V is continuous we can choose $\delta > 0$ such that $V(x) < m$ for all $x \in B_\delta(0)$. Our goal is to show that if $x_0 \in B_\delta(0)$, then $x(t, 0, x_0) \in B_\varepsilon(0)$ for all $t > 0$. Since $V' \leq 0$, then $V(x_0) < m$ implies that $V(x(t, 0, x_0)) < m$ for all $t > 0$ because the nonpositivity of the derivatives implies that V does not increase along the orbits with the initial condition x_0. This in turn implies that $x(t, 0, x_0)$ cannot cross S_ε since $V(x) \geq m$ for all $\in S_\varepsilon$. This implies that the origin is stable. This completes the proof. ☐

Example 7.3. Consider the planar system

$$\begin{aligned} x' &= -y + xy, \\ y' &= x - x^2. \end{aligned}$$

Clearly, $(0, 0)$ is an equilibrium point of the system. The linearization theorem, Theorem 6.2.1, is inconclusive since $(0, 0)$ is a center for the linear part

$$\begin{aligned} x' &= -y, \\ y' &= x. \end{aligned}$$

Moreover, using the method of polar coordinates gives no information about the stability of the origin. Let us try to construct a Lyapunov function. Let a and b be positive constants and set

$$V(x, y) = ax^2 + by^2,$$

Copyright Elsevier 2022

where a and b are to be determined. It is clear that $V(0,0) = 0$ and $V(x,y) > 0$ for $(x,y) \neq (0,0)$.

Computing the derivative of V along the solutions of the system using (7.1.8), we obtain

$$\begin{aligned} V'(x,y) &= axx' + byy' \\ &= ax(-y + xy) + by(x - x^2) \\ &= 0 \end{aligned}$$

by setting $a = b > 0$. The simplest choice would be $a = \frac{1}{2} = b$. In this case,

$$V(x,y) = \frac{1}{2}(x^2 + y^2).$$

Thus by Theorem 7.1.2 the origin is stable.

Here is another example.

Example 7.4. Consider the three-dimensional system

$$\begin{aligned} x' &= -2y + yz, \\ y' &= x - xz, \\ z' &= xy. \end{aligned}$$

Clearly, $(0,0,0)$ is an equilibrium solution of the system, and

$$J(0,0,0) = \begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus $J(0,0,0)$ has the eigenvalues $\lambda_1 = 0$ and $\lambda_2 = \lambda_3 = i\sqrt{2}$. The linearization theorem, Theorem 6.2.1, cannot be applied. Let us try to construct a Lyapunov function. Let a, b, and c be positive constants and set

$$V(x,y) = ax^2 + by^2 + cz^2,$$

where a and b are to be determined. It is clear that $V(0,0) = 0$ and $V(x,y,z) > 0$ for $(x,y,z) \neq (0,0,0)$. Computing the derivative of V along the solutions of the system using (7.1.8), we obtain

$$\begin{aligned} V'(x,y) &= axx' + byy' + czz' \\ &= ax(-2y + yz) + by(x - xz) + cz(xy) \\ &= 2[(a - b + c)xyz + (-2a + c)xy]. \end{aligned}$$

Copyright Elsevier 2022

Hence if $b = 2a$ and $c = a > 0$, then $V'(x, y, z) = 0$ for all $(x, y, z) \in \mathbb{R}^3$, and therefore by Theorem 7.1.2 the origin $(0, 0, 0)$ is stable. Furthermore, choosing $a = c = 1$ and $b = 2$, we see that the trajectories of this system lie on the ellipsoids

$$x^2 + 2y^2 + z^2 = k^2$$

for positive constants k.

Example 7.3 reveals the advantage of the use of Lyapunov functions over linearization since it showed that a center for a linear system is stable for the accompanying nonlinear system.

The next theorem provides criteria for asymptotic stability.

Theorem 7.1.3. *Let $V(x)$ be a strict Lyapunov function in D, that is, $V : D \to [0, \infty)$,*

$$V(0) = 0, \quad V(x) > 0, \text{ and } V' < 0 \text{ for } x \neq 0.$$

Then the zero solution of (7.1.11) *is asymptotically stable.*

Proof. Since the origin is stable and $V' < 0$ along the solutions of (7.1.11), it follows that $V(x)$ decreases to the value c as $t \to \infty$. The proof will be complete if we show that $c = 0$. Assume for contradiction that $c > 0$. Then there exists $\beta < \varepsilon$ (ε from the stability theorem) such that $V(x) < c$ for all $x \in B_\beta(0)$. However, then $x(t)$ cannot enter $B_\beta(0)$. Let

$$m = \min\{-V' : \beta \leq |x| \leq \varepsilon\}.$$

As $-V' > 0$, we have that $V' \leq -m$ for all $t \geq 0$. An integration of $V' \leq -m$ from 0 to t yields

$$V(x(t)) - V(x_0) \leq -mt.$$

Thus $\lim_{t \to \infty} V(x(t)) = -\infty$, which contradicts the assumption that V is positive definite in D and tends to c as $t \to \infty$. Hence c must vanish.

We next claim that $x(t) \to 0$ as $t \to \infty$. If this is not the case, then there are constants $\alpha > 0$ and a sequence $t_k \to \infty$ such that $|x(t)| \geq \alpha$ for all k. Since the closure $\bar{B}_\beta(0)$ is compact and all the points $x(t_k) \in \bar{B}_\beta(0)$, there is a subsequence that converges to a point $z \in B_\beta(0)$. We may assume that $x(t_k) \to z$, and we must have $0 \neq |z| \geq \alpha$. By the continuity of V we have that $V(x(t_k)) \to V(z) > 0$. However, since $t_k \to \infty$, we must have $V(x(t_k)) \to c = 0$, which is a contradiction. Thus the origin is asymptotically stable. This completes the proof. □

Example 7.5. Consider the planar system

$$\begin{aligned} x' &= -x - xy^2, \\ y' &= -y + 3x^2 y. \end{aligned}$$

Copyright Elsevier 2022

Clearly, $(0,0)$ is an equilibrium point of the system. The linearization theorem, Theorem 6.2.1, tells us that $(0,0)$ is AS. We verify the result by constructing a strict Lyapunov function. Let a and b be positive constants and set

$$V(x,y) = ax^2 + by^2,$$

where a and b are to be determined. Clearly, $V(0,0) = 0$ and $V(x,y) > 0$ for $(x,y) \neq (0,0)$. Computing the derivative of V along the solutions of the system using (7.1.8), we obtain

$$\begin{aligned}
V'(x,y) &= axx' + yy' \\
&= ax(-x - xy^2) + by(-y + 3x^2y) \\
&= -2ax^2 - 2ax^2y^2 - 2by^2 + 6bx^2y^2 \\
&= -(6x^2 + 2y^2) < 0
\end{aligned}$$

unless $x = y = 0$, where we choose $a = 3$, $b = 1$ to eliminate the mixed terms. Thus by Theorem 7.1.3 the origin is asymptotically stable.

Note that the origin $(0,0)$ of the system in Example 7.1 is asymptotically stable by Theorem 7.1.3.

Let $V(x)$ be a strict Lyapunov function in D. Then *domain of attraction* of an equilibrium consists of all points such that a solution starting at them tends to the equilibrium. In engineering applications, it is often important to get an estimate of the size of the domain of attraction. Thus if V is a strict Lyapunov function on a set $D \in \mathbb{R}^n$, then the origin is asymptotically stable, and the *domain of attraction* is

$$\bar{N} = \{x \in D : V(x) \leq C\}$$

for some positive constant C.

Example 7.6. Consider the planar system

$$\begin{aligned}
x' &= -x - y, \\
y' &= 2x - y + y^3.
\end{aligned}$$

Clearly, $(0,0)$ is an equilibrium point of the system. Set

$$V(x,y) = x^2 + \frac{y^2}{2}.$$

Clearly, $V(0,0) = 0$, and $V(x,y) > 0$ for $(x,y) \neq (0,0)$. Computing the derivative of V along the solutions of the system, we obtain

$$\begin{aligned}
V'(x,y) &= 2xx' + yy' \\
&= 2x(-x - y) + y(2x - y + y^3) \\
&= -2x^2 - y^2 + y^4.
\end{aligned}$$

Copyright Elsevier 2022

To keep $V'(x,y) < 0$, we must restrict y so that $-y^2 + y^4 = y^2(-1+y^2) < 0$, that is, $-1 < y < 1$. On the other hand, to keep the ellipses $V(x,y) = x^2 + \frac{y^2}{2} = C$ inside $D = \{(x,y) : x \in \mathbb{R},\ -1 < y < 1\}$, we must take $0 < C < \frac{1}{2}$. Thus for any value of such C, the origin is asymptotically stable, and the domain of attraction is

$$\bar{N} = \{x \in D : V(x) \le C < \frac{1}{2}\}.$$

The next theorem only asks for the existence of a Lyapunov function to obtain the asymptotic stability but with a catch.

Theorem 7.1.4. *Let $V(x)$ be a Lyapunov function in D. Let*

$$Z = \{x : V'(x) = 0, x \in D\}.$$

Suppose that the only bounded solution of (7.1.11) *that remains inside Z all the time is the equilibrium solution* 0. *Then the origin of* (7.1.11) *is asymptotically stable.*

Proof. Theorem 7.1.4 is a particular case of Theorem 7.4.2, and we ask the reader to either wait or flip the pages. ☐

As an application of Theorem 7.1.4, we provide the following example.

Example 7.7. Consider the planar system

$$\begin{aligned} x' &= -x - 2y, \\ y' &= 3x - 2y. \end{aligned}$$

Clearly, $(0,0)$ is an equilibrium point of the system. Set

$$V(x,y) = \frac{x^2}{2} + y^2.$$

Clearly, $V(0,0) = 0$ and $V(x,y) > 0$ for $(x,y) \neq (0,0)$. Computing the derivative of V along the solutions of the system, we obtain

$$\begin{aligned} V'(x,y) &= xx' + 2yy' \\ &= x(-x-2y) + 2y(3x-2y) \\ &= -(x-2y)^2. \end{aligned}$$

It is clear that $V'(x,y) = 0$ if $x = 2y$. Thus $V(x,y)$ is a Lyapunov function (not strict) on the set

$$Z = \{(x,y) \in \mathbb{R}^2 : x = 2y\}.$$

Let (x,y) be a bounded solution (trajectory) in Z, that is, there exists a positive constant K such that $|x|,\ |y| \le K$. Substitute $x = 2y$ into the original system to get $x' = -2x$ and $y' = 4y$. By straightforward calculations we obtain the solutions

$$x(t) = x(0)e^{-2t} \text{ and } y(t) = y(0)e^{4t},$$

Copyright Elsevier 2022

respectively, for $t \in \mathbb{R}$. These functions cannot satisfy $|x|, |y| \leq K$, unless $x(0) = y(0) = 0$. To have you more convinced, the solutions must also satisfy $x = 2y$, or

$$x(0)e^{-2t} = 2y(0)e^{4t},$$

which can only hold if $x(0) = y(0) = 0$. Thus the only bounded solution in Z is the origin, and hence by Theorem 7.1.4 the origin is asymptotically stable.

The next theorem provides criteria for instability of the zero solution of (7.1.11).

Theorem 7.1.5. *Let D be an open set in $\mathbb{R}^n$ containing* 0*. Let $V : D \to [0, \infty)$ be continuous real-valued scalar differentiable function that is positive definite on D. If $V'(x) > 0$ for $x \neq 0$, then the origin of* (7.1.11) *is unstable.*

Proof. Assume that the zero solution is stable. Then, for $\varepsilon = 1$ and $t_0 = 0$, there exists $\delta > 0$ such that $|x_0| < \delta$ implies that $|x(t, 0, x_0)| \leq 1$. Hence for $x(t) = x(t, 0, x_0)$, $V(x(t))$ is bounded for $t \geq 0$ when $|x_0| \leq \delta$. Fix x_0 with $|x_0| = \frac{\delta}{2}$ such that $V(x_0) > 0$. Since $V' > 0$, it follows that V is increasing along the solutions. Therefore we have $V(x(t)) \geq V(x_0) > 0$, $t \geq 0$. Since V is continuous and $V(0) = 0$, there exists $p \in (0, 1)$ such that $p \leq |x(t)| \leq 1$, $t \geq 0$. Since $V' > 0$ along the solutions of (7.1.11) and positive on the closed and bounded set $p \leq |x(t)| \leq 1$, we have $l = \min\limits_{p \leq |x| \leq 1} V'(x) > 0$. Thus

$$V(x(t)) - V(x_0) = \int_0^t \frac{d}{ds} V(x(s))ds \geq \int_0^t l ds = l(t - t_0) \to \infty, \ t \to \infty.$$

This contradicts the fact that $V(x(t))$ is bounded for $t \geq 0$ when $|x_0| < \delta$. It follows that the zero solution $x(t) = 0$ is unstable. ☐

Example 7.8. Consider the planar system

$$x' = 2xy + x^3,$$
$$y' = -x^2 + y^5.$$

Clearly, $(0, 0)$ is an equilibrium point of the system. Linearization will not work here since there are no linear terms at $(0, 0)$. Set

$$V(x, y) = x^2 + 2y^2.$$

Clearly, $V(0, 0) = 0$ and $V(x, y) > 0$ for $(x, y) \neq (0, 0)$. Computing the derivative of V along the solutions of the system, we obtain

$$\begin{aligned} V'(x, y) &= 2xx' + 4yy' \\ &= 2x(2xy + x^3) + 4y(-x^2 + y^5) \\ &= 2x^4 + 4y^6 > 0, \ \ (x, y) \neq (0, 0). \end{aligned}$$

It follows from Theorem 7.1.5 that the origin is unstable.

Copyright Elsevier 2022

Example 7.9. We examine the stability and asymptotic stability for different values of a for the planner system

$$\begin{aligned} x' &= ax(x^2+y^2) - xy^2, \\ y' &= ay(x^2+y^2) + x^2y. \end{aligned}$$

Clearly, $(0,0)$ is an equilibrium point of the system. Linearization will not work here since there are no linear terms in the neighborhood of $(0,0)$. Set

$$V(x,y) = x^2+y^2.$$

Clearly, $V(0,0)=0$ and $V(x,y)>0$ for $(x,y)\neq(0,0)$. Set $V(t) := V(x(t),y(t))$. Then along the solutions of the system we have that

$$\begin{aligned} V'(t) &= 2xx' + 2yy' \\ &= 2x[ax(x^2+y^2) - xy^2] + 2y[ay(x^2+y^2) + x^2y] \\ &= 2a(x^2+y^2)^2. \end{aligned} \tag{7.1.12}$$

It is clear from relation (7.1.12) that the origin is stable if $a=0$, asymptotically stable if $a<0$, and unstable if $a>0$. In the following, we will use expression (7.1.12) and try to understand the behavior of the Lyapunov function along the solutions.

We observe that (7.1.12) is equivalent to

$$V'(t) = 2aV^2(t),$$

which is a separable first-order differential equation in $V(t)$. Thus separating the variables and then integrating from 0 to t gives

$$V(t) = \frac{V(0)}{1-2aV(0)t}, \ t \geq 0. \tag{7.1.13}$$

Case 1. If $a=0$, then from (7.1.13) we have that $V(t)=V(0)=x(0)^2+y(0)^2>0$. In this case, we have a periodic solution surrounding the origin

$$x^2+y^2=k^2$$

for nonzero constant k. Hence the origin is stable.

Case 2. If $a<0$, then from (7.1.13) we have that $V(t)\to 0$ as $t\to\infty$, that is, $x^2+y^2\to 0$ as $t\to\infty$, which means that the solutions spiral toward the origin, and we conclude that the origin is asymptotically stable.

Case 3. If $a>0$, then by (7.1.13) $V(t)$ is not defined for all $t\geq 0$. In fact, $V(t)\to\infty$ as $t\to\frac{1}{2aV(0)}$ from the left. We conclude that the origin is unstable.

The next theorem provides criteria of the instability of the zero solution, which is different from that given in Theorem 7.1.5.

Copyright Elsevier 2022

Theorem 7.1.6. *Let x^* be an equilibrium solution of (7.1.11). Let U be a neighborhood of x^*. Let $V : U \to \mathbb{R}$ be a continuously differentiable function on U, except perhaps at x^*. If $V(x^*) = 0$, but there are points arbitrarily close to x^* where V is strictly positive and $V' > 0$ on $U \setminus \{x^*\}$, then x^* is unstable for (7.1.11).*

Proof. Without loss of generality, we take $x^* = 0$. Let $B_\beta(0) \subseteq U$, $\beta > 0$. We may choose a point $x_0 \in B_\beta(0)$ close to 0 such that $V(x_0) > 0$. Let (a, b) be the interval of existence of the solution $x(t) = x(t, x_0)$. The proof will be completed if we can show that $x(t)$ is outside the ball $\bar{B}_\beta(0)$. Assume the contrary, that is, solutions stay in $B_\beta(0)$ for all $t \geq 0$. Since $\bar{B}_\beta(0)$ is compact, V achieves its maximum, say $V_0 \in \bar{B}_\beta(0)$. Thus

$$V(x(t)) \leq V_0, \quad t \geq 0. \tag{7.1.14}$$

Since $V' > 0$ and V is increasing along the solutions and $V' \geq V(x_0) > 0$, we can find $\rho > 0$ such that $V(x) < V(x_0)$ for $x \in B_\rho(0)$. Then $x(t)$ is trapped in the compact annulus $\rho \leq |x| \leq \beta$. Let

$$m = \min\{V' : \rho \leq |x| \leq \beta\}.$$

An integration of $V' \geq m$ from 0 to t yields

$$V(x(t, x_0)) \geq V(x_0) + mt,$$

which contradicts (7.1.14) since the right side goes to infinity as t goes to infinity. Thus the solution $x(t)$ must escape $B_\beta(0)$, and the origin is unstable. This completes the proof. □

Lyapunov functions can also be used to determine limit cycles, as the next example demonstrates.

Example 7.10. In Example 6.15 we considered the planar autonomous system

$$x' = -y - x(x^2 + y^2 - \mu), \quad y' = x - y(x^2 + y^2 - \mu)$$

with $\mu > 0$ and used polar coordinates to show the existence of a stable limit cycle. In this example, we use a Lyapunov function to arrive at the same result. Let

$$V(x, y) = x^2 + y^2.$$

Then along the solutions of the system we have the expression

$$\begin{aligned} V'(t) &= 2xx' + 2yy' \\ &= 2x[-y - x(x^2 + y^2 - \mu)] + 2y[x - y(x^2 + y^2 - \mu)] \\ &= -2(x^2 + y^2 - \mu)(x^2 + y^2) \\ &= 2\big(\mu - V(x, y)\big)V(x, y). \end{aligned}$$

Copyright Elsevier 2022

Now since $V > 0$, the sign of V' depends on the term $\mu - V(x, y)$. We have $V' = 0$ if $V = \mu$ or $x^2 + y^2 = \mu$, which is a circle centered at the origin with radius μ. If $0 < V < \mu$, then $V' > 0$. In this case the trajectories start inside the circle with radius μ and move upward toward it but do not cross it due to uniqueness. If $V > \mu$, then $V' < 0$. In this case the trajectories start outside the circle with radius μ and move toward it but do not cross it due to uniqueness. We conclude that $x^2 + y^2 = \mu$ is a stable limit cycle.

The next example shows the importance of Theorem 7.1.6 and illustrates its usage as Theorem 7.1.5 is not applicable.

Example 7.11. Considered the planar autonomous system

$$x' = 3x + y^2, \quad y' = -2y + x^3.$$

Clearly, the origin is the only equilibrium point of the system. We easily conclude that $(0, 0)$ is unstable using the linearization method. Let us see how we can implement Theorem 7.1.6. We consider the Lyapunov function

$$V(x, y) = \frac{x^2}{2} - \frac{y^2}{2}.$$

There are points in the neighborhood of the origin where V is positive, for example, the points on the x-axis. Since we are only interested what happens in the neighborhood of $(0, 0)$, we limit ourselves to $|x| < 1$ and $|y| < 1$. To better understand the derivative of V, we will utilize the inequality

$$0 < (a - b)^2 = a^2 - 2ab + b^2$$

for all real numbers a and b, from which it follows that

$$ab \le \frac{a^2}{2} + \frac{b^2}{2}.$$

It readily follows that along the solutions of the system,

$$V' = (3x^2 + 2y^2) + (xy^2 - x^3 y). \tag{7.1.15}$$

Next, we compare the term $(3x^2 + 2y^2)$ with $(xy^2 - x^3 y)$:

$$\begin{aligned} \left|xy^2 - x^3 y\right| &\le |xy||y - x^2| \\ &\le \frac{1}{2}(x^2 + y^2)(x^2 + |y|) \\ &\le \frac{1}{2}(x^2 + y^2)(|x| + |y|) \\ &\le \frac{1}{2}(|x| + |y|)(|x| + |y|) \end{aligned}$$

Copyright Elsevier 2022

$$
\begin{aligned}
&= \frac{1}{2}(|x|^2 + |y|^2) + |x||y| \\
&\le \frac{1}{2}(|x|^2 + |y|^2) + \frac{1}{2}(|x|^2 + |y|^2) \\
&= x^2 + y^2 \\
&\le 3x^2 + 2y^2.
\end{aligned}
$$

Thus a substitution into (7.1.15) gives

$$V'(t) \ge \left|xy^2 - x^3y\right| + (xy^2 - x^3y) > 0$$

for $(x, y) \neq (0, 0)$. We conclude that $V' > 0$ on the set $|x|, |y| < 1$ minus the origin, and by Theorem 7.1.6 the origin is unstable. Note that the set U of Theorem 7.1.6 is defined by

$$U = \{x, y \in \mathbb{R} : |x| < 1,\ |y| < 1\}.$$

Theorem 7.1.6 also has its limitation since there may be some points near the equilibrium solution x^* that approach x^*, which makes it difficult to have $V' > 0$ on a whole neighborhood of x^*. The next generalization theorem is due to Četaev.

Theorem 7.1.7. *(Četaev) Let x^* be an equilibrium solution of (7.1.11) contained in the closure of an open set U, and let E be a neighborhood of x^*. Let $V : E \to \mathbb{R}$ be a continuously differentiable function on E such that*

(1) V and V' are both positive on $(U \cap E) \setminus \{x^*\}$.
(2) $V(x) = 0$ on the part of the boundary of U that is in E.

Then the equilibrium solution x^* is unstable.

Proof. Exercise. □

Next, we briefly discuss a special class of Lyapunov functions that includes the absolute value of the unknown solution. Let $x : \mathbb{R} \to \mathbb{R}$ be continuous. Observing that

$$|x| = \sqrt{x^2} = (x^2)^{\frac{1}{2}}$$

and using the chain rule, we arrive at

$$
\begin{aligned}
\frac{d}{dt}|x(t)| &= \frac{1}{2}(x^2(t))^{-\frac{1}{2}}(2x(t)x'(t)) \\
&= \frac{x(t)}{(x^2(t))^{\frac{1}{2}}}x'(t) \\
&= \frac{x(t)}{|x(t)|}x'(t).
\end{aligned}
$$

We have the following lemma.

Copyright Elsevier 2022

Lemma 7.2. *Let $x : \mathbb{R} \to \mathbb{R}$ be differentiable. Then*

$$\frac{d}{dt}|x(t)| = \frac{x(t)}{|x(t)|}x'(t).$$

Example 7.12. Consider the scalar nonlinear Volterra integro-dynamic equation

$$x'(t) = ax(t) + b\frac{x(t)}{1 + \int_0^t x^2(s)ds}, \quad t \in [0, \infty). \tag{7.1.16}$$

Suppose there exists a constant β such that

$$a + |b| = \beta.$$

Consider the function

$$V(t, x) = |x(t)|.$$

Then by Lemma 7.2, we have

$$\begin{aligned} V'(t,x) &= \frac{d}{dt}|x(t)| = \frac{x(t)}{|x(t)|}x'(t) \\ &= \frac{x(t)}{|x(t)|}\Big(ax(t) + b\frac{x(t)}{1 + \int_0^t x^2(s)ds}\Big) \\ &= ax^2\frac{1}{|x(t)|} + \frac{x}{|x|}b\frac{x(t)}{1 + \int_0^t x^2(s)ds} \\ &\le a|x| + |b||x|\frac{1}{1 + \int_0^t x^2(s)ds} \\ &\le a|x| + |b||x| \\ &= (a + |b|)|x|. \end{aligned}$$

Then it follows from Theorems 7.1.2–7.1.5 that the zero solution of (7.1.16) is stable if $\beta = 0$, asymptotically stable if $\beta < 0$, and unstable if $\beta > 0$.

7.1.2 Time-varying systems; non-autonomous

Now we turn our attention to systems that explicitly include the variable time t. Such systems are referred to as non-autonomous systems. Thus we consider the non-autonomous system

$$x' = f(t, x), \tag{7.1.17}$$

where $f \in C\big([0, \infty) \times D, \mathbb{R}^n\big)$, where $D \subset \mathbb{R}^n$ is open with $0 \in D$. We say that a vector $x^* \in \mathbb{R}^n$ is an *equilibrium solution*, or *constant solution*, or *equilibrium point* of (7.1.17) if

$$f(t, x^*) = 0.$$

Copyright Elsevier 2022

This is the perfect place to make clear that the existence of equilibrium points for non-autonomous systems may depend on the initial time. To see this, we consider

$$x' = \begin{pmatrix} \frac{1}{t} & t \\ 1 & t \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ t > 0, \ x(0) = x_0.$$

The system has a unique equilibrium solution $(0, 0)$, provided that the matrix is non-singular, which is the case for $t \neq 1$. Thus, in this case, the system attains an equilibrium solution if and only if the initial time $t_0 > 1$.

Definition 7.1.6 of stability carries over to (7.1.17). In addition, the uniform asymptotic stability for (7.1.17) is equivalent to the asymptotic stability for autonomous system. We begin by considering the initial value time-varying scalar differential equation

$$x' = (6t \ \sin(t) - 2t)x, \ \ x(t_0) = x_0, \ \ t_0 \geq 0.$$

Its unique solution passing through (t_0, x_0) is given by

$$\begin{aligned} x(t) &= x_0 \ \exp\Big(\int_{t_0}^{t} (6s\sin(s) - 2s)ds\Big) \\ &= x_0 \ \exp\Big(6\sin(t) - 6t\cos(t) - t^2 - 6\sin(t_0) + 6t_0 + 6t_0\cos(t_0) + t_0^2\Big). \end{aligned}$$

Set

$$c(t_0) = \exp\Big(6\sin(t) - 6t\cos(t) - t^2 - 6\sin(t_0) + 6t_0 + 6t_0\cos(t_0) + t_0^2\Big).$$

Then

$$|x(t)| \leq |x_0| c(t_0)|, \ \ t \geq t_0.$$

Thus for any $\varepsilon > 0$, choose $\delta(t_0, \varepsilon) = \frac{\varepsilon}{c(t_0)}$ so that

$$|x_0| < \delta \Longrightarrow |x(t)| < \varepsilon,$$

and hence the zero solution is stable but not uniformly stable. Note that $c(t_0)$ increases as t_0 increases, and hence δ shrinks as t_0 increases and t is fixed. To be more specific, suppose $t_0 = 2n\pi$, $n = 0, 1, 2, \ldots$, and $x(t)$ is evaluated π seconds later in each case. Then

$$x(t_0 + \pi) = x(t_0)e^{(4n+1)(6-\pi)},$$

and if $x(t_0) \neq 0$, then we get

$$\frac{x(t_0 + \pi)}{x(t_0)} \to \infty \ \text{ as } \ t \to \infty.$$

Thus there is no δ independent of t_0 that would satisfy the requirement for uniform stability.

Copyright Elsevier 2022

Definition 7.1.9. A strictly increasing continuous function $W : [0, \infty) \to [0, \infty)$ with $W(0) = 0$ is called a wedge. (In this book, wedges are always denoted by W or W_i, where i is a positive integer.)

Definition 7.1.10. Let D be an open set in $\mathbb{R}^n$ containing 0. A continuous autonomous function $V : [0, \infty) \times D \to [0, \infty)$ is *positive definite* if

$$V(t, 0) = 0 \text{ and } V(t, x) > 0 \text{ for all } x \neq 0 \text{ and all } t \in [0, \infty).$$

Thus $V(t, x)$ is positive definite if and only if there is a wedge independent of time t such that $V(t, x) \geq W(|x|)$ for all $x \in D$.

Definition 7.1.11. Let D be an open set in $\mathbb{R}^n$ containing 0. Let

$$V : [0, \infty) \times D \to [0, \infty)$$

have continuous first partial derivatives. If V is positive definite and

$$V'(t, x(t)) = \frac{\partial V(t, x(t))}{\partial x_1} f_1(t, x(t)) + \cdots + \frac{\partial V(t, x(t))}{\partial x_n} f_n(t, x(t)) + \frac{\partial V(t, x(t))}{\partial t} \leq 0$$

for $(t, x) \in [t_0, \infty) \times D$ with $x \neq 0$, then V is called *Lyapunov function* for system (7.1.17).

If the inequality is strict, that is, $V'(t, x(t)) < 0$ then V is said to be a *strict Lyapunov function*.

Definition 7.1.12. Let D be an open set in $\mathbb{R}^n$ containing 0. Let

$$V : [0, \infty) \times D \to [0, \infty).$$

We say $V(t, x)$ is:

(a) *decrescent* in D if there is a wedge W_1 such that

$$V(t, x) \leq W_1(|x|) \text{ for all } x \in D,$$

(b) *radially unbounded* if $V(t, x) \to \infty$ as $|x| \to \infty$, which means that for each $M > 0$, there is $N > 0$ such that $V(t, x) > M$ for all $t \in [0, \infty)$ and all x such that $|x| > N$,

(c) both positive definite and decrescent if

$$W_1(|x|) \leq V(t, x) \leq W_2(|x|).$$

Theorem 7.1.8. *Let $D \subset \mathbb{R}^n$ be an open set containing the origin, and let $V(t, x)$: $[0, \infty) \times D \to [0, \infty)$ be a given continuously differentiable function satisfying*

$$V(t, 0) = 0, \tag{7.1.18}$$

Copyright Elsevier 2022

$$W_1(|x|) \le V(t,x), \tag{7.1.19}$$

and

$$V'(t,x) \le 0 \text{ for all } t \ge 0 \text{ and } x \ne 0. \tag{7.1.20}$$

Then the zero solution of (7.1.17) *is stable.*

Proof. By (7.1.20), we have $V'(t,x) \le 0$, V is continuous, $V(t,0) = 0$, and $W_1(|x|) \le V(t,x)$. Let $\varepsilon > 0$ and $t_0 \ge 0$. We must find δ such that $|x_0| < \delta$ and $t \ge t_0$ imply $|x(t,t_0,x_0)| < \varepsilon$. (Throughout these proofs, we assume that ε is small enough so that $|x(t,t_0,x_0)| < \varepsilon$ implies that $x \in D$.) As V is continuous in t and x and $V(t,0) = 0$, there is $\delta > 0$ such that $|x_0| < \delta$ implies $V(t_0,x_0) < W_1(\varepsilon)$. Thus if $t \ge t_0$, $|x_0| < \delta$, and $x = x(t,t_0,x_0)$, then we have

$$W_1(|x(t)|) \le V(t,x) \le V(t_0,x_0) < W_1(\varepsilon).$$

Applying W_1^{-1} to both sides of the above inequality gives

$$|x(t)| \le \varepsilon.$$

This completes the proof. □

Theorem 7.1.9. *Let $D \subset \mathbb{R}^n$ be an open set containing the origin, and let $V(t,x) : [0,\infty) \times D \to [0,\infty)$ be a continuously differentiable function. Assume that* (7.1.18), (7.1.19), *and* (7.1.20) *hold. If, in addition, V satisfies*

$$V(t,x) \le W_2(|x|), \tag{7.1.21}$$

then the zero solution of (7.1.17) *is uniformly stable.*

Proof. For a given ε, we choose $\delta > 0$ such that $W_2(\delta) < W_1(\varepsilon)$, where $W_1(|x|) \le V(t,x) \le W_2(|x|)$. If $t_0 \ge 0$, then we have

$$\begin{aligned} W_1(|x(t)|) \le V(t,x) &\le V(t_0,x_0) \\ &\le W_2(|x_0|) < W_2(\delta) < W_1(\varepsilon), \end{aligned}$$

or $|x(t)| < \varepsilon$, as required. This completes the proof. □

The next example shows that $V'(t,x) \le 0$ is not enough to drive solutions to zero.

Example 7.13. Let $g : [0,\infty) \to (0,\beta]$ with $g(0) = 1$ and $0 < \beta \le 1$. Consider the non-autonomous first-order differential equation

$$x'(t) = \left[g'(t)/g(t)\right]x(t). \tag{7.1.22}$$

It is clear that $x(t) = g(t)$ is a solution of (7.1.22). Our goal is to construct a function

$$V(t,x) = a(t)x^2(t)$$

Copyright Elsevier 2022

such that $V'(t,x) = -\alpha(t)x^2(t)$, where $a(t), \alpha(t) > 0$ for $t \in \mathbb{R}^+$, and $\int_0^\infty \alpha(s)ds < \infty$. Along the solutions of (7.1.22) we have

$$V'(t,x) = \big[a'(t) + 2a(t)\frac{g'(t)}{g(t)}\big]x^2(t).$$

By substituting $V'(t,x) = -\alpha(t)x^2(t)$ we arrive at

$$a'(t) + 2a(t)\frac{g'(t)}{g(t)} = -\alpha(t).$$

It readily follows that the solution is

$$\begin{aligned} a(t) &= [a(0)g^2(0) - \int_0^t \alpha(s)g^2(s)ds]g^2(t) \\ &\geq [a(0)g^2(0) - \beta^2 \int_0^t \alpha(s)ds]/g^2(t). \end{aligned}$$

Since $0 < g \leq 1$ and $\int_{s=0}^\infty \alpha(s) < \infty$, we may choose $a(0)$ large enough to imply that $a(t) > 1$ on $[0,\infty)$. Thus we have shown that $V \geq 0$ and $V' \leq 0$ do not imply that solutions tend to zero. As a matter of fact, there is no wedge W_2 with $V(t,x) \leq W_2(|x|)$. Notice that V is not decrescent.

Theorem 7.1.10. *Let $D \subset \mathbb{R}^n$ be an open set containing the origin, and let $V(t,x) : [0,\infty) \times D \to [0,\infty)$ be a continuously differentiable function satisfying* (7.1.18) *and* (7.1.19)*. If, in addition,*

$$V'(t,x) \leq -W_3(|x|) \text{ for all } t \geq 0, \tag{7.1.23}$$

then the zero solution of (7.1.17) *is asymptotically stable.*

Proof. By Theorem 7.1.8, the zero solution is stable. Therefore given $\varepsilon > 0$, there are $\delta > 0$ and $\gamma > 0$ such that

$$\gamma \leq |x(t,t_0,x_0)| < \varepsilon, \quad t \geq t_0, \quad |x_0| < \delta.$$

Due to (7.1.23) and $|x(t,t_0,x_0)| \geq \gamma$ for $t \geq t_0$, there exists a constant $d > 0$ such that

$$V'(t,x) \leq -d < 0 \text{ for all } t \geq t_0.$$

An integration of $V' \leq -d$ from t_0 to t yields

$$V(t, x(t,t_0,x_0)) \leq V(t_0,x_0) - d(t - t_0).$$

Thus $\lim_{t\to\infty} V(x(t)) = -\infty$, which contradicts the assumption that V is positive definite. Hence no such γ exists, and since $V(t,x)$ is a positive decreasing function, it

Copyright Elsevier 2022

follows that $\lim_{t\to\infty} V(x(t)) = 0$. Therefore

$$\lim_{t\to\infty} x(t, t_0, x_0) = 0$$

which implies that the zero solution is asymptotically stable. This completes the proof. □

Theorem 7.1.11. *Let* $D \subset \mathbb{R}^n$ *be an open set containing the origin, and let* $V(t,x) : [0,\infty) \times D \to [0,\infty)$ *be a continuously differentiable function satisfying*

$$V(t,0) = 0,$$

$$W_1(|x|) \le V(t,x) \le W_2(|x|),$$

and

$$V'(t,x) \le -W_3(|x|) \text{ for all } t \ge 0.$$

Then the zero solution of (7.1.17) *is uniformly asymptotically stable.*

Proof. By Theorem 7.1.9, the zero solution is uniformly stable. Let $\varepsilon = 1$, and find $\delta > 0$ of uniform stability and call it η. Let $\gamma > 0$ be given. We must find $T > 0$ such that

$$|x_0| < \eta, \quad t_0 \ge 0, \quad \text{and } t \ge t_0 + T$$

imply $|x(t, t_0, x_0)| < \gamma$. Pick $\mu > 0$ with $W_2(\mu) < W_1(\gamma)$, so that there is $t_1 \ge t_0$ with $|x(t_1)| < \mu$. Then we have, for $t \ge t_1$,

$$\begin{aligned} W_1(|x(t)|) &\le V(t,x) \le V(t_1, x_1) \\ &\le W_2(|x_1|) < W_2(\eta) < W_1(\gamma), \end{aligned}$$

or $|x(t_1)| < \gamma$. Since $V'(t,x) \le -W_3(|x|)$, and so as long as $|x(t)| > \mu$, it follows that $V'(t,x) \le -W_3(\mu)$. Thus

$$\begin{aligned} V(t, x(t)) &\le V(t_0, x_0) - \int_{t_0}^{t} W_3(|x(s)|)ds \\ &\le W_2(|x_0|) - W_3(\mu)(t - t_0) \\ &\le W_2(\eta) - W_3(\mu)(t - t_0), \end{aligned}$$

which vanishes at

$$t = t_0 + \frac{W_2(\eta)}{W_3(\mu)} \ge t_0 + T,$$

where $T \ge \frac{W_2(\eta)}{W_3(\mu)}$. Hence, if $T > \frac{W_2(\eta)}{W_3(\mu)}$, then $|x(t)| > \mu$ fails, and we have $|x(t)| < \gamma$ for all $t \ge t_0 + T$. This proves UAS. □

Copyright Elsevier 2022

Example 7.14. Consider the second-order differential equation

$$u'' + e(t)u = 0,$$

where e is continuously differentiable with

$$e'(t) \le 0 \text{ and } \lim_{t\to\infty} e(t) = e_0 > 0.$$

Making the substitution $x = u$, $y = u'$, we arrive at the equivalent system

$$x' = y, \;\; y' = -e(t)x.$$

Let

$$V(t, \mathbf{x}) = e(t)x^2 + y^2,$$

where $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$. Then along the solutions we have

$$\begin{aligned} V'(t, \mathbf{x}) &= e'(t)x^2 + 2e(t)xy - 2e(t)xy \\ &= e'(t)x^2 \\ &\le 0. \end{aligned}$$

It is clear that $V(t, 0) = 0$. We will show that the zero solution is uniformly stable using Theorem 7.1.9 if we show that $W_1(|\mathbf{x}|) \le V(t, \mathbf{x}) \le W_2(|\mathbf{x}|)$. Since e is decreasing and its limit at infinity exists and is equal to e_0, we may let

$$c = \min\{e_0, 1\} \text{ and } k = \max\{e(0), 1\}.$$

Then we can define

$$W_1(|\mathbf{x}|) = c(x^2 + y^2) = c|\mathbf{x}|, \text{ and } W_2(|\mathbf{x}|) = k(x^2 + y^2) = k|\mathbf{x}|,$$

so that we have $W_1(|\mathbf{x}|) \le e(t)x^2 + y^2 \le W_2(|\mathbf{x}|)$. This implies that the zero solution is uniformly stable.

Example 7.15. Consider the system

$$x_1' = -x_1 - e^{-2t}x_2, \;\; x_2' = x_1 - x_2.$$

Here $(0, 0)$ is the only equilibrium solution. Let

$$V(x_1, x_2) = x_1^2 + (1 + e^{-2t})x_2^2.$$

Then along the solutions we have by Definition 7.1.11 that

$$V'(t, x_1, x_2) = -2\big[x_1^2 - x_1x_2 + x_2^2(1 + 2e^{-2t})\big].$$

Copyright Elsevier 2022

Using $x_1x_2 \le \frac{x_1^2}{2} + \frac{x_2^2}{2}$, we arrive at

$$\begin{aligned} V'(t, x_1, x_2) &= -2\big(x_1^2 + x_2^2(1 + 2e^{-2t})\big) + 2x_1x_2 \\ &\le -(x_1^2 + x_2^2 + 4x_2^2e^{-2t}) \\ &\le -(x_1^2 + x_2^2) := -W_3(|x|). \end{aligned}$$

In addition, if we set

$$W_1(|x|) = \frac{1}{2}(x_1^2 + x_2^2), \;\; W_2(|x|) = x_1^2 + x_2^2,$$

then we have

$$W_1(|x|) \le V(t, x) \le W_2(|x|),$$

and by Theorem 7.1.11, the origin is uniformly asymptotically stable.

Remark 7.2. Theorems 7.1.8–7.1.10 hold for the autonomous system (7.1.11) with autonomous $V(x)$.

7.2 Global asymptotic stability

In this section we briefly discuss the notion of global asymptotic stability.

Definition 7.2.1. Assume that the zero solution of (7.1.11) is stable. If $|x(t, t_0, x_0)| \to 0$ as $t \to \infty$ for every $x(t_0) = x_0$, then we say that the zero solution is *globally asymptotically stable* (GAS).

Remark 7.3. Theorem 7.1.10 implies the zero solution is GAS if $D = \mathbb{R}^n$.
In addition, for the next theorem we note that (7.1.19) implies

$$V(x) \to \infty \;\text{ as }\; ||x|| \to \infty,$$

since $W_1(|x|) \to \infty$ as $|x| \to \infty$.

Example 7.16. Consider the system

$$x_1' = -x_1(1 - 2x_1x_2), \;\; x_2' = -x_2.$$

It is clear that $(0, 0)$ is the only equilibrium solution. Let

$$V(x_1, x_2) = \frac{1}{2}x_1^2 + x_2^2.$$

Then along the solutions we have

$$V'(x_1, x_2) = -x_1^2(1 - 2x_1x_2) - 2x_2^2.$$

Copyright Elsevier 2022

Now $V'(x_1, x_2) < 0$ if $1 - 2x_1x_2 > 0$, that is, if $x_1x_2 < \frac{1}{2}$. Thus the origin is asymptotically stable but not GAS since we are confined to the bounded region that consists of all points (x_1, x_2) such that $x_1x_2 < \frac{1}{2}$. Therefore we cannot conclude that the zero solution is GAS.

We state the following theorem regarding GAS.

Theorem 7.2.1. *Suppose the hypothesis of Theorem 7.1.3 holds with $D = \mathbb{R}^n$. In addition, if $V(x)$ is radially unbounded, that is,*

$$V(x) \to \infty \ \textit{as} \ ||x|| \to \infty, \tag{7.2.1}$$

then the zero solution of (7.1.11) is GAS.

Proof. By Theorem 7.1.2, the zero solution is stable. Suppose there is a solution or trajectory that does not converge to zero. Since V is decreasing and nonnegative, it converges to some γ as $t \to \infty$. Since x does not converge to zero, we have $\gamma \leq V(x(t)) \leq V(x_0)$. Let

$$L = \{z : \gamma \leq V(z) \leq V(x_0)\}.$$

Then L is closed and bounded and hence compact. Therefore V attains its supremum on L, that is, $\sup_{z \in L} V' = -d < 0$, or $V'(x(t)) \leq -d$ for all $t \geq t_0$. An integration of $V' \leq -d$ from 0 to T yields

$$V(x(T)) \leq V(x_0) - dT.$$

This implies that $V(x_0) < 0$ for $T > V(x_0)/d$, which is a contradiction, and hence every trajectory converges to zero. This completes the proof. ☐

Basically, the conditions on V imply that every trajectory $x(t)$ tends to the origin as $t \to \infty$. Since $x(t)$ is continuous, this requires that $x(t)$ remains bounded for all $t \geq t_0$ for arbitrary initial conditions $x(t_0)$. Hence the purpose of the radial unboundedness condition on V is to ensure that $x(t)$ remains at all times within the bounded region defined by $V(x) \leq V(x_0)$. If V were not radially unbounded, then not all level curves given by $V(x) \leq c$ for positive constant c would be closed curves, and it would be possible for $x(t)$ to drift away from the equilibrium even though $V' < 0$.

Example 7.17. The zero solution of the system in Example 7.5 is GAS.

Example 7.18. Consider the system

$$x_1' = (x_2 - 1)x_1^3, \quad x_2' = -\frac{x_1^4}{(1 + x_1^2)^2} - \frac{x_2}{1 + x_2^2}.$$

It is clear that $(0, 0)$ is the only equilibrium solution. Let

$$V(x_1, x_2) = \frac{x_1^2}{1 + x_1^2} + x_2^2.$$

Copyright Elsevier 2022

Then along the solutions we have

$$V'(x_1,x_2) = -2\frac{x_1^4}{(1+x_1^2)^2} - 2\frac{x_2}{1+x_2^2},$$

so that $V(x_1,x_2)$ is a strict Lyapunov function on the set $D=\mathbb{R}^n$, as required by Theorem 7.1.3, and the zero solution is AS. However, V is not radially unbounded since the level curves for $V=c$ are not closed curves for $c \geq 1$. Thus it is not possible to conclude that the zero solution is GAS. Neither is it possible to conclude that the zero solution is *not* GAS without further analysis, since we might find another function V that satisfies all the requirements for the GAS.

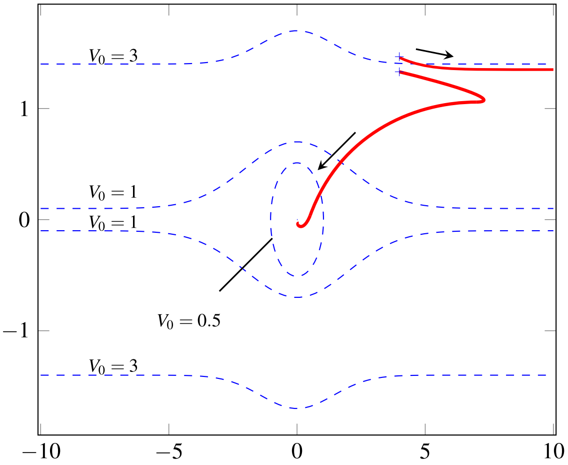


FIGURE 7.2

Level curves escaping to infinity.

As a matter of fact, according to Fig. 7.2, the zero solution is not GAS since solutions escape to infinity with the choice of the initial condition $(x_1(t_0), x_2(t_0)) = (3, 3/2)$.

Remark 7.4. We examine in detail why

$$V(x_1,x_2) = \frac{x_1^2}{1+x_1^2} + x_2^2$$

is not radially unbounded. As mentioned above, the condition that $V(x)$ is radially unbounded means that all level sets are closed. Letting $x_2 = 0$, we try to determine the level set $V(x) = 1$ or $V(x_1, 0) = \dfrac{x_1^2}{1+x_1^2} = 1$, which cannot hold for finite x_1. Thus V is not radially unbounded. Next, we examine the definition of being radially

Copyright Elsevier 2022

unbounded in the context of

$$V(x) \to \infty \text{ as } ||x|| \to \infty.$$

This means that any combinations of x_1, x_2 that make $||x|| \to \infty$ also have to make $V(x) \to \infty$. The Euclidean norm

$$||(x_1, 0)||_2 = \sqrt{x_1^2 + 0^2} = \sqrt{\infty^2 + 0^2} \to \infty.$$

However, in this case, we get for $V(x)$ that

$$\lim_{x_1 \to \infty} V(x_1, 0) = 1 \neq \infty.$$

It follows that V is not radially unbounded.

7.3 Instability

We consider once more the non-autonomous system

$$x' = f(t, x), \quad f(t, 0) = 0, \tag{7.3.1}$$

where $f \in C\big([0, \infty) \times D, \mathbb{R}^n\big)$, and $D \subset \mathbb{R}^n$ is open with $0 \in D$. We have discussed the instability for autonomous systems, but let us recall that if we negate the definition of stability, then there are $\varepsilon > 0$ and $t_0 \geq 0$ such that for any $\delta > 0$, there are x_0 with $|x_0| < \delta$ and $t_1 > t_0$ such that

$$|x(t_1, t_0, x_0)| \geq \varepsilon.$$

Then the zero solution is unstable. We make the following formal definition.

Definition 7.3.1. The zero solution of (7.3.1) is unstable if there are $\varepsilon > 0$ and $t_0 \geq 0$ such that for any $\delta > 0$, there are x_0 with $|x_0| < \delta$ and $t_1 > t_0$ such that $|x(t_1, t_0, x_0)| \geq \varepsilon$.

Theorem 7.3.1. *Suppose* $V : [0, \infty) \times D \to [0, \infty)$ *is a continuous function locally Lipschitz in x such that*

$$W_1(|x|) \leq V(t, x) \leq W_2(|x|), \tag{7.3.2}$$

and along the solutions of (7.3.1) *we have*

$$V'(t, x) \geq W_3(|x|). \tag{7.3.3}$$

Then the zero solution of (7.3.1) *is unstable.*

Copyright Elsevier 2022

Proof. Suppose not. Then for $\varepsilon = \min\{1, d(0, \partial D)\}$, where ∂D denotes the boundary of the set D, we can find $\delta > 0$ such that $|x_0| < \delta$ and $t \geq 0$ imply that $|x(t, 0, x_0)| < \varepsilon$. We may pick x_0 such that $|x_0| = \delta/2$ and find $\gamma > 0$ with $W_2(\gamma) = W_1(\delta/2)$. Then for $x(t) = x(t, 0, x_0)$, we have $V'(t, x) \geq 0$ by (7.3.3), so that

$$W_2(|x(t)|) \geq V(t, x(t)) \geq V(0, x_0) \geq W_1(\delta/2) = W_2(\gamma),$$

from which we conclude that $\gamma \leq |x(t)|$ for $t \geq 0$. Thus

$$V'(t, x) \geq W_3(|x(t)|) \geq W_3(\gamma).$$

It follows that

$$W_2(|x(t)|) \geq V(t, x(t)) \geq V(0, x_0) + tW_3(\gamma),$$

from which we conclude that $|x(t)| \to \infty$, a contradiction. This completes the proof. □

We ask the curious reader to compare condition (7.3.2) with the hypothesis of Theorem 7.1.6 and try to have an appreciation and some understanding of the differences between autonomous systems and non-autonomous systems.

In Theorem 7.3.1 we asked for $|x_0| = \delta/2$, which is too much to ask. This will be relaxed in the next theorem.

Theorem 7.3.2. *Suppose that* $V : [0, \infty) \times D \to (-\infty, \infty)$ *is a continuous function locally Lipschitz in* x *and that there is a sequence* $x_n \in D$ *starting at* $t_0 \geq 0$ *with* $V(t_0, x_n) > 0$ *and* $x_n \to 0$ *as* $n \to \infty$. *If*

$$W_1(|x|) \leq V(t, x) \leq W_2(|x|)$$

and if along the solutions of (7.3.1) *we have*

$$V'(t, x) \geq W_3(|x|),$$

then the zero solution of (7.3.1) *is unstable.*

Proof. Suppose not. Then there are ε and δ as in the proof of Theorem 7.3.1. For δ, we may find t_n and x_n with $|x_n| < \delta$ and $V(t_n, x_n) > 0$. The rest of the proof is identical to that of Theorem 7.3.1. This completes the proof. □

7.4 ω-limit set

We discuss the Krasovskii–LaSalle invariance principle, which provides a way to prove the asymptotic stability when V' is negative *semi-definite*. It is usually much easier to construct such a function V. We begin with the following definition.

Copyright Elsevier 2022

Definition 7.4.1. Let D be an open subset of $\mathbb{R}^n$ with $0 \in D$. A continuous autonomous function $V : D \to [0, \infty)$ is

(a) *positive definite* if

$$V(0) = 0 \text{ and } V(x) > 0 \text{ for } x \neq 0,$$

(b) *positive semi-definite* if

$$V(0) = 0 \text{ and } V(x) \geq 0 \text{ for } x \neq 0.$$

To illustrate the differences, consider the two functions

$$V_1(x) = x_1^2 \text{ and } V_2(x) = x_1^2 + x_2^2.$$

Clearly, both functions are nonnegative. However, it is possible for V_1 to be zero even when $x \neq 0$. Particularly, if we set $x = (0, d)$ where $d \in \mathbb{R}$ is any non-zero number, then $V_1(x) = V_1(0, d) = 0$, whereas $V_2(x) = 0$ if and only if $x = 0$. We conclude that V_1 is positive semi-definite and V_2 is positive definite.

Remark 7.5. If the signs on V' are reversed in Definition 7.4.1, then V is said to be *negative definite* and *negative semi-definite*, respectively.

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuous with continuous first-order partial derivatives with respect to x_i, $i = 1, 2, \ldots, n$. Consider the initial value autonomous system

$$x' = f(x), \; x(t_0) = x_0, \; t \geq t_0. \tag{7.4.1}$$

Let x^* be an equilibrium solution of (7.4.1).

Let $x(t, t_0, x_0)$ be a trajectory of (7.4.1) at time t starting from x_0 at the initial time t_0. Denote the set of all points lying along the trajectory by $x(\cdot, t_0, x_0)$

Definition 7.4.2. (ω-limit set) The *ω-limit set* of a trajectory $x(\cdot, t_0, x_0)$ is the set of all points $z \in \mathbb{R}^n$ such that

$$\lim_{n \to \infty} x(t_n, t_0, x_0) = z,$$

where the sequence t_n is strictly increasing.

Definition 7.4.3. (Invariant set) The set $M \in \mathbb{R}^n$ is said to be an *invariant set* if for all $y \in M$ and $t_0 \geq 0$, we have

$$x(t, t_0, y) \in M \quad \text{for all } t \geq t_0.$$

Theorem 7.4.1. *Let V be a Lyapunov function of* (7.4.1) *on a bounded open set $U \in \mathbb{R}^n$ with $x^* \in U$. Let c be a positive constant, and set*

$$S = \{x \in U : V(x) \leq c\},$$

which is closed in $\mathbb{R}^n$. Then S is invariant for $t \geq t_0 \geq 0$.

Copyright Elsevier 2022

Proof. Let $x \in S$. Since V is a Lyapunov function on U, we have that $V(x(t,t_0,x_0)) \leq V(x_0) \leq c$ for $t \geq t_0 \geq 0$ as long as the solutions exist. Now since S is closed and bounded, the existence of solutions holds for all $t \geq t_0 \geq 0$. As a result, $x(t,t_0,x_0)$ remains in S for all $t \geq t_0 \geq 0$. This shows that the set S is invariant. This completes the proof. □

In Example 7.3 we considered $V(x,y) = x^2 + y^2$ and obtained $V' = 0$. As a result of Theorem 7.4.1, we have that the set

$$S = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq c\} \text{ for } c > 0$$

is invariant.

Theorem 7.4.2. *(Krasovskii–LaSalle invariance principle) Let D be an open subset of $\mathbb{R}^n$ containing x^*. Let $V : D \to \mathbb{R}$ be a positive definite function such that on the compact set*

$$\Omega_r = \{x \in D : V(x) \leq r, \ \ r > 0\},$$

we have $V'(x) \leq 0$. Define

$$S = \{x \in \Omega_r : V'(x) = 0\}.$$

Then, as $t \to \infty$, the trajectory tends to the largest invariant set inside S, that is, its ω-limit set is contained inside the largest invariant set in S. If S contains no invariant sets other than $x = x^*$, then the equilibrium solution x^* of (7.4.1) is asymptotically stable.

Proof. Let M be the largest invariant set in S, and let L^+ be the ω-limit set for the trajectory $x(t,t_0,x_0)$, which is invariant. Since $V'(x) \leq 0$ and is continuous in Ω_r, we have that $V(x) \geq b = \min_{x \in \Omega_r} V(x) > 0$. Thus there is a positive constant a such that

$$\lim_{t \to \infty} V(x(t,t_0,x_0)) = a.$$

Now $x(t,t_0,x_0) \in \Omega_r$ implies $x(t,t_0,x_0)$ is bounded, which implies that L^+ exists. Moreover, $L^+ \in \Omega_r$, and $x(t,t_0,x_0)$ approaches L^+ as $t \to \infty$. For any $z \in L^+$, there is an increasing sequence $\{t_n\}$ such that

$$\lim_{n \to \infty} x(t_n,t_0,x_0) = z.$$

As $V(x)$ is continuous, we have that

$$V(z) = \lim_{n \to \infty} V(x(t_n,t_0,x_0)) = a.$$

Since $V(x) = a$ on L^+ and L^+ is invariant, we get that $V'(x) = 0$ for all $x \in L^+$ and

$$L^+ \subset M \subset S \subset \Omega_r.$$

Copyright Elsevier 2022

Thus $x(t, t_0, x_0) \to L^+$, and hence $x(t, t_0, x_0) \to M$ as $t \to \infty$. This completes the proof. □

We provide a couple of examples.

Example 7.19. Consider the system

$$x' = y,$$
$$y' = -(x^3 - x^5) - y^3.$$

Let $D = \{(x, y) : x^2 + y^2 \leq 1\}$ and define

$$V(x, y) = \frac{y^2}{2} + \frac{x^4}{4} - \frac{x^6}{6}.$$

Then V is positive definite on D. We can easily verify that $V' = -y^4$. Thus $V' = 0$ along the x-axis, and V' is negative semi-definite. Note that on D, we have $V(x, y) \leq \frac{y^2}{2} + \frac{x^4}{4} \leq \frac{1}{2}(\frac{1}{\sqrt{2}})^2 + \frac{1}{4}(\frac{1}{\sqrt{2}})^4 = \frac{5}{16}$. Thus $\Omega_r = \{(x, y) \in D : \frac{y^2}{2} + \frac{x^4}{4} \leq \frac{5}{16}\}$. Moreover,

$$S = \{(x, y) \in \Omega_r : V'(x) = 0\} = \{(x, y) \in \Omega_r : y = 0\}.$$

Now V is constant only on solutions that lie on the x-axis, but by the uniqueness of solutions, the only such solution is the equilibrium $(0, 0)$. Thus $M = \{(x, y) \in S : x = 0, y = 0\}$ is the largest invariant subset of S. By Theorem 7.4.2, we have that all trajectories in S go to $(0, 0)$ as $t \to \infty$. Note that the other two equilibrium solutions, $(-1, 0)$ and $(1, 0)$, are outside the set Ω_r.

Corollary 7.1. *If the hypotheses of Theorem 7.4.2 hold with $D = \mathbb{R}^n$ and $V(x)$ is radially unbounded, then the origin is GAS.*

Example 7.20. Consider the system

$$x' = y,$$
$$y' = -h_1(x) - h_2(y),$$

where the functions h_1 and h_2 are continuous and satisfy

$$h_i(0) = 0, \;\; uh_i(u) > 0, \;\; i = 1, 2, \;\; \text{for } \; 0 < |u| < a.$$

Let $D = \{(x, y) : |x| < a, \; |y| < a\}$ and set

$$V(x, y) = \frac{y^2}{2} + \int_0^x h_1(s)ds.$$

Copyright Elsevier 2022

Then $V(0,0)=0$ and $V(x,y)>0$ on D. In addition,

$$\begin{aligned} V'(x,y) &= h_1(x)x' + yy' \\ &= h_1(x)y + y\big[-h_1(x) - h_2(y)\big] \\ &= -yh_2(y). \end{aligned}$$

Now $V'(x,y)=0$ implies $yh_2(y)=0$, and hence $y=0$. Thus $V'=0$ along the x-axis, and V' is negative semi-definite. Note that on D, we have

$$V(x,y) \le \frac{a^2}{2} + \int_0^a a\,ds = \frac{3}{2}a^2.$$

Thus $\Omega_a = \{(x,y) \in D : V(x,y) \le \frac{3}{2}a^2\}$. Moreover,

$$S = \{(x,y) \in \Omega_a : V'(x) = 0\} = \{(x,y) \in \Omega_a : y = 0\}.$$

As $y(t)=0$, we have $y'(t)=0$. The second equation of the system gives

$$h_1(x(t)) + h_2(0) = h_1(x(t)) = 0,$$

from which we have $x(t)=0$. Hence the only solution that can stay identically in S is $(x(t),y(t))=(0,0)$, that is, $M=\{(x,y)\in S : x=0,\ y=0\}$ is the largest invariant subset of S. Thus the origin is asymptotically stable by Theorem 7.4.2.

Now we suppose $a=\infty$ and $\displaystyle\int_0^u h_1(s)ds \to \infty$ as $|u|\to\infty$. Then $D=\mathbb{R}^2$, and $V(x,y)=\frac{y^2}{2}+\int_0^x h_1(s)ds$ is radially unbounded. Moreover,

$$S = \{(x,y) \in \Omega_a : V'(x) = 0\} = \{(x,y) \in \Omega_a : y = 0\},$$

and the only solution that can stay identically in S is $(x(t),y(t))=(0,0)$. Thus the origin is GAS. Note that we did not have to worry about the set Ω_a since $a=\infty$ or $D=\mathbb{R}^2$.

We end this section by revisiting the pendulum with friction that we considered in Examples 3.5 and 6.9.

Example 7.21. In Example 3.5 we considered a pendulum with friction and arrived at the planar autonomous system

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -\frac{g}{L}\sin(x_1) - \frac{b}{m}x_2. \end{aligned} \tag{7.4.2}$$

We are interested in examining the stability of the zero solution by constructing a suitable Lyapunov function. Let $D=\{(x_1,x_2)\in\mathbb{R}^2 : |x_1|<\pi\}$ and set

$$V(x_1,x_2) = \frac{g}{L}\big(1-\cos(x_1)\big) + \frac{1}{2}x_2^2.$$

Copyright Elsevier 2022

Then $V(0,0)=0$, and $V(x_1,x_2)>0$ on $D-(0,0)$. Moreover, along the solutions of (7.4.2) we have

$$\begin{aligned} V'(x_1,x_2) &= \frac{g}{L}\sin(x_1)x_1' + x_2x_2' \\ &= \frac{g}{L}x_2\sin(x_1) + x_2\big(-\frac{g}{L}\sin(x_1) - \frac{b}{m}x_2\big) \\ &= -\frac{k}{m}x_2^2. \end{aligned}$$

Now $V'=0$ along the x_1-axis, and V' is negative semi-definite in D. Let

$$S=\{(x,y)\in D : V'(x)=0\}.$$

As $x_2(t)=0$, this implies that $x_1'(t)=0$, and $x_1(t)$ is constant. Similarly, for $x_2(t)=0$, this implies that $x_2'(t)=0$, and hence $\sin(x_1)=0$. However, the only point on the segment $-\pi<x_1<\pi$ rendering $\sin(x_1)=0$ is $x_1=0$. Hence the only solution that can forever stay in S is $(x_1(t),x_2(t))=(0,0)$. Thus $M=\{(x,y)\in S : x=0, y=0\}$ is the largest invariant subset of S, and the origin is asymptotically stable by Theorem 7.4.2.

Again, there was no need to construct the set Ω_a since $a=\infty$.

We end this section by giving an example in which we expose the difficulties of using Theorem 7.1.3, which requires V to be a strict Lyapunov function, and the advantage of using Theorem 7.4.2. First, we state the following lemma.

Lemma 7.3. *(Young's inequality) For any two nonnegative real numbers w and z, we have*

$$wz \le \frac{w^e}{e} + \frac{z^f}{f} \quad \textit{with } \frac{1}{e}+\frac{1}{f}=1,\ e,f\in(1,\infty).$$

Example 7.22. Consider the autonomous system

$$x'=-y-x^3,\quad y'=x^5.$$

We see that $(0,0)$ is the only equilibrium solution. Let $D=\{(x,y)\in\mathbb{R}^2 : |y|^2\le 1/2\}$ and set

$$V(x,y)=x^6+xy^3+3y^2.$$

We need to show V is a strict Lyapunov function on the set D. Clearly, $V(0,0)=0$. We need to worry about the mixed term xy^3. Let $e=6$ and $f=6/5$. Then by Lemma 7.3, for $|y|\le 1$, we have

$$|xy^3| = |x||y^3| \le \frac{|x|^6}{6} + \frac{5|y|^{18/5}}{6} \le \frac{1}{6}x^6 + \frac{5}{6}y^2.$$

Copyright Elsevier 2022

This implies that

$$3y^2 \geq \frac{18}{5}|xy^3| - \frac{3}{5}x^6.$$

Substituting $3y^2$ into V gives

$$\begin{aligned} V(x, y) &\geq x^6 + xy^3 - \frac{3}{5}x^6 + \frac{18}{5}|xy^3| \\ &= \frac{2}{5}x^6 + xy^3 + |xy^3| + \frac{13}{5}|xy^3| \\ &\geq \frac{2}{5}x^6 + \frac{13}{5}|xy^3|. \end{aligned}$$

Along the solutions of the system we have

$$V'(x, y) = -6x^8 - x^3y^3 + 3x^6y^2 - y^4.$$

Next, we apply Lemma 7.3 to reduce the two mixed terms. Let $(x, y) \in D$. Let $e = 8/3$ and $f = 8/5$. Then

$$|-x^3y^3| = |x^3||y^3| \leq \frac{3|x|^8}{8} + \frac{5|y|^{24/5}}{8} \leq \frac{3}{8}x^8 + \frac{5}{8}y^4,$$

and similarly, if we let $e = 8/6$ and $f = 8/2$, then

$$|3x^6y^2| = 3|x|^6|y|^2 \leq 3[\frac{3|x|^8}{4} + \frac{|y|^8}{4}] = \frac{9}{4}x^8 + \frac{3}{4}y^8.$$

Thus

$$\begin{aligned} V'(x, y) &= -6x^8 + \frac{3}{8}x^8 + \frac{9}{4}x^8 - y^4 + \frac{5}{8}y^4 + \frac{3}{4}y^8 \\ &\leq -\frac{27}{8}x^8 - y^4 + \frac{5}{8}y^4 + \frac{3}{4}y^4(\frac{1}{4}) \\ &= -\frac{27}{8}x^8 - \frac{3}{16}y^4 < 0 \end{aligned}$$

on $D - \{(0, 0)\}$. By Theorem 7.1.3, the zero solution is asymptotically stable. Next, we use Theorem 7.4.2. Let

$$V(x, y) = x^6 + 3y^2.$$

Clearly, $V(0, 0) = 0$, and $V(x, y) > 0$ for $(x, y) \neq (0, 0)$. Moreover,

$$V'(x, y) = -6x^8 \leq 0.$$

Thus $V' = 0$ along the x-axis, and hence V is a Lyapunov function (semi-definite). Here $D = \mathbb{R}^2$, and

$$S = \{(x, y) \in D : V'(x) = 0\}.$$

Copyright Elsevier 2022

As $y = 0$, this implies that $y' = 0$, from which we have $x^5 = 0$, that is, $x = 0$. Thus $M = \{(x, y) \in S : x = 0, y = 0\}$ is the largest invariant subset of S, and the origin is asymptotically stable by Theorem 7.4.2.

7.5 Connection between eigenvalues and Lyapunov functions

Consider the homogeneous system

$$x'(t) = Ax(t), \tag{7.5.1}$$

where A is an $n \times n$ constant matrix. We are interested in relations between the existence of a Lyapunov function for system (7.5.1) and its eigenvalues. It turns out that if all the eigenvalues of A have negative real parts, then there is a strict Lyapunov function. We begin with the following definition.

Definition 7.5.1. Let $x \in \mathbb{R}^n$, and let B be an $n \times n$ constant matrix. We say that B is *positive definite* (*negative definite*) if

$$x^T Bx > 0(< 0) \quad \text{for } x \neq 0.$$

Example 7.23. Let

$$B = I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Then

$$\begin{aligned} x^T Bx &= (x_1, x_2, \dots, x_n) \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= x_1^2 + x_2^2 + \ldots + x_n^2 > 0 \quad \text{for } x \neq 0. \end{aligned}$$

Thus the $n \times n$ identity matrix I is positive definite.

To better understand where we are heading, we assume the existence of a positive definite symmetric $n \times n$ constant matrix B and define

$$V(x) = x^T Bx. \tag{7.5.2}$$

Copyright Elsevier 2022

Then $V(x)$ is positive definite, and, moreover, along the solutions of (7.5.1) we have that

$$\begin{aligned} V'(x) &= (x^T)' Bx + x^T Bx' \\ &= (x')^T Bx + x^T Bx' \\ &= (Ax)^T Bx + x^T B(Ax) \\ &= x^T A^T Bx + x^T BAx \\ &= x^T \Big(A^T B + BA\Big) x. \end{aligned} \tag{7.5.3}$$

Notice that

$$\begin{aligned} \Big(A^T B + BA\Big)^T &= B^T A + A^T B^T \\ &= A^T B + BA. \end{aligned}$$

Thus $A^T B + BA$ is symmetric, and so is $A^T B + BA = -C$ for a symmetric matrix C; we wish $-C$ to be negative definite to obtain stability results concerning system (7.5.3).

Remark 7.6. If (7.5.2) holds, then there are positive constants α_1 and α_2 such that

$$\alpha_1 x^T x \le x^T Bx \le \alpha_2 x^T x. \tag{7.5.4}$$

Then $W_1(|x|) = \alpha_1 x^T x$ and $W_2(|x|) = \alpha_2 x^T x$ could serve as two wedges such that

$$\alpha_1 x^T x \le V(x) \le \alpha_2 x^T x.$$

We have the following theorem.

Theorem 7.5.1. *Given an $n \times n$ constant matrix A and a positive definite symmetric matrix C, there exists a positive definite symmetric matrix B such that*

$$A^T B + BA = -C$$

if and only if the zero solution of (7.5.1) *is asymptotically stable.*

Proof. ($\Rightarrow$) Let V be given by (7.5.2) for a positive definite symmetric matrix B. Substituting $A^T B + BA = -C$ into (7.5.3) gives $V'(x) = -x^T Cx < 0$. Then by Theorem 7.1.3, the zero solution is AS.

($\Leftarrow$) Given A, suppose C is symmetric positive definite matrix. Assume that $x = 0$ is AS and let $\Phi(t)$ be the fundamental matrix of (7.5.1). Then by Theorem 4.2.3, there exist positive constants K and α such that

$$||\Phi(t)\Phi^{-1}(s)|| \le Ke^{-\alpha(t-s)}, \ \ t > s.$$

Copyright Elsevier 2022

In particular, $||e^{At}e^{-As}|| \le Ke^{-\alpha(t-s)}$, $t > s$, or

$$||e^{A(t-s)}|| \le Ke^{-\alpha(t-s)}, \quad t > s.$$

Now $\left(e^{A}\right)^{T} = e^{A^T}$, and so the same above estimate holds for $e^{A^T(t-s)}$. Define

$$B(t) = \int_t^\infty e^{A^T(s-t)}Ce^{A(s-t)}ds,$$

which exists since $||e^{At}e^{-As}|| \le Ke^{-\alpha(t-s)}$, $t > s$. Differentiate both sides to get

$$\begin{aligned}\frac{d}{dt}B(t) &= -C - A^T\int_t^\infty e^{A^T(s-t)}Ce^{A(s-t)}ds - \int_t^\infty e^{A^T(s-t)}Ce^{A(s-t)}dsA\\ &= -C - A^TB(t) - B(t)A.\end{aligned}$$

Let $u = s - t$ and rewrite the integral as

$$B(t) = \int_0^\infty e^{A^Tu}Ce^{Au}du,$$

which does not depend on t at all, that is, $\frac{d}{dt}B(t) = 0$, or

$$-C - A^TB(t) - B(t)A = 0.$$

Finally, B is symmetric positive definite because C is. To see this,

$$\begin{aligned}B^T &= \int_0^\infty \left(e^{A^Tu}Ce^{Au}\right)^T du\\ &= \int_0^\infty (e^{Au})^TC^T(e^{A^Tu})^Tdu\\ &= \int_0^\infty e^{A^Tu}Ce^{Au}du = B.\end{aligned}$$

Thus B is symmetric. It remains to show that B is positive definite. Let $x \in \mathbb{R}^n$ and consider the integrand. Then, for each t, we have

$$x^Te^{A^Tt}Ce^{At}x = \left(e^{At}x\right)^TC\left(e^{At}x\right) := y^TCy,$$

and if $x \neq 0$, then $y \neq 0$; hence, as C is positive definite, $y^TCy > 0$ for $x \neq 0$. Thus B is positive definite. This completes the proof. □

In Theorem 5.1.4 we showed if $Re(\sigma(A)) < 0$, then the zero solution of $x' = Ax$ is AS. Thus we may reformulate Theorem 7.5.1 in terms of the eigenvalues of the matrix A.

Copyright Elsevier 2022

Theorem 7.5.2. *Given an $n \times n$ matrix A and a positive definite symmetric matrix C, there exists a positive definite symmetric matrix B such that*

$$A^T B + BA = -C$$

if and only if every eigenvalue of the matrix A has a negative real part.

Example 7.24. Let us find a strict Lyapunov function to show that the zero solution of

$$x' = \begin{pmatrix} -1 & 0 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ t > 0,$$

is asymptotically stable. Let $A = \begin{pmatrix} -1 & 0 \\ 2 & -3 \end{pmatrix}$. We will find a positive definite symmetric matrix B such that $A^T B + BA = -I$. If $B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, then the relation $A^T B + BA = -I$ implies that

$$\begin{pmatrix} -2a + 4b & -4b + 2c \\ -4b + 2c & -6c \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The above equations give $a = \frac{2}{3}$, $b = \frac{1}{12}$, and $c = \frac{1}{6}$. Define

$$\begin{aligned} V(x_1, x_2) = x^T Bx = (x_1 \ \ x_2) \begin{pmatrix} \frac{2}{3} & \frac{1}{12} \\ \frac{1}{12} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \frac{2}{3}x_1^2 + \frac{x_1 x_2}{6} + \frac{x_2^2}{6} \\ &\geq \frac{1}{12}x_1^2 + \frac{x_1 x_2}{6} + \frac{x_2^2}{12} \\ &= \frac{1}{12}(x_1 + x_2)^2. \end{aligned}$$

Thus $V(x_1, x_2)$ is positive definite. On the other hand, after simple calculations, we arrive at

$$\begin{aligned} V'(x_1, x_2) &= -x_1^2 - \frac{2}{3}x_2^2 \\ &\leq -\frac{2}{3}(x_1^2 + x_2^2). \end{aligned}$$

Thus V is a strict Lyapunov function, and by Theorem 7.5.2, the zero solution of the system is AS. We could avoid verifying that V is a strict Lyapunov function and using Theorem 7.5.2 by finding a matrix B satisfying $A^T B + BA = -I$.

Copyright Elsevier 2022

Remark 7.7. To ensure that the symmetric matrix B is positive definite, where $B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, it is sufficient for its coefficients to satisfy

$$a > 0, \;\; c > 0, \;\; ac - b^2 > 0.$$

In the investigation of stability and boundedness for nonlinear non-autonomous systems, the effective approach is to study equations related in some way to the linear equations whose behavior is known to be covered by the known theory. We consider the perturbed nonlinear system

$$x'(t) = Ax(t) + g(t, x(t)), \;\; x(t_0) = x_0, \;\; t \geq t_0 \geq 0, \tag{7.5.5}$$

where A is an $n \times n$ constant matrix, and $g : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$ is a function continuous in t and x. Here $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $g(t, x) = \begin{pmatrix} g_1(t, x) \\ g_2(t, x) \\ \vdots \\ g_n(t, x) \end{pmatrix}$. The perturbation term g is assumed to be small in some sense. The next theorem pertains to the stability of the perturbed nonlinear system (7.5.5).

By $|x|$ we denote the Euclidean norm, and for an $n \times n$ matrix C, by $|C|$ we denote a compatible norm such that $|Cx| \leq |C||x|$.

Theorem 7.5.3. *Suppose there exists a positive definite symmetric matrix B such that*

$$A^T B + BA = -I \tag{7.5.6}$$

and

$$\lim_{x \to 0} \frac{|g(t, x)|}{|x|} = 0 \;\; \textit{uniformly in } t \in [0.\infty). \tag{7.5.7}$$

Then the zero solution of (7.5.5) *is uniformly asymptotically stable.*

Proof. Write x for $x(t)$ and define the function V by

$$V = x^T Bx.$$

Using (7.5.6) and (7.5.7), along the solutions of (7.5.5) we have

$$\begin{aligned} V' &= (Ax + g(t, x))^T B + x^T B(Ax + g(t, x)) \\ &= x^T(-I)x + 2x^T Bg(t, x) \\ &\leq -|x|^2 + 2|x||B||g(t, x)|. \end{aligned}$$

Let $\varepsilon > 0$. Due to (7.5.7), we may choose ε sufficiently small so that for $|x| < \varepsilon$, we have

$$2|B||g(t, x)| < \frac{|x|}{2}.$$

Copyright Elsevier 2022

Thus

$$V' \le -|x|^2 + \frac{|x|^2}{2} = -\frac{|x|^2}{2}.$$

Since B is symmetric and positive definite, there are positive constants α and β such that

$$\alpha^2 x^T x \le V(x) \le \beta^2 x^T x, \tag{7.5.8}$$

from which we get

$$-|x|^2 \le -\frac{1}{\beta^2}V(x) \le -\frac{\alpha^2}{\beta^2}|x|^2.$$

As a consequence,

$$V' \le -\frac{1}{2\beta^2}V(x).$$

An integration of the above inequality from t_0 to t yields

$$V(x(t)) \le V(x_0)e^{-\frac{1}{2\beta^2}(t-t_0)}.$$

We have by (7.5.8)

$$\alpha^2|x|^2 \le V(x) \le V(x_0)e^{-\frac{1}{2\beta^2}(t-t_0)}.$$

It follows that

$$\alpha^2|x|^2 \le x_0^T B x_0 e^{-\frac{1}{2\beta^2}(t-t_0)} \le |x_0^T||B||x_0|e^{-\frac{1}{2\beta^2}(t-t_0)}.$$

For $\varepsilon > 0$ and $|x_0| < \delta$, we have

$$|x| \le \frac{1}{\alpha}(\delta|B|\delta)^{\frac{1}{2}} = \frac{\delta}{\alpha}|B|^{\frac{1}{2}} < \varepsilon$$

for $\delta = \dfrac{\varepsilon\alpha}{|B|^{1/2}}$. Hence the zero solution is uniformly stable. Let $\delta = 1$ from the uniform stability, so that for $|x(t_0)| = |x_0| < \delta$, we need to show that $|x(t)| < \varepsilon$ for $t \ge t_0 + T(\varepsilon)$.

Set $T(\varepsilon) > -4\beta^2 \mathrm{Ln}\left(\frac{\alpha\varepsilon}{|B|^{1/2}}\right)$. Then

$$\begin{aligned} |x| &\le \frac{|B|^{1/2}}{\alpha}e^{-\frac{1}{4\beta^2}(t-t_0)} \\ &\le \frac{|B|^{1/2}}{\alpha}e^{-\frac{1}{4\beta^2}T(\varepsilon)} \\ &< \varepsilon. \end{aligned}$$

This completes the proof. □

Copyright Elsevier 2022

7.6 Exponential stability

We return to the non-autonomous system

$$x' = f(t,x), \quad f(t,0) = 0, \tag{7.6.1}$$

where $f \in C\big([0,\infty) \times D, \mathbb{R}^n\big)$, and $D \subset \mathbb{R}^n$ is an open set with $0 \in D$.

Definition 7.6.1. We say that solutions of system (7.6.1) are *bounded* if any solution $x(t,t_0,x_0)$ of (7.6.1) satisfies

$$\|x(t,t_0,x_0)\| \le C\Big(|x_0|, t_0\Big) \quad \text{for all } t \ge t_0,$$

where $C : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is a constant that depends on t_0 and x_0. We say that solutions of system (7.6.1) are uniformly bounded if C is independent of t_0.

Definition 7.6.2. Suppose $f(t,0) = 0$. We say the zero solution of (7.6.1) is α-*exponentially asymptotically stable* if there exists a continuous function $\alpha(t)$ such that $\int_{t_0}^t \alpha(s)ds \to \infty$ as $t \to \infty$ and constants d and $C \in \mathbb{R}^+$ such that for any solution $x(t,t_0,\varphi)$ of (7.6.1),

$$\|x(t,t_0,\varphi)\| \le C\Big(|\varphi|, t_0\Big)\big(e^{-\int_{t_0}^t \alpha(s)ds}\big)^d \quad \text{for all } t \in [t_0,\infty).$$

The zero solution of (7.6.1) is said to be α-uniformly exponentially asymptotically stable if C is independent of t_0.

Definition 7.6.3. (Exponential stability) The zero solution of system (7.6.1) said to be *exponentially stable* if any solution $x(t,t_0,x_0)$ of (7.6.1) satisfies

$$\|x(t,t_0,x_0)\| \le C\Big(|x_0|, t_0\Big)e^{-d(t-t_0)} \quad \text{for all } t \ge t_0,$$

where d is a positive constant, and $C : [0,\infty) \times [0,\infty) \to [0,\infty)$. The zero solution of (7.6.1) said to be *uniformly exponentially stable* if C is independent of t_0.

Theorem 7.6.1. *Assume that $D \subset \mathbb{R}^n$ and $V : D \to [0,\infty)$ is a continuous function such that for all $(t,x) \in [0,\infty) \times D$,*

$$W_1(||x||) \le V(t,x), \tag{7.6.2}$$

$$V'(t,x) \le -\alpha(t)V^q(t,x) + \beta(t), \tag{7.6.3}$$

$$0 \le V(t,x) - V^q(t,x) \le \gamma(t), \tag{7.6.4}$$

where $\alpha(t)$, $\gamma(t)$, and $\beta(t)$ are nonnegative continuous functions, and p, q are positive constants. Then all solutions of (7.6.1) that start in D satisfy

Copyright Elsevier 2022

$$\begin{aligned}||x|| \le W_1^{-1}\Big\{ & V(t_0, x_0)e^{-\int_{t_0}^{t}\alpha(s)ds} \\ & + \int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\, e^{-\int_u^t \alpha(s)ds}\, du\Big\}\end{aligned} \tag{7.6.5}$$

for all $t \ge t_0$.

In case $q = 1$, the proof is a simple application of the variation of parameters formula. On the other hand, if $q \neq 1$, then the variation of parameters formula will not be of much help. Now we prove the theorem.

Proof. For any initial time $t_0 \ge 0$ and $x_0 \in D$, let $x(t, t_0, x_0)$ be a solution of (7.6.1) with $x_{t_0} = x_0$. Consider

$$\begin{aligned}\frac{d}{dt}\Big(e^{\int_{t_0}^{t}\alpha(s)ds}V(t,x)\Big) &= \Big[V'(t,x)+\alpha(t)V(t,x)\Big]e^{\int_{t_0}^{t}\alpha(s)ds} \\ &\le \Big(-\alpha(t)V^q(t,x)+\beta(t)+\alpha(t)V(t,x)\Big)e^{\int_{t_0}^{t}\alpha(s)ds} \\ &\le \Big(\alpha(t)(V(t,x)-V^q(t,x))+\beta(t)\Big)e^{\int_{t_0}^{t}\alpha(s)ds} \\ &\le \Big(\alpha(t)\gamma(t)+\beta(t)\Big)e^{\int_{t_0}^{t}\alpha(s)ds}.\end{aligned}$$

Integrating both sides from t_0 to t, we obtain, for $t \in [t_0, \infty)$,

$$\begin{aligned}V(t,x)e^{\int_{t_0}^{t}\alpha(s)ds} \le\; & V(t_0, x_0) \\ & + \int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\, e^{\int_{t_0}^{u}\alpha(s)ds}\, du.\end{aligned} \tag{7.6.6}$$

It follows that

$$V(t,x) \le V(t_0,x_0)e^{-\int_{t_0}^{t}\alpha(s)ds} + \int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\, e^{-\int_u^t \alpha(s)ds}\, du$$

for all $t \in [t_0, \infty)$. Inequality (7.6.2) implies that

$$||x(t)|| \le W^{-1}\Big\{V(t_0,x_0)e^{-\int_{t_0}^{t}\alpha(s)ds} + \int_{t_0}^{t}\Big(\gamma\alpha(u)+\beta(u)\Big)\, e^{-\int_u^t \alpha(s)ds}\, du\Big\}$$

for all $t \ge t_0$. This concludes the proof. □

It is easy to see that if $W_1(||x||) = ||x||^p$ for some positive p, then inequality (7.6.5) becomes

$$\begin{aligned}||x|| \le \Big\{ & V(t_0, x_0)e^{-\int_{t_0}^{t}\alpha(s)ds} \\ & + \int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\, e^{-\int_u^t \alpha(s)ds}\, du\Big\}^{1/p}.\end{aligned} \tag{7.6.7}$$

Copyright Elsevier 2022

Corollary 7.2. *Assume that the hypotheses of Theorem 7.6.1 hold.*

(i) If

$$\int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\,e^{-\int_u^t \alpha(s)ds}\,du \le M, \forall t \ge t_0 \ge 0, \tag{7.6.8}$$

for some positive constant M, then all solutions of (7.6.1) are uniformly bounded.

(ii) If

$$\int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\,e^{\int_{t_0}^u \alpha(s)ds}\,du \le M \tag{7.6.9}$$

for some positive constant M, and

$$\int_{t_0}^{t}\alpha(s)ds \to \infty \;\; as \;\; t\to\infty \;\; for\ all \;\; t \ge t_0, \tag{7.6.10}$$

then every solution of solution (7.6.1) decays α-exponentially asymptotically to zero with $d = 1/p$.

(iii) If

$$f(t,0)=0$$

and (7.6.9) holds for q positive constant α, then the zero solution of (7.6.1) is exponentially stable.

Proof. Let x be a solution to (7.6.1) that starts in D for all $t \ge t_0 \ge 0$. Consequently, the proof of *(i)* is an immediate consequence of inequality (7.6.7). For the proof of *(ii)*, we consider the inequality from the proof of Theorem 7.6.1

$$V(t,x_t)e^{\int_{t_0}^t \alpha(s)ds} \le V(t_0,x_0) + \int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\,e^{\int_{t_0}^u \alpha(s)ds}\,du$$

for all $t \in [t_0,\infty)$. This yields

$$V(t,x_t) \le \Big[V(t_0,x_0) + \int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\,e^{\int_{t_0}^u \alpha(s)ds}\,du\Big]e^{-\int_{t_0}^t \alpha(s)ds}$$

for all $t \in [t_0,\infty)$. Using $W_1(||x||) = ||x||^p$, we have

$$|x| \le \Big[V(t_0,x_0) + \int_{t_0}^{t}\Big(\gamma(u)\alpha(u)+\beta(u)\Big)\,e^{\int_{t_0}^u \alpha(s)ds}\,du\Big]^{1/p} e^{-\frac{1}{p}\int_{t_0}^t \alpha(s)ds}. \tag{7.6.11}$$

The proof of *(iii)* is an immediate consequence of inequality (7.6.11). This completes the proof. □

Remark 7.8. If $f(t,0) \neq 0$ and (7.6.9) and (7.6.10) hold, then all solutions of (7.6.1) decay exponentially to zero.

Copyright Elsevier 2022

Example 7.25. For $a(t), b(t), h(t) \geq 0$, consider the scalar differential equation

$$x'(t) = -\big(a(t) + \frac{7}{6}\big)x(t) + b(t)x^{\frac{1}{3}}(t) + h(t) \quad t \geq t_0 \geq 0, \tag{7.6.12}$$

with initial condition $x(t_0) = x_0 \in \mathbb{R}$. Let $V(t,x) = x^2$. Then along any solution $x := x(t)$ of (7.6.12), we have

$$\begin{aligned} V'(t,x) &= 2xx' \\ &= -2\big(a(t) + \frac{7}{6}\big)x^2 + 2b(t)x^{4/3} + 2xh(t) \\ &\leq -2\big(a(t) + \frac{7}{6}\big)x^2 + 2b(t)x^{4/3} + x^2 + h^2(t). \end{aligned}$$

To further simplify $V'(t,x)$, we use Young's inequality, which says for any two nonnegative real numbers w and z, we have

$$wz \leq \frac{w^e}{e} + \frac{z^f}{f} \quad \text{with } \frac{1}{e} + \frac{1}{f} = 1.$$

Thus for $e = 3/2$ and $f = 3$, we obtain

$$2|b(t)|x^{4/3} \leq 2\Big[\frac{1}{3}|b(t)|^3 + \frac{(x^{4/3})^{3/2}}{3/2}\Big] = \frac{4}{3}x^2 + \frac{2}{3}|b(t)|^3.$$

As a result, we have

$$V'(t,x) \leq -2a(t)x^2 + \frac{2}{3}b^3(t) + h^2(t). \tag{7.6.13}$$

It follows that

$$V'(t,x) \leq -\alpha(t)x^2 + \beta(t), \tag{7.6.14}$$

where $\alpha(t) = 2a(t)$ and $\beta(t) = \frac{2}{3}b^3(t) + h^2(t)$. We can easily check that conditions (7.6.2)–(7.6.4) of Theorem 7.6.1 are satisfied with $W_1(|x|) = x^2$, $q = 2$, and $\gamma = 0$. Next, we choose some convenient values for our variables.

Let $a(t) = \frac{t}{2}$, $b(t) = t^{\frac{1}{3}}$, and $h(t) = t^{1/2}$. Then $\alpha(t) = t$, $\beta(t) = \frac{5}{3}t$, and (7.6.14) becomes $V'(t,x) \leq -tx^2 + \frac{5}{3}t$. We note that

$$\begin{aligned} \int_{t_0}^{t} \big(\gamma\alpha(u) + \beta(u)\big)e^{-\int_u^t \alpha(s)ds}\, du &= \frac{5}{3}\int_{t_0}^{t} ue^{-\int_u^t s ds}\, du \\ &\leq \frac{5}{3}, \end{aligned}$$

for all $t \geq t_0 \geq 0$. Thus condition (7.6.8) holds. By Corollary 7.2 all solutions of

$$x'(t) = -\big(\frac{t}{2} + \frac{7}{6}\big)x(t) + t^{\frac{1}{3}}x^{\frac{1}{3}}(t) + t^{1/2}, \; x(t_0) = x_0, \; t \geq t_0 \geq 0, \tag{7.6.15}$$

are uniformly bounded.

Copyright Elsevier 2022

On the other hand, if we take $a(t) = 1/2$, $b(t) = e^{-\kappa_1 t/3}$, and $h(t) = 0$, where $\kappa_1 > 1$, then we have $\alpha(t) = 1$, $\beta(t) = \frac{2}{3}e^{-\kappa_1 t}$, and

$$\begin{aligned}\int_{t_0}^{t} \big(\gamma\alpha(u) + \beta(u)\big)e^{\int_{t_0}^{u}\alpha(s)\,ds}\,du &= \int_{t_0}^{t} \frac{2}{3}e^{-\kappa_1 u}e^{\int_{t_0}^{u} ds}\,du \\ &\leq \frac{2}{3(\kappa_1 - 1)}e^{t_0(1-\kappa_1)}\end{aligned}$$

for all $t \geq t_0 \geq 0$. Hence condition (7.6.9) is satisfied. Finally, (7.6.11) gives the inequality

$$\begin{aligned}|x(t)| &\leq \big(x_0^2 + \frac{2}{3(\kappa_1 - 1)}e^{t_0(1-\kappa_1)}\big)e^{\kappa_1(t-t_0)} \\ &\leq \big(x_0^2 + \frac{2}{3(\kappa_1 - 1)}\big)e^{\kappa_1(t-t_0)},\end{aligned}$$

from which we arrive at the uniform exponential stability of the zero solution of

$$x'(t) = -\frac{5}{3}x(t) + e^{-\frac{\kappa_1}{3}t}x^{\frac{1}{3}}(t), \;\; x(t_0) = x_0, \; t \geq t_0 \geq 0.$$

7.7 Exercises

Exercise 7.1. Use Lyapunov function to investigate the stability of the origin for each of the following systems and compare with the linearization method:
(a) $x' = y$, $y' = -x$.
(b) $x' = -xy^2$, $y' = 3yx^2$.
(c) $x' = y - 2x$, $y' = 2x - y - x^3$.
(d) $x' = x + 3y$, $y' = 2x$.
(e) $x' = x^3 + y$, $y' = x + y^3$.
(f) $x' = y$, $y' = -x - y - x^3$.

Exercise 7.2. Find the two equilibrium solutions and use Lyapunov functions to investigate their stability of

$$x' = y - x^2, \;\; y' = x - y^2.$$

Exercise 7.3. Consider the planar system

$$\begin{aligned}x' &= -xy^2, \\ y' &= -y - yx^2.\end{aligned}$$

Examine the stability of the origin using:
(a) Theorem 6.2.1,
(b) a Lyapunov function.

Copyright Elsevier 2022

Exercise 7.4. Consider the planar system

$$\begin{aligned} x' &= ax(x+y) - xy^3, \\ y' &= ay(x+y) + x^2y^2. \end{aligned}$$

Examine the stability of the origin for different values of a using:
(a) a Lyapunov function,
(b) the linearization method.

Exercise 7.5. Put the differential equation

$$x'' + (x')^2 + x' + x(x-2) = 0$$

into a system and find its two equilibrium solutions and examine their stability using:
(a) a Lyapunov function,
(b) the linearization method.

Exercise 7.6. Use the indicated Lyapunov function V and Theorem 7.1.4 to show that the origin is asymptotically stable. Show that all solutions tend to the origin as $t \to \infty$.
(a) $x' = 2y$, $y' = -3x - 3y^3$, $V = 3x^2 + 2y^2$.
(b) $x' = -x + 6y$, $y' = -x - 2y$, $V = x^2 + 2y^2$.
(c) $x' = -2x - y$, $y' = 6x - y$, $V = 2x^2 + y^2$.
(d) $x' = y$, $y' = -x^2y - x^3$, $V = x^4 + 2y^2$.

Exercise 7.7. Study the stability of the origin for the planar system

$$\begin{aligned} x' &= x(y-1), \\ y' &= -\frac{x^2}{1+x^2}, \end{aligned}$$

by considering the Lyapunov function $V(x, y) = \mathrm{Ln}(1 + x^2) + y^2$.

Exercise 7.8. Use a Lyapunov function to show the existence of a stable limit cycle for the system

$$x' = y + x(\mu^2 - x^2 - y^2), \quad y' = -x + y(\mu^2 - x^2 - y^2)$$

for $\mu \neq 0$.

Exercise 7.9. Consider the planar system

$$x' = x + g_1(x, y), \quad y' = -y + g_2(x, y),$$

where g_1 and g_2 are continuous on a subset D of $\mathbb{R}^2$ that contains $(0, 0)$. Suppose that for a positive constant k,

$$||g_i|| \leq k(x^2 + y^2), i = 1, 2,$$

with $g_1(0, 0) = g_2(0, 0) = 0$. Use Theorem 7.1.6 to show that the origin is unstable.

Copyright Elsevier 2022

Exercise 7.10. Consider the second-order differential equation

$$x'' + g(x) = 0,$$

where g is continuously differentiable, $g(0) = 0$, and $xg(x) > 0$. Write the equation as a system and use a Lyapunov function to show that its zero solution is stable.

Exercise 7.11. Consider the second-order differential equation

$$x'' + f(x)x' + g(x) = 0,$$

where $f(x)$ and $g(x)$ are even and odd polynomials, respectively.
(a) Show that the equation is equivalent to the system

$$x' = y - F(x), \quad y' = g(x),$$

where $F(x) = \int_0^x f(s)ds$.
(b) Let $G(x) = \displaystyle\int_0^x g(s)ds$ and suppose there are positive constants α and β such that $g(x)F(x) > 0$ for $0 \le |x| < \alpha$ and $G(x) < \beta$ for $|x| < \alpha$.

Show that

$$V(x, y) = \frac{y^2}{2} + G(x)$$

is a Lyapunov function in the region $A = \{\alpha, \beta \in \mathbb{R} : |x| < \alpha, y^2 < \beta\}$ and that the origin is asymptotically stable.

Exercise 7.12. Consider the predator–prey model with carrying capacity

$$x' = x[a - dx - by], \quad y' = y[-c + kx],$$

where $x(t)$ and $y(t)$ are the numbers of preys and predators at time t, respectively, with positive constants a, b, c, d, and k.
(a) Find the equilibrium point (x^*, y^*) located in the first quadrant.
(b) Use the transformation

$$u = \mathrm{Ln}(\frac{x}{x^*}) \text{ and } v = \mathrm{Ln}(\frac{y}{y^*})$$

to obtain a system in (u, v).
(c) Show that

$$V(u, v) = kx^*(e^u - u) + by^*(e^v - v) - kx^* - by^*$$

is a Lyapunov function and that the equilibrium solution (x^*, y^*) is stable.

Exercise 7.13. Prove Theorem 7.1.7.

Copyright Elsevier 2022

Exercise 7.14. Consider the second-order differential equation

$$x'' + (\mu + \nu x^2)x' + x + x^3 = 0$$

(a) If $\nu = 0$ and $\mu > 0$, then show that the zero solution is asymptotically stable.
(b) If $\mu = 0$ and $\nu > 0$, then show that the zero solution is stable using an appropriate Lyapunov function.

Exercise 7.15. Consider the planar system

$$\begin{aligned} x' &= (x - by)(x^2 + y^2 - 1), \\ y' &= (ax + y)(x^2 + y^2 - 1). \end{aligned}$$

Show that the origin is stable if $a, b > 0$.

Exercise 7.16. Show that the origin is unstable for the planar system

$$\begin{aligned} x' &= x + y, \\ y' &= x - y + xy. \end{aligned}$$

Exercise 7.17. Suppose f is continuously differentiable and consider the planar system

$$\begin{aligned} x' &= y - xf(x, y), \\ y' &= -x - yf(x, y). \end{aligned}$$

Show that the origin is asymptotically stable if $f(x, y) > 0$ and unstable if $f(x, y) < 0$.

Exercise 7.18. Let $V = x^2 + y^2$ and consider the system

$$\begin{aligned} x' &= (x - y)(x^2 + y^2 - 1), \\ y' &= (x + y)(x^2 + y^2 - 1). \end{aligned}$$

Show that the origin is asymptotically stable and find the region of attraction.

Exercise 7.19. Let $V = x^2 + y^2$ and consider the system

$$\begin{aligned} x' &= -x + 2x^2 + y^2, \\ y' &= -y + y^2. \end{aligned}$$

Show that the origin is asymptotically stable and find the region of attraction.

Exercise 7.20. Let $V = x^2 + y^4$ and consider the system

$$\begin{aligned} x' &= -x + 2y^3 - 2y^4, \\ y' &= -x - y + xy. \end{aligned}$$

Show that the origin is globally asymptotically stable.

Copyright Elsevier 2022

Exercise 7.21. Let $V = ax^{2m} + by^{2m}$, $a, b > 0$ and consider the system

$$\begin{aligned} x' &= -3x^3 - y, \\ y' &= -x^5 - 2y^3. \end{aligned}$$

Determine the values of a, b, m, and n, where m and n are positive integers, such that the origin is globally asymptotically stable.

Exercise 7.22. Decide if the functions given below are radially unbounded or not.
(a)

$$V(x_1, x_2) = (x_1 - x_2)^2.$$

(b)

$$V(x_1, x_2) = \frac{x_1^2 + x_2^2}{1 + x_1^2 + x_2^2} + (x_1 - x_2)^2.$$

Exercise 7.23. Consider the equation

$$y'' + By' + y + y^3 = 0, \quad B \geq 0,$$

which is a particular case of mass-spring nonlinear damper.
(a) Use the transformation $v = y'$ and obtain a corresponding system in y and v.
(b) Show that

$$V(y, v) = \frac{y^2 + y^4}{2} + \frac{v^2}{2}$$

is a Lyapunov function.
(c) Decide whether the origin is stable or asymptotically stable.

Exercise 7.24. Consider the non-autonomous system

$$\begin{aligned} x' &= -x - g(t)y, \\ y' &= x - y, \end{aligned}$$

where $g(t)$ is continuously differentiable and, for positive constants k_1, k_2, satisfies

$$k_1 \leq g(t) \leq k_2 \quad \text{and} \quad g'(t) \leq g(t)$$

for all $t \geq 0$. Use a Lyapunov function to show that the zero solution is uniformly stable.

Exercise 7.25. Consider the scalar non-autonomous differential equation

$$x' = -\Big(1 + g(t)\Big)x^3,$$

Copyright Elsevier 2022

where $g(t)$ is continuous, and

$$g(t) \geq 0 \quad \text{for all} \quad t \geq 0.$$

Use a Lyapunov function to show that the zero solution is uniformly asymptotically stable.

Exercise 7.26. Let $D = \{x \in \mathbb{R}^n : |x| \leq \delta\}$ for $\delta > 0$, and let $V(t, x) : [0, \infty) \times D \to [0, \infty)$ be a continuously differentiable function satisfying

$$V(t, 0) = 0,$$
$$W_1(|x|) \leq V(t, x) \leq W_2(|x|),$$

and

$$V'(t, x) \leq -W_3(V(t, x)) \ \text{ for all } \ t \geq 0, \ |x| \leq \delta.$$

Show that the zero solution of (7.1.17) is uniformly asymptotically stable.

Exercise 7.27. Put the scalar equation

$$x'' - x^3$$

into a system and find a function that satisfies the requirements of Theorem 7.3.1.

Exercise 7.28. Write the scalar equation

$$x'' + f(x') + g(x) = 0$$

as a system. Assume that

$$xf(x') > 0, \ \ x' \neq 0, \ \text{ and } \ xg(x) > 0, \ \ x \neq 0.$$

Use Theorem 7.3.1 to show that the origin is asymptotically stable.

Exercise 7.29. Use Exercise 7.28 to show that the zero solutions of given differential equations are asymptotically stable.
(a) $x'' + x'|x'| + x - \frac{x^3}{3} = 0$.
(b) $x'' + (x')^3 + x^5 - x^4 \sin^2(x) = 0$.

Exercise 7.30. Consider the system

$$x_1' = -x_1 + g(x_2),$$
$$x_2' = -x_2 + h(x_1),$$

where g and h are continuous and satisfy

$$|g(u)| \leq |u|/2, \ \ |h(u)| \leq |u|/2.$$

Use the Lyapunov function $V(x_1, x_2) = \frac{1}{2}(x_1^2 + x_2^2)$ to show that the origin is uniformly exponentially stable.

Copyright Elsevier 2022

Exercise 7.31. Let $C = I$ and $A = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix}$. Find a symmetric matrix B such that $A^T B + BA = -I$.

Exercise 7.32. Let B be chosen from Exercise 7.31 and set $V = x^T Bx$. Show that V is an autonomous Lyapunov function in some open set D containing $x = 0$ for the system

$$\begin{aligned} x_1' &= x_2 + f_1(t, x), \\ x_2' &= -2x_1 - 3x_2 + f(t, x), \end{aligned}$$

where

$$\lim_{|x| \to 0} \frac{|f_i(t, x)|}{|x|} = 0, \; i = 1, 2,$$

uniformly for $t \in [0, \infty) \times B_1(0)$. Then find a wedge W such that $V'(x) \leq -W(|x|)$ on D.

Exercise 7.33. Consider the pendulum of Example 7.21 and define the function

$$V(x_1, x_2) = \frac{g}{L}\big(1 - \cos(x_1)\big) + \frac{1}{2}x^T Bx$$

with positive definite matrix B that is to be found.
Hint: See Remark 7.7 and then compute V' along the solutions. Pick the entries of B to ensure that V is a strict Lyapunov function on the set $D - \{(0, 0)\}$.

Exercise 7.34. Use Theorem 7.4.2 to show that the zero solution is asymptotically stable for

$$x' = -x^3 + 2y^3, \quad y' = -2xy^2.$$

Hint: See Example 7.22.

Exercise 7.35. (Harmonic oscillator) Let

$$x'' + \sin(x) = 0.$$

(a) Can you prove the stability of the origin using linearization? Use an appropriate Lyapunov function to show that the origin is a stable equilibrium solution.
(b) Add the damping term

$$x'' + \varepsilon x' + \sin(x) = 0.$$

Study the stability of the origin for $\varepsilon > 0$.

Exercise 7.36. Let $x(t) = (x_1(t), x_2(t))^T \in \mathbb{R}^2$ and consider

$$\begin{aligned} x_1'(t) &= x_2(t) - x_1(t)|x_1(t)|, \\ x_2'(t) &= -x_1(t) - x_2(t)|x_2(t)| \end{aligned}$$

Copyright Elsevier 2022

with initial condition $x(t_0) = x_0 \in \mathbb{R}^2$. Set $D = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \le \frac{8}{9}\}$. Let $V : [0, \infty) \times D \to [0, \infty)$ be given by $V(t, x) = \frac{1}{2}(x_1^2 + x_2^2)$. Show that:

(a)

$$V'(t, x) \le -2V^{3/2}(t, x);$$

(b)

$$V(t, x) - V^q(t, x) = V(t, x) - V^{3/2}(t, x) \le \frac{4}{27};$$

for $(x_1, x_2) \in D$.

(c) Refer to a specific theorem in the text to conclude that the zero solution is uniformly exponentially stable for any $x_0 \in D$.

(d) Explain why linearization around the origin cannot be used here.

Hint: Use the inequality

$$\Big(\frac{a+b}{2}\Big)^e \le \frac{a^e}{2} + \frac{b^e}{2}, a, b > 0, e > 1.$$

Exercise 7.37. Let $x(t) = (x_1(t), x_2(t))^T \in \mathbb{R}^2$ and consider

$$x_1'(t) = x_2 - g(t)x_1(x_1^2 + x_2^2),$$

$$x_2'(t) = -x_1 - g(t)x_2(x_1^2 + x_2^2)$$

with initial condition $x(t_0) = x_0 \in \mathbb{R}^2$, where $g(t)$ is continuously differentiable and bounded, and $g(t) \ge k > 0$ for all $t \ge 0$. Set $D = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \le 1\}$. Let $V : [0, \infty) \times D \to [0, \infty)$ be given by $V(t, x) = \frac{1}{2}(x_1^2 + x_2^2)$. Show that the hypotheses of Theorem 7.6.1 are satisfied for $(x_1, x_2) \in D$ and conclude that the zero solution is uniformly exponentially stable for any $x_0 \in D$.

Exercise 7.38. Consider

$$\begin{aligned} x_1'(t) &= x_2 + x_1x_3, \\ x_2'(t) &= -x_1 - x_2 + x_2x_3, \\ x_3'(t) &= -x_1^2 - x_2^2. \end{aligned}$$

Use Theorem 7.4.2 to show that

$$x_1(t),\ x_2(t) \to 0 \text{ as } t \to \infty, \text{ and } x_3(t) \to c \text{ as } t \to \infty,$$

for some constant c.

Exercise 7.39. Repeat Example 7.24 for the system

$$x' = \begin{pmatrix} -1 & 0 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ t > 0.$$

Copyright Elsevier 2022

Exercise 7.40. Consider the SIR epidemic model of Section 6.3 and let

$$T = \{(S, I) : S \geq 0, I \geq 0, \ S + I < 1\}.$$

Define the function V on T by

$$V(S, I) = S - S^* - S^* \mathrm{Ln}(\frac{S}{S^*}) + I - I^* - I^* \mathrm{Ln}(\frac{I}{I^*}).$$

(a) Show that along the solutions we have

$$V'(S, I) \leq -\mu \frac{(S - S^*)^2}{SS^*}.$$

(b) Show that V is positive definite on T.
Hint: Clearly, $V(S^*, I^*) = 0$. To show that $V(S, I) > 0$ for $(S, I) \neq (S^*, I^*)$, consider the function

$$f(x) = x - x^* - x^* \mathrm{Ln}(\frac{x}{x^*}), \ 0 < x < 1, \ 0 < x^* < 1$$

and show it has a minimum of zero at x^* and $f(x) > 0$ for $x \neq x^*$.

(c) Use Theorem 7.4.2 to conclude that (S^*, I^*) is asymptotically stable.

Exercise 7.41. Consider the Lorenz system

$$\begin{aligned} x'(t) &= \rho(y - x), \\ y'(t) &= rx - y - xz, \\ z'(t) &= xy - bz, \end{aligned}$$

where ρ, r, and b are positive constants.
1) Show that, for $0 < r < 1$, the origin $(0, 0, 0)$ is the only equilibrium solution.
2) Let

$$V(x, y, z) = \frac{1}{2\rho}\left(x^2 + \rho y^2 + \rho z^2\right).$$

Show that the origin $(0, 0, 0)$ is asymptotically stable for $0 < r < 1$.

Copyright Elsevier 2022

CHAPTER

Delay differential equations

8

This chapter is introductory to the study of *delay differential equations*. We will touch on the existence and uniqueness of solutions and stability using the Lyapunov method. Toward the end of the chapter, we use the fixed point theory on Banach spaces developed in Chapter 2 to deduce boundedness and stability results for neutral delay differential equations. We end the chapter by utilizing Lyapunov functionals and obtain the exponential stability for totally nonlinear finite delay integro-differential equations.

8.1 Introduction

Delay differential equations are a type of equation in which the derivative of the unknown function at a certain time is given in terms of the values of the function at previous times. In the previous seven chapters, we considered finite-dimensional differential systems, where the state at any time can be obtained from a finite set of values. On the other hand, delay differential equations are *infinite-dimensional systems*. As a consequence, delay differential equations present us with some difficulties when we analyze them. Nevertheless, some known theories for differential equations can be extended to cover delay differential equations.

For the moment, we consider the scalar delay differential equation

$$x' = f(t, x(t), x(t-\tau)), \tag{8.1.1}$$

where τ is a positive constant known as the delay. If we set $t = 0$ in (8.1.1), then we must know the value of the dependent variable x at s, where $-\tau \le s \le 0$, that is, we need to define an initial function $\varphi : [-\tau, 0] \to \mathbb{R}$. In other words, the initial function φ gives the behavior of the system prior to time 0 (assuming that we start at time $t = 0$). In general, for system (8.1.1) to be *well posed*, we must specify an initial function and work with the delay system

$$x' = f(t, x(t), x(t-\tau)), \quad 0 \le t_0 \le t \le t_f, \tag{8.1.2}$$

$$x(s) = \varphi(s), \quad -\tau \le s \le t_0 \le 0,$$

where $f : [t_0, t_f] \times \mathbb{R}^2 \to \mathbb{R}$. Note that system (8.1.2) is referred to as a delay differential equation with finite delay.

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00014-0
Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

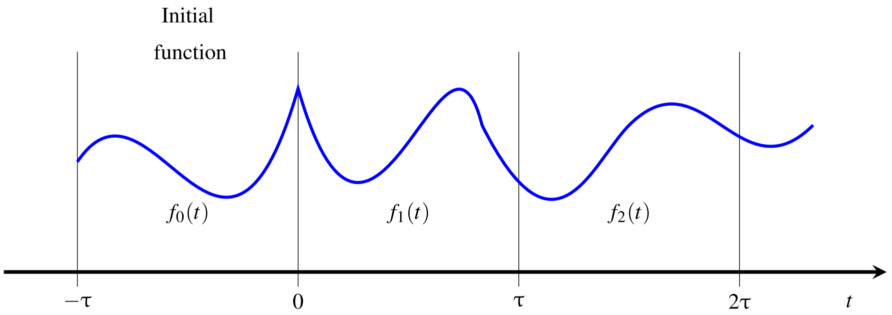


FIGURE 8.1

The effect of the initial function and the delay and the solution space.

Definition 8.1.1. We say that $x : [t_0 - \tau, t_f] \to \mathbb{R}$ is a solution of (8.1.2) if it is continuous and satisfies (8.1.2) for $t_0 \le t \le t_f$ and $x(s) = \varphi(s)$, $-\tau \le s \le t_0$. If every two solutions of (8.1.2) agree with each other as far as both are defined, then the solution is said to be *unique*.

We may think of the solution of (8.1.2) as a mapping from functions on the interval $[t - \tau, t]$ into functions on the interval $[t, t + \tau]$. In other words, the solutions can be thought of as a sequence of functions $f_0(t), f_1(t), f_2(t), \ldots$ defined over intervals of length τ. The points $t = j\tau$, $j = 0, 1, 2, \ldots$, where the solution segments meet are called knots. Fig. 8.1 illustrates this mapping.

It is to be understood that $x'(t_0)$ is the *right-hand derivative* of x at t_0. In most cases the solution is not differentiable at $t = t_0$ as seen in Fig. 8.1. In the case of differential equations, we only have to give an initial condition at a single point, known as the initial time.

For the delay system (8.1.2), we had to specify an initial function on the whole interval $[-\tau, t_0]$, $t_0 \le 0$. The system inherits being infinite-dimensional from the fact that the initial function takes infinitely many values at infinitely many points in the interval $[-\tau, t_0]$.

There are different types of delay differential equations:

$$x' = f(t, x(t), x(t - \tau_1), \ldots, x(t - \tau_m)), \quad 0 \le t_0 \le t \le t_f, \tag{8.1.3}$$

$$x(s) = \varphi(s), \quad -\tau \le s \le t_0 \le 0,$$

where $\tau = \max\{\tau_j,\ 1 \le j \le m\}$. We can try to solve (8.1.3) for $0 \le t_0 \le t \le t_0 + r$, where $r = \min\{\tau_j,\ 1 \le j \le m\}$. System (8.1.3) is known as a differential equation with *multiple delays*. For (8.1.3), the initial function $\varphi : [-\tau, t_0] \to \mathbb{R}^m$, $t_0 \le 0$. On the other hand, if $a(t)$ is a continuous function with $a : [0, \infty) \to (-\infty, \infty)$, then the system

Copyright Elsevier 2022

$$x' = a(t)x(t) + \int_{-\infty}^{t} f(s, x(s))ds, \quad 0 \le t_0 \le t \le t_f, \tag{8.1.4}$$

$$x(s) = \varphi(s), \ \ -\infty < s \le t_0 \le 0$$

is referred to as a *Volterra integro-differential equation with infinite delay*. System (8.1.4) is called a *Volterra integro-differential equation with finite delay* if the lower limit of integration is $t - \tau$, where $\tau > 0$, instead of $-\infty$. In this case the initial function is the same as in (8.1.2). Others may refer to it as a *Volterra integro-differential equation with distributed delay*. Finally, if the delay τ of system (8.1.2) depends on time, then the system is referred to as a *delay differential equation with time-varying delay or functional delay*.

8.2 Method of steps

The *Method of steps* can be used as a tool for simple delay differential equations to obtain a solution, given some initial data. The method of steps is best illustrated by looking at an example.

Example 8.1. (Method of steps) Consider the delay problem

$$x' = -x(t-1), \ t > 0, \tag{8.2.1}$$

$$x(s) = \varphi(s) = 10, \ \ -1 \le s \le 0.$$

We will try to piece together the solution on [0, 3].

Step 1. In this step, we find the solution on $0 \le t \le 1$. For $0 \le t \le 1$, we have $-1 \le t - 1 \le 0$, and hence $x(t-1) = 10$. Thus from (8.2.1) we obtain

$$\frac{dx}{dt} = -10,$$

which has the solution

$$x(t) = x(0) + \int_0^t (-10)du = 10 - 10t.$$

Step 2. Now we find the solution on $1 \le t \le 2$. For $1 \le t \le 2$, we have $0 \le t - 1 \le 1$, and hence $x(t-1) = 10 - 10(t-1)$, where we have made use of the solution from Step 1. Thus from (8.2.1) we obtain

$$\frac{dx}{dt} = -10 + 10(t-1),$$

Copyright Elsevier 2022

which has the solution

$$\begin{aligned} x(t) &= x(1) + \int_1^t [-10 + 10(u-1)]du \\ &= 0 - 10u + 5(u-1)^2\Big|_1^t \\ &= -10(t-1) + 5(t-1)^2, \end{aligned}$$

where we have used $x(1) = 0$ from the solution in Step 1.

Step 3. Now we find the solution on $2 \le t \le 3$. For $2 \le t \le 3$, we have $1 \le t-1 \le 2$, and hence $x(t-1) = -10(t-2) + 5(t-2)^2$, where we utilized the solution from Step 2. Thus from (8.2.1) we obtain

$$\frac{dx}{dt} = 10(t-2) - 5(t-2)^2,$$

which has the solution

$$\begin{aligned} x(t) &= x(2) + \int_1^t [10(u-2) - 5(u-2)^2]du \\ &= -5 + 5(t-2)^2 - \frac{5}{3}(t-2)^3, \end{aligned}$$

where we have used $x(2) = -5$ from the solution in Step 2.

In Example 8.1

$$\lim_{t\to 0^-} x'(t) = \lim_{t\to 0^-} \varphi'(t) = 10,$$

and

$$\lim_{t\to 0^+} x'(t) = \lim_{t\to 0^+} (-10t + 10)' = -10.$$

Therefore the solution is not differentiable at the initial time $t_0 = 0$.

Now we consider the basic delay differential equation

$$x' = f(t,\ x(t-\tau)), \quad t \ge 0, \quad x(s) = \varphi(s), \quad -\tau \le s \le 0. \tag{8.2.2}$$

If f and φ are continuous on $\mathbb{R}^2$ and $\mathbb{R}$, respectively, then a unique solution defined by

$$x(t) = \varphi(0) + \int_0^t f(s, \varphi(s-\tau))ds \tag{8.2.3}$$

exists for $0 \le t \le \tau$. We may argue using the method of steps to show that a unique solution exists for all $t \ge 0$, which is valid regardless of how small the delay τ is.

Copyright Elsevier 2022

Note that if $\tau = 0$, then our delay problem becomes an ordinary differential equation. Consider the delay problem

$$x' = x^2(t-\tau),\ t > 0,\ \ x(s) = 1,\ \ -\tau \le s \le 0. \tag{8.2.4}$$

Using the method of steps, we may find a unique solution. However, if we set $\tau = 0$ in (8.2.4) with $x(0) = 1$, then the solution of $x' = x^2(t)$, $t > 0$, exists only for $t < 1$. This example should cement that the existence of solutions in finite delay differential equations is much easier to obtain, unlike ordinary differential equations.

Example 8.2. Consider the predator–prey model with carrying capacity

$$x' = ax - dx^2 - bxy, \quad y' = -cy + kxy,$$

where $x(t)$ and $y(t)$ are the numbers of preys and predators at time t, respectively, with positive constants a, b, c, d, and k. The term kxy represents the predator's utilization of the prey it captures. Hence there is certainly a delay, say $r > 0$. As a consequence, we have the predator–prey delay system

$$x' = ax - dx^2 - bxy, \quad y' = -cy + kx(t-r)y(t-r).$$

Suppose we are given an initial continuous function $(\phi_1(t), \phi_2(t))$ with $x(s) = \phi_1(s)$, $y(s) = \phi_2(s)$, $-r \le s \le 0$. Then by integrating the linear equation

$$y' = -cy + k\phi_1(t-r)\phi_2(t-r),\ 0 \le t \le r,$$

we arrive at the explicit solution of y given by $y(t) := \zeta(t)$. Substituting y back into the equation with x' gives the homogeneous Bernoulli differential equation

$$x' = ax - dx^2 - bx\zeta(t),\ \ x(0) = \phi_1(0), \tag{8.2.5}$$

which can be easily solved using techniques from basic differential equations. To see this, we let $c(t) = a - b\zeta(t)$ and $v = x^{-1}$. Then we have the resulting differential equation in v,

$$v'(t) + c(t)v = b.$$

Multiplying by the integrating factor, integrating, and simplifying, we arrive at the solution

$$x(t) = \frac{\phi_1(0)}{e^{-\int_0^t c(s)ds} + b\phi_1(0)\int_0^t e^{-\int_u^t c(s)ds}du},$$

where we have used $v = x^{-1}$.

Copyright Elsevier 2022

8.3 Existence and uniqueness

We examine the existence and uniqueness of systems with constant delays since they are involved in many applications. For simplicity of notation, we assume that $t_0 = 0$. This should not cause any difficulties or take away from the meaning of the problem.

Let x be an $n \times 1$ vector-valued function, let D be an open set in $\mathbb{R}^n$, and let f be defined on $[0, \infty) \times D^m$. Define an initial function $\varphi : [-\tau, 0] \to \mathbb{R}^m$, where $\tau = \max\{\tau_j,\ 1 \le j \le m\}$. Let $r = \min\{\tau_j,\ 1 \le j \le m\} > 0$ and consider the system with multiple constant delays

$$x' = f(t, x(t-\tau_1), \dots, x(t-\tau_m)), \quad 0 \le t \le t_f, \tag{8.3.1}$$

$$x(s) = \varphi(s), \quad -\tau \le s \le 0. \tag{8.3.2}$$

We may rename the delays so that they are listed in an increasing order so that $\tau = \tau_m$ and $r = \tau_1$:

$$0 \le \tau_1 < \tau_2 < \cdots < \tau_m.$$

Definition 8.3.1. We say that $x : [t_0 - \tau, t_\beta] \to D$ is a solution of (8.3.1)–(8.3.2), where $0 < t_\beta \le t_f$, if it is continuous and satisfies (8.3.1) for $0 \le t < t_\beta$, and $x(s) = \varphi(s)$, $-\tau \le s \le 0$. The solution of (8.3.1)–(8.3.2) is said to be unique if every two solutions agree with each other as far as both are defined.

Theorem 8.3.1. *Let f be continuous on $[0, t_f) \times D^m$, and let $\varphi : [-\tau, 0] \to D$ be continuous on $[-\tau, 0]$. Then system (8.3.1)–(8.3.2) has a unique solution x on $[-\tau, \beta_1)$, where $0 < \beta_1 \le t_f$. Moreover, if $\beta_1 < t_f$, then $x(t)$ approaches the boundary of D as $t \to \beta_1$.*

Proof. Let $t \in [0, r]$. Then every solution resides over the domain of the initial function, that is,

$$x' = f(t, \varphi(t-\tau_1), \dots, \varphi(t-\tau_m))$$

with $x(0) = \varphi(0)$. An integration from 0 to t gives

$$x(t) = \varphi(0) + \int_0^t f(u, \varphi(u-\tau_1), \dots, \varphi(u-\tau_m))du,$$

which uniquely defines $x(t)$ for $0 \le t \le r$, provided that $r < t_f$ and $x(t)$ remains in D. Now that we know $x(t)$ on $[-\tau, r]$, we may continue in this fashion on $[r, 2r]$ as long as $2r < t_f$ and $x(t)$ remains in D. We continue with this method of steps until we reach t_f or until $x(t)$ reaches the boundary of D. This completes the proof. ☐

Remark 8.1. Suppose at least one of the $\tau_j = 0$, $j = 1, 2, \dots, m$. To be specific, suppose only $\tau_1 = 0$ and the rest of the delays are distinct and greater than zero. Then

Copyright Elsevier 2022

system (8.3.1) will include a differential equation. In this case proving the existence and uniqueness requires a Lipschitz condition on f of the form

$$|f(t, y_{(1)}, y_{(2)}, \dots, y_{(m)}) - f(t, z_{(1)}, z_{(2)}, \dots, z_{(m)})| \le K|y_{(1)} - z_{(1)}| \qquad (8.3.3)$$

for $y_{(j)}, z_{(j)} \in \mathbb{R}^n$ and positive constant K.

Corollary 8.1. *(Uniqueness) In addition to the hypotheses of Theorem 8.3.1, assume* (8.3.3). *Then* (8.3.1)–(8.3.2) *has at most one solution on any interval* $[-\tau, \beta_1)$*, where* $0 < \beta_1 \le t_f$.

Proof. The case $r > 0$ is already covered by Theorem 8.3.1, so we let $r = 0$. Then for $t \in [0, \tau_2]$,

$$x' = f(t, x(t), \varphi(t - \tau_2), \dots, \varphi(t - \tau_m))$$

with $x(0) = \varphi(0)$. Then by Theorem 2.2.1, it has at most one solution. We may repeat the above process for $t \in [\tau_2, 2\tau_2]$, for $\tau_2 < t_f$. We keep repeating the process until we reach t_f. This completes the proof. ☐

Recall our remark in Chapter 3 that a multivariable function f satisfies a Lipschitz condition on some region if all its first partial derivatives are continuous on that region. We have the following example.

Example 8.3. Consider the sun flower model

$$u''(t) + bu'(t) + c \sin u(t - r) = 0, \; t \ge 0,$$

where b, c, and r are positive constants. Using the transformation $x_1 = u$, $x_2 = u'$, we arrive at the system

$$x_1'(t) = x_2,$$

$$x_2'(t) = -c \; \sin x_1(t - r) - bx_2.$$

Therefore

$$f(t, y_{(1)}, y_{(2)}) = \begin{pmatrix} y_{(1)2} \\ -c \sin y_{(2)1} - by_{(1)2} \end{pmatrix}.$$

Clearly, f is continuous, and so are the partial derivatives

$$\frac{\partial f_1}{\partial y_{(1)1}} = 0, \qquad \frac{\partial f_1}{\partial y_{(1)2}} = 1, \qquad \frac{\partial f_2}{\partial y_{(1)2}} = -b.$$

Thus by Corollary 8.1, given any continuous function $\varphi : [-r, 0] \to \mathbb{R}^2$, the delay system has at most one solution on $[-r, \beta_1)$ for any $\beta_1 > 0$.

Copyright Elsevier 2022

8.4 Stability using Lyapunov functions

In this section, we consider general forms of delay differential equations that fall into the category of *functional delay differential equations*. A functional delay differential equation is a delay differential equation where the unknowns are functions. We will adopt the following notation. For any $r \geq 0$, any continuous function $x(u)$ defined on $-r \leq u \leq A$, $A > 0$, and any fixed t, $0 \leq t \leq A$, by x_t we denote the function $x_t(s) = x(t+s)$, $-r \leq s \leq 0$, that is, $x : [a,\ b] \to \mathbb{R}^n$ is continuous and is the segment of the function $x(s)$ defined by letting s range in the interval $t - r \leq s \leq t$. We consider the functional delay differential equation

$$x'(t) = F(t, x_t),\ t \geq t_0. \tag{8.4.1}$$

We assume that F is continuous in both arguments and that $F : \mathbb{R} \times C \to \mathbb{R}^n$, where C is the set of continuous functions $\phi : [-\alpha, 0] \to \mathbb{R}^n$, $\alpha > 0$. If $t_0 = -\infty$, then $t \geq t_0$ means that $t > t_0$. Let

$$C(t) = \{\phi : [t-\alpha, t] \to \mathbb{R}^n\}.$$

It is to be understood that $C(t)$ is C for $t = 0$. Also, ϕ_t denotes $\phi \in C(t)$, and $||\phi_t|| = \max\limits_{t-\alpha \leq s \leq t} |\phi(t)|$, where $|\cdot|$ is a convenient norm on $\mathbb{R}^n$. A solution is denoted by $x(t_0, \phi_{t_0})$, and its value by $x(t, t_0, \phi_{t_0})$ or, if there is no confusion, by $x(t_0, \phi)$ and $x(t, t_0, \phi)$.

Definition 8.4.1. Let $x(t) = 0$ be a solution of (8.4.1).

(a) The zero solution of (8.4.1) is stable if for all $\varepsilon > 0$ and $t_1 \geq t_0$, there exists $\delta > 0$ such that $[\phi \in C(t_1),\ \|\phi\| < \delta, t \geq t_1]$ imply that $|x(t, t_1, \phi)| < \varepsilon$.

(b) The zero solution of (8.4.1) is uniformly stable if it is stable and if δ is independent of $t_1 \geq t_0$.

(c) The zero solution of (8.4.1) is asymptotically stable if it is stable and if for each $t_1 \geq t_0$, there is $\eta > 0$ such that $[\phi \in C(t_1),\ \|\phi\| < \eta]$ imply that $|x(t, t_1, \phi)| \to 0$ as $t \to \infty$. Note that if this is true for every $\eta > 0$, then $x = 0$ is asymptotically stable in the large or globally asymptotically stable.

(d) The zero solution of (8.4.1) is uniformly asymptotically stable if it is uniformly stable and if there is $\eta > 0$ such that for each $\gamma > 0$, there exists $S > 0$ such that $[t_1 \geq t_0, \phi \in C(t_1),\ \|\phi\| < \eta, t \geq t_1 + S]$ imply that $|x(t, t_1, \phi)| < \gamma$. We also note that if this is true for every $\eta > 0$, then $x = 0$ is uniformly asymptotically stable in the large.

Let $W_i : [0, \infty) \to [0, \infty)$ be continuous with $W_i(0) = 0$, $W_i(r)$ strictly increasing, and $W_i(r) \to \infty$ as $r \to \infty$, $i = 1, 2, 3$.

Consider the totally nonlinear differential equation with bounded delay

$$x'(t) = -ax^3(t) + bx^3(t-r),\ t \geq t_0 \geq 0,$$

where a, b, and r are positive constants. Consider the Lyapunov-type functional

$$V(x_t) = |x(t)| + b\int_{t-r}^{t} |x^3(s)|ds.$$

Let $a \geq b$. Then along the solutions we have

$$\begin{aligned} V'(x_t) &= \frac{x}{|x|}x' - b|x^3(t-r)| + b|x^3(t)| \\ &= \frac{x}{|x|}\big(-ax^3(t) + bx^3(t-r)\big) - b|x^3(t-r)| + b|x^3(t)| \\ &\leq -a|x^3(t)| + b|x^3(t-r)| - b|x^3(t-r)| + b|x^3(t)| \\ &= [-a+b]|x^3| \leq 0. \end{aligned} \tag{8.4.2}$$

Let $t_1 \in \mathbb{R}$, $\varphi \in C(t_1)$, and $x(t) = x(t, t_1, \varphi)$. Using the fact that $|x(t)| \leq V(x_t)$ and integrating (8.4.2) from $t_1 - r$ to t_1 yield

$$\begin{aligned} |x(t)| \leq V(x_t) &\leq V(\varphi_{t_1}) \\ &= |\varphi(t_1)|^3 + b\int_{t_1-r}^{t_1} |\varphi^3(s)|ds \\ &\leq \big(1+br\big)||\varphi_{t_1}||^3. \end{aligned} \tag{8.4.3}$$

Suppose $a > b$, let $\varepsilon > 0$, and choose $\delta = \Big(\frac{\varepsilon}{1+br}\Big)^{1/3}$. Then for $||\varphi_{t_1}|| < \delta$, from (8.4.3) we have

$$|x(t)| \leq \big(1+br\big)\Big[\Big(\frac{\varepsilon}{1+br}\Big)^{1/3}\Big]^3 = \varepsilon,$$

and the zero solution is uniformly stable. Note that stability is independent of the size of the delay r.

Integrate (8.4.2) for 0 to t and obtain $V(x_t) - V(\varphi_0) \leq -\int_0^t |x(s)|^3 ds$, which implies that

$$\int_0^t |x(s)|^3 ds \leq V(\varphi_0) - V(x_t) \leq V(\varphi_0) < \infty$$

for all $t \geq 0$. Hence $x^3(t) \in L^1[0, \infty)$.

Next, we prove parallel theorems to Theorems 7.1.8–7.1.11 concerning system (8.4.1).

Theorem 8.4.1. *Let $D > 0$. Suppose $V(t, \psi_t)$ is a scalar functional continuous in ψ and locally Lipschitz in ψ_t when $t \geq t_0$ and $\psi_t \in C(t)$ with $||\psi_t|| < D$. Suppose also that $V(t, 0) = 0$ and*

$$W_1(|\psi(t)|) \leq V(t, \psi_t). \tag{8.4.4}$$

Copyright Elsevier 2022

(a) *If*

$$V'(t,\psi_t) \le 0 \text{ for } t_0 \le t < \infty \text{ and } ||\psi_t|| \le D, \tag{8.4.5}$$

then the zero solution of (8.4.1) *is stable.*

(b) *If in addition to (a),*

$$V(t,\psi_t) \le W_2(||\psi_t||), \tag{8.4.6}$$

then the zero solution of (8.4.1) *is uniformly stable.*

(c) *If there is* $M > 0$ *such that* $|F(t,\psi_t)| \le M$ *for* $t_0 \le t < \infty$ *and* $||\psi_t|| \le D$, *and if*

$$V'(t,\psi_t) \le -W_2(|\psi(t)|), \tag{8.4.7}$$

then the zero solution of (8.4.1) *is asymptotically stable.*

Proof. Let $\varepsilon > 0$ be such that $\varepsilon < D$. Let $t_1 \ge t_0$. Since V is continuous and $V(t,0) = 0$, there exists $\delta > 0$ such that $\phi \in C(t_1)$ with $||\phi_{t_1}|| < \delta$ implies that $V(t_1,\phi_{t_1}) < W_1(\varepsilon)$. By condition (8.4.5) we have

$$\begin{aligned} W_1(|x(t,t_1,\phi_{t_1})|) &\le V(t,x(t,t_1,\phi_{t_1})) \le V(t_1,\phi_{t_1}) \\ &\le W_1(\varepsilon), \end{aligned}$$

from which it follows that

$$|x(t,t_1,\phi_{t_1})| \le W_1^{-1}\big(W_1(\varepsilon)\big) = \varepsilon.$$

This concludes the proof of (a).

As for the proof of (b), we let $\varepsilon > 0$ be such that $\varepsilon < D$. We find $\delta > 0$ with $W_2(\delta) < W_1(\varepsilon)$. Let $t_1 \ge t_0$ and $\phi_{t_1} \in C(t_1)$ with $||\phi_{t_1}|| < \delta$. Then

$$\begin{aligned} W_1(|x(t,t_1,\phi_{t_1})|) &\le V(t,x_t(t_1,\phi_{t_1})) \le V(t_1,\phi_{t_1}) \\ &\le W_2(\delta) < W_1(\varepsilon), \end{aligned}$$

from which it follows that

$$|x(t,t_1,\phi_{t_1})| \le W_1^{-1}\big(W_1(\varepsilon)\big) = \varepsilon.$$

This concludes the proof of (b).

To prove (c), let $t_1 \ge t_0$, and let $0 < \varepsilon < D$. Find δ as in part (b) and take $\eta = \delta$. Let $\phi_{t_1} \in C(t_1)$ with $||\phi_{t_1}|| < \delta$. For simplicity, we write $x(t) = x(t,t_1,\phi_{t_1})$. We will establish the proof by contradiction. Assume that $x(t) \nrightarrow 0$ as $t \to \infty$. Then there are $\varepsilon_1 > 0$ and a sequence $\{t_n\} \to \infty$ with $|x(t_n)| \ge \varepsilon_1$. Since $|F(t,\psi_t)| \le M$ for $t_0 \le t < \infty$ and $||\psi_t|| \le D$, there are $T > 0$ and $\varepsilon_2 < \varepsilon_1$ with $|x(t_n)| \ge \varepsilon_2$ for $t_n \le t \le t_n + T$.

Copyright Elsevier 2022

From this and condition (8.4.7) it follows that

$$
\begin{aligned}
0 \le V(t, x_t) \le V(t_1, \phi_{t_1}) &- \int_{t_1}^{t} W_2(|x(s)|) \\
\le V(t_1, \phi_{t_1}) &- \sum_{i=2}^{n} \int_{t_i}^{t_i+T} W_2(|x(s)|) \\
\le V(t_1, \phi_{t_1}) &- \sum_{i=2}^{n} \int_{t_i}^{t_i+T} W_2(\varepsilon_2) \\
= V(t_1, \phi_{t_1}) &- (n-1)T W_2(\varepsilon_2) \to -\infty \text{ as } n \to \infty,
\end{aligned}
$$

a contradiction. This concludes the proof of (c). ☐

Theorem 8.4.2. *Let* $D > 0$. *Suppose* $V(t, \psi_t)$ *is a scalar functional continuous in* ψ *and locally Lipschitz in* ψ_t *when* $t \ge t_0$ *and* $\psi_t \in C(t)$ *with* $||\psi_t|| < D$. *In addition, we assume that if* $x : [t_0 - \alpha, \infty) \to \mathbb{R}^n$ *is bounded, then* $F(t, x_t)$ *is bounded on* $[t_0, \infty)$. *Suppose* V *satisfies* $V(t, 0) = 0$,

$$W_1(|\psi(t)|) \le V(t, \psi_t) \le W_2(||\psi_t||), \tag{8.4.8}$$

and

$$V'(t, \psi_t) \le -W_3(|\psi(t)|). \tag{8.4.9}$$

Then the zero solution of (8.4.1) *is uniformly asymptotically stable.*

Proof. We refer to [15]. ☐

In the next theorem, we use a Lyapunov function and obtain boundedness results concerning non-homogeneous delay differential equations.

Theorem 8.4.3. *Let* $h > 0$ *and consider the finite delay differential equation*

$$x'(t) = a(t)x(t) + b(t)x(t-h) + g(t), \; t \ge 0, \tag{8.4.10}$$

with $x(t_0) = \varphi(t_0)$, $-h \le t_0 \le 0$. *Suppose* $a(t)$, $b(t)$, *and* $g(t)$ *are continuous for all* $t \ge 0$.

Assume that

$$2a(t) + 3 + |b(t)| \le 0, \tag{8.4.11}$$

$$|b(t)| - e^{-h} \le 0, \tag{8.4.12}$$

and there exists a positive constant L *such that*

$$|g(t)| \le L, \; t \ge 0.$$

Then all solutions of (8.4.10) *are uniformly bounded. Moreover, if* $g(t) = 0$ *for all* $t \ge 0$, *then the zero solution is exponentially stable.*

Copyright Elsevier 2022

Proof. For $t \geq 0$, let

$$V(t, x_t) = x^2(t) + \int_{t-h}^{t} e^{-(t-s)} x^2(s)\, ds.$$

Then along solutions of (8.4.10) we have

$$\begin{aligned} V'(t,x) &= 2a(t)x^2(t) + 2b(t)x(t)x(t-h) + 2x(t)g(t) + x^2(t) \\ &\quad - e^{-h}x^2(t-h) - \int_{t-h}^{t} e^{-(t-s)} x^2(s)\, ds \\ &\leq (2a(t) + 2 + |b(t)|)x^2(t) + (|b(t)| - e^{-h})x^2(t-h) \\ &\quad - \int_{t-h}^{t} e^{-(t-s)} x^2(s)\, ds + g^2(t). \end{aligned}$$

Using conditions (8.4.11) and (8.4.12), we have

$$\begin{aligned} V(t,x_t) + V'(t,x_t) &\leq (2a(t) + 3 + |b(t)|)x^2(t) + (|b(t)| - e^{-h})x^2(t-h) \\ &\quad - \int_{t-h}^{t} e^{-(t-s)} x^2(s)\, ds + \int_{t-h}^{t} e^{-(t-s)} x^2(s)\, ds + g^2(t) \\ &\leq g^2(t) \leq L^2. \end{aligned}$$

So we have

$$V'(t, x_t) \leq -V(t, x_t) + L^2.$$

It follows from the variation of parameters formula that

$$\begin{aligned} |x| &\leq \left[V(t_0, \varphi)e^{-(t-t_0)} + L^2\right]^{1/2} \\ &\leq \left[\left(1 + \int_{t_0-h}^{t_0} e^{-(t_0-s)}\, ds\right)||\varphi||^2\, e^{-(t-t_0)} + L^2\right]^{1/2} \\ &\leq \left[\left(2 - e^{-h}\right)||\varphi||^2\, e^{-(t-t_0)} + L^2\right]^{1/2}. \end{aligned}$$

This gives boundedness. Now if $g(t) = 0$ for all $t \geq 0$, then $L = 0$, and hence the above inequality gives the exponential stability. ☐

We end this section by noting that despite the results of Exercises 8.8 and 8.9, researchers are still trying to substitute delay terms with terms that are independent of delays and arrive at ordinary differential equations that can be explicitly solved. This could be a good prospect for research using numerical analysis.

Copyright Elsevier 2022

8.5 Stability using fixed point theory

Lyapunov functions and functionals have been successfully used to obtain boundedness, stability, and the existence of periodic solutions of differential equations, differential equations with functional delays, and functional differential equations. In the study of differential equations with functional delays by using Lyapunov functionals, many difficulties arise if the delay is unbounded or if the differential equation in question has unbounded terms. This section is mainly concerned with the asymptotic stability of the zero solution of the scalar neutral differential equation

$$x'(t) = -a(t)x(t) + c(t)x'(t-g(t)) + q\big(t, x(t), x(t-g(t))\big), \tag{8.5.1}$$

where $a(t)$, $c(t)$, $g(t)$, and q are continuous in their respective arguments. What makes Eq. (8.5.1) interesting to study is the fact that it cannot be put in the form of

$$\frac{d}{dt}[x(t) + c(t)x(t-g(t))] = -a(t)x(t) + q\big(t, x(t), x(t-g(t))\big),$$

and hence a direct integration cannot be performed to arrive at the desired mapping. We could not construct a suitable Lyapunov functional that yielded any meaningful results about the asymptotic stability of the zero solution. Hence the construction of such a Lyapunov functional for Eq. (8.5.1) remains open. Moreover, in this study, we concentrate on the stability of the zero solution of nonlinear neutral differential equations with bounded and unbounded delays. To achieve our goal, we make use of the contraction mapping principle. To justify the need of using fixed point theory when studying stability, we consider the time-varying delay differential equation

$$x'(t) = -a(t)x(t) + b(t)x(t-g(t)), \tag{8.5.2}$$

where a and b are bounded continuous functions, and g is continuously differentiable and nonnegative.

Theorem 8.5.1. *Suppose*

$$g'(t) \text{ is bounded}, \; 1 - g'(t) > 0, \tag{8.5.3}$$

$$-2a(t) + 1 \le -(1 - g'(t)), \tag{8.5.4}$$

and

$$|b(t)| \le \alpha(1 - g'(t)) \text{ for some } \alpha \in (0,1). \tag{8.5.5}$$

Then the zero solution of (8.5.2) *is uniformly asymptotically stable.*

Proof. Define the Lyapunov functional

$$V(t) = x^2(t) + \int_{t-g(t)}^{t} x^2(s)\, ds.$$

Copyright Elsevier 2022

Then along solutions of (8.5.2) we have

$$\begin{aligned}
V'(t) &= 2x(t)\Big(-a(t)x(t) + b(t)x(t-g(t))\Big) + x^2(t) - (1-g'(t))x^2(t-g(t)) \\
&\le -2a(t)x^2(t) + |b(t)|x^2(t) + |b(t)|x^2(t-g(t)) \\
&+ x^2(t) - (1-g'(t))x^2(t-g(t)) \\
&= \Big(-2a(t) + |b(t)| + 1\Big)x^2(t) + \Big(|b(t)| - (1-g'(t))\Big)x^2(t-g(t)) \\
&\le (\alpha-1)(1-g'(t))\Big(x^2(t) + x^2(t-g(t))\Big) \\
&\le (\alpha-1)(1-g'(t))x^2(t).
\end{aligned}$$

The result follows from Theorem 8.4.2. □

We remark that (8.5.4) and (8.5.5) imply that $|b(t)| \le \alpha(2a(t)-1)$. Thus we see that $|b(t)|$ must be bounded by $a(t)$. In addition, $a(t)$ must satisfy $a(t) \ge 1/2$.

8.5.1 Neutral differential equations

Conditions (8.5.3) and (8.5.4) are severe since a and b must be bounded and both a and b must satisfy an upper bound in terms of $1-g'$. To relax the restrictions, we resort to the use of fixed point theory, namely, the contraction mapping principle. To better illustrate our procedure, we first consider the scalar linear neutral differential equation (neutral since there is a delay in the derivative) with unbounded delay

$$x'(t) = -a(t)x(t) + b(t)x(t-g(t)) + c(t)x'(t-g(t)), \tag{8.5.6}$$

where $a(t)$ and $b(t)$ are continuous, $c(t)$ is continuously differentiable, and $g(t) > 0$ for all $t \in \mathbb{R}$ and is twice continuously differentiable. We note that Eq. (8.5.6) is more complicated than that we considered in Theorem 8.5.1 since the derivative on the right side has a delay. Hence the Lyapunov functional that was used in the proof of Theorem 8.5.1 will not work, and we have to construct a suitable one. However, the author believes that conditions similar to (8.5.3)–(8.5.5) have to be imposed. Using fixed point theory, a suitable mapping has to be formulated. To obtain the desired map, we invert Eq. (8.5.6). During the process of inverting (8.5.6), we have to integrate by parts the term involving $x'(t-g(t))$. We begin by requiring that

$$g'(t) \ne 1, \quad \forall\ t \in \mathbb{R}. \tag{8.5.7}$$

Definition 8.3.1 naturally extends to suit Eq. (8.5.6).

Lemma 8.1. *Suppose* (8.5.7) *holds. Then* $x(t)$ *is a solution of Eq.* (8.5.6) *if and only if*

Copyright Elsevier 2022

$$x(t) = \left(x(0) - \frac{c(0)}{1-g'(0)}x(-g(0))\right)e^{-\int_0^t a(s)ds} + \frac{c(t)}{1-g'(t)}x(t-g(t))$$
$$- \int_0^t (r(u) - b(u))x(u-g(u))e^{-\int_u^t a(s)ds}du, \tag{8.5.8}$$

where

$$r(u) = \frac{\left(c'(u) + c(u)a(u)\right)\left(1-g'(u)\right) + g''(u)c(u)}{(1-g'(u))^2}. \tag{8.5.9}$$

Proof. Multiplying both sides of (8.5.6) by $e^{\int_0^t a(s)ds}$ and then integrating from 0 to t, we obtain

$$\int_0^t \left[x(u)e^{\int_0^u a(s)ds}\right]' du = \int_0^t \left[b(u)x(u-g(u)) + c(u)x'(u-g(u))\right]e^{\int_0^u a(s)ds}du.$$

As a consequence, we arrive at

$$x(t)e^{\int_0^t a(s)ds} - x(0) = \int_0^t \left[b(u)x(u-g(u)) + c(u)x'(u-g(u))\right]e^{\int_0^u a(s)ds}du.$$

Dividing both sides of the above equation by $e^{\int_0^t a(s)ds}$, we obtain

$$x(t) = x(0)e^{-\int_0^t a(s)ds}$$
$$+ \int_0^t \left[b(u)x(u-g(u)) + c(u)x'(u-g(u))\right]e^{-\int_u^t a(s)ds}du. \tag{8.5.10}$$

Rewrite

$$\int_0^t c(u)x'(u-g(u))e^{-\int_u^t a(s)ds}du$$
$$= \int_0^t \frac{c(u)x'(u-g(u))(1-g'(u))}{(1-g'(u))}e^{-\int_u^t a(s)ds}du.$$

Integrating by parts the above integral with

$$U(u) = \frac{c(u)}{1-g'(u)}e^{-\int_u^t a(s)ds}$$

and

$$dV = x'(u-g(u))(1-g'(u))du,$$

we have

$$\int_0^t c(u)x'(u-g(u))e^{-\int_u^t a(s)ds}du = \frac{c(t)}{1-g'(t)}x(t-g(t))$$
$$- \frac{c(0)}{1-g'(0)}x(-g(0))e^{-\int_0^t a(s)ds} - \int_0^t r(u)e^{-\int_u^t a(s)ds}x(u-g(u))du, \tag{8.5.11}$$

Copyright Elsevier 2022

where $r(u)$ is given by (8.5.9). Finally, substituting (8.5.11) into (8.5.10) completes the proof. $\square$

Next, let $\psi(t) : (-\infty, 0] \to \mathbb{R}$ be a continuous bounded initial function. We say that $x(t) := x(t, 0, \psi)$ is a solution of (8.5.6) if $x(t) = \psi(t)$ for $t \le 0$ and satisfies (8.5.6) for $t \ge 0$.

We say that the zero solution of (8.5.6) is stable at t_0 if for each $\varepsilon > 0$, there is $\delta = \delta(\varepsilon) > 0$ such that $\big[\psi : [-\infty, t_0] \to \mathbb{R}$ with $|\psi(t)| < \delta$ on $(-\infty, t_0],\ t \ge t_0\big]$ implies $|x(t, t_0, \psi)| < \varepsilon$.

Let C be the space of all continuous functions from $\mathbb{R} \to \mathbb{R}$ and define the set

$$S = \Big\{\varphi : \mathbb{R} \to \mathbb{R} \,\Big|\, \varphi(t) = \psi(t) \text{ if } t \le 0,\ \varphi(t) \to 0 \text{ as } t \to \infty,\ \varphi \in C, \\ \text{and } \varphi \text{ is bounded}\Big\}.$$

Then $\big(S, ||\cdot||\big)$ is a complete metric space, where $||\cdot||$ is the supremum norm.

For the next theorem, we impose the following conditions:

$$e^{-\int_0^t a(s)ds} \to 0 \text{ as } t \to \infty, \tag{8.5.12}$$

there is $\alpha > 0$ such that

$$\left|\frac{c(t)}{1-g'(t)}\right| + \int_0^t |r(u) - b(u)| e^{-\int_u^t a(s)ds} du \le \alpha < 1, t \ge 0, \tag{8.5.13}$$

and

$$t - g(t) \to \infty \text{ as } t \to \infty. \tag{8.5.14}$$

Theorem 8.5.2. *If* (8.5.7) *and* (8.5.12)–(8.5.14) *hold, then every solution* $x(t, 0, \psi)$ *of* (8.5.6) *with small continuous initial function* $\psi(t)$ *is bounded and goes to zero as* $t \to \infty$. *Moreover, the zero solution is stable at* $t_0 = 0$.

Proof. Define the mapping $P : S \to S$ by

$$\big(P\varphi\big)(t) = \psi(t) \text{ if } t \le 0$$

and

$$\big(P\varphi\big)(t) = \Big(\psi(0) - \frac{c(0)}{1-g'(0)}\psi(-g(0))\Big)e^{-\int_0^t a(s)ds} + \frac{c(t)}{1-g'(t)}\varphi(t-g(t)) \\ - \int_0^t \Big[r(u) - b(u)\Big]\varphi(u - g(u))e^{-\int_u^t a(s)ds} du,\ t \ge 0.$$

It is clear that for $\varphi \in S$, $P\varphi$ is continuous. Let $\varphi \in S$ with $||\varphi|| \le K$ for some positive constant K. Let $\psi(t)$ be a small continuous initial function with $|\psi| < \delta$, $\delta > 0$. Then

Copyright Elsevier 2022

using (8.5.13) in the definition of $(P\varphi)(t)$, we have

$$\begin{aligned}
||(P\varphi)(t)|| &\le |(1-\frac{c(0)}{1-g'(0)})|K+|\frac{c(t)}{1-g'(t)}|K \\
&\quad +\int_0^t |r(u)-b(u)|e^{-\int_u^t a(s)ds}du\,K \\
&\le |(1-\frac{c(0)}{1-g'(0)})|\delta+\alpha K, \qquad (8.5.15)
\end{aligned}$$

which implies that $||\big(P\varphi\big)(t)|| \le K$ for the right δ. Thus (8.5.15) implies that $(P\varphi)(t)$ is bounded. Next, we show that $\big(P\varphi\big)(t) \to 0$ as $t \to \infty$. The first term on the right side of $\big(P\varphi\big)(t)$ tends to zero by condition (8.5.12). Also, the second term on the right side tends to zero because of (8.5.14) and the fact that $\varphi \in S$. It remains to show that the integral term goes to zero as $t \to \infty$.

Let $\varepsilon > 0$, and let $\varphi \in S$ with $||\varphi|| \le K$, $K > 0$. Then there exists $t_1 > 0$ such that $|\varphi(t-g(t))| < \varepsilon$ for $t > t_1$. Due to condition (8.5.12), there exists $t_2 > t_1$ such that $e^{-\int_{t_1}^t a(s)ds} < \frac{\varepsilon}{\alpha K}$ for $t > t_2$. Thus for $t > t_2$, we have

$$\begin{aligned}
&\Big|\int_0^t (r(u)-b(u))\varphi(u-g(u))e^{-\int_u^t a(s)ds}du\Big| \\
&\le K\int_0^{t_1} |r(u)-b(u)|e^{-\int_u^t a(s)ds}du+\varepsilon\int_{t_1}^t |r(u)-b(u)|e^{-\int_u^t a(s)ds}du \\
&\le Ke^{-\int_{t_1}^t a(s)ds}\int_0^{t_1} |r(u)-b(u)|e^{-\int_u^{t_1} a(s)ds}du+\alpha\varepsilon \\
&\le \alpha Ke^{-\int_{t_1}^t a(s)ds}+\alpha\varepsilon \\
&\le \varepsilon+\alpha\varepsilon.
\end{aligned}$$

Hence $\big(P\varphi\big)(t) \to 0$ as $t \to \infty$. It remains to show that $\big(P\varphi\big)(t)$ is a contraction under the supremum norm. Let $\zeta, \eta \in S$. Then

$$\begin{aligned}
\Big|(P\zeta)(t)-(P\eta)(t)\Big| &\le \Big\{|\frac{c(t)}{1-g'(t)}|+\int_0^t |r(u)-b(u)|e^{-\int_u^t a(s)ds}du\Big\}||\zeta-\eta|| \\
&\le \alpha||\zeta-\eta||.
\end{aligned}$$

Thus by the contraction mapping principle, P has a unique fixed point in S, which solves (8.5.6), is bounded, and tends to zero as t tends to infinity. The stability of the zero solution at $t_0 = 0$ follows from the above work by simply replacing K by ε. This completes the proof. □

Example 8.4. Consider the linear neutral differential equation

$$x'(t) = -2x(t) + c_0\,x'(t-(\frac{t}{2})), \qquad (8.5.16)$$

Copyright Elsevier 2022

where c_0 is a constant. Then $r(u) = -4c_0$, and condition (8.5.13) is satisfied for $|c_0| \leq \frac{\alpha}{4}$, $\alpha \in (0,1)$. Let $\psi(t)$ be a given continuous initial function with $|\psi(t)| \leq \delta$. Let

$$S = \Big\{\varphi : \mathbb{R} \to \mathbb{R} \,\Big|\, \varphi(t) = \psi(t) \text{ if } t \leq 0,\ \varphi(t) \to 0 \text{ as } t \to \infty,\ \varphi \in C, \\ \text{and } \varphi \text{ is bounded}\Big\}.$$

Define

$$\big(P\varphi\big)(t) = \psi(t) \ \text{ if } \ t \leq 0$$

and

$$\big(P\varphi\big)(t) = \Big((1-2c_0)\psi(0)\Big)e^{-2t} + 2c_0\varphi(\frac{t}{2}) \\ + \int_0^t 4c_0\varphi(\frac{u}{2})e^{-2(t-u)}du,\ t \geq 0.$$

Then, for $\varphi \in S$ with $||\varphi|| \leq K$, where $K \geq \frac{(1-\alpha)}{|1-2c_0|\delta}$, we have $||\big(P\varphi\big)(t)|| \leq K$. It is obvious that conditions (8.5.12) and (8.5.14) are satisfied. To see that P defines a contraction mapping, let $\zeta, \eta \in S$. Then

$$\Big|(P\zeta)(t) - (P\eta)(t)\Big| \leq 2|c_0|\,||\zeta-\eta|| + 2|c_0|(1-e^{-2t})||\zeta-\eta|| \\ \leq 4|c_0|\,||\zeta-\eta|| \leq \alpha||\zeta-\eta||.$$

Hence by Theorem 8.5.2 every solution $x(t,0,\psi)$ of (8.5.16) with small continuous initial function $\psi(t) : (-\infty,\ 0] \to \mathbb{R}$ is in S, is bounded, and goes to zero as $t \to \infty$.

Next, we turn our attention to the nonlinear neutral differential equation with unbounded delay

$$x'(t) = -a(t)x(t) + c(t)x'(t-g(t)) + q\Big(x(t), x(t-g(t))\Big), \tag{8.5.17}$$

where $a(t)$, $c(t)$, and $g(t)$ are defined as before. Here we assume that $q(0,0) = 0$ and q is locally Lipschitz continuous in x and y in the sense that there is $K > 0$ such that if $|x|, |y|, |z|, |w| \leq K$, then

$$|q(x,y) - q(z,w)| \leq L|x-z| + E|y-w| \tag{8.5.18}$$

for some positive constants L and E.

Note that

$$|q(x,y)| = |q(x,y) - q(0,0) + q(0,0)| \\ \leq |q(x,y) - q(0,0)| + |q(0,0)| \\ \leq L|x| + E|y|.$$

Copyright Elsevier 2022

Let

$$S=\Big\{\varphi:\mathbb{R}\to\mathbb{R}\,\Big|\,||\varphi||\le K,\ \varphi(t)=\psi(t)\text{ if }t\le 0,\ \varphi(t)\to 0\text{ as }t\to\infty,\ \varphi\in C\Big\}.$$

Define the map $P:S\to S$ by

$$\big(P\varphi\big)(t)=\psi(t)\quad\text{if }\ t\le 0$$

and

$$\big(P\varphi\big)(t)=\Big(\psi(0)-\frac{c(0)}{1-g'(0)}\psi(-g(0))\Big)e^{-\int_0^t a(s)ds}+\frac{c(t)}{1-g'(t)}\varphi(t-g(t))$$
$$+\int_0^t\Big[-r(u)\varphi(u-g(u))+q(\varphi(u),\varphi(u-g(u)))\Big]e^{-\int_u^t a(s)ds}du,\ t\ge 0.$$

It is clear that $P\varphi$ is continuous for $\varphi\in S$. If P has a fixed point, say φ, then φ is a solution of (8.5.17). For P to be a contraction, we assume that there is $\alpha>0$ such that

$$\Big|\frac{c(t)}{1-g'(t)}\Big|+\int_0^t(|r(u)|+L+E)e^{-\int_u^t a(s)ds}du\le\alpha<1,t\ge 0. \tag{8.5.19}$$

Theorem 8.5.3. *If* (8.5.7), (8.5.12), (8.5.14), *and* (8.5.19) *hold, then every solution* $x(t,0,\psi)$ *of* (8.5.17) *with small continuous initial function* $\psi(t):(-\infty,0]\to\mathbb{R}$ *goes to zero as* $t\to\infty$. *Moreover, the zero solution is stable at* $t_0=0$.

Proof. Let $\varphi\in S$, and let t_1 and t_2 be as in the proof of Theorem 8.5.2. Then for $t>t_2$,

$$\begin{aligned}&\Big|\int_0^t q(\varphi(u),\varphi(u-g(u)))e^{-\int_u^t a(s)ds}du\Big|\\&\le K(L+E)\int_0^{t_1}e^{-\int_u^t a(s)ds}du+\varepsilon(L+E)\int_{t_1}^t e^{-\int_u^t a(s)ds}du\\&\le K(L+E)e^{-\int_{t_1}^t a(s)ds}\int_0^{t_1}e^{-\int_u^{t_1}a(s)ds}du+\varepsilon(L+E)\alpha\\&\le\alpha Ke^{-\int_{t_1}^t a(s)ds}+\varepsilon(L+E)\alpha\\&\le\varepsilon+\varepsilon(L+E)\alpha.\end{aligned}$$

This, along with the proof of Theorem 8.5.2, shows that $\big(P\varphi\big)(t)\to 0$ as $t\to\infty$. Let $\psi(t)$ be a small continuous initial function with $|\psi|<\delta$, $\delta>0$. Then by using

Copyright Elsevier 2022

(8.5.19) we arrive at

$$\begin{aligned}
||\big(P\varphi\big)(t)|| &\leq |(1-\frac{c(0)}{1-g'(0)})|\delta + |\frac{c(t)}{1-g'(t)}|K \\
&\quad + \int_0^t (|r(u)|+L+E)e^{-\int_u^t a(s)ds}du K \\
&\leq |(1-\frac{c(0)}{1-g'(0)})|\delta + \alpha K,
\end{aligned} \tag{8.5.20}$$

which implies that $||\big(P\varphi\big)(t)|| \leq K$ with the right choice of δ and α. It remains to show that $\big(P\varphi\big)(t)$ is a contraction. Let $\zeta, \eta \in S$. Then

$$\begin{aligned}
&\Big|(P\zeta)(t) - (P\eta)(t)\Big| \\
&\leq \Big\{|\frac{c(t)}{1-g'(t)}|\,||\zeta-\eta|| + \int_0^t \big(|r(u)||\zeta(u-g(u)) - \eta(u-g(u))|e^{-\int_u^t a(s)ds}du \\
&+ \int_0^t |q(\zeta(u),\zeta(u-g(u))) - q(\eta(u),\eta(u-g(u)))|e^{-\int_u^t a(s)ds}du\Big\} \\
&\leq \Big\{|\frac{c(t)}{1-g'(t)}| + \int_0^t (|r(u)|+L+E)e^{-\int_u^t a(s)ds}du\Big\}||\zeta-\eta|| \\
&\leq \alpha||\zeta-\eta||.
\end{aligned}$$

Thus by the contraction mapping principle, P has a unique fixed point in S, which solves (8.5.17) and tends to zero as t tends to infinity. This completes the proof. ☐

8.5.2 Neutral Volterra integro-differential equations

Now we turn our attention to the scalar neutral Volterra integro-differential equation

$$x'(t) = -a(t)x(t) + c(t)x'(t-g(t)) + \int_{t-g(t)}^t k(t,s)h(x(s))ds, \tag{8.5.21}$$

where $0 \leq g(t) \leq g_0$ for some constant g_0. Here we assume that $g, a, c, h : \mathbb{R} \to \mathbb{R}$ and $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are continuous. For the next theorem, we make the following assumptions:

$$\Big|\frac{c(t)}{1-g'(t)}\Big| + \int_0^t \Big[|r(s)| + \int_{s-g(s)}^s |k(s,u)|du\Big]e^{-\int_s^t a(u)du}ds \leq \alpha < 1, t \geq 0; \tag{8.5.22}$$

for each $\varepsilon > 0$, there exist $t_1 > 0$ and $T > 0$ such that for $t_2 \geq t_1$ and $t \geq t_2 + T$, we have

$$e^{-\int_{t_2}^t a(s)ds} < \varepsilon \text{ and } e^{-\int_0^t a(s)ds} \to 0 \text{ as } t \to \infty; \tag{8.5.23}$$

Copyright Elsevier 2022

and there is $L > 0$ such that $|x|, |y| \le L$ imply that

$$|h(x) - h(y)| \le |x - y| \text{ and } h(0) = 0. \tag{8.5.24}$$

Theorem 8.5.4. *If* (8.5.22)–(8.5.24) *hold, then the zero solution of* (8.5.21) *is asymptotically stable at* $t_0 = 0$.

Proof. Let $\psi : [-g_0, 0] \to \mathbb{R}$ be a bounded initial function with $|\psi| < \delta$ for some positive constant δ. Define

$$S = \Big\{\varphi : [-g_0, \infty) \to \mathbb{R} \,\Big|\, \varphi(t) = \psi(t) \text{ if } -g_0 \le t \le 0,\ ||\varphi|| \le L,$$
$$\varphi \in C,\ \varphi(t) \to 0 \text{ as } t \to \infty\Big\},$$

where $||\cdot||$ is the supremum norm.

Define the mapping $P : S \to S$ by

$$\big(P\varphi\big)(t) = \psi(t) \text{ if } -g_0 \le t \le 0$$

and

$$\big(P\varphi\big)(t) = \Big(\psi(0) - \frac{c(0)}{1-g'(0)}\psi(-g(0))\Big)e^{-\int_0^t a(s)ds} + \frac{c(t)}{1-g'(t)}\varphi(t-g(t))$$
$$+ \int_0^t \Big[-r(s)\varphi(s-g(s)) + \int_{s-g(s)}^s k(s,u)h(x(u))du\Big]e^{-\int_s^t a(u)du}ds,\ t \ge 0,$$

where $r(u)$ is defined by (8.5.9). For $\varphi \in S$ with $||\varphi|| \le L$, we have that for any $\varepsilon > 0$, there exists $t_1 > 0$ such that $t \ge t_1 - g_0$ implies that $|\varphi(t)| < \varepsilon$. Similarly, $t_2 \ge t_1$ and $t \ge t_2 + T$ imply that $e^{-\int_{t_2}^t a(s)ds} < \varepsilon$. Thus $t \ge t_2 + T$ and t_1 large enough give

$$\begin{aligned}
||\big(P\varphi\big)(t)|| &\le |(1 - \frac{c(0)}{1-g'(0)})|\delta\varepsilon + |\frac{c(t)}{1-g'(t)}|\varepsilon \\
&\quad + \int_0^{t_2}\Big[|r(s)|L + \int_{s-g(s)}^s k(s,u)Ldu\Big]e^{-\int_s^{t_2} a(u)du}e^{-\int_{t_2}^t a(u)du}ds \\
&\quad + \int_{t_2}^t \Big[|r(s)|\varepsilon + \int_{s-g(s)}^s |k(s,u)|\varepsilon du\Big]e^{-\int_s^t a(u)du}ds \\
&\quad \le |(1 - \frac{c(0)}{1-g'(0)})|\delta\varepsilon + |\frac{c(t)}{1-g'(t)}|\varepsilon + \varepsilon\alpha L + \varepsilon\alpha.
\end{aligned} \tag{8.5.25}$$

Thus $\big(P\varphi\big)(t) \to 0$ as $t \to \infty$. Also, as in the proof of (8.5.15), with the right choice of δ, we have that $||P(\varphi)|| \le L$ for $||\varphi|| \le L$. It remains to show that P is

Copyright Elsevier 2022

a contraction. For $\zeta, \eta \in S$, we have

$$\begin{aligned}\Big|(P\zeta)(t)-(P\eta)(t)\Big| &\le \Big|\frac{c(t)}{1-g'(t)}\Big|\,||\zeta-\eta|| \\ &+\int_0^t \Big[|r(s)|+\int_{s-g(s)}^{s}|k(s,u)|du\Big]e^{-\int_s^t a(u)du}ds||\zeta-\eta|| \\ &\le \alpha||\zeta-\eta||.\end{aligned}$$

Hence P has a unique fixed point in S. This completes the proof. □

We conclude this section by considering Eq. (8.5.21) when there is no delay in the derivative, that is, $c(t)$ is identically zero. We use the Lyapunov functional method to show that the zero solution is asymptotically stable. As we try to show that the derivative term of the Lyapunov functional is less than zero along the solutions of the desired equation, an unpleasant condition relating the size of $a(t)$ to the size of $g(t)$ arises, which limits the types of equations that can be discussed. In particular, we consider the scalar Volterra integro-differential delay equation

$$x'(t) = -a(t)x(t) + \int_{t-g(t)}^{t} k(t,s)h(x(s))ds, \tag{8.5.26}$$

where $0 \le g(t) \le g_0$ for some constant g_0.

Suppose there exists a positive constant β such that

$$|k(t,s)| \le \beta \text{ for all } t \ge 0, \tag{8.5.27}$$

$$g'(t) \le 1-\beta, \tag{8.5.28}$$

and

$$a(t) \ge \frac{(\beta+1)}{2}g(t) \text{ for all } t \ge 0. \tag{8.5.29}$$

Theorem 8.5.5. *If (8.5.24) and (8.5.27)–(8.5.29) hold, then the zero solution of (8.5.26) is asymptotically stable.*

Proof. Define

$$V(t) = x^2(t) + \int_{t-g(t)}^{t}\int_s^t h^2(x(v))dvds.$$

It follows from (8.5.24) and (8.5.27) that along the solutions of (8.5.26)

$$\begin{aligned}V'(t) &= 2x(t)\Big[-a(t)x(t)+\int_{t-g(t)}^{t}k(t,s)h(x(s))ds\Big] \\ &\quad -(1-g'(t))\int_{t-g(t)}^{t}h^2(x(v))dv + g(t)h^2(x(t))\end{aligned}$$

Copyright Elsevier 2022

$$\begin{aligned}
&\leq -2a(t)x^2(t) + 2\beta|x(t)| \int_{t-g(t)}^{t} |h(x(s))|ds \\
&- (1-g'(t)) \int_{t-g(t)}^{t} h^2(x(s))ds + g(t)h^2(x(t)) \\
&\leq -2a(t)x^2(t) + \beta \int_{t-g(t)}^{t} \Big[x^2(t) + h^2(x(s))ds\Big] \\
&- (1-g'(t)) \int_{t-g(t)}^{t} h^2(x(v))dv + g(t)h^2(x(t)) \\
&\leq x^2(t)\Big[-2a(t) + \beta g(t) + g(t)\Big] + (1-g'(t)+\beta) \int_{t-g(t)}^{t} h^2(x(s))ds \\
&\leq -\gamma x^2(t) \text{ for some positive constant } \gamma.
\end{aligned}$$

Then by (c) of Theorem 8.4.1, the zero solution is asymptotically stable. ☐

8.6 Exponential stability

In this section we consider the scalar and totally nonlinear Volterra integro-differential equation with finite delay

$$x'(t) = -\int_{t-r}^{t} a(t,s)g(x(s))ds, \tag{8.6.1}$$

where $r > 0$ is a constant, and $a : [0,\infty) \times [0,\infty] \to (-\infty,\infty)$. The function $g(x)$ is continuous in x.

The Lyapunov method allowed us to deduce inequalities that all solutions must satisfy and from which we deduce the exponential stability and instability.

Let $\psi : [-r, 0] \to (-\infty,\infty)$ be a continuous initial function with

$$||\psi|| = \max_{-r\leq s\leq 0} |\psi(s)|.$$

It should cause no confusion to denote the norm of a continuous function $\varphi : [-r,\infty) \to (-\infty,\infty)$ by

$$||\varphi|| = \sup_{-r\leq s<\infty} |\varphi(s)|.$$

The notation x_t means that $x_t(\tau) = x(t+\tau)$, $\tau \in [-r, 0]$, as long as $x(t+\tau)$ is defined. Thus x_t is a function mapping an interval $[-h, 0]$ into $\mathbb{R}$. We say that $x(t) \equiv x(t, t_0, \psi)$ is a solution of (8.6.1) if $x(t)$ satisfies (8.6.1) for $t \geq t_0$ and $x_{t_0} = x(t_0 + s) = \psi(s)$, $s \in [-r, 0]$.

Now we turn our attention to the totally nonlinear equation (8.6.1). We will construct a Lyapunov functional $V(t,x) := V(t)$ and show that for some positive α,

Copyright Elsevier 2022

under suitable conditions, $V'(t) \leq -\alpha V(t)$ along the solutions of (8.6.1). To rewrite (8.6.1) so that a suitable Lyapunov functional can be displayed, we let

$$A(t,s) := \int_{t-s}^{r} a(u+s,s)du, \ t,s \geq 0,$$

and assume that

$$A(t,t) = \int_{0}^{r} a(u+t,t)du \geq 0. \tag{8.6.2}$$

In preparation of the main results, we assume that

$$xg(x) \geq x^2 \ \text{ if } \ x \neq 0 \tag{8.6.3}$$

and there exists a positive constant λ such that

$$|g(x)| \leq \lambda |x|. \tag{8.6.4}$$

It is clear that conditions (8.6.3) and (8.6.4) imply that $g(0) = 0$. In addition to the above assumptions, we require that

$$A(t,s)\frac{\partial A(t,s)}{\partial t} \leq 0 \ \text{ for all } \ (t,s) \in [0,\infty) \times [t-r,t]. \tag{8.6.5}$$

Finally, we assume that for $1 < \alpha \leq 2$,

$$A^2\left(t-(\alpha-1)r/\alpha, z\right) \geq A^2(t,z) \tag{8.6.6}$$

for all $t \in [0,\infty)$ and $z \in [t-r/\alpha, t-(\alpha-1)r/\alpha]$. As a consequence of (8.6.5), we have

$$\begin{aligned}
&\int_{-r}^{0}\int_{t+s}^{t} A(t,z)\frac{\partial A(t,z)}{\partial t} g^2(x(z))dz\,ds \\
&= \int_{t-r}^{t}\int_{-r}^{z-t} A(t,z)\frac{\partial A(t,z)}{\partial t} g^2(x(z))ds\,dz \\
&= \int_{t-r}^{t} A(t,z)\frac{\partial A(t,z)}{\partial t} g^2(x(z))(z-t+r)dz \\
&\leq 0,
\end{aligned} \tag{8.6.7}$$

which plays an essential role in the proof of the next lemma. We note that $g(x) = x(\sin^2(x)+1)$ satisfies (8.6.3) and (8.6.4). To construct a suitable Lyapunov functional, we put (8.6.1) in the form

$$x'(t) = -A(t,t)g(x(t)) + \frac{d}{dt}\int_{t-r}^{t} A(t,s)g(x(s))ds. \tag{8.6.8}$$

Copyright Elsevier 2022

Lemma 8.2. *Let* (8.6.2)–(8.6.5) *hold, and suppose that* $0 < r \le 1/2$ *with*

$$\big[(r+1)\lambda^2+1\big]\Big(\int_0^r a(u+t,t)du\Big)^2 \le \int_0^r a(u+t,t)du \le \frac{1-2r}{2r}. \tag{8.6.9}$$

If

$$\begin{aligned} V(t) &= \Big(x(t) - \int_{t-r}^t A(t,s)g(x(s))ds\Big)^2 \\ &\quad + \int_{-r}^0 \int_{t+s}^t A^2(t,z)g^2(x(z))dz\,ds, \end{aligned} \tag{8.6.10}$$

then along the solutions of (8.6.1) *we have*

$$V'(t) \le -A(t,t)V(t).$$

Proof. Let $x(t) = x(t,t_0,\psi)$ be a solution of (8.6.1) and define $V(t)$ by (8.6.10). Then along solutions of (8.6.1) we have

$$\begin{aligned} V'(t) &= 2\Big(x(t) - \int_{t-r}^t A(t,s)g(x(s))ds\Big)[-A(t,t)g(x(t))] \\ &\quad + rA^2(t,t)g^2(x(t)) - \int_{-r}^0 A^2(t,t+s)g^2(x(t+s))ds \\ &\quad + \int_{-r}^0 \int_{t+s}^t 2A(t,z)\frac{\partial A(t,z)}{\partial t}g^2(x(z))dz\,ds \\ &\quad \le -A(t,t)\big[x^2(t) - 2x(t)\int_{t-r}^t A(t,s)g(x(s))ds\big] + r\lambda^2A^2(t,t)x^2(t) \\ &\quad - \int_{-r}^0 A^2(t,t+s)g^2(x(t+s))ds - A(t,t)x^2(t) \ \text{ by (8.5.6)} \\ &\quad - 2A(t,t)x(t)\int_{t-r}^t A(t,s)g(x(s))ds + 2A(t,t)g(x(t))\int_{t-r}^t A(t,s)g(x(s))ds \\ &\quad = -A(t,t)V(t) + A(t,t)\Big(\int_{t-r}^t A(t,s)g(x(s))ds\Big)^2 \\ &\quad - 2A(t,t)x(t)\int_{t-r}^t A(t,s)g(x(s))ds + A(t,t)\int_{-r}^0 \int_{t+s}^t A^2(t,z)g^2(x(z))dz\,ds \\ &\quad + 2A(t,t)g(x(t))\int_{t-r}^t A(t,s)g(x(s))ds \\ &\quad + \big(r\lambda^2A^2(t,t)x^2(t) - A(t,t)x^2(t)\big) \\ &\quad - \int_{-r}^0 A^2(t,t+s)g^2(x(t+s))ds. \end{aligned} \tag{8.6.11}$$

Copyright Elsevier 2022

In what follows, we perform some calculations to simplify (8.6.11). First, if we let $u = t + s$, then

$$-\int_{-r}^{0} A^2(t, t+s)g^2(x(t+s))ds = -\int_{t-r}^{t} A^2(t, s)g^2(x(s))ds. \tag{8.6.12}$$

Also, by Holder's inequality we have that

$$A(t,t)\Big(\int_{t-r}^{t} A(t,s)g(x(s))ds\Big)^2 \le A(t,t)r\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds. \tag{8.6.13}$$

Finally, we easily observe that

$$A(t,t)\int_{-r}^{0}\int_{t+s}^{t} A^2(t,z)g^2(x(z))dz\,ds \le A(t,t)r\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds \tag{8.6.14}$$

and

$$\begin{aligned} -2A(t,t)x(t)\int_{t-r}^{t} A(t,s)g(x(s))ds &\le A^2(t,t)x^2(t) \\ &\quad + r\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds. \end{aligned} \tag{8.6.15}$$

Similarly,

$$\begin{aligned} 2A(t,t)g(x(t))\int_{t-r}^{t} A(t,s)g(x(s))ds &\le \lambda^2 A^2(t,t)x^2(t) \\ &\quad + r\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds. \end{aligned} \tag{8.6.16}$$

Invoking (8.6.12) and substituting expressions (8.6.12)–(8.6.16) into (8.6.11) yield

$$\begin{aligned} V'(t) &\le -A(t,t)V(t) + \big[(r+1)\lambda^2 A^2(t,t) + A^2(t,t) - A(t,t)\big]x^2(t) \\ &\quad + [2rA(t,t) + 2r - 1]\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds \\ &\le -A(t,t)V(t) \text{ by (8.6.9).} \end{aligned} \tag{8.6.17}$$

□

Theorem 8.6.1. *Assume that the hypotheses of Lemma 8.2 and (8.6.6) hold, and let* $1 < \alpha \le 2$. *Then any solution* $x(t) = x(t, t_0, \psi)$ *of (8.6.1) satisfies the exponential*

Copyright Elsevier 2022

inequality

$$|x(t)| \leq \sqrt{2\frac{1+\frac{\alpha-1}{\alpha}}{\frac{\alpha-1}{\alpha}}V(t_0)}e^{-\frac{1}{2}\int_{t_0}^{t-(\frac{\alpha-1}{\alpha})r} A(s,s)ds} \tag{8.6.18}$$

for $t \geq t_0 + (\dfrac{\alpha-1}{\alpha})r$.

Proof. By changing the order of integrations we have

$$\begin{aligned}\int_{-r}^{0}\int_{t+s}^{t} A^2(t,z)g^2(x(z))dz\,ds &= \int_{t-r}^{t}\int_{-r}^{z-t} A^2(t,z)g^2(x(z))ds\,dz \\ = \int_{t-r}^{t} A^2(t,z)g^2(x(z))(z-t+r)dz.\end{aligned} \tag{8.6.19}$$

For $1 < \alpha \leq 2$, if $t - \dfrac{r}{\alpha} \leq z \leq t$, then $(\dfrac{\alpha-1}{\alpha})r \leq z - t + r \leq r$. Expression (8.6.19) yields

$$\begin{aligned}&\int_{-r}^{0}\int_{t+s}^{t} A^2(t,z)g^2(x(z))dz\,ds = \int_{t-r}^{t} A^2(t,z)g^2(x(z))(z-t+r)dz \\ &= \int_{t-r}^{t-\frac{r}{\alpha}} A^2(t,z)g^2(x(z))(z-t+r)dz + \int_{t-\frac{r}{\alpha}}^{t} A^2(t,z)g^2(x(z))(z-t+r)dz \\ &\geq \int_{t-\frac{r}{\alpha}}^{t} A^2(t,z)g^2(x(z))(z-t+r)dz \\ &\geq (\frac{\alpha-1}{\alpha})r\int_{t-\frac{r}{\alpha}}^{t} A^2(t,z)g^2(x(z))dz.\end{aligned} \tag{8.6.20}$$

Let $V(t)$ be given by (8.6.10). Then

$$\begin{aligned}V(t) &\geq \int_{-r}^{0}\int_{t+s}^{t} A^2(t,z)g^2(x(z))dz\,ds \\ &\geq (\frac{\alpha-1}{\alpha})r\int_{t-\frac{r}{\alpha}}^{t} A^2(t,z)g^2(x(z))dz.\end{aligned} \tag{8.6.21}$$

Using (8.6.6), this implies that, for $1 < \alpha \leq 2$,

$$\begin{aligned}V(t-(\alpha-1)r/\alpha) &\geq (\alpha-1)r/\alpha\int_{t-r}^{t-r+r/\alpha} A^2(t,z)g^2(x(z))dz \\ &\geq (\alpha-1)r/\alpha\int_{t-r}^{t-r/\alpha} A^2(t,z)g^2(x(z))dz.\end{aligned} \tag{8.6.22}$$

Copyright Elsevier 2022

Note that since $V'(t) \le 0$, for $t \ge t_0 + (\frac{\alpha-1}{\alpha})r$ we have that

$$0 \le V(t) + V\big(t - (\frac{\alpha-1}{\alpha})r\big) \le 2V\big(t - (\frac{\alpha-1}{\alpha})r\big).$$

In summary, (8.6.11), (8.6.21), and (8.6.22) imply that

$$\begin{aligned}
&V(t) + V\big(t - (\frac{\alpha-1}{\alpha})r\big) \ge \Big(x(t) - \int_{t-r}^{t} A(t,s)g(x(s))ds\Big)^2 \\
&+ \int_{-r}^{0}\int_{t+s}^{t} A^2(t,z)g^2(x(z))dz\,ds + (\frac{\alpha-1}{\alpha})r\int_{t-r}^{t-\frac{r}{\alpha}} A^2(t,z)g^2(x(z))dz \\
&\ge \Big(x(t) - \int_{t-r}^{t} A(t,s)g(x(s))ds\Big)^2 + (\frac{\alpha-1}{\alpha})r\int_{t-\frac{r}{\alpha}}^{t} A^2(t,z)g^2(x(z))dz \\
&+ (\frac{\alpha-1}{\alpha})r\int_{t-r}^{t-\frac{r}{\alpha}} A^2(t,z)g^2(x(z))dz \\
&= \Big(x(t) - \int_{t-r}^{t} A(t,s)g(x(s))ds\Big)^2 + (\frac{\alpha-1}{\alpha})r\int_{t-r}^{t} A^2(t,z)g^2(x(z))dz \\
&\ge \Big(x(t) - \int_{t-r}^{t} A(t,s)g(x(s))ds\Big)^2 \\
&+ (\frac{\alpha-1}{\alpha})\Big(\int_{t-r}^{t} A(t,s)g(x(s))ds\Big)^2 \text{ (by Hölder's inequality)} \\
&= \frac{\frac{\alpha-1}{\alpha}}{1+\frac{\alpha-1}{\alpha}}x^2(t) + \Big[\frac{1}{\sqrt{1+\frac{\alpha-1}{\alpha}}}x(t) - \sqrt{1+\frac{\alpha-1}{\alpha}}\int_{t-r}^{t} A(t,s)g(x(s))ds\Big]^2 \\
&\ge \frac{\frac{\alpha-1}{\alpha}}{1+\frac{\alpha-1}{\alpha}}x^2(t).
\end{aligned} \tag{8.6.23}$$

Thus (8.6.23) shows that

$$\begin{aligned}
\frac{\frac{\alpha-1}{\alpha}}{1+\frac{\alpha-1}{\alpha}}x^2(t) &\le V(t) + V\big(t - (\frac{\alpha-1}{\alpha})r\big) \\
&\le 2V\big(t - (\frac{\alpha-1}{\alpha})r\big).
\end{aligned}$$

An integration of (8.6.17) from t_0 to t yields the inequality

$$V(t) \le V(t_0)e^{-\int_{t_0}^{t} A(s,s)ds}.$$

As a consequence,

$$V\big(t - (\frac{\alpha-1}{\alpha})r\big) \le V(t_0)e^{-\int_{t_0}^{t-(\frac{\alpha-1}{\alpha})r} A(s,s)ds},$$

Copyright Elsevier 2022

and

$$|x(t)| \le \sqrt{2\frac{1+\frac{\alpha-1}{\alpha}}{\frac{\alpha-1}{\alpha}} V(t_0) e^{-\frac{1}{2}\int_{t_0}^{t-(\frac{\alpha-1}{\alpha})r} A(s,s)ds}}$$

for $t \ge t_0 + (\dfrac{\alpha-1}{\alpha})r$. This completes the proof. ☐

Remark 8.2. It is clear that inequality (8.6.18) implies that the zero solution of (8.6.1) is asymptotically stable, provided that

$$\int^{\infty} A(t,t)dt \to \infty,$$

and exponentially stable if

$$\int_{t_0}^{t-(\frac{\alpha-1}{\alpha})r} A(s,s)ds \ge \beta(t-t_0)$$

for all $t \ge t_0 + (\dfrac{\alpha-1}{\alpha})r$ and a constant $\beta > 0$.

As an example, we take $g(x) = x(\dfrac{\sin^2(x)}{4} + 1)$. Then $g(0) = 0$, $xg(x) > x^2$, and $|g(x)| \le \dfrac{5}{4}|x|$. Choose $a(t,s) = \dfrac{1}{2}$ and $r = \dfrac{1}{3}$. Next, we make sure that (8.6.7) is satisfied:

$$\int_{-r}^{0}\int_{t+s}^{t} A(t,z)\frac{\partial A(t,z)}{\partial t} g^2(x(z))dz\,ds = -a^2\int_{t-r}^{t}(z-t+r)^2 g^2(x(z))dz \le 0.$$

Thus we have shown that the zero solution of the nonlinear Volterra integro-differential equation

$$x'(t) = -\frac{1}{2}\int_{t-\frac{1}{3}}^{t} x(s)(\frac{\sin^2(x(s))}{4} + 1)ds$$

is exponentially stable.

8.6.1 Instability

We turn our attention to the instability of the zero solution of (8.6.1). We start with the following lemma.

Lemma 8.3. *Suppose (8.6.3), (8.6.4), and (8.6.5) hold and there is a positive constant $D > r$ such that*

$$\frac{r}{r-D} \le A(t,t) \le -A^2(t,t)\big[(1+\lambda)^2 + D\lambda^2\big]. \qquad (8.6.24)$$

Copyright Elsevier 2022

If

$$V(t)=\Big(x(t)-\int_{t-r}^{t}A(t,s)g(x(s))ds\Big)^2$$
$$-D\int_{t-r}^{t}A^2(t,z)g^2(x(z))dz, \tag{8.6.25}$$

then along the solutions of (8.6.1) we have

$$V'(t)\geq -A(t,t)V(t).$$

Proof. Let $x(t)=x(t,t_0,\psi)$ be a solution of (8.6.1) and define $V(t)$ by (8.6.25). Then along solutions of (8.6.1) we have

$$\begin{aligned}
V'(t)&=2\Big(x(t)-\int_{t-r}^{t}A(t,s)g(x(s))ds\Big)[-A(t,t)g(x(t))]\\
&\quad -DA^2(t,t)g^2(x(t))-D\int_{t-r}^{t}2A(t,z)\frac{\partial A(t,z)}{\partial t}g^2(x(z))dz\\
&\geq 2\Big(x(t)-\int_{t-r}^{t}A(t,s)g(x(s))ds\Big)[-A(t,t)g(x(t))]-DA^2(t,t)g^2(x(t))\\
&=-A(t,t)V(t)\\
&\quad -A(t,t)\Big[-\big(\int_{t-r}^{t}A(t,s)g(x(s))ds\big)^2+D\int_{t-r}^{t}A^2(t,z)g^2(x(z))dz\Big]\\
&\quad -2A(t,t)[x(t)-g(x(t))]\int_{t-r}^{t}A(t,s)g(x(s))ds\\
&\quad +\big(-A(t,t)-\lambda^2DA^2(t,t)\big)x^2(t).
\end{aligned}\tag{8.6.26}$$

First, we remark that, as a consequence of (8.6.5), we have

$$-D\int_{t-r}^{t}A(t,z)\frac{\partial A(t,z)}{\partial t}g^2(x(z))dz\geq 0.$$

We note that

$$\begin{aligned}
&2A(t,t)[g(x(t))-x(t)]\int_{t-r}^{t}A(t,s)g(x(s))ds\\
&\leq 2|A(t,t)|(|g(x)|+|x|)|\int_{t-r}^{t}A(t,s)g(x(s))ds|\\
&\leq 2|A(t,t)||x|(1+\lambda)|\int_{t-r}^{t}A(t,s)g(x(s))ds|
\end{aligned}$$

Copyright Elsevier 2022

$$\leq A^2(t,t)(1+\lambda)^2 x^2(t) + \Big(\int_{t-r}^{t} A(t,s)g(x(s))ds\Big)^2$$
$$\leq A^2(t,t)(1+\lambda)^2 x^2(t) + r\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds.$$

Similarly,

$$A(t,t)\Big(\int_{t-r}^{t} A(t,s)g(x(s))ds\Big)^2 \geq rA(t,t)\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds.$$

Hence (8.6.26) reduces to

$$\begin{aligned}V'(t) &\geq -A(t,t)V(t) - \{A(t,t) + (1+\lambda)^2A^2(t,t) + D\lambda^2A^2(t,t)\}x^2(t)\\ &+ \big[A(t,t)(r-D) - r\big]\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds\\ &\geq -A(t,t)V(t). \end{aligned} \tag{8.6.27}$$

This completes the proof. □

Theorem 8.6.2. *Suppose hypotheses of Lemma 8.3 hold. Then the zero solution of* (8.6.1) *is unstable, provided that*

$$-\int_{t_0}^{\infty} A(s,s)\,ds = \infty.$$

Proof. An integration of (8.6.27) from t_0 to t yields

$$V(t) \geq V(t_0)e^{-\int_{t_0}^{t} A(s,s)\,ds}. \tag{8.6.28}$$

Let $V(t)$ be given by (8.6.25). Then

$$\begin{aligned}V(t) &= x^2(t) - 2x(t)\int_{t-r}^{t} A(t,s)g(x(s))ds + \Big[\int_{t-r}^{t} A(t,s)g(x(s))ds\Big]^2\\ &- D\int_{t-r}^{t} A^2(t,z)g^2(x(z))dz. \end{aligned} \tag{8.6.29}$$

Let $\beta = D - r$. Then from

$$\Big(\frac{\sqrt{r}}{\sqrt{\beta}}a - \frac{\sqrt{\beta}}{\sqrt{r}}b\Big)^2 \geq 0$$

we have

$$2ab \leq \frac{r}{\beta}a^2 + \frac{\beta}{r}b^2.$$

Copyright Elsevier 2022

With this in mind, we arrive at

$$\begin{aligned}-2x(t)\int_{t-r}^{t} A(t,s)g(x(s))ds &\leq 2|x(t)||\int_{t-r}^{t} A(t,s)g(x(s))ds| \\ &\leq \frac{r}{\beta}x^2(t)+\frac{\beta}{r}\Big[\int_{t-r}^{t} A(t,s)g(x(s))ds\Big]^2 \\ &\leq \frac{r}{\beta}x^2(t)+\frac{\beta}{r}r\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds.\end{aligned}$$

A substitution of the above inequality into (8.6.29) yields

$$\begin{aligned}V(t) &\leq x^2(t)+\frac{r}{\beta}x^2(t)+(\beta+r-D)\int_{t-r}^{t} A^2(t,s)g^2(x(s))ds \\ &= \frac{\beta+r}{\beta}x^2(t) \\ &= \frac{D}{D-r}x^2(t).\end{aligned}$$

Using inequality (8.6.28), we get

$$\begin{aligned}|x(t)| &\geq \sqrt{\frac{D-r}{D}}\,V^{1/2}(t) \\ &= \sqrt{\frac{D-r}{D}}\,V^{1/2}(t_0)e^{-\frac{1}{2}\int_{t_0}^{t} A(s,s)\,ds}.\end{aligned}$$

This completes the proof. □

As an example, we take $g(x)=x(\frac{\sin^2(x)}{100}+1)$. Then $g(0)=0$, $xg(x)>x^2$, and $|g(x)|\leq \frac{101}{100}|x|$. Choose $a(t,s)=-\frac{1}{2}$, $r=\frac{1}{8}$, and $D=\frac{1}{4}$. Then (8.6.24) is satisfied. Also, it is clear that (8.6.7) holds.

Thus we have shown that the nonlinear Volterra integro-differential equation

$$x'(t)=\frac{1}{2}\int_{t-\frac{1}{8}}^{t} x(s)(\frac{\sin^2(x(s))}{100}+1)ds$$

is unstable.

We end this section with the following open problem.

Open Problem. In light of this research, what can be said about the exponential stability and instability of the zero solution of the nonlinear Volterra integro-differential equation with infinite delay

$$x'(t)=-\int_{-\infty}^{t} a(t,s)g(x(s))ds?$$

Copyright Elsevier 2022

8.7 Existence of positive periodic solutions

We consider the nonlinear neutral differential equation with functional delay

$$x'(t) = -a(t)x(t) + c(t)x'(t - g(t)) + q(t, x(t - g(t))), \tag{8.7.1}$$

which arises in food-limited population models. For system (8.7.1), there may be a stable equilibrium point of the population. In the case the equilibrium point becomes unstable, there may exist a nontrivial periodic solution. Then the oscillation of solutions occurs. The existence of such a stable periodic solution is of quite fundamental importance biologically since it concerns the long-time survival of species. The study of such phenomena has become an essential part of the qualitative theory of differential equations. One of the most used models, a prototype of (8.7.1), is the system of Volterra integro-differential equations

$$\dot{N}(t) = -\gamma(t)N(t) + \alpha(t)\int_0^\infty B(s)e^{-\beta(t)N(t-s)}ds,$$

where $N(t)$ is the number of red blood cell at time t, $\alpha, \beta, \gamma \in C(\mathbb{R}, \mathbb{R})$ are T-periodic, and $B \in L^1(\mathbb{R}^+)$ is piecewise continuous. This is a generalized model of the red cell system introduced by Wazewska-Czyzewska and Lasota

$$\dot{n}(t) = -\gamma n(t) + \alpha e^{-\beta n(t-r)},$$

where α, β, γ, and r are constants with $r > 0$. The existence of positive periodic solutions was established for the neutral logistic equation with distributed delays

$$x'(t) = x(t)\Big[a(t) - \sum_{i=1}^{n} a_i(t)\int_{-T_i}^{0} x(t+\theta)\,\mathrm{d}\mu_i(\theta) - \sum_{j=1}^{m} b_j(t)\int_{-\hat{T}_j}^{0} x'(t+\theta)\,\mathrm{d}\nu_j(\theta)\Big], \tag{8.7.2}$$

where the coefficients a, a_i, and b_j are continuous periodic functions with the same period. The values $T_i, \hat{T}_j$ are positive, and the functions μ_i, ν_j are nondecreasing with $\int_{-T_i}^{0} \mathrm{d}\mu_i = 1$ and $\int_{-\hat{T}_j}^{0} \mathrm{d}\nu_j = 1$. Eq. (8.7.2) is of logistic form, and hence the method used to obtain the existence of positive periodic solutions will not work for our model (8.7.1). For example, in the above equation the transformation $x(t) = e^{N(t)}$ was used to put (8.7.2) in the form

$$\begin{aligned} N'(t) = a(t) - \sum_{i=1}^{n} a_i(t)\int_{-T_i}^{0} e^{N(t+\theta)}\,\mathrm{d}\mu_i(\theta) \\ - \sum_{j=1}^{m} b_j(t)\int_{-\hat{T}_j}^{0} N'(t+\theta)e^{N(t+\theta)}\,\mathrm{d}\nu_j(\theta). \end{aligned} \tag{8.7.3}$$

Eq. (8.7.1) represents a generalization of the hematopoiesis and blood cell production models.

Copyright Elsevier 2022

The Krasnoselskii fixed point theorem has been extensively used in differential and functional differential equations by Burton [16] to prove the existence of periodic solutions. Also, Burton was the first to use the theorem to obtain stability results regarding solutions of integral equations and functional differential equations. For a collection of different type of results, we refer the reader to [16] and the references therein. The author is unaware of any results regarding the use of Krasnoselskii to prove the existence of a positive periodic solution.

Theorem 8.7.1. *(Krasnoselskii) Let* $\mathbb{M}$ *be a closed convex nonempty subset of a Banach space* $\big(\mathbb{B}, \|\cdot\|\big)$. *Suppose that* A *and* B *map* $\mathbb{M}$ *into* $\mathbb{B}$ *and satisfy the following conditions:*

- **(i)** A *is compact and continuous,*
- **(ii)** B *is a contraction mapping,*
- **(iii)** *If* $x, y \in \mathbb{M}$, *then* $Ax + By \in \mathbb{M}$.

Then there exists $z \in \mathbb{M}$ *such that* $z = Az + Bz$.

For $T > 0$, define $P_T = \{\phi \in C(\mathbb{R}, \mathbb{R}), \phi(t+T) = \phi(t)\}$, where $C(\mathbb{R}, \mathbb{R})$ is the space of all real-valued continuous functions. Then P_T is a Banach space when endowed with the supremum norm

$$\|x\| = \max_{t\in[0,T]} |x(t)| = \max_{t\in\mathbb{R}} |x(t)|.$$

We assume that

$$a(t+T) = a(t), \;\; c(t+T) = c(t), \;\; g(t+T) = g(t), \;\; g(t) \geq g^* > 0 \tag{8.7.4}$$

with continuously differentiable $c(t)$, twice continuously differentiable $g(t)$, and constant g^*. Usually, $a(t)$ is assumed to be positive, but here we only require that

$$\int_0^T a(s)ds > 0. \tag{8.7.5}$$

It is interesting to note that Eq. (8.7.1) becomes of advanced type when $g(t) < 0$. Since we are searching for periodic solutions, it is natural to require that $q(t, x)$ is continuous in both arguments and periodic in t. Also, we assume that for all $0 \leq t \leq T$,

$$g'(t) \neq 1. \tag{8.7.6}$$

Lemma 8.4. *Suppose* (8.7.4)–(8.7.6) *hold. If* $x(t) \in P_T$, *then* $x(t)$ *is a solution of Eq.* (8.7.1) *if and only if*

$$\begin{aligned} x(t) = {} & \frac{c(t)}{1-g'(t)} x(t-g(t)) \\ & + \int_t^{t+T} \Big[-r(u)x(u-g(u)) + q(u, x(u-g(u)))\Big] \frac{e^{\int_t^u a(s)ds}}{e^{\int_0^T a(s)ds} - 1}\, du, \end{aligned} \tag{8.7.7}$$

Copyright Elsevier 2022

where

$$r(t) = \frac{\Big(c'(t) + c(t)a(t)\Big)\Big(1 - g'(t)\Big) + g''(t)c(t)}{(1 - g'(t))^2}. \tag{8.7.8}$$

Proof. Let $x(t) \in P_T$ be a solution of (8.7.1). Multiply both sides of (8.7.1) with $e^{\int_0^t a(s)ds}$ and then integrate from t to $t + T$ to obtain

$$\int_t^{t+T} \Big[x(u)e^{\int_0^u a(s)ds}\Big]' du = \int_t^{t+T} \Big[c(u)x'(u - g(u)) \\ + q(u, x(u), x(u - g(u)))\Big] e^{\int_0^u a(s)ds} du.$$

As a consequence, we arrive at

$$x(t + T)e^{\int_0^{t+T} a(s)ds} - x(t)e^{\int_0^t a(s)ds} \\ = \int_t^{t+T} \Big[c(u)x'(u - g(u)) + q(u, x(u - g(u)))\Big] e^{\int_0^u a(s)ds} du.$$

By dividing both sides of the above equation by $e^{\int_0^{t+T} a(s)ds}$ and the fact that $x(t + T) = x(t)$ we obtain

$$x(t) = \Big(1 - e^{-\int_t^{t+T} a(s)ds}\Big)^{-1} \int_t^{t+T} \Big[c(u)x'(u - g(u)) \\ + q(u, x(u), x(u - g(u)))\Big] e^{-\int_u^{t+T} a(s)ds} du. \tag{8.7.9}$$

Rewrite

$$\int_t^{t+T} c(u)x'(u - g(u))e^{-\int_u^{t+T} a(s)ds} du \\ = \int_t^{t+T} \frac{c(u)x'(u - g(u))(1 - g'(u))}{(1 - g'(u))} e^{-\int_u^{t+T} a(s)ds} du.$$

Integrating by parts the above integral with

$$U = \frac{c(u)}{1 - g'(u)} e^{-\int_u^{t+T} a(s)ds}$$

and

$$dV = x'(u - g(u))(1 - g'(u))du,$$

Copyright Elsevier 2022

we obtain

$$\int_t^{t+T} c(u)x'(u-g(u))e^{-\int_u^{t+T} a(s)ds}du = \frac{c(t)}{1-g'(t)}x(t-g(t))\Big(1-e^{-\int_t^{t+T} a(s)ds}\Big)$$
$$-\int_t^{t+T} r(u)e^{-\int_u^{t+T} a(s)ds}x(u-g(u))du,$$

where $r(u)$ is given by (8.7.8). Finally, due to the integration over one period and the periodicity of all functions, we have that

$$\frac{e^{-\int_u^{t+T} a(s)ds}}{1-e^{-\int_t^{t+T} a(s)ds}} = \frac{e^{-\int_u^{t+T} a(s)ds}}{e^{-\int_t^{t+T} a(s)ds}(e^{\int_t^{t+T} a(s)ds}-1)}$$
$$= \frac{e^{\int_t^u a(s)ds}}{e^{\int_t^{t+T} a(s)ds}-1} = \frac{e^{\int_t^u a(s)ds}}{e^{\int_0^T a(s)ds}-1}.$$

This completes the proof. □

To simplify notation, we let

$$M = \frac{e^{\int_0^{2T} |a(s)|ds}}{e^{\int_0^T a(s)ds}-1} \tag{8.7.10}$$

and

$$m = \frac{e^{-\int_0^{2T} |a(s)|ds}}{e^{\int_0^T a(s)ds}-1}. \tag{8.7.11}$$

Let

$$G(t,u) = \frac{e^{\int_u^t a(s)ds}}{e^{\int_0^T a(s)ds}-1}. \tag{8.7.12}$$

It is easy to see that for all $(t,u) \in [0,2T] \times [0,2T]$,

$$m \le G(t,u) \le M,$$

and for all $t, u \in \mathbb{R}$, we have

$$G(t+T, u+T) = G(t,u).$$

We obtain the existence of a positive periodic solution by considering two cases:

(a) $0 \le \dfrac{c(t)}{1-g'(t)} < 1,$

and

(b) $-1 \le \dfrac{c(t)}{1-g'(t)} \le 0.$

Copyright Elsevier 2022

For some nonnegative constant L and positive constant K, we define the set

$$\mathbb{M} = \{\phi \in P_T : L \le ||\phi|| \le K\},$$

which is a closed convex bounded subset of the Banach space P_T. In addition, we assume that there are constants $0 \le \beta \le \alpha < 1$ such that

$$0 \le \beta \le \frac{c(t)}{1 - g'(t)} \le \alpha < 1, \tag{8.7.13}$$

and for all $u \in \mathbb{R}$ and $\rho \in \mathbb{M}$,

$$\frac{(1-\beta)L}{mT} \le q(s,\rho) - r(s)\rho \le \frac{(1-\alpha)K}{MT}, \tag{8.7.14}$$

where M and m are defined by (8.7.10) and (8.7.11), respectively. To apply Theorem 8.7.1, we need to construct two mappings: one is a contraction, and the other is compact. Thus we define the map $\mathbf{A} : \mathbb{M} \to P_T$ by

$$(\mathbf{A}\varphi)(t) = \int_t^{t+T} G(t,s)[q(s,\varphi(s-g(s))) - r(s)\varphi(s-g(s))]\,ds, t \in \mathbb{R}. \tag{8.7.15}$$

In a similar way, we set the map $\mathbf{B} : \mathbb{M} \to P_T$ by

$$(\mathbf{B}\varphi)(t) = \frac{c(t)}{1-g'(t)}\varphi(t-g(t)), t \in \mathbb{R}. \tag{8.7.16}$$

It is clear from condition (8.7.13) that $\mathbf{B}$ defines a contraction mapping under the supremum norm.

Lemma 8.5. *If* (8.7.4)–(8.7.6)*,* (8.7.13)*, and* (8.7.14) *hold, then the operator* $\mathbf{A}$ *is completely continuous on* $\mathbb{M}$*.*

Proof. For $t \in [0,T]$, which implies that $u \in [t, t+T] \subseteq [0, 2T]$, and for $\varphi \in \mathbb{M}$, we have by (8.7.13) that

$$\begin{aligned}|(\mathbf{A}\varphi)(t)| &\le \| \int_t^{t+T} G(t,s)[q(s,\varphi(s-g(s))) - r(s)\varphi(s-g(s))]\,ds\| \\ &\le TM\frac{(1-\alpha)K}{MT}.\end{aligned}$$

From the estimate of $|\mathbf{A}\varphi(t)|$ it follows that

$$||\mathbf{A}\varphi(t)|| \le (1 - \frac{c(t)}{1-g'(t)})K \le Q_1$$

for some positive constant Q_1. This shows that $\mathbf{A}(\mathbb{M})$ is uniformly bounded. It remains to show that $\mathbf{A}(\mathbb{M})$ is equicontinuous. Let $\varphi \in \mathbb{M}$. Then a differentiation of

Copyright Elsevier 2022

(8.7.13) with respect to t yields

$$(\mathbf{A}\varphi)'(t) = G(t,t+T)[q(t,\varphi(t-g(t))) - r(t)\varphi(t-g(t))] + a(t)(\mathbf{A}\varphi)(t).$$

Hence by taking the supremum norm in the above expression we have

$$||(\mathbf{A}\varphi)'|| \le \frac{Q_1}{T} + ||a(t)||Q_1.$$

Thus the estimation on $|(\mathbf{A}\varphi)'(t)|$ implies that $\mathbf{A}(\mathbb{M})$ is equicontinuous. Then by the Ascoli–Arzelà theorem, we obtain that A is a compact map. Due to the continuity of all terms in (8.7.13) for $t \in [0,T]$, we have that $\mathbf{A}$ is continuous. This completes the proof. □

Theorem 8.7.2. *If* (8.7.4)–(8.7.6), (8.7.13), *and* (8.7.14) *hold, then Eq.* (8.7.1) *has a positive periodic solution* z *satisfying* $L \le z \le K$.

Proof. Let $\varphi, \psi \in \mathbb{M}$. Then by (8.7.13) and (8.7.16) we have that

$$\begin{aligned}(\mathbf{B}\varphi)(t) + (\mathbf{A}\psi)(t) &= \frac{c(t)}{1-g'(t)}\varphi(t-g(t)) \\ &\quad + \int_t^{t+T} G(t,s)[q(s,\varphi(s-g(s))) - r(s)\varphi(s-g(s))]\,ds \\ &\le \alpha K + MT\frac{(1-\alpha)K}{MT} = K.\end{aligned}$$

On the other hand,

$$\begin{aligned}(\mathbf{B}\varphi)(t) + (\mathbf{A}\psi)(t) &= \frac{c(t)}{1-g'(t)}\varphi(t-g(t)) \\ &\quad + \int_t^{t+T} G(t,s)[q(s,\varphi(s-g(s))) - r(s)\varphi(s-g(s))]\,ds \\ &\ge \beta L + m\int_t^{t+T}[q(s,\varphi(s-g(s))) - r(s)\varphi(s-g(s))]\,ds \\ &\ge \beta L + mT\frac{(1-\beta)L}{mT} = L.\end{aligned}$$

This shows that $\mathbf{B}\varphi + \mathbf{A}\psi \in \mathbb{M}$. All the hypotheses of Theorem 8.7.1 are satisfied, and therefore Eq. (8.7.1) has a periodic solution, say z, residing in $\mathbb{M}$. This completes the proof. □

For the next theorem, we assume that there are constants $-1 < \beta \le \alpha \le 0$ such that

$$-1 < \beta \le \frac{c(t)}{1-g'(t)} \le \alpha \le 0, \tag{8.7.17}$$

Copyright Elsevier 2022

and for all $u \in \mathbb{R}$ and $\rho \in \mathbb{M}$,

$$\frac{(L-\beta K)}{mT} \le q(s,\rho) - r(s)\rho \le \frac{(K-\alpha L)}{MT}, \tag{8.7.18}$$

where M and m are defined by (8.7.10) and (8.7.11), respectively.

Theorem 8.7.3. *If* (8.7.4)–(8.7.6), (8.7.17), *and* (8.7.18) *hold, then Eq.* (8.7.1) *has a positive periodic solution z satisfying $L \le z \le K$.*

Proof. The proof follows along the lines of Theorem 8.7.2, and hence we omit it. ☐

Example 8.5. The neutral differential equation

$$x'(t) = -\frac{1}{2}\sin^2(t)x(t) + \frac{1}{50}x'(t-\pi) + \frac{\cos^2(t)}{x^2(t-\pi)+100} + \frac{1}{25} \tag{8.7.19}$$

has a positive π-periodic solution x satisfying

$$\frac{1}{10} \le x \le 2.$$

To see this, we let

$$q(s,\rho) = \frac{\cos^2(s)}{\rho^2+100} + \frac{1}{25}, \ r(s) = \frac{1}{2}\sin^2(s), \quad \text{and} \quad T = g(t) = \pi.$$

Then

$$\frac{c(t)}{1-g'(t)} = \frac{1}{50} < 1,$$

and

$$r(t) = \frac{1}{100}\sin^2(t).$$

A simple calculation yields

$$4.030 < M < 4.032 \ \text{ and } \ 1.74 < m < 1.75.$$

Let $K=2$ and $L=\frac{1}{10}$, and define the set $\mathbb{M} = \{\frac{1}{10} \le \upsilon \le 2\}$. Then for $\rho \in [\frac{1}{10}, 2]$, we have

$$\begin{aligned} q(s,\rho) - r(s)\rho &= \frac{\cos^2(s)}{\sigma^2+100} + \frac{1}{100}\sin^2(s)\rho + \frac{1}{25} \\ &\le \frac{1}{100} + \frac{1}{50} + \frac{1}{25} = 0.07 < \frac{(1-\frac{c(t)}{1-g'(t)})K}{MT}. \end{aligned}$$

Copyright Elsevier 2022

On the other hand,

$$q(s,\rho) - r(s)\rho = \frac{\cos^2(u)}{\sigma^2+100} + \frac{1}{100}\sin^2(u)\rho + \frac{1}{25}$$
$$> \frac{1}{25} > \frac{(1-\frac{c(t)}{1-g'(t)})L}{mT}.$$

We see that all the conditions of Theorem 8.7.2 are satisfied, and hence (8.7.19) has a positive π-periodic solution x satisfying $\frac{1}{10} \le x \le 2$.

8.8 Exercises

Exercise 8.1. Use the method of steps to find the solution of the delay differential equation with delay 1,

$$x' = -tx(t-1),\ t > 0;\ x(s) = \varphi(s) = 12,\ \ -1 \le s \le 0,$$

on the interval [0, 3] by considering the subintervals [0, 1], [1, 2], and [2, 3].

Exercise 8.2. Consider the delay differential equation with delay 1,

$$x' = ax(t) + bx(t-1),\ t > 0,$$

with initial function

$$x(s) = \varphi(s) = 1 + t,\ \ -1 \le s \le 0,$$

where a and $b \neq 0$ are constants. Use the method of steps to find the solution on the interval $[0, 2]$ by considering the subintervals $[0, 1]$ and $[1, 2]$. Is the solution differentiable at the initial time $t_0 = 0$? What about the differentiability at $t = 1$?

Exercise 8.3. Consider the delay differential equation with delay 1,

$$x' = x(t) + x(t-1),\ t > 0,$$

with initial function

$$x(s) = \varphi(s) = t,\ \ -1 \le s \le 0.$$

Use the method of steps to find the solution on the interval $[0, 3]$. Is the solution differentiable at the initial time $t_0 = 0$? What about the differentiability at $t = 1$?

Exercise 8.4. Solve the Bernoulli equation (8.2.5).

Exercise 8.5. Repeat Example 8.3 for the delay second-order differential equation

$$mu''(t) + bu'(t) + qu'(t-r) + cu(t) = 0,\ t \ge 0,$$

where m, b, q, c, and r are positive constants.

Copyright Elsevier 2022

Exercise 8.6. Consider the delay second-order differential equation

$$u''(t)+au'(t)+bu'(t-r)=0,\ t\geq 0,$$

where a, b, and r are positive constants. Write the delay equation into a system in x and y and use the Lyapunov-type functional

$$V(x_t, y_t)=y^2+bx^2+\gamma\int_{-r}^{0}\int_{t+s}^{t} y^2(u)du\,ds,\quad \gamma>0.$$

(i) Show that if

$$\gamma=b \text{ and } -a+br<0,$$

then the zero solution is uniformly stable.
(ii) Choose

$$\gamma>b \text{ and } \frac{-2a}{r}+\gamma+b<0$$

and show that the zero solution is asymptotically stable and uniformly asymptotically stable.

Exercise 8.7. Give an example of application of Theorem 8.7.3.

Exercise 8.8. Consider the delay differential equation

$$x'(t)=-ax(t)+bx(t-r),\ t\geq t_0\geq 0,$$

where a, b, and r are positive constants. Consider the Lyapunov-type functional

$$V(x_t)=x^2(t)+b\int_{t-r}^{t} x^2(s)ds.$$

(i) Show that if $a\geq b$, then all solutions are bounded, and the zero solution is uniformly stable.
(ii) Show that if $a>b$, then the zero solution is uniformly asymptotically stable, and $x^2(t)\in L^1[0,\infty)$.
(iii) Use part (i) to show that all solutions of

$$x'(t)=-2x(t)+x(t-\tau),\ x(t)=\varphi(s),\ -\tau\leq s\leq 0,\ t\geq 0,$$

are bounded.
(iv) Replace the delay term $x(t-\tau)$ in the delay differential equation given in (iii) with

$$x(t)-\tau x'(t)+\frac{1}{2}\tau^2 x''$$

and show that the solutions of the resulting equation are unbounded regardless of the size of the delay τ.

Copyright Elsevier 2022

Exercise 8.9. Consider the finite delay differential equation

$$x'(t) = -3x(t) + \frac{1}{e}x(t-1), \ t \geq 0, \tag{8.8.1}$$

with $x(t) = \varphi(s)$, $-1 \leq s \leq 0$.
(a) Use Theorem 8.4.3 to show that all solutions of (8.8.1) are bounded and its zero solution is exponentially stable.
(b) Replace the delay term $x(t-1)$ in (8.8.1) with $x(t) - x'(t)$. Are the solutions of the resulting equation bounded?
(c) Now replace the delay term $x(t-1)$ in (8.8.1) with $x(t) - x'(t) + \frac{1}{2}x''$ and show that the solutions of the resulting equation are unbounded.

Exercise 8.10. Construct an example that satisfies the hypotheses of Theorem 8.5.5.

Exercise 8.11. Let $h > 0$ and consider the finite delay differential equation

$$x'(t) = a(t)x(t) + b(t)x(t-h) + g(t), \ t \geq 0, \tag{8.8.2}$$

with $x(t_0) = \varphi$. Suppose $a(t)$, $b(t)$, and $g(t)$ are continuous for all $t \geq 0$. For some $\tau > 0$, define the continuous function

$$\xi(t) = \frac{e^{\int_0^t c(s)ds}}{1 + 2h\int_t^{t+\tau} e^{\int_0^u c(s)ds} du} \geq 0 \tag{8.8.3}$$

with

$$|b(t)| \leq h\xi(t), \tag{8.8.4}$$

where $c(t) := 2a(t) + |g(t)|$. Use the functional

$$V(t, x_t) = x^2(t) + h\xi(t)\int_{t-h}^{t} x^2(s)\, ds$$

to show that every solution of (8.8.2) with $x(t_0) = \varphi$ satisfies the inequality

$$||x(t)|| \leq \left[V(t_0, \varphi) + \int_{t_0}^{t} |g(u)|\, e^{\int_{t_0}^{u} \alpha(s)ds}\, du\right]^{1/2} e^{-\frac{1}{2}\int_{t_0}^{t} \alpha(s)ds},$$

where $\alpha(t) = c(t) + 2h\xi(t)$ and $V(t_0, \varphi) = \varphi^2(0) + h\xi(t_0)\displaystyle\int_{t_0-h}^{t_0} \varphi^2(s)ds$.

Exercise 8.12. (Big project) Develop similar results as in Section 8.6 for the totally delayed differential equation

$$x' = ax(t-r), \ t > 0,$$

where $a \in \mathbb{R}$, and r is a positive constant.

Copyright Elsevier 2022

CHAPTER

New variation of parameters

9

In this chapter we introduce *new variation of parameters formula* and apply the idea to ordinary and delay differential equations. The results of this chapter are totally new and should serve as foundations for future research.

9.1 Applications to ordinary differential equations

It is customary in a nonlinear differential equation to add and subtract a convenient term that allows us to invert the equation in question and obtain a variation of parameters formula, which can be used to obtain different results on the solutions. However, the added term will cause restrictions on the coefficients, and as a result, limit the class of equations that can be considered. In this section we adopt the terminologies of Chapter 2 with respect to definitions of solutions and other important matters. We just try to avoid redundancies and repetitions. Consider the simplest ordinary differential equation

$$x'(t) = ax(t), \quad x(0) = x_0, \tag{9.1.1}$$

which has the solution

$$x(t) = x_0 e^{at} \to 0 \text{ as } t \to \infty,$$

provided that

$$a < 0.$$

Suppose $a : \mathbb{R} \to \mathbb{R}$ is continuous and consider the ordinary differential equation with variable coefficient

$$x'(t) = a(t)x(t), \quad x(0) = x_0, \tag{9.1.2}$$

which has the solution

$$x(t) = x_0 e^{\int_0^t a(s)ds} \to 0 \text{ as } t \to \infty,$$

provided that

$$\int_0^t a(s)ds \to -\infty. \tag{9.1.3}$$

Advanced Differential Equations. https://doi.org/10.1016/B978-0-32-399280-0.00015-2

Copyright © 2023 Elsevier Inc. All rights reserved.

Copyright Elsevier 2022

Condition (9.1.3) implies that the function $a(t)$ can be positive or oscillates for a short time. Now let

$$v : [0, \infty) \to \mathbb{R}$$

be a continuous function. Multiply both sides of (9.1.1) by

$$e^{\int_0^t v(s)ds}$$

and then integrate from 0 to any $t \in [0, T)$:

$$\int_0^t e^{\int_0^u v(s)ds}\, x'(u)du = \int_0^t ax(u)e^{\int_0^u v(s)ds}du.$$

Integrating by parts the left side and simplifying, we get

$$x(t) = x_0 e^{-\int_0^t v(s)ds} + \int_0^t x(u)\big(v(u) + a\big)e^{-\int_u^t v(s)ds}du. \tag{9.1.4}$$

Expression (9.1.4) is a *new variation of parameters formula for* (9.1.1) and of Volterra-type integral equation. Note that if

$$v(t) = -a,$$

then (9.1.4) becomes the regular solution $x(t) = x_0 e^{at}$ of (9.1.1). In a similar fashion,

$$x'(t) = a(t)x(t), \ \ x(0) = x_0,$$

has the solution

$$x(t) = x_0 e^{-\int_0^t v(s)ds} + \int_0^t x(u)\big(v(u) + a(u)\big)e^{-\int_u^t v(s)ds}du. \tag{9.1.5}$$

Again, letting

$$v(t) = -a(t),$$

we get the regular known solution $x(t) = x_0 e^{\int_0^t a(s)ds}$, and $x(t) \to 0$ as $t \to \infty$, provided that

$$\int_0^t a(s)ds \to -\infty.$$

Again, (9.1.5) is a new variation of parameters formula that we will analyze when considering nonlinear equations. In the mean time, for (9.1.5), by setting up the proper spaces and using the contraction mapping principle, we can show that $x(t) \to 0$ as $t \to \infty$, provided that

$$\int_0^t \big|v(u) + a(u)\big|e^{-\int_u^t v(s)ds}du \le \alpha, \ 0 < \alpha < 1,$$

Copyright Elsevier 2022

and

$$\int_0^t v(s)ds \to \infty.$$

Suppose $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous and consider the nonlinear differential equation

$$x'(t) = f(t, x(t)),\ x(0) = x_0,\ \text{for a given constant } x_0. \tag{9.1.6}$$

Then multiplying by a function $\int_0^t v(s)ds$ the solution of (9.1.6) is given by

$$x(t) = x_0 e^{-\int_0^t v(s)ds} + \int_0^t \big(x(u)v(u) + f(u, x(u))\big)e^{-\int_u^t v(s)ds}du. \tag{9.1.7}$$

Next, we will use (9.1.7) to define a mapping on the proper space and show that the zero solution is AS. Let $\mathscr{C}$ be the set of all real-valued continuous functions. Define the space

$$\mathscr{S} = \{\phi : [0, \infty) \to \mathbb{R}/\phi \in \mathscr{C},\ |\phi(t)| \le L,\ \phi(t) \to 0 \text{ as } t \to \infty\}.$$

Then

$$(\mathscr{S}, ||\cdot||)$$

is a complete metric space under the uniform metric

$$\rho(\phi_1, \phi_2) = ||\phi_1 - \phi_2||,$$

where

$$||\phi|| = \sup\{|\phi(t)| : t \ge 0\}.$$

Let f be locally Lipschitz on the set $\mathscr{S}$, such that

$$f(t, 0) = 0, \tag{9.1.8}$$

that is, for any ϕ_1 and $\phi_2 \in \mathscr{S}$, we have

$$|f(t, \phi_1) - f(t, \phi_2)| \le \lambda(t)||\phi_1 - \phi_2|| \tag{9.1.9}$$

with continuous $\lambda : [0, \infty) \to (0, \infty)$. Assume that for $\phi \in \mathscr{S}$, we have that

$$|x_0|e^{-\int_0^t v(s)ds} + L\int_0^t \big(|v(u)| + \lambda(u)\big)e^{-\int_u^t v(s)ds}du \le L. \tag{9.1.10}$$

Note that (9.1.10) implies that

$$\int_0^t \big(|v(u)| + \lambda(u)\big)e^{-\int_u^t v(s)ds}du \le \alpha < 1.$$

Copyright Elsevier 2022

Theorem 9.1.1. *Assume* (9.1.8)–(9.1.10). *Suppose there exists a positive constant* k *such that*

$$e^{-\int_0^t v(s)ds} \leq k. \tag{9.1.11}$$

Then the unique solution of (9.1.6) *is bounded, and its zero solution is stable.*

If, in addition,

$$\int_0^t v(s)ds \to \infty, \tag{9.1.12}$$

then the zero solution of (8.5.2) *is asymptotically stable.*

Proof. For $\phi \in \mathscr{S}$, define the mapping $\mathfrak{P} : \mathscr{S} \to \mathscr{S}$ by

$$(\mathfrak{P}\phi)(t) = x_0 e^{-\int_0^t v(s)ds} + \int_0^t \big(\phi(u)v(u) + f(u, \phi(u))\big)e^{-\int_u^t v(s)ds}du. \tag{9.1.13}$$

It is clear that $(\mathfrak{P}\phi)(0) = x_0$. Now for $\phi \in \mathscr{S}$, we have that

$$\big|(\mathfrak{P}\phi)(t)\big| \leq |x_0|k + \int_0^t \big(|\phi(u)||v(u)| + \lambda(u)|\phi(u)|\big)e^{-\int_u^t v(s)ds}du.$$

Consequently,

$$\|\mathfrak{P}\phi\| \leq |x_0|k + \int_0^t \big(|v(u)| + \lambda(u)\big)e^{-\int_u^t v(s)ds}du||\phi||,$$

and

$$\|\mathfrak{P}\phi\| \leq |x_0|k + \alpha\|\phi\| \leq L, \text{ due to (9.1.10) and } \phi \in \mathscr{S}. \tag{9.1.14}$$

Since $\mathfrak{P}$ is continuous, we have that $\mathfrak{P} : \mathscr{S} \to \mathscr{S}$. Next, we show that $\mathfrak{P}$ is a contraction.

For $\phi_1, \phi_2 \in \mathscr{S}$, from (9.1.13) we have that

$$\begin{aligned}\big|(\mathfrak{P}\phi_1)(t) - (\mathfrak{P}\phi_2)(t)\big| &\leq \int_0^t \big(|v(u)| + \lambda(u)\big)e^{-\int_u^t v(s)ds}du||\phi_1 - \phi_2|| \\ &\leq \alpha||\phi_1 - \phi_2||.\end{aligned}$$

This shows that $\mathfrak{P}$ is a contraction. By Banach's contraction mapping principle, $\mathfrak{P}$ has a unique fixed point $x \in \mathscr{S}$, which is a bounded continuous function. Moreover, the unique fixed point is a solution of (9.1.6) on $[0, \infty)$. Let x be the unique solution. Let $\varepsilon > 0$ and choose $\delta = \varepsilon\frac{1-\alpha}{k}$. If $|x_0| < \delta$, then by (9.1.14) we have that

$$(1-\alpha)||x|| \leq |x_0|k < \delta k$$

or

$$||x|| \leq \varepsilon.$$

Copyright Elsevier 2022

The proof of

$$|x(t)| \to 0 \text{ as } t \to \infty$$

rests on condition (9.1.12), and we refer to the proof of Theorem 8.5.2. This completes the proof. □

To see the benefits of our new inversion method, we consider a particular nonlinear equation and rewrite it so that we can invert the usual way. Consequently, the *contraction mapping principle* will no longer work. Let

$$f(t,x) = -x^3 + h(t,x),$$

where $h(t,x)$ satisfies a smallness condition. Thus we consider

$$x' = -x^3 + h(t,x). \tag{9.1.15}$$

Due to the absence of a linear term, we borrow one and try to invert the old way. We write (9.1.15) in the form

$$x' = -x + (x - x^3) + h(t,x) \tag{9.1.16}$$

and then use the variation of parameters formula to obtain

$$x(t) = x_0 e^{-t} + \int_0^t e^{-(t-s)}[x(s) - x^3(s) + h(s,x(s))]ds. \tag{9.1.17}$$

It is naive to believe that most maps can be defined so that they are contractions, even with the strictest conditions. To that effect, consider

$$g(x) = x - x^3.$$

Then for $x, y \in \mathbb{R}$ with $|x|, |y| \leq \frac{\sqrt{3}}{3}$, we have that

$$|g(x) - g(y)| = |x - x^3 - y + y^3| \leq |x - y|\left(1 - \frac{x^2 + y^2}{2}\right),$$

and the contraction constant tends to 1 as $x^2 + y^2 \to 0$. As a consequence, the regular contraction mapping principle failed to produce any results. Using (9.1.7), (9.1.15) has the solution

$$x(t) = x_0 e^{-\int_0^t v(s)ds} + \int_0^t \big(x(u)v(u) - x^3(u) + h(u,x(u))\big)e^{-\int_u^t v(s)ds}du. \tag{9.1.18}$$

Now using the contraction mapping principle, we would have to show that a function of the form

$$f(x) = v(r)x - x^3$$

Copyright Elsevier 2022

is a contraction. This is true on some bounded and small set, provided that v is of sufficiently small magnitude.

Thus using (9.1.18) we can easily obtain stability and boundedness results regarding (9.1.15) under the right conditions.

9.1.1 Periodic solutions

Next, we apply our new method to linear or nonlinear differential equations to show the existence of periodic solutions without the requirement of some classic conditions. To better illustrate our approach, we let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $a : \mathbb{R} \to \mathbb{R}$ be continuous and consider the nonlinear differential equation

$$x' = a(t)x(t) + f(t, x), \tag{9.1.19}$$

where f is continuous in x. For $T \in R$, we assume the periodicity condition

$$a(t+T) = a(t) \text{ and } f(t+T, \cdot) = f(t, \cdot). \tag{9.1.20}$$

Let BC be the space of continuous bounded functions $\phi : \mathbb{R} \to \mathbb{R}$ with the maximum norm $||\cdot||$. Define

$$P_T = \{\phi \in BC, \phi(t+T) = \phi(t)\}.$$

Then P_T is a Banach space endowed with the maximum norm

$$\|x\| = \max_{t\in[0,T]} |x(t)|.$$

Also, we assume that

$$e^{\int_0^T a(s)ds} \neq 1. \tag{9.1.21}$$

Throughout this section, we assume that $a(t) \neq 0$ for all $t \in [0, T]$. Now (9.1.19) is equivalent to

$$\left[x(t)e^{-\int_0^t a(s)ds}\right]' = f(t, x(t))e^{-\int_0^t a(s)ds}.$$

Integrating this expression from $t-T$ to t and using the fact that $x(t-T) = x(t)$ give

$$x(t) = \left(1 - e^{\int_0^T a(s)ds}\right)^{-1} \int_{t-T}^{t} f(u, x(u))e^{\int_u^t a(s)ds} du. \tag{9.1.22}$$

Theorem 9.1.2. *Assume* (9.1.20) *and* (9.1.21). *Suppose the function f is Lipschitz continuous with Lipschitz constant k. If*

$$k\left|\left(1 - e^{\int_0^T a(s)ds}\right)^{-1}\right| \int_{t-T}^{t} e^{\int_u^t a(s)ds} du \leq \alpha$$

for $\alpha \in (0, 1)$, then (9.1.19) *has a unique periodic solution.*

Copyright Elsevier 2022

Proof. Define $\mathfrak{P}$ by the right side of (9.1.22). It is easily verified that $(\mathfrak{P}\phi)(t+T) = (\mathfrak{P}\phi)(t)$, and hence $\mathfrak{P} : P_T \to P_T$. The rest of the proof is a direct application of the contraction mapping principle on the set P_T. ☐

On the other hand, if we take a function $v(t)$ as before and assume that $v \in P_T$ with $v(t) \neq 0$ for all $t \in [0, T]$, then we obtain the following variation of parameters formula:

$$\begin{aligned} x(t) = \left(1 - e^{\int_0^T v(s)ds}\right)^{-1} \int_{t-T}^{t} \Big[& x(u)v(u) + x(u)a(u) \\ & + f(u, x(u)) e^{\int_u^t a(s)ds} du \Big]. \end{aligned} \tag{9.1.23}$$

Theorem 9.1.3. *Suppose $v(t) \neq 0$ for all $t \in [0, T]$ and the function f is Lipschitz continuous with Lipschitz constant k. If*

$$\left| \left(1 - e^{\int_0^T v(s)ds}\right)^{-1} \right| \int_{t-T}^{t} [|a(r)| + |v(r)| + k] \left| e^{\int_u^t a(s)ds} du \right| \leq \alpha$$

for $\alpha \in (0, 1)$, then (9.1.19) *has a unique periodic solution.*

Proof. The proof is similar to that of Theorem 9.1.2. ☐

Note that (9.1.23) can now handle equations of the form

$$x'(t) = \cos(t)x(t) + f(t, x(t)).$$

We easily see that

$$1 - e^{\int_0^T a(s)ds} = 1 - e^{\int_0^{2\pi} \cos(s)ds} = 0,$$

and hence (9.1.22) cannot be used.

9.2 Applications to delay differential equations

We begin by considering the following totally delayed nonlinear differential equation:

$$x'(t) = a(t)l(x_r), \tag{9.2.1}$$

where $x_r(t) = x(t - r)$ for $r > 0$ constant, and $l = l(t)$ is a continuous function satisfying some conditions to be imposed later. Let $\psi : [-r, 0] \to \mathbb{R}$ be a continuous initial function. We rewrite (9.2.1) in the form

$$x'(t) = a(t + r)l(x) - \frac{d}{dt} \int_{t-r}^{t} a(s + r)l(x(s))ds$$

Copyright Elsevier 2022

$$\begin{aligned}&= a(t+r)x(t) - a(t+r)[x(t) - l(x(t))]\\ &- \frac{d}{dt}\int_{t-r}^{t} a(s+r)l(x(s))ds. \qquad (9.2.2)\end{aligned}$$

Note that we added and subtracted $a(t+r)x$, so that the inversion is possible. Thus by the variation of parameters formula we obtain the integral

$$\begin{aligned}x(t) &= \psi(0)e^{\int_0^t a(s+r)ds}\\ &- \int_0^t e^{\int_s^t a(u+r)du} a(s+r)[x(s) - l(x(s))]ds\\ &- \int_0^t e^{\int_s^t a(u+r)du} \frac{d}{ds}\int_{s-r}^{s} a(u+r)l(x(u))du\, ds. \qquad (9.2.3)\end{aligned}$$

The appearance of the term $x(s) - l(x(s))$ in (9.2.3) is a direct consequence of the borrowed term. In addition, to get any meaningful results, we would have to assume that $l(x(t))$ is odd. Otherwise, $x(s) - l(x(s))$ will not define a contraction. Next, we invert our way by assuming $\nu : [0, \infty) \to \mathbb{R}$ to be a nonnegative continuous function such that $0 < \int_0^\infty \nu(s)\,ds = m < \infty$. To solve (9.2.1), we apply an inversion technique that starts by multiplying both sides by $e^{\int_0^t \nu}$ and then integrates them. We simplify the notation here and drop the ds at the end of the exponent:

$$\int_0^t e^{\int_0^s \nu} x'(s)ds = x(t)e^{\int_0^t \nu} - x(0) - \int_0^t x\nu e^{\int_0^s \nu}\, ds.$$

Therefore (9.2.1) becomes

$$x(t)e^{\int_0^t \nu} = x(0) + \int_0^t x\nu e^{\int_0^s \nu}\, ds + \int_0^t e^{\int_0^s \nu} a(s)l(x_r)\, ds,$$

$$x(t) = x(0)e^{-\int_0^t \nu} + \int_0^t x\nu e^{-\int_s^t \nu}\, ds + \int_0^t e^{-\int_s^t \nu} a(s)l(x_r)\, ds. \qquad (9.2.4)$$

By imposing adequate conditions on $\nu(t)$, $a(t)$, and $l(t)$, we will prove the existence of bounded solutions to (9.2.1). To this end, we will use the right-hand side to define a contraction map on a complete metric space $\mathscr{S}$. The resulting unique fixed point will be the solution we are looking for. Let $K > 0$ be a constant and fix an initial continuous function $\Psi : [-r, 0] \to \mathbb{R}$ with $|\Psi(t)| < K$ for $t \in [-r, 0]$ and $|\Psi(0)| > 0$. Define the space

$$\mathscr{S} := \{x : [-r, \infty) \to \mathbb{R} \mid x \in \mathscr{C}^1[-r, \infty),\ \|x\|_\infty \le K,\ x \equiv \Psi \text{ on } [-r, 0]\}, \qquad (9.2.5)$$

where $\mathscr{C}^1[-r, \infty)$ denotes the space of continuously differentiable functions, and $\|\cdot\|_\infty$ is the supremum norm. By general principles, it follows that $\mathscr{S}$ is a complete

Copyright Elsevier 2022

metric space. Define the map $\mathfrak{P}$ on $\mathscr{S}$ by

$$\mathfrak{P}(x) := x(0)e^{-\int_0^t \nu} + \int_0^t x\nu e^{-\int_s^t \nu}\, ds + \int_0^t e^{-\int_s^t \nu} a(s) l(x_r)\, ds. \tag{9.2.6}$$

The following lemma can be adapted to several different situations.

Lemma 9.1. *Let* $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$ *be a triple of positive numbers such that* $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \leq 1 - \delta$ *for fixed* $0 < \delta < 1$. *Assume that* $|\Psi(0)| \leq \varepsilon_1 K$ *and* m *is small enough such that* $|1 - e^{-m}| < \varepsilon_2$. *Assume that* $l(y)$ *is Lipschitz on* $[-K, K]$ *and satisfies* $|l(y)| \leq C_l |y|$ *for* $y \in [-K, K]$ *and for some positive constant* C_l. *Suppose that* $a \in L^1([0, \infty))$ *and* $\int_0^\infty |a(s)|\, ds \leq \frac{\varepsilon_3}{C_l}$. *Then the map* $\mathfrak{P}$ *has range in* $\mathscr{S}$ *and is a contraction. This implies that* (9.2.1) *has a unique solution in* $\mathscr{S}$.

Proof. We use the hypotheses of the theorem to bound each of the summands on the right-hand side of (9.2.6). The first term satisfies the bound

$$\left| x(0)e^{-\int_0^t \nu} \right| \leq \varepsilon_1 K,$$

since $|e^{-\int_s^t \nu}| \leq 1$. The second summand can be bounded by

$$\left| \int_0^t x\nu e^{-\int_s^t \nu}\, ds \right| \leq \|x\|_\infty \int_0^t \nu e^{-\int_s^t \nu}\, ds \leq (1 - e^{-m})\|x\|_\infty \leq \varepsilon_2 K.$$

The conditions stated above allow us to bound the third summand by

$$\left| \int_0^t e^{-\int_s^t \nu} a(s) l(x_r)\, ds \right| \leq \varepsilon_3 \|x\|_\infty \leq \varepsilon_3 K.$$

Therefore

$$|\mathfrak{P}(x)| \leq K\varepsilon_1 + K\varepsilon_2 + K\varepsilon_3 \leq (1 - \delta)K. \tag{9.2.7}$$

This implies $\|\mathfrak{P}(x)|\|_\infty \leq (1 - \delta)K$ so that $\mathfrak{P}$ has the range in $\mathscr{S}$.
The contraction part follows similarly:

$$\begin{aligned} |\mathfrak{P}(x) - \mathfrak{P}(y)| &\leq |x - y| \left(\int_0^t \nu e^{-\int_s^t \nu} ds + C_l \int_0^t |a(s)|\, ds \right) \\ &\leq |x - y| \left(\varepsilon_2 + C_l \frac{\varepsilon_3}{C_l} \right) \\ &\leq (1 - \delta)|x - y|. \end{aligned} \tag{9.2.8}$$

Therefore $\|\mathfrak{P}(x) - \mathfrak{P}(y)\|_\infty \leq (1 - \delta)\|x - y\|_\infty$, and $\mathfrak{P} : \mathscr{S} \to \mathscr{S}$ is a contraction. ☐

Now we consider a slightly more general equation

$$x'(t) = a(t) g(x_r) + b(t) G(x_r), \tag{9.2.9}$$

Copyright Elsevier 2022

where a, b, g, G are all continuous on $[-r, \infty)$. Let us denote the right-hand side of (9.2.9) by $L(x_r)$. On the same space $\mathscr{S}$ as before, we define the new map

$$\hat{\mathfrak{P}}(x) := x(0)e^{-\int_0^t \nu} + \int_0^t x\nu e^{-\int_s^t \nu}\,ds + \int_0^t e^{-\int_s^t \nu} L(x_r)\,ds. \tag{9.2.10}$$

Impose the same conditions on ν as in Lemma 9.1. The only difference now is the third term. We impose the following conditions:

C1- The functions g and G are Lipschitz on $[-K, K]$. Explicitly, there are positive constants C_g and C_G such that $|g(y)| \le C_g|y|$ and $|G(y)| \le C_G|y|$ for all $y \in [-K, K]$.

C2- $a, b \in L^1([0, \infty))$, $\int_0^\infty |a(s)|\,ds \le \frac{\varepsilon_3}{2C_g}$, and $\int_0^\infty |b(s)|\,ds \le \frac{\varepsilon_3}{2C_G}$.

Lemma 9.2. *Using the notation and definitions from Lemma 9.1, under conditions C1–C2, the map $\hat{\mathfrak{P}} : \mathscr{S} \to \mathscr{S}$ is a contraction.*

Proof. The only difference is the third summand in the definition of $\hat{\mathfrak{P}}$. We bound it as follows:

$$\left|\int_0^t e^{-\int_s^t \nu} L(x_r)\,ds\right| \le K\left(C_g \frac{\varepsilon_3}{2C_g} + C_G \frac{\varepsilon_3}{2C_G}\right) = K\varepsilon_3,$$

and the rest follows similarly to the proof of Lemma 9.1. □

As a consequence, we find that (9.2.9) has a unique bounded solution in $\mathscr{S}$.

9.2.1 The main inversion

Now we present a more general way of inverting (9.2.9), which starts by rewriting it as

$$\begin{aligned} x'(t) &= (a+b)g(x_r) + b(G(x_r) - g(x_r)) \\ &= -\frac{d}{dt}\int_{t-r}^t c(p+r)g(x(p))\,dp + c(t+r)g(x(t)) + bl(x_r), \end{aligned} \tag{9.2.11}$$

where $c = a + b$ and $l = G - g$. Now multiply both sides by $e^{\int_0^t \nu}$ and integrate as before to solve for $x(t)$. We get

$$\begin{aligned} x(t) =& x(0)e^{-\int_0^t \nu} + \int_0^t x\nu e^{-\int_s^t \nu} ds \\ &- \int_0^t e^{-\int_s^t \nu}\left(\frac{d}{ds}\int_{s-r}^s c(p+r)g(x(p))\,dp\right) ds \\ &+ \int_0^t c(s+r)g(x)e^{-\int_s^t \nu} ds + \int_0^t e^{-\int_s^t \nu} bl(x_r) ds. \end{aligned} \tag{9.2.12}$$

Copyright Elsevier 2022

Apply integration by parts in the middle line to get

$$\begin{aligned} x(t) = x(0)e^{-\int_0^t \nu} &+ \int_0^t x\nu e^{-\int_s^t \nu} ds \\ &- F(t) + e^{-\int_0^t \nu} F(0) + \int_0^t \nu e^{-\int_s^t \nu} F(s)\, ds \\ &+ \int_0^t c(s+r)g(x)e^{-\int_s^t \nu} ds + \int_0^t e^{-\int_s^t \nu} bl(x_r) ds, \end{aligned} \tag{9.2.13}$$

where $F(t) = \int_{t-r}^t c(p+r)g(x(p))dp$.

Again, we consider the same space $\mathscr{S}$ and define the new map

$$\begin{aligned} \mathfrak{P}(x) = x(0)e^{-\int_0^t \nu} &+ \int_0^t x\nu e^{-\int_s^t \nu} ds \\ &- F(t) + e^{-\int_0^t \nu} F(0) + \int_0^t \nu e^{-\int_s^t \nu} F(s)\, ds \\ &+ \int_0^t c(s+r)g(x)e^{-\int_s^t \nu} ds + \int_0^t e^{-\int_s^t \nu} bl(x_r) ds. \end{aligned} \tag{9.2.14}$$

We need to impose a different set of conditions on the summands on the right-hand side of (9.2.14) to get a contraction mapping of $\mathscr{S}$. First, label the terms on the right-hand side as 1–7. The new bounds on the absolute values the terms now are as follows:

D1- The first two summands are bounded exactly as before, that is, $x(0) = \Psi(0) \le \varepsilon_1 K$ and $\|\nu\|_{L^1} = m > 0$, so that $1 - e^{-m} \le \varepsilon_2$.

D2- To bound $F(t)$, assume that $\int_0^\infty |c|\, dt \le \int_0^\infty |a|\, dt + \int_0^\infty |b|\, dt \le \frac{\varepsilon_3}{C_g + C_G}$. Here C_g and C_G are the Lipschitz constants of the delayed functions. The third and fourth terms are each bounded above by $\varepsilon_3 K$.

D3- The definitions above produce an upper bound for the fifth term equal to $m\varepsilon_3 K$, where $m = \int_0^\infty |\nu| dt$.

D4- The sixth term is again bounded above by $\varepsilon_3 K$.

D5- The upper bound on b above implies that the seventh term is also bounded above by $\varepsilon_3 K$.

Therefore

$$|\mathfrak{P}(x)| \le K(\varepsilon_1 + \varepsilon_2 + (m+4)\varepsilon_3). \tag{9.2.15}$$

Theorem 9.2.1. *If the bounds D1–D5 above hold and there is* $0 < \delta < 1$ *such that* $\varepsilon_1 + \varepsilon_2 + (m+4)\varepsilon_3 \le 1 - \delta$, *then the map* $\mathfrak{P} : \mathscr{S} \to \mathscr{S}$ *is a contraction with a unique fixed point. This fixed point is a solution to* (9.2.9) *that belongs to* $\mathscr{S}$.

Example 9.1. Let $\alpha, \beta > 0$ be such that

$$\dot{x} = e^{-\alpha t} g(x_r) + e^{-\beta t} G(x_r), \tag{9.2.16}$$

Copyright Elsevier 2022

where α and β are big enough so that $\alpha^{-1}+\beta^{-1} \leq \frac{\varepsilon_3}{C_g+C_G}$. This equation satisfies the previous conditions. Eq. (9.2.16) is a first-order differential equation with exponentially damped time-delayed terms. Theorem 9.2.1 implies the existence of a solution to (9.2.16) in $\mathscr{S}$.

9.2.2 Variable time delay

Here we consider $r = r(t)$ variable. We impose conditions on it later on. We want to find solutions to

$$x'(t) = a(t)g(x(t-r(t))) + b(t)G(x(t-r(t))). \tag{9.2.17}$$

We still denote $x(t-r(t))$ by x_r. The same strategy as before gives

$$\begin{aligned} x'(t) &= (a+b)g(x_r) + b(G(x_r) - g(x_r)) \\ &= c(t)g(x_r) + bl(x_r) \\ &= \frac{cg(x_r)}{1-r'}(1-r') + bl(x_r). \end{aligned} \tag{9.2.18}$$

Denote $f = \frac{c}{1-r'}$ and $f_r = f(t-r(t))$. Now we rewrite (9.2.18) as follows:

$$\begin{aligned} x' &= f_r g(x_r)(1-r') + (f(t) - f_r)g(x_r)(1-r') + bl(x_r) \\ &= \big(f_r g(x_r)(1-r') - f(t)g(x(t))\big) + f(t)g(x(t)) \\ &\quad + (f(t) - f_r)g(x_r)(1-r') + bl(x_r) \\ &= -\Big(\frac{d}{dt}\int_{t-r(t)}^{t} f(s)g(x(s))ds\Big) + f(t)g(x(t)) \\ &\quad + (f(t) - f_r)g(x_r)(1-r') + bl(x_r). \end{aligned} \tag{9.2.19}$$

To simplify the upcoming expression, assume that $r(0) = 0$. The inversion gives

$$\begin{aligned} x(t) &= x(0)e^{-\int_0^t \nu} + \int_0^t x\nu e^{-\int_s^t \nu} ds \\ &\quad - \int_{t-r(t)}^{t} f(p)g(x(p))dp + \int_0^t \nu e^{-\int_s^t \nu} \int_{s-r(s)}^{s} f(p)g(x(p))dpds \\ &\quad + \int_0^t e^{-\int_s^t \nu}\big(f(s)g(x(s)) + (f(s) - f_r)g(x_r)(1-\dot{r}) + bl(x_r)\big)ds. \end{aligned} \tag{9.2.20}$$

The most important condition to be imposed on r is $|r'(t)| \leq \kappa < 1$ for some κ small and positive. This condition allows us to get uniform pointwise upper bounds on the function $f = \frac{c}{1-r'}$. From the definition of $\mathscr{S}$ we impose that $r(t) \leq r$ for all $t \geq 0$. Now we impose bounds on the different terms.

E1- The first two summands are bounded exactly as before, that is, $x(0) = \Psi(0) \leq \varepsilon_1 K$ and $\|\nu\|_{L^1} = m > 0$, so that $1 - e^{-m} \leq \varepsilon_2$.

Copyright Elsevier 2022

E2- To bound $f(t)$, assume that

$$\int_{-r}^{\infty} |f|\,dt \le (1-\kappa)^{-1} \int_{-r}^{\infty} |c|\,dt \le (1-\kappa)^{-1} \left(\int_{0}^{\infty} |a|\,dt + \int_{0}^{\infty} |b|\,dt \right)$$
$$\le \frac{\varepsilon_3}{(1-\kappa)(C_g + C_G)}.$$

This suffices to bound all the terms in the last line as well.

Theorem 9.2.2. *Define the map* $\mathfrak{P}$ *on* $\mathscr{S}$ *by the right-hand side of* (9.2.20). *Suppose there are constants* $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$ *such that bounds E1–E2 above are satisfied and such that* $\varepsilon_1 + \varepsilon_2 + \frac{2\kappa+4}{1-\kappa}\varepsilon_3 \le 1 - \delta$ *for some* $0 < \delta < 1$. *Then* $\mathfrak{P} : \mathscr{S} \to \mathscr{S}$ *is a contraction. This implies that* (9.2.17) *has a unique solution in* $\mathscr{S}$.

9.3 Exercises

Exercise 9.1. Give the details for the proof of Theorem 9.1.3.

Exercise 9.2. Use our new inversion to show the existence of periodic solutions to

$$x'(t) = a(t)g(x) + b(t)G(t, x_L), \tag{9.3.1}$$

where $x_L = x(t-L)$, and $a(t), b(t)$ are continuous periodic functions with period $L > 0$. We assume that $g = g(t,x) = g(t+L,x)$ and $G = G(t,x) = G(t+L,x)$. We also assume that g and G are Lipschitz on $[0, L]$ (and extended periodically) with constants C_g and C_G, respectively.

Exercise 9.3. Consider the equation

$$x'(t) = -a(t)x(t) + c(t)x'(t-g(t)) + q(t, x(t-g(t))). \tag{9.3.2}$$

For $T > 0$, define $P_T = \{\phi \in C(\mathbb{R},\mathbb{R}), \phi(t+T) = \phi(t)\}$, where $C(\mathbb{R},\mathbb{R})$ is the space of all real-valued continuous functions. Then P_T is a Banach space endowed with the supremum norm

$$\|x\| = \max_{t\in[0,T]} |x(t)| = \max_{t\in\mathbb{R}} |x(t)|.$$

Assume that

$$a(t+T) = a(t), \quad c(t+T) = c(t), \quad g(t+T) = g(t), \quad g(t) \ge g^* > 0 \tag{9.3.3}$$

with continuously differentiable $c(t)$, twice continuously differentiable $g(t)$, and constant g^*. We suppose that $q(t,x)$ is continuous in both arguments and periodic in t. Usually, $a(t)$ is assumed to be positive, but here we only assume that

$$\int_0^T a(s)ds > 0.$$

Copyright Elsevier 2022

Choose a positive function $\nu(t) = \nu(t+T)$. Multiply both sides of (9.3.2) by $e^{\int_0^t \nu}$ and integrate by parts between $t-T$ and t.
(i) Obtain a variation of parameters formula.
(ii) Use the results of part (i) to define a mapping $\mathfrak{P}$ on $P_T = \{\phi \in C(\mathbb{R}, \mathbb{R}), \phi(t+T) = \phi(t)\}$ and show that $\mathfrak{P} : P_T \to P_T$.
(iii) Impose additional conditions and prove theorems parallels to Theorems 8.7.2 and 8.7.3.

Copyright Elsevier 2022

Bibliography

[1] M. Adivar, H. Koyuncuoğlu, Y. Raffoul, Existence of periodic and nonperiodic solutions of systems of nonlinear Volterra difference equations with Infinite delay, Journal of Difference Equations and Applications 19 (12) (December 2013) 1927–1939.

[2] M. Adivar, Y. Raffoul, Stability, Periodicity and Boundedness in Functional Dynamical Systems on Time Scales, Springer Nature, Switzerland AG, 2020.

[3] M. Adivar, Y. Raffoul, Existence of resolvent for Volterra integral equations on time scales, Bulletin of the Australian Mathematical Society 82 (2010) 139–155.

[4] M. Adivar, Y. Raffoul, Qualitative analysis of nonlinear Volterra integral equations on time scales using resolvent and Lyapunov functionals, Applied Mathematics and Computation 273 (2016) 258–266.

[5] M. Adivar, M. Islam, Y. Raffoul, Separate contraction and existence of periodic solutions in totally nonlinear delay differential equations, Hacettepe Journal of Mathematics and Statistics 41 (1) (2012) 1–13.

[6] R. Agarwal, Difference Equations and Inequalities. Theory, Methods and Applications, second ed., Monographs and Textbooks in Pure and Applied Mathematics, Marcel Decker, Inc., New York, 2000.

[7] R. Agarwal, D. O'Regan, Infinite Interval Problems for Differential, Difference, and Integral Equations, Kluwer Academic, The Netherlands, 2001.

[8] S. Alsahafi, Y. Raffoul, A. Sanbo, Qualitative analysis of solutions in Volterra nonlinear systems of difference equations, International Journal of Mathematical Analysis 8 (31) (2014) 1505–1515.

[9] S. Banach, Théorie des Opération Linéairs (reprint of the 1932 ed.), Chelsea, New York, 1929.

[10] R. Bellman, Stability Theory of Differential Equations, McGraw-Hill Book Company, New York, London, 1953.

[11] L. Berezansky, E. Braverman, Exponential stability of difference equations with several delays: recursive approach, Advances in Difference Equations 2009 (2009) 104310, 13 pages.

[12] M. Bohner, A. Peterson, Dynamic Equations on Time Scales, An Introduction with Applications, Birkhäuser, Boston, 2001.

[13] H. Brunner, P.J. Van der Houwen, The Numerical Solution of Volterra Equations, SIAM, Philadelphia, 1985.

[14] T.A. Burton, Volterra Integral and Differential Equations, Academic Press, New York, 1983.

[15] T.A. Burton, Stability and Periodic Solutions of Ordinary and Functional Differential Equations, Academic Press, New York, 1985.

[16] T.A. Burton, Stability by Fixed Point Theory for Functional Differential Equations, Dover, New York, 2006.

[17] F. Brauer, J.A. Nohel, Qualitative Theory of Ordinary Differential Equations, Dover, New York, 1969.

[18] E.A. Coddington, N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill Book Company, New York, London, 1955.

Copyright Elsevier 2022

[19] J.M. Cushing, Integro-differential Equations and Delay Models in Population Dynamics, Lecture Notes in Biomathematics, vol. 20, Springer, Berlin, New York, 1977.

[20] R.D. Driver, Introduction to Ordinary Differential Equations, Harper & Row, New York, 1978.

[21] G. Eid, B. Ghalayani, Y. Raffoul, Lyapunov functional and stability in nonlinear finite delay Volterra discrete systems, International Journal of Difference Equations 10 (1) (2015) 77–90.

[22] S.E. Elaydi, An Introduction to Difference Equations, second ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1999.

[23] P. Eloe, M. Islam, Stability properties and integrability of the resolvent of linear Volterra equations, Tohoku Mathematical Journal 47 (1995) 263–269.

[24] J. Hale, H. Kocak, Dyanmics and Bifurcations, Springer-Verlag, New York, 1991.

[25] J. Hale, S. Verduyn Lunel, Introduction to Functional Differential Equations, Springer-Verlag, New York, 1993.

[26] P. Hartman, Ordinary Differential Equations, John Wiley & Sons, Inc., New York, 1964.

[27] P. Holmes, A Nonlinear oscillator with a strange attractor, Philosophical Transactions of the Royal Society of London A 292 (1979) 419–448.

[28] M. Islam, Y. Raffoul, Stability properties of linear Volterra integro-differential equations with nonlinear perturbations, Communications in Applied Analysis 7 (3) (2003) 405–416.

[29] J.L. Goldberg, A.J. Schwartz, Systems of Ordinary Differential Equations: An Introduction, Harper's Series in Mathematics, Harper & Row, 1972.

[30] W. Kelley, A. Peterson, Difference Equations an Introduction with Applications, Academic Press, San Diego, CA, 2000.

[31] W. Kelley, A. Peterson, The Theory of Differential Equations, Classical and Qualitative, Pearson Prentice Hall, New York, 2004.

[32] H. Lauwerier, Mathematical Models of Epidemics, Mathematisch Centrum, Amsterdam, Netherlands, 1981.

[33] H.J. Liu, A First Course in the Qualitative Theory of Differential Equations, Prentice Hall, Upper Saddle River, New Jersey, 2003.

[34] J. Marsden, M. McCracken, The Hopf Bifurcation and Its Applications, Springer-Verlag, New York, 1976.

[35] R.K. Miller, Nonlinear Volterra Integral Equations, Benjamin, New York, 1971.

[36] L. Perko, Differential Equations and Dynamical Systems, Springer-Verlag, New York, 1991.

[37] Y. Raffoul, Qualitative Theory of Volterra Difference Equations, Springer Nature Switzerland AG, 2018.

[38] Y. Raffoul, T-periodic solutions and a priori bound, Mathematical and Computer Modelling 32 (2000) 643–652.

[39] Y. Raffoul, Positive periodic solutions of nonlinear functional difference equations, Electronic Journal of Differential Equations 55 (2002) 1–8.

[40] Y. Raffoul, General theorems for stability and boundedness for nonlinear functional discrete systems, Journal of Mathematical Analysis and Applications 279 (2003) 639–650.

[41] Y. Raffoul, Uniform asymptotic stability in linear Volterra systems with nonlinear perturbation, International Journal of Differential Equations 6 (1) (2002) 19–28.

[42] Y. Raffoul, Periodicity in general delay nonlinear difference equations using fixed point theory, Journal of Difference Equations and Applications 10 (2004) 1229–1242.

[43] Y. Raffoul, Stability and periodicity in discrete delay equations, Journal of Mathematical Analysis and Applications 324 (2006) 1356–1362.

Copyright Elsevier 2022

[44] Y. Raffoul, Periodic solutions in neutral nonlinear differential equations with functional delay, Electronic Journal of Differential Equations and Applications 2003 (102) (2003) 1–7.

[45] Y. Raffoul, Periodicity in nonlinear systems with infinite delay, Advances of Dynamical Systems and Applications 3 (1) (2008) 185–194.

[46] Y. Raffoul, Boundedness in nonlinear differential equations, Nonlinear Studies 10 (4) (2003) 343–350.

[47] Y. Raffoul, Stability in neutral nonlinear differential equations with functional delays using fixed point theory, Mathematical and Computer Modelling 40 (2004) 691–700.

[48] Y. Raffoul, Boundedness in nonlinear functional differential equations with applications to Volterra Integro-Differential Equations, Journal of Integral Equations and Applications 16 (4) (2004).

[49] Y. Raffoul, Stability in functional difference equations using fixed point theory, Communications of the Korean Mathematical Society 29 (1) (2014) 195–204.

[50] Y. Raffoul, Existence of positive periodic solutions in neutral nonlinear equations with functional delay, Rocky Mountain Journal of Mathematics 42 (6) (2012) 1983–1993.

[51] P.A. Robinson, C.J. Rennie, J.J. Wright, Propagation and stability of waves of electrical activity in the cerebral cortex, Physical Review E 56 (1997) 826–840.

[52] L. Roger, A Nonstandard discretization method for Lotka-Volterra models that preserves periodic solutions, Journal of Difference Equations and Applications 11 (8) (2005) 721–733.

[53] Y. Saito, M. Ma, T. Hara, A necessary and sufficient condition for permanence of a Lotka-Volterra discrete system with delays, Journal of Mathematical Analysis and Applications 256 (2001) 162–174.

[54] D.A. Sanchez, Ordinary Differential Equations and Stability Theory: An Introduction, W. H. Freeman and Company, New York, 1968.

[55] H. Schaefer, Uber die Method der a priori Scranken, Mathematische Annalen 129 (1955) 45–416.

[56] M. Scheffer, Fish nutrient interplay determines algal biomass: a minimal model, Oikos 62 (1991) 271–282.

[57] M.F. Scud, Vito Volterra and theoretical ecology, Theoretical Population Biology 2 (1971) 1–23.

[58] D.R. Smart, Fixed Point Theorems, Cambridge University Press, London, 1980.

[59] M.J. Smith, M.B. Wisten, A continuous day-to-day traffic assignment model and the existence of a continuous dynamic user equilibrium, Annals of Operations Research 60 (1) (1995) 59–79, https://doi.org/10.1007/BF02031940.

[60] Y. Song, C. Baker, Qualitative behavior of numerical approximations to Volterra integro-differential equations, Journal of Computational and Applied Mathematics 172 (2004) 101–115.

[61] X.H. Tang, X. Zou, Global attractivity of non-autonomous Lotka-Volterra competition system without instantaneous negative feedback, Journal of Differential Equations 192 (2003) 502–535.

[62] T. Taniguchi, Asymptotic behavior of solutions of non-autonomous difference equations, Journal of Mathematical Analysis and Applications 184 (2006) 342–347.

[63] P. Yang, R. Xu, Global attractivity of the periodic Lotka-Volterra system, Journal of Mathematical Analysis and Applications 233 (1999) 221–232.

[64] X. Yang, Uniform persistence and periodic solutions for a discrete predator-prey system with delays, Journal of Mathematical Analysis and Applications 316 (2006) 161–177.

Copyright Elsevier 2022

[65] E. Yankson, Stability in discrete equations with variable delays, Electronic Journal on the Qualitative Theory of Differential Equations (8) (2009) 1–7.

[66] E. Yankson, Stability of Volterra difference delay equations, Electronic Journal on the Qualitative Theory of Differential Equations 20 (2006) 1–14.

[67] W. Yin, Eigenvalue problems for functional differential equations, Journal of Nonlinear Differential Equations 3 (1997) 74–82.

[68] X. Wen, Global attractivity of positive solution of multispecies ecological competition-predator delay system (in Chinese), Acta Mathematica Sinica 45 (1) (2002) 83–92.

[69] J. Wiener, Differential equations with piecewise constant delays, in: Trends in Theory and Practice of Nonlinear Differential Equations: Proceedings of International Conference, University of Texas, Arlington, 1982, in: Lecture Notes in Pure and Applied Mathematics, vol. 90, Dekker, New York, 1984, pp. 547–552.

[70] K. Yosida, Lectures on Differential and Integral Equations, Interscience Publishers, New York, London, Sydney, 1960.

Copyright Elsevier 2022

Index

Symbols

α-exponentially asymptotically stable, 273
ω-limit set, 260, 261

A

A priori bound, 344
Antiderivative, 2
Ascoli–Arzelà, 11
Asymptotic stability, 120
Asymptotically stable, 107, 142, 156, 170, 194, 236–238, 241–243, 249, 253, 262, 264, 265, 268, 270, 279, 280, 294, 309, 327, 332
Attractor, 159
Autonomous systems, 139, 238, 261

B

Banach space, 10, 13–17, 43, 323
Bendixson–Dulac criteria, 182
Bifurcation, 161
Blue-sky bifurcation, 162
Bounded, 133, 134, 273
Boxcars, 91
Buckling rod, 167

C

Cauchy sequence, 10
Cauchy–Peano theorem, 36
Center manifold, 211, 213, 214
Closed orbit, 182
Compact, 10, 11, 15–17
Complete metric, 15–18, 331
Complex roots, 144
Conserved, 147
Constant solution, 105, 249
Continuation of solutions, 45
Contraction, 17, 18
Contraction mapping principle, 10, 330
Convex, 323

D

Decrescent, 251
Delay differential equations, 287
Dependence on initial conditions, 48
Differentiable manifold, 204
Displacement, 90, 92
Dissipated, 147
Distinct eigenvalues, 80
Distributed delay, 289
Domain of attraction, 242

E

Eigenpair, 78
Eigenvalues, 77, 267
Eigenvectors, 77
Energy function, 70, 71, 145
Epidemic, 8
Equicontinuous, 10–12, 15, 16, 323
Equilibrium solution, 105, 249
Error estimation, 67
Euler's buckling beam, 196
Existence and uniqueness, 27, 66, 292
Existence on Banach spaces, 41
Exponential matrix, 84, 85
Exponential stability, 273

F

Finite dimensional, 287
Fixed point, 17, 18, 344, 345
Floquet exponents, 115
Floquet multipliers, 114
Floquet theorem, 118
Floquet theory, 111
Focal point, 145
Fold bifurcation, 162
Functional, 294
Functional delay, 289, 308
Fundamental matrix, 71, 75, 98
Fundamental set of solutions, 74
Fundamental solutions, 73

G

Generalization of Bendixson–Dulac theorem, 183
Generalized eigenvector, 81
Generalized Gronwall's inequality, 54
Global asymptotic stability, 257
Global Lipschitz, 58
Global Lipschitz condition, 27, 63
Global manifolds, 210
Globally asymptotically stable, 159, 256, 263, 280, 281
Globally identically stable, 264
Gradient function, 153
Gradient system, 145, 153
Gronwall's inequality, 39

Copyright Elsevier 2022

H
Hamiltonian function, 147
Hamiltonian system, 145, 147
Harmonic motion, 235
Hartman–Grobman Theorem, 198, 218
Hill’s equation, 128
Homeomorphism, 204, 218
Hopf bifurcation, 194
Hyperbolic, 198

I
Implicit function theorem, 165
Improper node, 144
Infinite delay, 289
Infinite delay integro-differential equations, 318
Infinite dimensional, 287
Initial function, 287
Instability, 259
Integral equation, 29
Integro-differential equations, 315, 319
Intrinsic, 115
Invariant set, 261

J
Jacobian, 169
Jordan canonical form, 86

K
Kinetic energy, 71, 148
Krasnoselskii, 324
Krasovskii–LaSalle invariance principle, 260

L
Law of mass, 8
Leaky bucket, 5
Left-hand derivative, 46
Level curve, 147
Limit cycle, 182
Linearly dependent, 72
Linearly independent, 72
Linearly independent eigenvectors, 78
Lipschitz, 11
Local Lipschitz condition, 27
Locally asymptotically stable, 159
Logarithmic Matrix, 117
Lorenz system, 285
Lotka–Volterra, 105, 174
Lyapunov function, 70, 231, 235, 249, 251, 253, 254, 267, 277, 281, 282, 294
Lyapunov method, 231

M
Mass matrix, 90, 92
Mass-spring system, 91
Mathieu’s equation, 122
Matrix norm, 65
Maximal interval, 46, 47, 62, 70, 99
Method of steps, 289–291, 326
Multiple delays, 288

N
Near constant systems, 133
Negative definite, 235, 261
Negative definite matrix, 267
Negative semi-definite, 261
Neutral differential equations, 300
Neutral Volterra integro-differential equations, 306
New variation of parameters, 329
Newtonian system, 148
Non-autonomous, 136, 155, 249, 259, 273
Non-homogeneous differential equation, 18, 44
Norm, 14
Normal form, 167
Nullclines, 175

O
Onto, 204
Open ball, 238
Open covering, 204
Orbit, 140
Order of a differential equation, 1
Ordinary differential equations, 1

P
Pendulum with friction, 68
Periodic solutions, 335
Perturbed, 137
Picard’s theorem, 29
Picard’s theorem for systems, 59
Picard’s theorem global version, 41
Picard’s theorem local version, 43
Pitchfork bifurcation, 163
Poincaré–Bendixson, 184
Positive definite, 235, 251, 261
Positive definite matrix, 267
Positive periodic solutions, 319
Positive semi-definite, 261
Potential energy, 71, 148
Predator–prey delay model, 291
Predator–prey model, 105, 227, 279
Principal matrix, 75, 86
Proper node, 143
Pseudo-periodic, 124
Putzer’s algorithm, 87, 131

R
Radially bounded, 257
Radially unbounded, 251, 264

Copyright Elsevier 2022

Ratio test, 83
RC circuit, 20
Real and distinct roots, 143
Real and equal roots, 143
Reduced system, 213, 214
Repeller, 159
Representative, 167
Resolvent, 343
Right-hand derivative, 46, 288

S
Saddle, 143
Saddle node bifurcation, 162
SI epidemic model, 8
SIR epidemic model, 177
Solution, 8, 28
Solution space, 71
Spectrum, 129
Spiral, 141
Spring, 235
Stable, 107, 109, 121, 133, 142, 155, 236, 237, 239, 249, 251, 294
Stable limit cycle, 193, 247, 278
Stable manifold, 198, 200
Stable node, 143
Stiffness matrix, 90, 92
Strict Lyapunov function, 235, 251
Strong maximum, 152
Strong minimum, 151
Subcritical bifurcation, 164
Sun flower model, 293
Supercritical bifurcation, 164

T
Time-varying delay, 289
Trajectory, 140
Transcritical bifurcation, 162
Turning point, 162

U
Unbounded, 134
Uniform asymptotic stability, 110
Uniform exponential stability, 273
Uniform metric, 331
Uniform stability, 236
Uniformly asymptotically stable, 107, 135, 136, 138, 171, 237, 254, 256, 271, 282, 294, 299, 327
Uniformly bounded, 8, 10, 11, 15, 16, 121, 273
Uniformly exponentially stable, 282, 284
Uniformly stable, 107, 109, 110, 134, 137, 225, 237, 252, 255, 272, 281, 294, 295, 327
Unstable, 107, 111, 121, 134, 142, 237, 244, 248, 249, 259, 260
Unstable manifold, 200
Unstable node, 143

V
Variation of parameters, 18, 19, 92
Volterra integrodifferential equation, 289

W
Wedge, 251
Weierstrass M-test, 12
Well-posed, 287

Copyright Elsevier 2022